ACSIJ

WWW.ACSIJ.ORG

Advances in Computer Science: an International Journal

Vol. 5, Issue 1, January 2016

ISSN : 2322-5157

ACSIJ Reviewers Committee 2016

- Prof. José Santos Reyes, Faculty of Computer Science, University of A Coruña, Spain
- Dr. Dariusz Jacek Jakóbczak, Technical University of Koszalin, Poland
- Dr. Artis Mednis, Cyber-Physical Systems Laboratory Institute of Electronics and Computer Science, Latvia
- Dr. Heinz DOBLER, University of Applied Sciences Upper Austria, Austria
- Dr. Nguyen Phuoc Loc, Sunflower Soft Company, Ho Chi Minh city, Vietnam
- Dr. Ivan Simecek, Czech Technical University in Prague, Department of Computer Systems, Czech Republic
- Dr. Dimitra Anastasiou, Luxembourg Institute of Science and Technology, Luxembourg
- Dr. Yajie Miao, Carnegie Mellon University, United States
- Dr. Wei Huang, Rensselaer Polytechnic Institute, Troy, NY, United States
- Dr. Sugam Sharma, Center for Survey Statistics and Methodology, Iowa State University, Ames, Iowa, United States
- Dr. Ahlem Nabli, Faculty of sciences of Sfax,Tunisia
- Prof. Zhong Ji, School of Electronic Information Engineering, Tianjin University, Tianjin, China
- Prof. Noura AKNIN, Abdelmalek Essaadi University, Morocco
- Dr. Qiang Zhu, Geosciences Dept., Stony Brook University, United States
- Dr. Urmila Shrawankar, G. H. Raisoni College of Engineering, Nagpur, India
- Dr. Uchechukwu Awada, Network and Cloud Computing Laboratory, School of Computer Science and Technology, Dalian University of Technology, China
- Dr. Seyyed Hossein Erfani, Department of Computer Engineering, Islamic Azad University, Science and Research branch, Tehran, Iran
- Dr. Nazir Ahmad Suhail, School of Computer Science and Information Technology, Kampala University, Uganda
- Dr. Fateme Ghomanjani, Department of Mathematics, Ferdowsi University Of Mashhad, Iran
- Dr. Islam Abdul-Azeem Fouad, Biomedical Technology Department, College of applied Medical Sciences, SALMAN BIN ABDUL-AZIZ University, K.S.A
- Dr. Zaki Brahmi, Department of Computer Science, University of Sousse, Tunisia
- Dr. Mohammad Abu Omar, Information Systems, Limkokwing University of Creative Technology, Malaysia
- Dr. Kishori Mohan Konwar, Department of Microbiology and Immunology, University of British Columbia, Canada
- Dr. S.Senthilkumar, School of Computing Science and Engineering, VIT-University, INDIA
- Dr. Elham Andaroodi, School of Architecture, University of Tehran, Iran
- Dr. Shervan Fekri Ershad, Artificial intelligence, Amin University of Isfahan, Iran
- Dr. G.UMARANI SRIKANTH, S.A.ENGINEERING COLLEGE, ANNA UNIVERSTIY, CHENNAI, India
- Dr. Senlin Liang, Department of Computer Science, Stony Brook University, USA
- Dr. Ehsan Mohebi, Department of Science, Information Technology and Engineering, University of Ballarat, Australia
- Sr. Mehdi Bahrami, EECS Department, University of California, Merced, USA
- Dr. Sandeep Reddivari, Department of Computer Science and Engineering, Mississippi State University, USA
- Dr. Chaker Bechir Jebari, Computer Science and information technology, College of Science, University of Tunis, Tunisia
- Dr. Javed Anjum Sheikh, Assistant Professor and Associate Director, Faculty of Computing and IT, University of Gujrat, Pakistan
- Dr. ANANDAKUMAR.H, PSG College of Technology (Anna University of Technology), India
- Dr. Ajit Kumar Shrivastava, TRUBA Institute of Engg. & I.T, Bhopal, RGPV University, India

ACSIJ Published Papers are Indexed By:

Google Scholar
EZB, Electronic Journals Library (University Library of Regensburg, Germany)
DOAJ, Directory of Open Access Journals
Bielefeld University Library - BASE (Germany)
Academia.edu (San Francisco, CA)
Research Bible (Tokyo, Japan)
Academic Journals Database
Technical University of Applied Sciences (TH - WILDAU Germany)
AcademicKeys
WorldCat (OCLC)
TIB - German National Library of Science and Technology
The University of Hong Kong Libraries
Science Gate
OAJI Open Academic Journals Index. (Russian Federation)
Harvester Systems University of Ruhuna
J. Paul Leonard Library _ San Francisco State University
OALib _ Open Access Library
Université Joseph Fourier _ France
CIVILICA (Iran)
CiteSeerX _ Pennsylvania State University (United States)
The Collection of Computer Science Bibliographies (Germany)
Indiana University (Indiana, United States)
Tsinghua University Library (Beijing, China)
Cite Factor
OAA _ Open Access Articles (Singapore)
Index Copernicus International (Poland)
Scribd
QOAM _ Radboud University Nijmegen (Nijmegen, Netherlands)
Bibliothekssystem Universität Hamburg
The National Science Library, Chinese Academy of Sciences (NSLC)
Universia Holding (Spania)
Technical University of Denmark (Denmark)
JournalTOCs School of Mathematical and Computer Sciences Heriot-Watt University Edinburgh(UK)

TABLE OF CONTENTS

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

C_3 Effective features inspired from Ventral and dorsal stream of visual cortex for view independent face recognition

Somayeh Saraf Esmaili [1], Keivan Maghooli [2] and Ali Motie Nasrabadi [3]

[1] Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
s.saraf@srbiau.ac.ir

[2] Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
k_maghooli@srbiau.ac.ir

[3] Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran.
nasrabadi@shahed.ac.ir

Abstract

This paper presents a model for view independent recognition using features biologically inspired from dorsal and ventral stream of visual cortex. The presented model is based on the C_3 features inspired from Ventral stream and Itti's visual attention model inspired from the dorsal stream of visual cortex. The C_3 features, which are based on the higher layer of the HMAX of the ventral stream of visual cortex, are modified to extract important features from faces in various viewpoints. By Itti's visual attention model, visual attention points are detected from faces in various views of faces. Effective features are extracted from these visual attention points and the view independent C_3 effective features (C_3EFs) are created from faces. These C_3EFs are used to distinct multi-classes of different subjects in various views of faces. The presented model is tested using FERET face datasets with faces in various views of faces, and compared with C_2 features (C_2SMFs) and C_3 features of standard HMAX model (C_3SMFs). The results illustrated that our presented view independent face recognition model has high accuracy and speed in comparison with standard model features, and can recognize faces in various views by 97% accuracy.

Keywords: *view independent face recognition, HMAX model, Itti's visual attention model, C_3 effective features, ventral stream, dorsal stream.*

1. Introduction

In recent years, providing computational algorithms for face recognition in these different situations especially in various views of faces have been the fundamental issues [1-3]. Computational view independent face recognition is one challenging work that human visual system can do it with high-speed performance, easily. Thus based on need, recently the study of brain mechanisms of the human visual system have been more considered.

The visual system is organized in two functionally specialized processing pathways in the visual cortex. One pathway (from the primary visual cortex to the parietal cortex for controlling eye movements and visual attention) is named dorsal stream, and other pathway (from the primary visual cortex towards the inferior temporal lobe including V_1, V_2, V_4, Posterior Infer temporal (PIT), Anterior Infer temporal (AIT) and Fusiform Face Area (FFA)) is named ventral stream, which processes detail of objects and faces in different conditions [4-6]. The dorsal and ventral streams are not completely independent and there are interactions. For example, area V_4 is interconnected with some visual attention areas in dorsal stream [7-9].

Partial analogies of the ventral stream cortex have been used in many computing models in canonical computer vision. Among these models, HMAX is one powerful computational model that models the object recognition mechanism of human ventral visual stream in visual cortex [10-13], that first proposed by Poggio et al., is based on experimental results in neurobiology. The extracted bio-inspired C_2 features of this model can recognize and classify objects based on the mechanism of human ventral visual stream in visual cortex. Then Serre et al. established HMAX model and considered learning ability of the model for real-world object recognition in [13]. The features extracted by this model are C_2 standard model features (SMFs). The C_2SMFs of the ventral stream have been used for multi-class face recognition [14-16]. In recent years, different development models of the ventral stream HMAX model have been presented to enhance the efficiency of the model and in all these models, some feature extraction methods are considered [17-20]. Leibo et al. in [20], extended HMAX model and added new S_3 and C_3 layers. They found that the performance of the model on view independent within-category identification tasks on different objects was increased and the C_3 features of the extended HMAX model performed significantly

better than C_2 and it was independent to viewpoints. Also, there is some visual attention model which based on visual attention regions in the dorsal stream of the visual cortex and these models were applied in many applications such as target detection, object recognition, object segmentation and robotic localization [21-24]. These visual attention models use low-level visual features such as colour, intensity and orientation to form saliency maps and find focus of attention locations. The basic computational model of visual attention was proposed by Itti in 1998 which are the basic model for the most of the new models and used bottom-up visual cortex features in the where path [22].

In this paper, a model based on new C_3EFs inspired from Ventral and dorsal stream of visual cortex for the goal of view independent face recognition are presented. For experimental analysis, the FERET faces datasets in various views of faces are utilized and view independent face recognition task by the SVM classifiers on the new C_3EFs of them are done and they are compared with the results of other extracted features.

The rest of the paper is organized as follows. Section 2 illustrates the C_2SMFs and C_3SMFs features of ventral stream model and, also the visual attention model of dorsal stream for detecting important facial regions. Our proposed view independent face recognition model is presented in section 3. Section 4 is presented experimental evaluation including dataset description, results and detailed discussions. In the final section, we sum up with a conclusion.

2. Material and methods

2.1 The C_2SMFs features of ventral stream model

The C_2SMFs features of HMAX model are inspired by ventral stream of visual cortex, and it created by four layers (S_1, C_1, S_2, C_2) [12, 13]. S_1 features resemble the simple cells found in the V_1 area of the primate visual cortex and consists of Gabor filters. C_1 features imitate the complex cells in V_1&V_2 area of cortex and have the same number of feature types (orientations) as S_1. These features pools nearby S_1 features (of the same orientation) to reach the position and scale invariance over larger local regions, and as a result can also subsample S_1 to reduce the number of features. The values of C_1 features are the value of the maximum S_1 features (of that orientation) that comes within a max filter [13]. S_1 features imitate the visual area V_4 and posterior infer temporal (PIT) cortex. They contain RBF-like units which tuned to object-parts and compute a function of the distance between the input C_1 patches and the stored prototypes. In human visual system, these patches correspond to learning patterns of previously seen visual images and store in the synaptic weights of the neural cells. The S_2 features learn from the training set of K patches ($P = 1,..., K$) with various $n \times n$ sizes ($n \times n$ = 4×4, 8×8, 12×12 and 16×16) and all four orientations at random positions (Thus a patch P of size $n \times n$ contains $n \times n \times 4$ elements). Then S_2 features, acting as Gaussian RBF-units, compute the similarity scores (i.e., Euclidean distance) between an input pattern X and the stored prototype P : $f(X) = \exp(-\|X - P\|^2 / 2\sigma^2)$, with σ chosen proportional to patch size. The C_2 features imitate the inferotemporal cortex (IT) and perform a max operation over the whole visual field and provide the intermediate encoding of the stimulus. Thus, for each face image, the C_2 features vector is computed and used for face recognition. This vector has robustness properties. The lengths of C_2 features vector are equal to the number of random patches extracted from the images and have the property of shift and scale independent.

2.2 The C_3SMFs features of ventral stream model

In the implementation of the S_3 and C_3 layers from developed HMAX model which was proposed by Leibo et al. in [20], the response of a C_2 cell (associating templates w at each position t) was given by Eq. (1):

$$S_3 = \exp\left(-\frac{1}{2\sigma}\sum_{j=1}^{n}(w_{t,j} - x_j)^2\right) \qquad (1)$$

Then, the S_3 features corresponding to all layers were extracted and the maximum of these values were utilized as (Eq. (2)). The C_3 features imitate the view independent properties in the FFA of the IT [20]. These original features are named C_3SMFs in this paper.

$$C_3 = \max(S_3^{\,i}) \qquad (2)$$

In this paper, we used the C_2SMFs and C_3SMFs features vector for view independent face recognition and present a model which can be extracted effectively C_2 and C_3 features from face image in different viewpoints.

2.3 Visual attention model in dorsal stream of visual cortex

In face images, some facial regions are more attentive and helpful regions to face recognition, such as eyes, nose and mouth that have been demonstrated by the results of psychophysical studies in paper [25]. Human can find distinctive information from face images in a short time and with high accuracy. These detected features have so much local information to recognize the similarity of face images and can track and match to the similar face images. So, visual attention models inspired from human visual

systems in visual attention regions of dorsal stream can detect salient points from face images. Visual attention Itti's model biologically models the visual attention regions in the posterior cortex and specifies the locations of salient points from a colour image simulating saccadic eye movements of human vision [22]. We utilized this model in our proposed model to find automatically important regions of the face by adjusting its parameters.

3. Proposed view independent face recognition model

In the ventral stream HMAX model, the S_2 features learned from the randomly extracted patches. So, maybe some of the extracted special features from cropped face images such as extracted features from the forehead and cheek regions are not useful features. Since, these features caused CPU usage in the system and make the system very slow achieving the effective features vector. Then, it seems necessary to detect best features. Also, HMAX model only models the ventral stream and the connections between the visual attention regions in the posterior cortex to the ventral stream are not considered. In the proposed model, in order to model the human-like face recognition system, we extract the C_2 and C_3 features from visual attention points and achieve the C_2EFs and C_3EFs for view independent face recognition.

Generally, by combining the hierarchical ventral stream features with visual attention model, the feature extraction model for face recognition system is proposed as follows (Fig. 1 illustrates different visual cortex layers in two dorsal and ventral streams, and also the whole structure of our proposed feature extraction model for view independent face recognition inspired from them):

1) For each face image, the colour features, intensity features and orientation features are extracted (inspired from the regions in the primary visual cortex, which showed by red dashed line ▬ in Fig. 1) as represented in Itti's Visual attention model [22].

$$R = r - \frac{g+b}{2}$$
$$G = r - \frac{r+b}{2}$$
$$B = b - \frac{r+g}{2}$$
$$Y = \frac{g+b}{2} - b - \frac{|r-g|}{2}$$
$$I = \frac{R+G+B}{3} \quad (3)$$

2) colour saliency map, intensity saliency map and orientation saliency map are founded as Eq. (4)-(9). $M_f(s)$ represents of F feature map in s scale. F included I intensity features, C colour features. A function of $N(0)$ is used for created normalization map where the symbol Θ represents interpolation of the coarser image to the finer scale and point by point subtraction. In this system, $c = \{2,3\}$ and $s = c + d$, where $d = \{2,3\}$.

$$F_{I,c,s} = N(|M_I(c) \Theta M_I(s)|)$$
$$F_{C,c,s} = N(|M_C(c) \Theta M_C(s)|)$$
$$= N(|M_{RG}(c) \Theta M_{RG}(c)| + |M_{BY}(s) \Theta M_{BY}(s)|) \quad (4)$$

I and C are the saliency maps of intensity and colour, respectively (in Eq. (5)).

$$I_{c,s} = \sum_{c=2}^{3} \sum_{s=c+2}^{3} N(|M_I(c) \Theta M_I(s)|)$$
$$C_{c,s} = \sum_{c=2}^{3} \sum_{s=c+2}^{3} N(|M_{RG}(c) \Theta M_{BY}(s)|) \quad (5)$$

Also, the Gabor filters which generated in S_1 are used as orientation features where $O(c,s,\theta)$ denotes the orientation saliency map of θ by c and s operation of scales (Eq. (6)).

$$O(c,s,\theta) = |O(c,\theta) \Theta O(s,\theta)|$$
$$= \sum_{\theta=\{0^\circ,45^\circ,90^\circ,135^\circ\}} \sum_{c=2}^{3} \sum_{s=c+2}^{3} N(O(c,s,\theta)) \quad (6)$$

Then the features are combined by Eq. (7) to create salient points (SP) [22, 24].

$$SP = (C + I + O)/3 \quad (7)$$

Then a winner take all network (WTN) is used to detect N salient points and N attention points (inspired from the regions in the dorsal stream of the visual cortex, which showed by blue dotted line ▬ ▪ ▬ in Fig. 1).

3) Create S_1 and C_1 features from each face image (inspired from the regions in the primary visual cortex, which showed by red dashed line ▬ in Fig. 1).

4) Extract N patches p_i $(i = 1,..., N)$ in four orientations and $n \times n$ best patch sizes from C_1 features of each face image by using detected attention points as the central pixel of them to create effective S_2 features. During recognition, from each test face image, N patches are created as X_i $(i = 1,..., N)$ patches and the distance between the patches p_i and X_i are calculated according to the Eq. (8).

$$V_k = \exp\left(\frac{-\left\|X_i - P_i\right\|^2}{2\sigma^2}\right) \quad k = 1,2,..., N \quad (8)$$

Which σ is proportional to the patch size and N dimensional vectors create for each face image. The set of V_k ($k = 1,2,\ldots,N$) forms S_2EFs (inspired from the V4 and

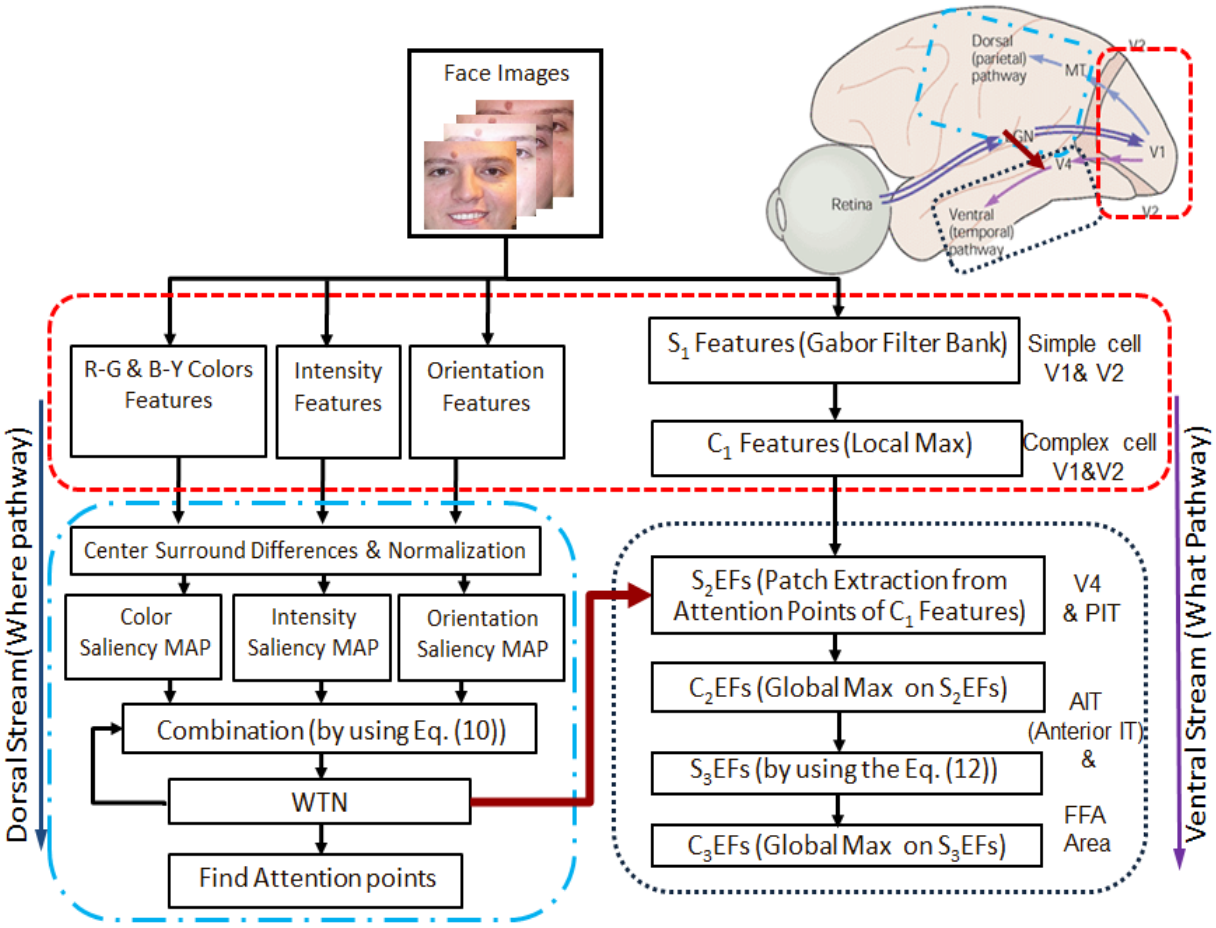


Fig 1 The structure of proposed view independent face recognition model.

PIT area in the ventral stream of the visual cortex, which showed by dark blue dotted line ▪▪▪, also the interaction between the V_4 area in ventral stream by the dorsal stream of the visual cortex has been showed by red arrows → in Fig. 1).

5) Obtain C_2EFs by general maximum on S_2EFs to create N -dimensional C_2EFs vector from distinctive regions of faces (inspired from AIT area in ventral stream visual cortex which showed by dark blue dotted dot ▪▪▪ in Fig. 1).

6) Create S_3EFs Features on C_2EFs (the response of a C_2EFs cell) by using the Eq. (9) to extract S_3EFs corresponding to all layers (inspired from the FFA area in the ventral stream visual cortex which showed by dark blue dotted dot ▪▪▪ in Fig. 1). During recognition, from each test face image, C_2EFs of them are created as x_i $(i = 1,..., N)$ responses and the distance between these responses and associating templates w at each position t are calculated according to the Eq. (9).

$$S_3 EFs = \exp(-\frac{1}{2\sigma}\sum_{j=1}^{n}(w_{t,j} - x_j)^2) \qquad (9)$$

7) Obtain C_3 EFs by using the maximum on the S_3EFs as following (Eq. (10)) (inspired from the view independent regions in the FFA area of the ventral stream visual cortex which showed by dark blue dotted dot - - - in Fig. 1).

$$C_3 EFs = \max(S_3^{\,i} EFs) \qquad (10)$$

8) Do step 1 to 6 for all images to extract C_3EFs vector from distinctive regions of faces.

9) Feed C_3EFs and C_2EFs vectors to SVM and classify face images with the goal of view independent face recognition.

4. Experimental Analysis

4.1 Image Dataset

To demonstrate the feasibility of our proposed face recognition system, experiments on the subset of the colour FERET database are organized [26]. The subset contains ten classes (unique subjects) of face images, with variations in pose, expressions and scales. It consist 200 face images (10 classes, each with 20 images). The proposed model is tested using a 10-fold cross validation strategy. So that, for each fold 18 images of each individual (180 face images) are selected as training samples and the rest 2 images of each individual as test images (20 face images). In the pre-processing method, all images are cropped manually to remove complex background and then they are resized to the dimensions of 140×140. Fig. 2 shows a series of 10 different classes of people from FERET database which used in this work. All experiments are performed entirely in a 32 Matlab2012 experimental environment (characteristic of a computer system is Intel core 2 duo processor (2.66 GHz) and 4 GB RAM). Fig. 3 shows twenty samples of images from one subject in varied view points and expression.

Fig. 2 Images from ten classes of cropped FERET dataset.

Fig. 3 Images of one subject in varied viewpoints and expressions.

4.2 Results and discussion

In proposing a view independent face recognition model, to achieve the effective features vector, at first the saliency maps of colour, intensity and orientations are extracted and the feature map's weights of them are adjusted to create the saliency maps from face images and select the attended locations of salient points as important regions of them. Fig. 4, shows the original images of faces with seventy attention points and saliency maps on them. Fig. 4(a-c) and also Fig. 4(d-f) shows the original face images of two individuals in three situations of viewpoints. In the saliency toolbox, local max is selected for normalization. As shown in Fig. 4, by visual attention model, important and important regions of faces such as nose, eyes, lip and mole are selected as attention points however the faces are in varied viewpoints.

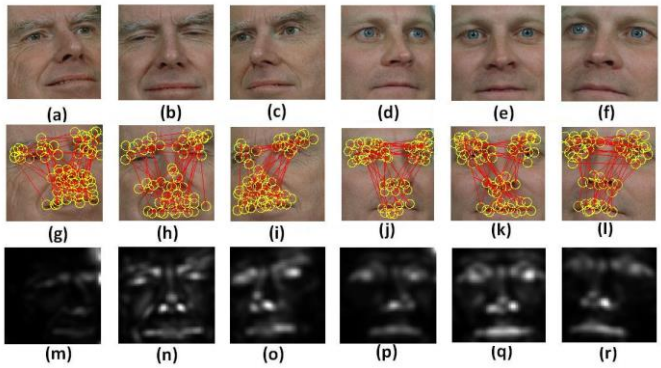


Fig. 4 (a) To (f) are original images. (g) To (l) are face images with selected sixty attention points. (m) To (k) are saliency maps of face images.

After finding attention points from face images, it is necessary to convert colour images to grey scale for extracting C_2 and C_3 features. The biological S_1 features of face images are created by convolving with 64 Gabor filters. So, there are 64 S_1 features for each face image. Fig. 5 shows these S_1 features for one original grey-scale face image after convolving with 64 Gabor filter. Also, Fig. 6 shows the C_1 features in band 1 and in four orientations ($\theta = 0^\circ$, 45°, 90° and 135°). For each orientation in band 1, there are two 7 × 7 and 9 × 9 filters. At first, Maximum responses to them are calculated for each pixel with 8 × 8 grid sizes. Then, the C_1 features are created by maximum on corresponding pixels from two images in each orientation.

The C_2EFs and C_3EFs of the images are extracted using the proposed view independent face recognition model that presented in last section. Then, SVM classifiers with RBF kernel are trained using these C_2 features vectors and the class labels, implemented using LIBSVM [27]. In this approach, an SVM is constructed for each class by

discriminating that class against the remaining 9 classes. The number of SVMs used in this approach is 10. In the testing phase, the features are extracted using proposed method and the classification is done on test data using the test SVM classifiers.

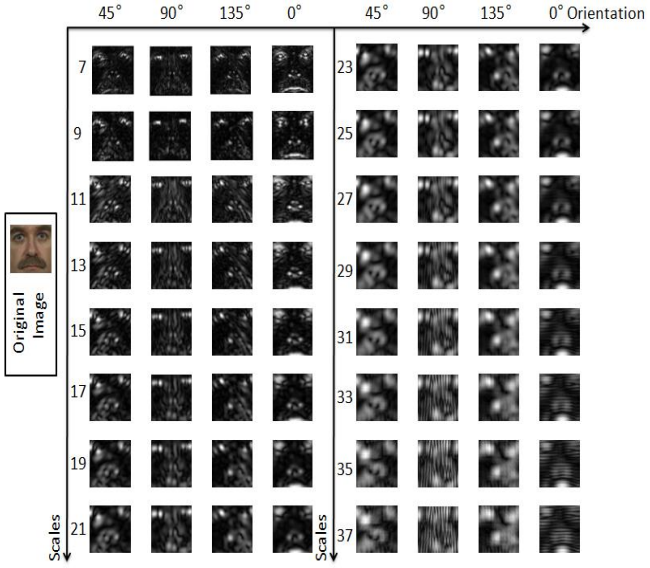


Fig. 5 S_1 features created by 64 Gabor filters.

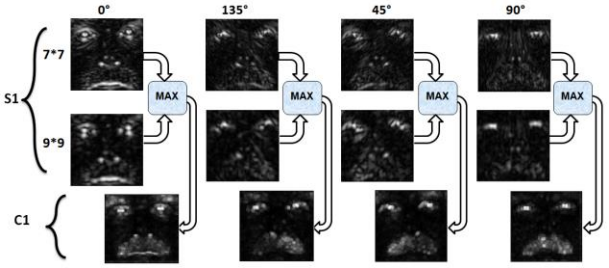


Fig. 6 C_1 features created in band 1 and in four orientations.

In order to find proper sizes of patches for extracting features, the representation of dissimilarity matrices (RDMs) are created. For obtaining this goal, we extract proper prototype patches with different patch sizes (4×4, 8×8, 12×12 and 16×16) and attain the best biological features. These features create the RDMs which have view independent identity specific between features of eight viewpoints of ten subjects. Scheme to extract hierarchical features from 10 subjects in 8 viewpoints and create RDMs is shown in Fig. 7, and Fig. 8 shows a comparison of the created RDMs in different patch sizes of C_3EFs in equal extracted features (N=40). The best RDMs can be created with the extraction prototype patches in 12×12 patch size from each of the cropped face images (Fig. 8(c)). It shows created RDMs on C_3EFs features in 12×12 patch size have high correlation in 15 diagonals parallel to the main diagonal that represent the view-independent specific of C_3EFs in 12×12 patch size in comparison with others patch sizes. As in other patch sizes (Figure 8(a), Figure 8(b) and Figure 8(d)), some pixels of these diagonals are not observable on the same background. In Figure 8 each pixel in vertical and horizontal of RDMs represents one subject in one viewpoint.

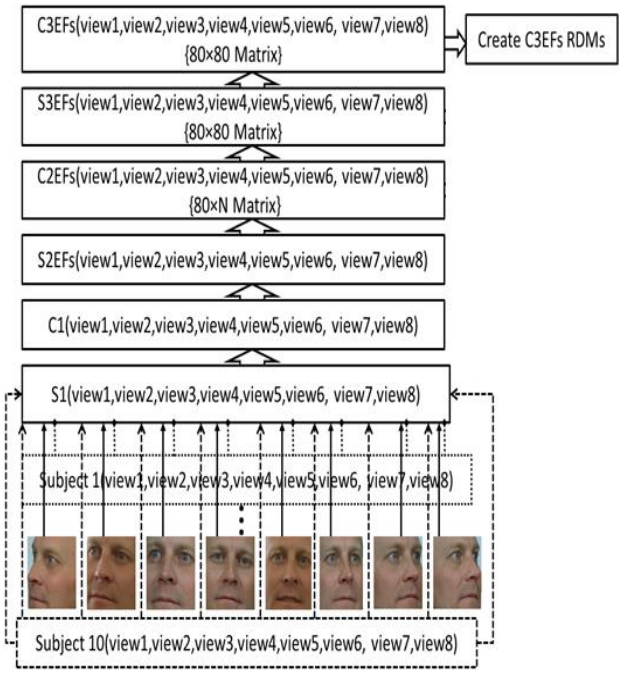


Fig. 7 Scheme to extract hierarchical features and create RDMs.

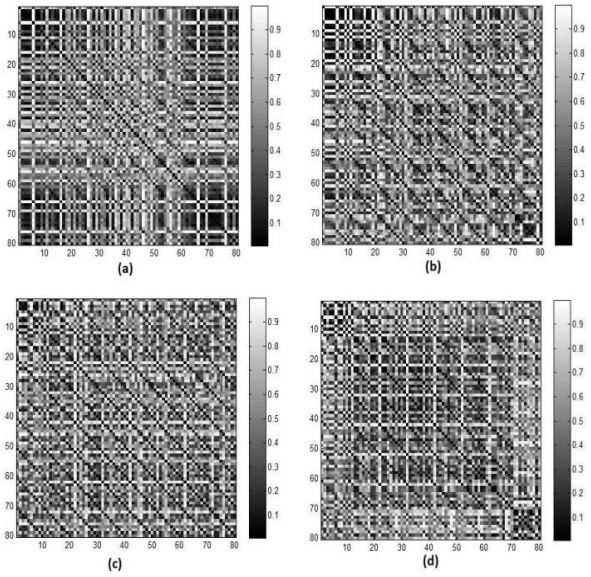


Fig. 8 RDMs on C_3EFs of 80 face images (8 viewpoints of 10 subjects) in (a) 4×4 patch size (b) 8×8 patch size (c) 12×12 patch size (d) 16×16 patch size.

Also, in order to demonstrate feasibility of the proposed view independent face recognition model, we compared the performance of the C_2EFs vector and C_3EFs vector (with selected attention points) of the proposed model in best patch size with the performances of the C_2SMFs and C_3SMFs (without attention points). So, *N* patches of C_2SMFs, C_3SMFs, C_2EFs and C_3EFs in best patch size from the same face of test and train images are extracted and fed into SVM classifiers to compare the recognition rate of them by ten-fold cross validation. Fig. 9 shows the accuracy rate of face recognition using SVM classifiers on the C_2SMFs, C_2EFs and the C_3EFs of proposed model in various numbers of extracting features. Given the results in Table 1 and Fig. 9, the recognition rates of C_3EFs and C_2EFs are better than C_2SMFs and C_2SMFs. Also, the results in table 1 show that 80 number (*N*=80) of the C_3EFs are enough to get good performance (around 97%) against; 300 number of the C_3SMFs are required to get this good face recognition rate. So, face recognition by C_2SMFs and C_3SMFs needs more extracted features and more extracting time.

The extracting time is the average computing time of each face image for extracting features that obtained by tic-toc function in Matlab2012 software. For example, the extracting time of C_3EFs is total computing time for selecting attention points and extracting C_3 features from them. Given by the results in the table 1, the extracting time to enhance the good recognition accuracy near 97% by using C_3SMFs (300 number of features) are more than others. So, by using C_2EFs and C_3EFs, the appropriate view independent face recognition rate in lesser time is achieved.

From these results, the C_2EFs of the proposed model showed a marginal improvement over the C_2SMFs on FERET face database which mainly deals with variances in viewpoints. The advantages of C_2 features intolerance to variations and the advantages of visual attention model to select the attention points are complementary to each other and mutually enhancing the C_2EFs to view independent face recognition.

Table 1: Accuracy rates of face recognition using SVM classifiers on extracted different features.

Features	**Recognition accuracy (mean ± standard deviation)%**	**Extracting time (Sec)**
C_2EFs (N=80)	94 ± 2.108	5.863
C_3EFs (N=80)	97± 2.852	6.052
C_2SMFs (N=80)	89±3.944	4.592
C_3SMFs (N=80)	91±3.162	4.923
C_2SMFs (N=300)	93.5 ± 2.838	9.624
C_3SMFs (N=300)	97 ± 3.162	10.357

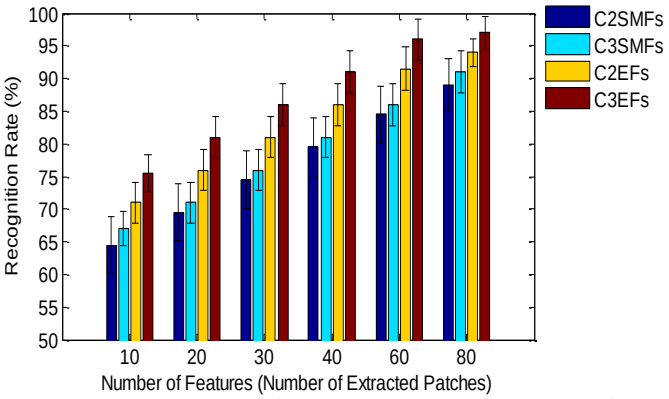


Fig. 9 The comparisons plot of computed recognition accuracy rate of C_3EFs, C_2EFs, C_2SMFs and C_2SMFs in different number of extracted features.

5. Conclusions

In this paper, we described a biologically-motivated framework for view independent face recognition, which the proposed C_3 EFs was inspired from the ventral and dorsal stream of cortex. In fact, we proposed a model to extract new view independent features, using visual attention model and ventral stream model for the goal of view independent face recognition. By visual attention model, we specified the set of attention points from salient points on a gallery of face images in varied viewpoints and by ventral stream features, we extracted proper view independent features from the set of attention points and then ran SVM classifiers on the vectors of features obtained from the input images.

Through the experimental results, we proved that the C_3EFs by using attention points in comparison to the same number of C_2SMFs improved the face classification accuracy under varying facial expressions and viewpoints. Likewise, in the proposed model, face recognition needs lesser time since we do not only use feature selection method on C_2 features, but also upgrade recognition rate with the appropriate number of attention points which selected from important regions of the face.

In the future work, we will propose the model that can detect and recognize multi-classes of faces with complicated and varied backgrounds. Also in this study, we used only SVM classifier and did not examine other classifiers. Then, in future we research directions for classification changes, the use of artificial neural network and fuzzy classification to enhance the efficiency measure.

Acknowledgments

Support of Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran is gratefully acknowledged. This study was extracted from Ph.D. thesis that was done in Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

References

[1] Z. Fan, and B. Lu, "Multi-View Face Recognition with Min-Max Modular Support Vector Machines", in Proceedings of the 1st International Conference on Advances in Natural Computation, Changsha, China, 2007, pp. 396-399.

[2] T.K. Kim, and J. Kittler, "Design and Fusion of Pose-Invariant Face-Identification Experts", IEEE Trans. Circuits and Systems for Video Technology, Vol. 16, No.9, 2006, pp. 1096–1106.

[3] K. R Singh., M. A. Zaveri, and M. M. Raghuwanshi M. M., "Illumination and Pose Invariant Face Recognition: A technical Review", IJCISIM, Vol. 2, 2010, pp. 029–038.

[4] R.H. Wurtz, and E.R. Kandel, Central visual pathways. In: E.R .Kandel, J.H. Schwartz, T.M. Jessell, editors. Principles of neural science. 4th ed. New York: McGraw-Hill, p. 523–547, 2000.

[5] E. Kobatake, and K. Tanaka, "Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex", Journal of Neurophysiology, Vol.71, No.3, 1994, pp. 856–867.

[6] Ungerleider L. G., J. V. Haxby, "‘‘What’’ and ‘‘where’’ in the human brain", Current Opinion in Neurobiology, Vol.4, No.2, pp. 157-165, 1994.

[7] L.G. Ungerleider, and L. Pessoa, "What and where pathways", Scholarpedia, Vol. 3, No.11, 2008, pp. 5342.

[8] J.M. Brown, "Visual streams and shifting attention", Progress in Brain Research, Vol. 176, 2009, pp. 47-63.

[9] R. Farivar, "Dorsal–ventral integration in object recognition", Brain Research Reviews, Vol.61, No.2, , 2009, pp. 144–153.

[10] M. Riesenhuber, and T. Poggio, "Hierarchical Models of Object Recognition in Cortex", nature neuroscience, Vol. 2, No.11, 1999, pp. 1019–1025.

[11] M. Riesenhuber, and T. Poggio, "Models of object recognition," nature neuroscience, Vol. 3, No. Suppl, 2000, pp. 1199–1204.

[12] T. Serre, and M. Kouh, C. Cadieu, U. Knoblich, G. Kreiman, and T. Poggio, A Theory of Object Recognition: Computations and Circuits in the Feedforward Path of the Ventral Stream in Primate Visual Cortex, Technical Report, MIT, Massa- chusetts, USA., 2005.

[13] T. Serre, and L. Wolf, S. Bileschi, S. Bileschi, M. Riesenhuber, "Robust object recognition with cortex-like mechanisms", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.29, No. 3, 2007, pp. 411–426.

[14] J. Lai, and W.X. Wang, "Face recognition using cortex mechanism and svm", in 1st international conference intelligent robotics and applications, Wuhan, 2008, pp. 625–632.

[15] E. Meyers, and L. Wolf, "Using Biologically Inspired Features for Face Processing", International Journal of Computer Vision, Vol. 76, 2008 ,pp. 93–104.

[16] P. Pramod Kumar, and P. Vadakkepat, "Hand posture and face recognition using a fuzzy-rough approach", International Journal of Humanoid Robotics, Vol. 7, No. 3, 2010, pp. 331–356.

[17] S. Dura-Bernal, T. Wennekers, and S. L. Denham, "Top-Down Feedback in an HMAX-Like Cortical Model of Object Perception Based on Hierarchical Bayesian Networks and Belief Propagation", PLoS ONE, Vol. 7, No. 11, 2012.

[18] J. Mutch, and D.G. Lowe, "Object Class Recognition and Localization Using Sparse Features with Limited Receptive Fields", International Journal of Computer Vision, Vol. 80, No. 1, 2008, pp. 45-57.

[19] C. Thériault, N. Thome, M. Cord, "Extended Coding and Pooling in the HMAX Model", IEEE Transactions on Image Processing, Vol. 22, No. 2, 2013, pp. 764 – 777.

[20] J. Z. Leibo, J. Mutch, and T. Poggio, "Why the Brain Separates Face Recognition from Object Recognition", in In Advances in Neural Information Processing Systems (NIPS), Granada, Spain, 2011.

[21] J. Han, and K.N. Ngan, "Unsupervised extraction of visual attention objects in color images", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 1, 2006, pp. 141-145.

[22] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual-attention for rapid scene analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, 1998, pp. 1254–1259.

[23] L. Itti, and C. Koch, "Computational modeling of visual attention", Nature Reviews Neuroscience, Vol. 2, No. 3, 2001, pp. 194-203.

[24] D. Walther, and C. Koch, "Modeling attention to salient proto-objects", Neural Networks, Vol. 19, No. 9, 2006, pp. 1395–1407.

[25] F. Kano, and M. Tomonaga, "Face scanning in chimpanzees and humans: Continuity and discontinuity", Animal Behaviour, Vol. 79, 2010, pp. 227.

[26] P.J. Phillips, H. Wechsler, J. Huang, P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms", Image and Vision Computing, Vol. 16, No. 5, 1998, pp. 295-306.

[27] C.C. Chang, C.J. Lin, LIBSVM: A library for support vector machines, Available at http://www.csie.ntu.edu.tw/cjlin/libsvm.

Somayeh Saraf Esmaili, was born in 1984. She received her B.S, M.S. and Ph.D degree in Bioelectric Engineering from Science and Research Branch of Islamic Azad University, Tehran, Iran in 2006, 2009 and 2015 respectively. Since 2011, she has been a lecturer in the Garmsar Branch of Islamic Azad University, Iran. Her current main research includes Bio-inspired Computing and Applications in face recognition.

Keivan Maghooli, is the Assistant Professor at Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. He was a Ph.D. Scholar at mentioned Department. He received M.S. degree in Biomedical Engineering from Tarbiat Modares University, Tehran, Iran. He has more than 10 years of experience, including teaching graduate and undergraduate classes. He also leads and teaches modules at B.Sc., M.Sc. and Ph.D. levels in Biomedical Engineering. His current research interests are pattern recognition, biometric and data mining.

Ali Motie Nasrabadi, received a B.S. degree in Electronic Engineering in 1994 and his M.S. and Ph.D. degrees in Biomedical Engineering in 1999 and 2004, respectively, from Amirkabir University of Technology, Tehran, Iran. Since 2005, he has been Associate Professor in the Biomedical Engineering Department at Shahed University, in Tehran, Iran. His current research interests are in the fields of biomedical signal processing, nonlinear time series analysis and evolutionary algorithms. Particular applications include: EEG signal processing in mental task activities, hypnosis, BCI, epileptic seizure prediction and visual attention models.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Providing a framework to improve the performance of business process management projects based on BPMN

Alireza Nikravanshalmani[1] and Mojtaba Ahmadi[2]

[1] Department of Computer Engineering, Karaj Branch, Islamic Azad University
Karaj, Alborz, Iran
Nikravan@kiau.ac.ir

[2] Department of Computer Engineering, Karaj Branch, Islamic Azad University
Karaj, Alborz, Iran
Ahmadi.lpi@gmail.com

Abstract

Modeling of business processes, based on Business Process Modeling Notation (BPMN), helps analysts and managers to understand business processes, and, identify their shortages. These models provide a context to make rational decision of organizing business processes activities in an understandable manner. The purpose of this paper is to provide a framework for better understanding of business processes and their problems by reducing the cognitive load of displayed information for their audience at different managerial levels while keeping the essential information which are needed by them. For this reason, we integrate business process diagrams across the different managerial levels to develop a framework to improve the performance of business process management (BPM) projects. This framework, which is referred to as "Business Process Improvement Framework Based on Managerial Levels (BPIML)" in this paper, considers three levels of management (Organizational level managers, Process /Departmental level managers and Activity level managers) for manager of an organization. Then, defines certain types of models based on BPMN, for each management level, by taking into the account the objectives and tasks of various managerial levels in organizations and their role in Business Process Management (BPM) projects. This framework will make us able to provide the necessary support for making decisions about business processes. The framework is evaluated with a case study in a real business process improvement project, to demonstrate its superiority over the conventional method. A questionnaire consisted of 10 questions using Likert scale was designed and given to the participants (three managerial levels). The results of this questionnaire suggested that, managers and senior experts of the organization considered utilization of the proposed framework improving for implementation of BPM projects and provide support for correct and timely decisions by increasing the clarity and transparency of the business processes which led to success in BPM projects.

Keywords: *Business Process Modeling Notation (BPMN), Business Process Management (BPM), Business Process Reengineering (BPR), business process optimizing*

1. Introduction

Business world is a space for competitiveness of organizations. Changing the market and customers, emergence of new competitors and changing the business rules of organizations, created the conditions that a need for a method and system for defining, managing, analyzing, and optimizing business process was coming into existence over time. Business process management provides an integrated approach to the definition, implementation and management of business processes of organizations and minimize the workload of information solutions development in organizations by using their own specific methods and tools. For modern economic enterprise, planning for changes is essential [1]. Continuous improvement and reengineering of business processes is a process that never ends. It is important to visualize the sequence of business process activities and their related information properly, or business process modeling in other words. The purpose of the modeling of business processes of an organization, is create a common conceptually and simple language (standard-based, in form of graphic shapes that have less volume and are also easily understandable by the user) between organization's managers, experts and analysts. Process modeling is an activity that used by analysts in all reengineering methodologies and strategies in order to extract current business processes and display new business processes. In this activity, analysts used modeling tools for model the current state (AS-IS) and the desired state (TO-BE) of organization[2]. Continuous improvement and management of business processes has become a critical strategy for organizations, due to the market competitive conditions and permanently changing of customer demands and technology [3]. Thus, there is a serious need to provide frameworks which support correctly making decisions by managers in business processes management projects.

The rest of this paper is organized as follows:

In the second part we will introduce Business Process Management (BPM), Business Process Reengineering (BPR) and BPMN. In the third part, we review related works and in the fourth part, the proposed framework will be introduced. Then, in the fifth part, we will evaluate and validate the proposed framework. Finally, the last section is the conclusions of this paper.

2. Literature review

Business Process Management (BPM), is an integrated approach to design, implementation and monitoring of business processes that staffs or softwares of organization may be involved in any part of these. Interactions between people, softwares and information flow of organization, gives life to it. The purpose of existence of BPM is management of organization's processes and provide tools to continuous improvement over time. BPM optimizes processes, efficiency and effectiveness of any organization with its business process automation [1][4].

Business process management, has been a development of workflow management, which began in the 90s and includes support for business processes with the use of methods, techniques and design software. BPM provides an opportunity for approval, monitoring and analysis of documentation and operational processes of staffs, organizations and applications [5].

The main purpose of business process management is to increase the ability of organizations in order to quickly respond to environment changes. Information technology plays a major role in support and control of today's business processes, and facilitates its management. Business process management has a long journey from analysis and design through implementation and deployment processes. BPM is a kind of t change and system deployment management that helps to continuous business process management [6][7].

Business Process Reengineering (BPR) means fundamental rethinking and redesigning of Business Processes in order to achieve significant improvements in critical measures of performance such as cost, quality, speed and services. BPR is an improvement philosophy which aims to achieve phased improvement by redesigning the process and in this redesign, the organization tries to maximize valuable activities and minimize other activities. This approach can be used at the single process level or entire the organization [8][9].

Business process modeling is one of the most basic steps in moving toward business process optimization. The purpose of business process modeling is documentation of activities implementation procedure in organizations. The existence of a common and standard language in order to business process modeling helps to maintain the effectiveness of this document in different time and place situation.

Business Process Model and Notation (BPMN) is a standard for defining business processes diagrams. Version 2.0 of BPMN that have introduced by OMG in 2011, includes a set of graphic shapes (geometric) that have been developed based on the flowchart[10].

The main feature of BPMN is the ability to convert it to executive languages that can be understood by the software systems. BPMN provides a chart, entitled "Business Process Diagram (BPD)" that has been used in order to design and management of the business process. BPD is actually a network of graphical objects that show the activities, control flow and how to arrange the implementation of activities [2][11].

3. Related Work

Continuous improvement and management of business processes has become a critical strategy for organizations, due to the market competitive conditions and permanently changing of customer demands and technology. Business processes improvement is challenging due to the complexity of make changes in processes. The organization may be get problems to choosing the starting point to begin changes as well as selecting the type of necessary reforms in the business process. The aim of complexity of business processes, include some matters such as: the dependencies between activities, business process stakeholder, the elements involved in the business process, business process characteristics, and finally the business process applications[12][13][14].

The nature of business process improvement and reengineering projects, impose extensive changes within organizations. These changes in the structure of the business processes in order to improve the performance of the organization. Implement such extensive changes require manager's serious and persistent support of BPM projects. The wrong decisions of managers about improvement of business processes, that mainly caused by incomplete and incorrect understanding of business processes and its problems is one of the important reason of business processes management and continuous improvement projects failure [11][15].

Research done on business processes improvement should provide an appropriate response to the concerns mentioned and propose a framework to support correct decisions in business processes improvement projects. We followed the theory of "pattern matching" for developing the proposed framework. Pattern matching includes an attempt to link a theoretical pattern and an operational one [16].

-Ref[11], proposed an alignment framework for business process management projects. They believed in order to overcoming on staffs resistance against business process continuous improvement projects, it is necessary to understand the existing organizational reality in order to chart a way for the creation of the new organizational reality aligned with business process reengineering. They believe that the successful BPM implementation need effective and structured participation of different levels managers in this project and therefore, they defined goals, measures and key tasks of all three levels as follows[11]:

3.1 Senior managers

Senior managers have been prepared and approved general policies, macro programs, long range and strategic plans of the organization and monitor and coordinate the all actions. Successful implementation of BPM in organizations requires effective participation of process and operational managers.

3.2 Process managers

Process managers are responsible for designing and monitoring of business processes and organizational structure based on strategic vision of the organization and existing resources and constraints. Therefore, the most important responsibilities of process managers to align the existing systems and process with the new design of the workflows and interactions.

3.3 Operational managers

The duty of operational managers is implementation of rules, roles, and determined procedures for business processes. They are the closest managerial level to implementation and operation of tasks and activities in the organization and report to process managers.
-Ref[13], have provided an extensions on the BPMN. They proposed a method to analysis activities in different dimensions by use of analytical data. They aimed that business objects can analyzed from various dimensions such as: time, cost, and quality. In evaluating and improving the business processes, only specific elements (business objects and entities) have considered. For example, in enterprise perspective, organization resources, such as: staffs and machines are considered as a dimension. Further, in literature, pools and lanes are used to illustrate organizational elements and their interaction within a business process. Lodhi et al. in their paper used pools and lanes to illustrate dimensions and different classes of these dimensions[13].

4. The proposed framework

One of the most important factor in the failure of the business processes continuous improvement is lack of manager's serious and persistent support of this projects and their wrong decisions about improvement of business processes, that mainly caused by incomplete and incorrect understanding of business processes and its problems [15][17].
Business process models extensions, helps analysts and managers to understand business processes and identify their defects. Business process modeling has increased the ability to understand business processes and to make rational decisions for organizing activities in a traceable and understandable way [13][18]. Regard to that elements and rules of BPMN is not able to singly provide necessary support to make decisions about business processes improvement, Therefore, following using the theory of pattern matching and linking between a theoretical and an operational perspectives, we provide a framework for business processes continuous improvement based on BPMN.
According to Ref[11], the successful BPM implementation, needs effective and structured participation of different levels of managers in the project to make correct decisions about business processes. On the other hand, according to Ref[13] and Ref[19], providing efficient extensions on BPMN cause to increase understanding of business processes and therefore a wise decision to organize the activities of business processes will be followed. The essential issue in providing a graphical models is filtering of available information according to the needs and status of audience and stakeholders of the project.
The proposed framework entitled "Business process improvement framework based on managerial levels (BPIML)", determines a certain type of business process diagrams (BPD) based on BPMN with respect to the objectives and tasks of the various managerial levels of organizations and their roles in business process management (BPM) projects. This framework also improves understanding of business processes and their problems by reducing the cognitive load of displayed information for their audience at different managerial levels and provide details and complete information. This framework able to provide the necessary support to making decisions about improving business processes. Following we introduce the proposed framework.

4.1 Senior managers

Senior managers have played a leading role in BPM projects and make the final decision on the choice of methods and procedures to business processes improvement. Due to the multiple duties of senior managers and multiple business processes within an organization, study of the exact details of the business processes are not required for them. So at this managerial level, BPMN conventional diagrams are not applied, it is necessary to generate diagrams for this level that includes only informations such as: the efficiency of business processes and level of business processes readiness to improvement, by filtering of informations. For example, in the case of product manufacturing process that there are five activities that executed to manufacture a product within it (Fig. 1) we assume three diagrams for this manager level (Fig. 2-4). We define three classes in cost dimension to arrange activities of processes and their involved elements in the Fig. 2. It should be noted that time efficiency of the process comes from comparison of operational time and idle time of the process. Similarly, efficiency of the process comes from comparison of time efficiency and cost of the process. Also the level of business processes readiness to improvement can be displayed by colors in diagrams. In most processes, usually there are activities that according to the organization's policies, national laws, religious ordinances, etc. changes in them are impossible or are low. This also shown in Fig. 4.

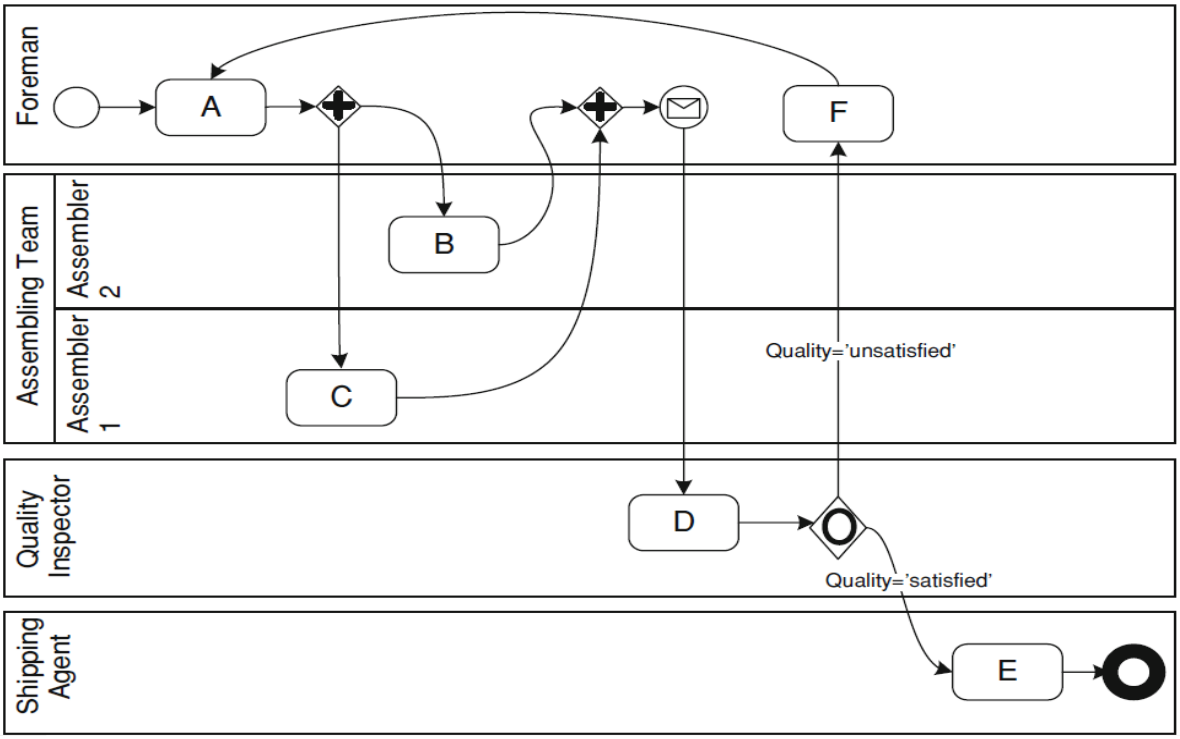


Figure 1. Product manufacturing process diagram using BPMN[13]

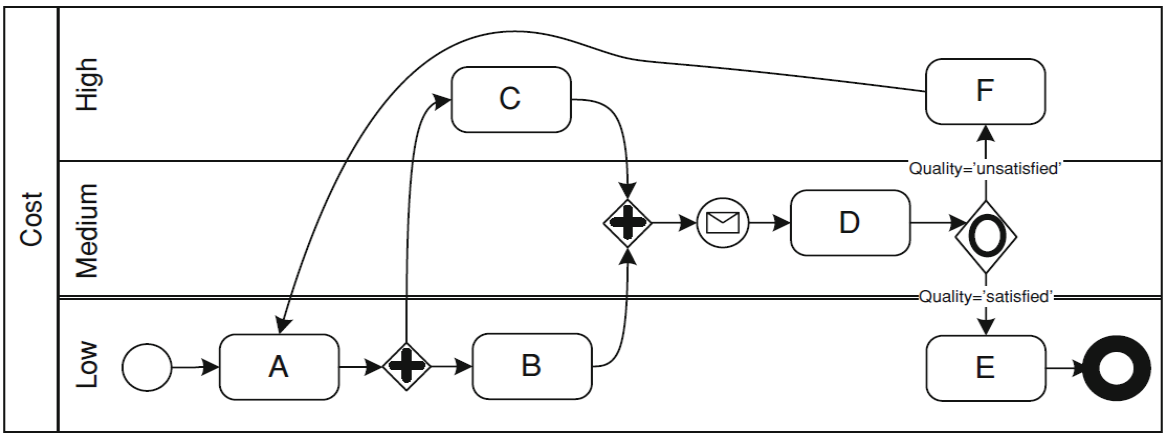


Figure 2. Product manufacturing process diagram based on cost[13]

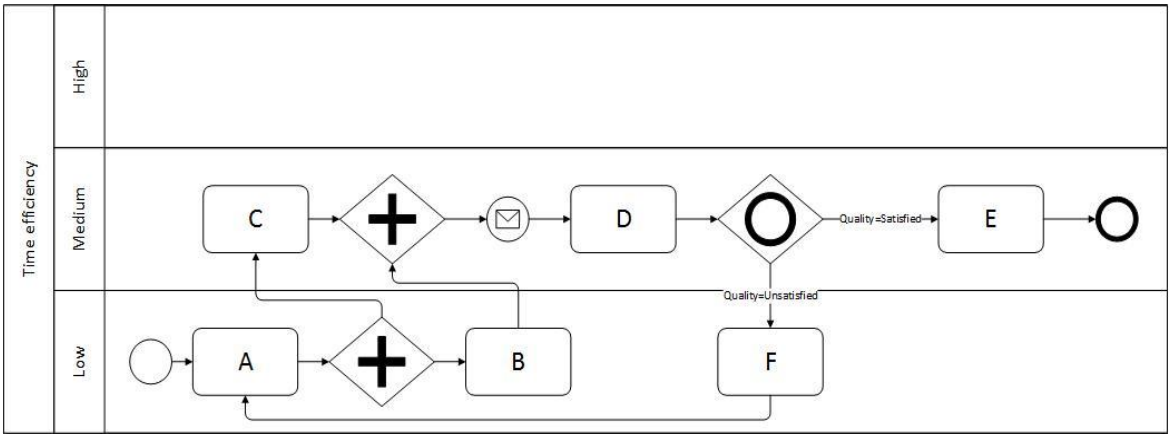


Figure 3. Product manufacturing process diagram based on Time efficiency

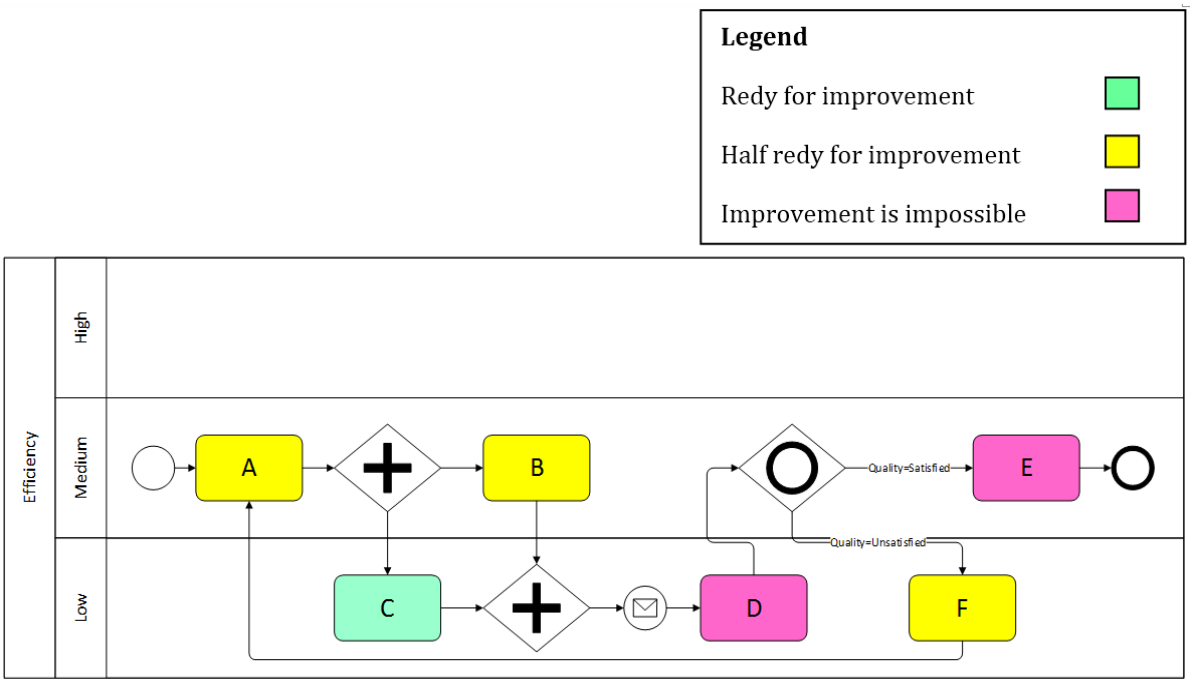


Figure 4. Product manufacturing process diagram based on Efficiency

4.2 Process managers

As previously mentioned, these managers are responsible for designing and monitoring of business processes and organizational structure based on strategic vision of the organization and existing resources and constraints. So, they are project manager in BPM projects and naturally knowing the details of business processes are necessary for them. Therefore, BPMN conventional diagrams (e.g. Fig. 1) are useful at this level.

4.3 Operational managers

Operational managers play two key role in BPM project. First, identify, analysis and drawing of each business process activities, Second, efforts to successful build to make the changes in the structure of business processes that senior managers concerning about them and follow the strategic shared value. So at this managerial level, BPMN conventional diagrams are not applied, it is necessary to generate diagrams for this level that more focus on the details of the implementation of the business processes activities. For example, in the process of Fig. 1, during the implementation of the Activity D, assembled parts inspection done. These activities may include steps such as: Initial evaluation, Practical test, Comparison test results with quality requirements and finally approval or disapproval parts. So, we consider a diagram such as Fig.5 that only model the implementation of Activity D.

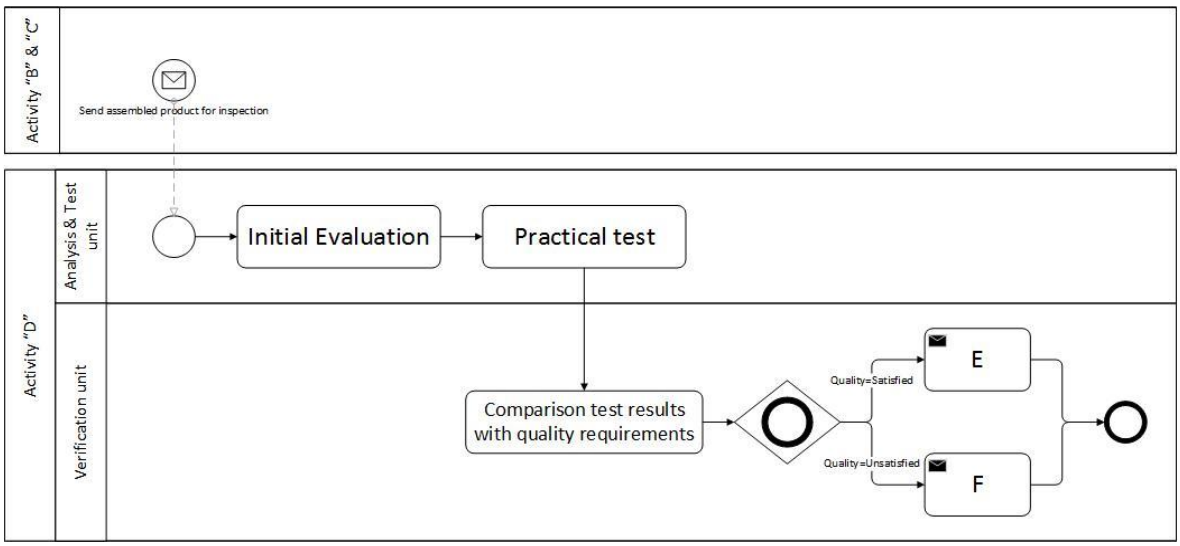


Figure 5. Diagram of activity D

In Table 1, we compare conventional method and the proposed framework (BPIML). In the next section, we evaluate and validate the proposed framework and the contents of the following table.

Table 1. Compare conventional method and the proposed framework (BPIML)

No.	Parameter	Conventional method			Proposed framework (BPIML)		
		Low	Medium	High	Low	Medium	High
1	Cost and time			*		*	
2	Ability to understand		*				*
3	Complexity		*			*	
4	Support for the correct and timely decisions	*					*
5	Ability to be used in different areas			*			*
6	Functionality for all managerial levels	*					*

5. Evaluation and Validation of the proposed framework

In order to evaluate and validate the proposed framework we used it in Bank Refah Kargaran with the respect to some BPMN exceptions measurement methods such as: González et al (2011) method, Indulska et al(2011) method, Recker et al (2009) method and Rolón et al (2006) method [20][21][22][23]. Bank Refah Kargaran is one of Iran's major banks that its headquarters in Tehran, Iran. Considering that according to the ideas of the Bank's managers and experts, the main process of the Bank is the credit process (Bank loans), thus, this process was selected as a priority for analysis and improvement. For this purpose, we modeled the current state of the credit process by conventional method (like Fig. 1), then we modeled the current state of the credit process according to BPIML framework and presented those diagrams to managers of the bank.

After that we designed a questionnaire consists of 10 questions using Likert scale and was given to the participants (managers of Bank Refah Kargaran three managerial levels). The purpose of this questionnaire was assessing strengths, weaknesses and the success rate of the BPIML framework. The questions were designed as propositions that the interviewees had to declare explicit feedback. These people could respond to questions with five options, 1-strongly disagree, 2-disagree, 3-neither agree nor disagree, 4-agree and 5-strongly agree. In Table2, we describe the results, calculate the average and the number of positive responses (agree and strongly agree responses).

Given the results of this questionnaire, it can be concluded that position of the BPIML framework in the life cycle of enterprise solutions has been clearly determined, and has a certain quota in advancing of business process management projects. Also as expected, the framework is very simple and understandable for the operation and the audience by reduce the complexity of the models based on BPMN and make them more applied. The proposed framework is not customization for specific applications or business and is a public model. Also according to breadth of business processes in organizations, the proposed framework has high usability in large scale organizations.

As previously been mentioned according to Ref[11], the successful BPM implementation need effective and structured participation of different levels managers in this project and makes correct decision about business processes. By examining the results of the questionnaire, it can be said that the proposed framework provide support for correct and timely decisions by increasing the clarity and transparency of the business processes which led to success in BPM projects.

According to opinions of the participants in this poll, the proposed framework with a special focus on each business process activities is semi useful to make the changes in the structure of business processes that senior managers concerning about them.

Table 2. The proposed framework (BPIML) evaluation and validation questionnaire

No.	Questions	1- Strongly disagree	2- Disagree	3- Neither Agree nor Disagree	4- Agree	5- Strongly Agree	Number of Positive responses	Present of Positive responses	Average
1	The position of the framework in life cycle of enterprise solutions is clearly identified.	0	0	1	5	4	9	90%	4/30
2	The proposed framework is simple and understandable.	0	0	1	4	5	9	90%	4/40
3	This framework, with consider to the different levels of management reduce complexity of models based on the BPMN and make them more practical.	0	1	1	2	6	8	90%	4/30
4	The proposed framework can be used in different areas. (Different Business)	0	0	0	4	6	10	100%	4/60
5	Given the wide range of business processes, the proposed framework can be used in organizations with large scale.	0	0	0	2	8	10	100%	4/80
6	The proposed framework with increases the success of BPM, provide platform for ERP implementation in organization.	0	0	1	3	6	9	90%	4/50
7	The proposed framework, reduce time and cost of the BPM projects.	1	1	2	3	3	6	60%	3/60
8	The proposed framework with increased clarity and transparency of modeled processes, support correct and timely decisions.	0	1	0	4	5	9	90%	4/30
9	The proposed framework, with a specific focus on each of the activities of business processes, facilitate to make changes.	2	3	0	3	2	5	50%	3/00
10	This framework, helps organizations to achieving the goals and strategic vision, and enhance its position among competitors.	1	4	1	2	2	4	40%	3/00

6. Conclusions and future work

In this paper we proposed an extension of BPMN. Emergence extended business process models based on BPMN, helps analysts and managers to understand business processes and identify their defects. These models provide the context for rational decision to organize business processes activities in an understandable manner. In this research, we reviewed the literature and study the theoretical and operational perspectives in this regard. We also proposed a framework entitled "Business process improvement framework based on managerial levels (BPIML)". This framework, considers three levels of management (Organizational level managers, Process/Departmental level managers and Activity level managers) for manager of an organization. Then, defines certain types of models based on BPMN, for each management level, by taking into the account the objectives and tasks of various

managerial levels in [organizations and their role in Business Process Management (BPM) projects. For Organizational level managers, the model generates information such as: business processes efficiency, cost and etc. For Process/Departmental level managers, according to their role in BPM projects as a project manager, knowing the details of the business process is essential for them. So, conventional models of BPMN are considered for this level. For Activity level managers, models which focuses on details of implementation of business process activities are generated. This framework improved understanding of business processes and their problems by reducing the cognitive load of displayed information for their audience at different managerial levels and provided details and complete information. This framework able to provide the necessary support to making decisions about improving business processes. In the end, we evaluated and validated the proposed framework. By examining the results of the questionnaire, it can be said that the proposed framework improved the effective and structured mangers participation and provide support for correct and timely decisions by increasing the clarity and transparency of the business processes which led to success in BPM projects.
Business process management and business processes reengineering have a special importance because of extensiveness of the workspace, so with respect to the limitations of this study the future works as follows:

A. The case study we carried out involves a single organization, albeit a large and complex one. We therefore believe that further case studies need to be undertaken to improve external validity of our results, and demonstrate BPIML superiority over the conventional method.
B. The proposed framework introduce an extension of BPMN that have a significant impact on the success of business process management projects in organizations, so, we recommend further studies carried out to meet the business and technical perspectives in these projects.

References

[1] Weske, M. "Business Process Management: Concepts, Languages, Architectures, Second Edition", Springer Inc, 2012.
[2] Rosing, M, White, S, Cummins, F, and Man, H. "Business Process Model and Notation—BPMN", A chapter of “The Complete Business Process Handbook”, Elsevier Inc., 1st Edition, pp 429–453, 2015.
[3] Abreu, J, Martins, P, Fernandes, S, and Zacarias, M."Business Processes Improvement on Maintenance Management:a Case Study",CENTERIS 2013 -Conference on ENTERprise Information Systems ,Procedia Technology , Elsevier B.V., Vol. 9, pp 320–330, 2013.
[4] Chen, F., Wang, Q., Wei, Q., Ren, C., Shao, B. and Li, J. "Integrate ERP system into Business Process Management system", IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), Dongguan, China, 2013.
[5] Van der Aalst,M. Ter Hofstede, A. and Weske, M. "Business Process Management :A Survey", Verlag Berlin Heidelberg, Springer Inc. Vol. 2678, pp. 1-12, 2003.
[6] Chang, J. "Business process management systems : Strategy and Implementation", New Yourk , USA, Auerbach Publications,2006.
[7] Idron, N. "A Business process management approach to ERP implementation", Master thesis, School of Economics and Management, Lund University,Lund, Sweden, 2008.
[8] Al-Mashari, M, and Zairi , M. Revisiting BPR : a holistic review of practice and development, Business Process Management Journal , Vol. 6 No. 1,2000, pp.10 – 42, 2000.
[9] Weerakkody, V, Janssen, M, Dwivedi, Y. "Transformational change and business process reengineering (BPR): Lessons from the British and Dutch public sector", Elsevier Inc. Vol. 28, pp. 320-328, 2011.
[10] Silver, B. "BPMN 2.0 Handbook :Methods, Concepts, Case Studies and Standards in Business Process Management Notation", Future Strategies Inc, 2011.
[11] Sikdar, A and Payyazhi, J. "A process model of managing organizational change during business process redesign", Business Process Management Journal, Vol. 20 No. 6, pp. 971-998,2014.
[12] Raschke, R, and Sen, S. "A value-based approach to the ex-ante evaluation of IT enabled business process improvement projects", Information & Management Journal, Elsevier B.V., Vol. 50, pp 446–456, 2013.
[13] Lodhi, A, Köppen, V and Saake G. "Business Process Improvement Framework and Representational Support", Third International Conference on Intelligent Human Computer Interaction, Prague, Czech Republic, Vol. 179, 2011, pp 155-167, 2011.
[14] Damij, N, Damij, T, Grad, D, and Jelenc, F. "A methodology for business process improvement and IS development", Information and Software Technology Journal, Elsevier B.V., Vol. 50, pp 1127–1141, 2008.
[15] Habib, M. "Understanding Critical Success and Failure Factors of Business Process Reengineering", International Review of Management and Business Research Journal, Vol. 2 No. 1, pp. 1-10, 2013.
[16] Trochim, W.M.K. "Outcome pattern matching and program theory", Evaluation and Program Planning, Vol. 12 No. 4, pp. 355-366, 1989.
[17] Hongjun, L. and Nan, L. "Process Improvement Model and It’s Application for Manufacturing Industry Based on the BPM-ERP Integrated Framework", Verlag Berlin Heidelberg, Springer Inc. Vol. 231, pp. 533-542, 2011.
[18] Geambasu, C. "BPMN VS. UML activity diagram for business process modeling", Accounting and Management Information Systems Journal, Vol. 11, No. 4, pp. 637–651, 2012.
[19] Braun, R and Esswein, W. "Classification of Domain-Specific BPMN Extensions", 7th IFIP WG 8.1 Working Conference, PoEM, Manchester, UK, Vol. 197, 2014, pp 42-57, 2014.
[20] Rolón,E, Ruiz, F, García,F, and Piattini,M. "Evaluation Measures for Business Process Models", Proceedings of the 2006 ACM symposium on Applied computing, pp. 1567-1568, 2006 .
[21] González,L, Ruiz, F, García,F, Cardoso,J. "Towards thresholds of control flow complexity measures for BPMN models", Proceedings of the 2011 ACM symposium on Applied computing, pp. 1545-1450, 2011.
[22] Indulska, M, Muehlen, M, and Recker, J. "Measuring Method Complexity: The Case of the Business Process Modeling Notation", Fifteenth Americas Conference on Information Systems, Detroit, Michigan, USA, pp. 3–6, 2011.
[23] Recker, J, Muehlen, M, Siau, K, Erickson, J, and Indulska, M. "Measuring Method Complexity: UML versus BPMN", Fifteenth Americas Conference on Information Systems, San Francisco, California, USA, pp 8–12, 2009.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

A Survey on Applications of Augmented Reality

Federico Manuri[1] and Andrea Sanna[2]

[1] Dipartimento di Automatica e Informatica, Politecnico di Torino
Torino, 10129, Italy
federico.manuri@polito.it

[2] Dipartimento di Automatica e Informatica, Politecnico di Torino
Torino, 10129, Italy
andrea.sanna@polito.it

Abstract

The term Augmented Reality (AR) refers to a set of technologies and devices able to enhance and improve human perception, thus bridging the gap between real and virtual space. Physical and artificial objects are mixed together in a hybrid space where the user can move without constraints. This mediated reality is spread in our everyday life: work, study, training, relaxation, time spent traveling are just some of the moments in which you can use AR applications.

This paper aims to provide an overview of current technologies and future trends of augmented reality as well as to describe the main application domains, outlining benefits and open issues.

Keywords: Augmented reality; Computer augmented environment; User interfaces; Head-mounted display; Mobile and ubiquitous computing

1. Introduction

Virtual Reality (VR) and Augmented Reality (AR) are now well known concepts, but the relationship between real space, virtual space and all the intermediate forms of mixed space have been formalized by Milgram & Kishino in 1994 [1]. An augmented/mixed space "blends" together real and virtual objects; moreover, virtual elements are positioned and aligned in order to appear as part of the real world with respect to the user view.

Researchers and scientists have been discussing for a long time in order to establish a hierarchy between virtual and augmented reality and this discussion is not yet closed: it is out of the scope of this paper to compare these two paradigms. However, an unquestionable advantage can be attributed to AR: users of AR applications can keep a contact with the real world and this is an advantage both from the technical point of view (a part of the space to be presented already exists and it is not necessary a computer-generated model of it) and the physical point of view (a detachment form the real world can lead to physical and mind discomforts). If a sentence should characterize and define what the AR can do, a very good choice could be: "AR is able to bridge the gap between real and virtual objects". Therefore, every time it is necessary to represent real and computer-generated elements within the same space, augmented reality is the best solution.

The first prototype of an AR system can be dated back to 1968, when Sutherland [2] proposed an application based on a Head Mounted Device (HMD). A long time has passed from that first prototype and AR is now widespread in our everyday life. We use AR applications and technologies when we work, study, play, spend our free time and in many other situations. Moreover, AR technologies designed for very special purposes (e.g. military aircraft cockpits) are now available for commercial applications.

A first definition of what AR is has been provided in [3]: 1. AR combines real and virtual content, 2. AR is interactive in real time, 3. AR is registered in 3D. Virtual contents are usually named assets: a set of computer-generated information overlapped to the real word. Assets can be: text labels, audio messages, 3D models, animations and videos. The real world can be directly seen by the user or can be perceived through a camera. Other classifications of AR have been provided; for instance, a taxonomy based on functional purposes is presented in [4], a classification based on four axes (tracking, augmentation, content displayed and non-visual rendering) is presented in [5], whereas in [6] it is theorized that AR systems are composed by six sub-systems.

A lot of challenges have to be tackled to deploy an AR application (for more details see the next section) but it is immediately clear that the (exact) alignment of some assets (e.g. 3D objects) plays a key role in several AR-based applications. For instance, patient data are overlapped to the body in a lot of medical AR systems and the exact position-orientation is mandatory. The computation of the user's head with respect to the real world is the main problem an AR application has to tackle; this computation is usually in charge of the so called tracking system, which allows the application to know how the user is placed-oriented in the real world.

AR has found its initial application in six domains: medicine, manufacturing and repair, annotation and visualization, robot path planning, military applications and entertainment. Other important areas followed such as

tourism, architecture, cultural heritage, education and so on. This work aims to provide the readers a picture of AR technologies and the most important application fields in order to outline the potential impact in our everyday life. Although AR can enhance the user experience, several open problems have to be still effectively tackled; moreover, AR has introduced some issues concerning privacy and security, which have to be carefully considered.

Other AR survey papers have been published [3][7][8][9]; this manuscripts aims to provide an updated application-oriented vision, which should help readers to understand the potentially huge impact of AR on everyday life.

This paper is organized as follows: the next section shows the architecture of an AR application and discusses technologies available at the present moment, section Applications depicts the main application fields of mixed and augmented reality, whereas the last section presents future trends and emerging scenarios.

2. Technologies and architectures of AR systems

This section aims to analyze both the generic architecture of an AR system and the available technologies. High-level blocks characterizing an AR system are: a tracking system, an asset/scene generator and a combiner. A lot of AR systems rely on optical trackers; in this case, a camera is necessary to frame the real environment and then the video stream is received by a head tracking module. This module computes position and orientation of the head with respect to the framed objects; position and orientation are necessary to correctly align assets. When absolute values of the position are required, it is necessary to define a reference frame. Two approaches are used: marker-based and marker-less. Marker-based solutions require the camera to start by framing a well-known pattern (e.g. a QR code or an ARTag); this allows the tracker to define a World Coordinate System (WCS). Being the size of the marker known, the tracker is able to fix the camera position with respect to the WCS in an absolute way. On the other hand, marker-less approaches use environment features to correctly compute camera orientation and align assets. In this case, it is possible to compute relative positions, but the real sizes of a framed object have to be known in advance to determine absolute positions.

The output of the head tracking module is sent to the scene generator block. Head location and orientation information allow the scene generator to generate assets and accurately align them; of course, this is not an issue for audio messages or videos, but 3D objects, animations and also text labels have to be conveniently positioned. The last block, the combiner, overlaps assets to the user view; the combiner acts in different ways according to the used AR paradigm.

The first paradigm is based on see-through devices. A see-through device allows the user to see the real environment with his/her own eyes (see Fig. 1 left) and assets generated by the scene generator are overlapped by an optical effect. This solution allows the user to directly perceive the world surrounding him/her, but it requires special purpose devices such as AR glasses. Glasses can be monocular or binocular (in this second case also 3D assets can be correctly perceived) and assets can be projected over the lenses or displayed on semitransparent mini-monitors placed between eyes and lenses. Generally, AR glasses encompass a RGB camera, thus they directly feed the head tracker module. When an optical tracking is not available (AR glasses do not include a camera or the computation power is not sufficient to process in real time the video coming from the camera) other tracking systems can be used: inertial, mechanical and magnetic. Readers interested in a depth explanation about these systems are encouraged to read [10]; moreover, most handheld outdoor AR systems use GPS and compass sensors for tracking. Assets and real objects are mixed by an optical combiner when see-through devices are used to implement an AR system.

The second paradigm is based on hand-held (mobile) devices (see Fig. 1 right). Mobile devices (e.g. tablets and smartphones) encompass all the hardware necessary to implement an AR system: the camera to gather a video of the real world, a display to show the augmented environment and the computational power to: compute the camera position-orientation (by optical tracking), generate assets and combine virtual objects with the video. Mobile AR (MAR) has become very popular as cellular phones and personal digital assistants have been replaced by smart devices able to run computational intensive apps. With respect to the see-through paradigm, the MAR approach presents users a mediated reality where a direct perception of the surrounding world is not available. Assets and video of the virtual world are mixed together by a so called video combiner module.

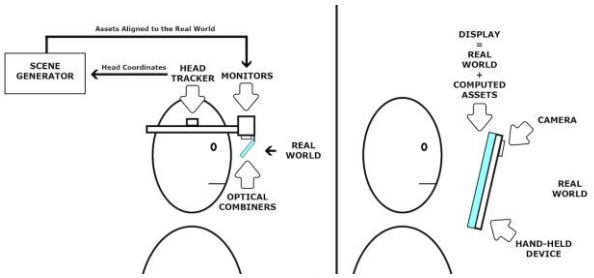


Fig. 1 The see-through paradigm (left): the user is able to see the real world with his/her own eyes and assets are overlapped by an optical combiner. The hand-held paradigm (right): the user perceives the real world through the video streaming coming from the camera and assets are overlapped by a video combiner.

The third paradigm, called monitor-based AR, is conceptually comparable with the previous one, but camera, computational power and display (a monitor) are not encompassed in a single device. This paradigm is used when either large displays are needed or the camera must be independent (e.g. in augmented endoscopy).
This high level description of an AR system clearly outlines possible challenges AR applications have to tackle. First of all, AR systems have to deal with performance, mainly concerning the real time computation of position and orientation of the camera. Then, a second issue concerns the precision of asset alignment: several applications can require a sub-millimeter accuracy (e.g. medical AR applications). The third issue is related to the user interface: users have to be able to interact with augmented contents but the traditional keyboard-mouse paradigm is not generally available. The last two challenges concern mobility and visualization; nowadays mobile devices seem to fully satisfy the best part of application fields. On the other hand, see-through visualization devices are not still fully compliant to contrast, resolution, brightness and field of view requirements.

3. Applications

Next sections provide an overview of the main application domains of AR. This analysis well depicts how AR is pervasive in our everyday life: we have to cope with AR applications when we work, play, travel, study and in many other activities.

3.1 Medicine

As their natural attitude is to bridge the gap between real and virtual, AR technologies were immediately identified as a valuable tool to bring patients and their medical data into the same space. The potential of AR in medical applications was foreseen by Steinhaus (an Austrian mathematician) in 1938 [11]: Steinhaus suggested a method to display a bullet inside a body by a very cumbersome overlay process. On the other hand, the first real application of AR in medicine can be dated back to 1986, when a system to integrate data from computer tomography into an operating microscope was proposed. Readers interested in examining in depth applications of AR in medicine can refer to [12].
Recent advances in medical imaging have provided scientists and physicians a huge amount of data that could be profitable used to support several activities. Anatomical and functional data can be a support in surgery as well as in diagnostic of preoperative and intraoperative data or in training tasks. The use of AR in surgery is strictly related to the display technology the surgeon opts for: computer-generated assets can be directly overlapped onto the operating microscope or can be displayed over a monitor (augmented endoscopes can be considered as a particular case of monitor-based AR). Fig. 2 shows an example of medical application: the patient's medical record is shown and the radiography image is overlapped to the face with precise positioning. The projection of assets directly on the patient is apart from a particular visualization technology, but it involves very complex setups, which need accurate calibrations.
Much more than other application fields, AR in medicine has to overcome three main issues: tracking precision, misperception and interaction with synthetic data. The precision required for several surgical operations is of the sub-millimeter order, therefore assets must be overlaid very accurately. On the other hand, medical AR generally involves very limited and controlled working indoor volumes and current tracking systems are able to provide the required precision under these conditions. The misperception is basically related to a wrong perception of depth (although assets are correctly aligned, the user perceives them in a wrong position), but this issue can be mitigated by using stereoscopic visualization devices. The interaction issue is more generally related to user interface design problems; for instance, a surgeon cannot interact with assets by using touch, therefore natural and multimodal user interfaces have to be implemented. Multimodal interfaces allow the user to choose among different input modes: gesture/pose recognition and speech recognition can be two alternatives to naturally interact with computer-generated contents. Unfortunately, these alternative input modes can introduce robustness issues, which have to be taken into account when safety-critical systems such as medical ones are designed.

Fig. 2 Medical application for dental surgery. The application shows the medical record for a patient, overlapping the radiography image to the face; points of interest are highlighted as well as doctor's annotations.

3.2 Assembly, maintenance and repair

Technicians involved in complex maintenance and repair tasks often need to refer to instruction manuals to correctly complete assigned procedures. This might entail a high cognitive load deriving from a continuous switch of the attention between the device under maintenance and the

manual. In other words, mistakes are more likely and repair times (and hence the costs) can grow up.
A first attempt to mitigate these issues was the introduction of Interactive Electronic Technical Manuals (IETMs) able to replace paper instructions. On the other hand, also IETMs cannot be completely part of the interaction between technician and equipment to be maintained.
AR can efficiently tackle this issue and manufacturing-repair has been immediately identified as one of the most promising application field of AR [3]. Assets are overlaid and correctly aligned with respect to the device to be maintained and can be conveyed to technicians while they are performing the procedure. Moreover, AR applications for maintenance and repair are often completed by tele-presence systems; in this way, a remote expert can interactively support maintainers when AR aid is not sufficient. Fig. 3 provides a technician's view of an AR maintenance application with the head-mounted paradigm: audio, 3d model animated and textual assets describe the instructions the technician should follow to perform the maintenance procedure. Benefits of the AR support in maintenance, repair and assembly tasks have been deeply analyzed in [13]. First attempts of supporting technicians by AR tools can be dated back to 1990s and they were all based on special purpose hardware (e.g., HMDs). On the other hand, recent advances of mobile technologies have opened new challenging and intriguing scenarios for this type of applications. Tablets and smart-phones allow users to perform easier some tasks such as routine maintenance of a car, furniture assembly, installation of electrical appliances and so on. If we consider the huge amount of money related to these activities, it is immediately clear the tremendous impact the AR might have on our everyday life. At the present time, the spread of AR in maintenance and repair tasks is very limited; this is mainly due to the time needed to create, change and improve the procedures. A first attempt to mitigate this issue has been presented in [14], where a procedure is managed as a set of sequential states: each state is related both to a tracking configuration and to a set of assets to be played when the tracking system recognizes the configuration in the real world. Existing states can be deleted and new states can be easily added; an asset library allows content makers to quickly associate an object configuration with a set of aids to be conveyed.

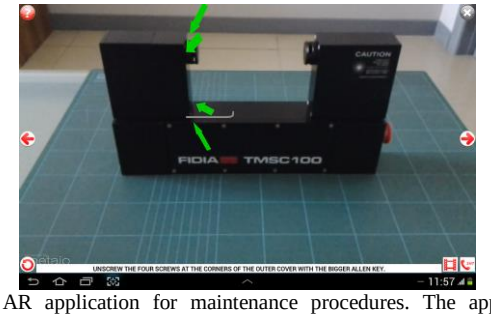


Fig. 3 AR application for maintenance procedures. The application supports the technician through the repair task of an industrial device; textual and 3D assets are shown.

3.3 Entertainment, sport and marketing

The entertainment industry is one of the most important drivers of ICT technology progress and a lot of improvements in augmented reality can be related to it. Video game players have a desire: they want to be part of the game. If the player is more involved in the game, then the game experience will be enhanced. This idea has inspired the design of all modern game consoles; in other words, players do not control characters but they are the characters.
AR aims to bridge the gap between real and virtual, therefore it is the best tool to provide users a new game experience. The real world can become the set for a game (see for instance ARQuake [15]) and the player can experience a completely new and exciting game modality; moreover, beyond a more natural and intuitive interaction, AR-based games can be played, in general, everywhere, without any space limitation. Another important benefit concerns the development step: all game scenarios have to be completely modeled and rendered in a traditional game; on the other hand, AR games use the real world as a scenario and only virtual characters and assets have to be created. Fig. 4 shows an example of this type of augmented game: a number of enemies is placed in the scenario and move towards the player, who have to shoot them down to gain points and to move to avoid the monsters and save is life.
Modern game consoles also implement AR by using different types of cameras to augment computer graphics onto live footage. Also the market of stickers can take advantage from AR: when a sticker is framed by a personal device running a specific AR application, the user can play on the device multimedia contents about the *star* (e.g. a football player) or the personality depicted on the sticker.

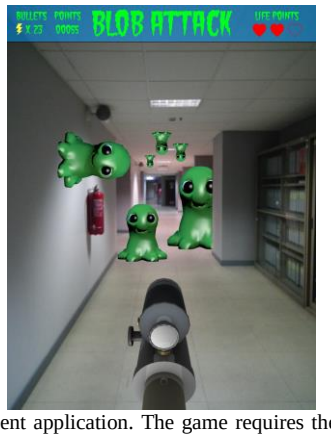


Fig. 4 Entertainment application. The game requires the player to move in the real space in order to aim and shoot at the enemies; alien enemies are placed in the real environment.

AR is often used for augmenting live broadcast of sport events [16]. For instance, computer-generated aids are added to the raw video images in order to show the off-side line in a football match or the trajectory of the puck in hockey games. Maybe, the most famous system for the augmentation of sport events is Hawk-Eye; Hawk-Eye is a system used in tennis matches, which provides the tracking and the visualization of balls trajectories, thus enabling players to challenge the Referees' decisions. In the augmentation of many sport events, assets have to be overlapped to raw images in real time; this means that the AR system has to be able to identify and track a given object in the scene in a very performing way.
AR plays also a very important role when advertisements and logos have to be inserted in live video broadcasting: messages as well as 3D objects/animations can be aligned to real objects framed by the set of different cameras used to film the event.
AR is also often used by marketers to promote new products [17]. Such systems are already widely used and they aim to present products in an engaging way. For instance, by a sort of magic mirror [9], customers can virtually wear clothes or shoes; in this way, the need to try on anything in stores it is replaced by a sort of virtual shop, thus saving time for clients.

3.4 Collaborative visualization space

Data visualization is a very broad discipline encompassing several different fields; for instance, information visualization aims to find new paradigms to efficiently display huge amount of data (e.g. the network traffic over the Internet), whereas scientific visualization aims to present users phenomena that are very hard (or impossible) to perceive (e.g. air flows around to a plane wing).
The intrinsic nature of AR provides visualization a worthwhile tool to display virtual objects within a physical space; moreover, AR is particularly suitable for collaborative visualization. Collaborating users communicate by using speech, gaze, gesture and other nonverbal cues. These "communication channels" are often inhibited or limited when traditional collaborative work tools are adopted: tele-presence and screen-based collaboration often create barriers between physical and virtual space. On the other hand, virtual collaborative spaces migrate both objects and users (represented by avatars) in a cyber space, thus altering usual communication channels. Face-to-face AR collaborative applications are able to avoid this problem, thus allowing users to communicate each other by usual verbal and nonverbal cues within a physical space [18]. Moreover, AR applications can provide the users with custom visualizations, thus enabling custom views of the dataset. This is basically obtained by organizing data to be displayed on layers: each layer should contain computer-generated objects, which share one or more attributes (e.g. they share the same material or they are sub-parts of the same object) and each layer can be activated/deactivated. A key role in collaborative work is often played by annotation: users have to be able to add their own textual comments to parts of the augmented space as well as they have to be able to share these annotations with other collaborating users. Figure 5 shows an example of collaborative application, based on a 3D model of a city's district: the users can share notes attaching them to the different parts of the 3D model (e.g. buildings, streets, squares). AR applications for collaborative visualization often provide users tangible interfaces. A tangible interface links a real object to a computer-generated one; for instance, a building could be related to a physical object such as a small box: transformations applied to the box (e.g. translations and rotations) will be propagated to the associated virtual object. Of course, the tracking system of the AR application has to be able to track both users' heads and artifacts selected as alter egos of virtual objects; at this purpose, marker-based approaches are mainly used.

Fig. 5 Collaborative application. The application allows users to visualize a virtual collaborative space and to add notes in order to share information with other users.

Features of an AR collaborative visualization application are: virtuality (computer-generated objects can be displayed and examined), augmentation (virtual objects

can be displayed in a physical space and real objects can be annotated), cooperation (several users can cooperate in a natural way), independence (each user can customize the dataset to be displayed) and individuality (each user can customize the graphic metaphors used to represent data). The first example of collaborative AR visualization is the Studierstube [19].
A main issue still concerns AR collaborative visualization tools: a better interaction form is obtained by using the see-through paradigm, but AR-glasses (e.g. binocular glasses and HMDs) often partially cover the user's eyes, thus inhibiting (or strongly limiting) gaze communication, which represents an important form of nonverbal cue. Monocular glasses can mitigate this issue, but they are not able to provide stereoscopic visualization of 3D objects.

3.5 Tourism

AR can play a key role in tourism mobilities [20] and a significant improvement has been obtained by new generation smartphones and tablets, which are often equipped with GPS sensors, and fast network connections. These devices are therefore able to support location-based AR services.
A Tourist experience can be enriched (and mediated) by adding multimedia and customized contents according to the tourist's needs [21]. Basically, three types of AR applications for tourism can be categorized.
Augmented guides are the first type of AR application for tourism; an augmented guide searches, retrieves and visualizes information gathered from several Internet sources (e.g., touristic portals). Information are arranged in order to provide the users with all the support necessary to: organize travels, reserve tours, rent cars, and so on. The first example of an augmented guide is Tuscany+: the official augmented reality application of the Italian region Tuscany. Information gathered by the AR application can be also used to generate a sort of *hybrid* space: the physical space (e.g. a square of a city) is filled up with information coming from the cyber space (e.g. the Net). Fig. 6 shows a Tourism application that provides the user his current location and a set of nearby Points of Interest (POIs); an arrow near the POI tag indicates the direction to reach it; each POI is grouped by color and icon to simplify recognition.
Several augmented guides allow the user to mark and share a POI; moreover, multimedia contents can be completed by manual annotations and comments. If the AR application allows the user to share POIs and annotations, it is said a *social* application. In this case, the impact of AR technologies might be magnified by the Net as users generally ask for a direct link between applications and their own profiles on social networks.
Sometimes, AR applications for tourism and entertainment have a strong relationship. In this second type of AR applications, a tour is organized as a multi-level game and users have to solve mysteries and answer questions related to the tour itself; users receive information about the next stage of the tour only when they are able to complete the current level. This approach provides the so called space gamification.

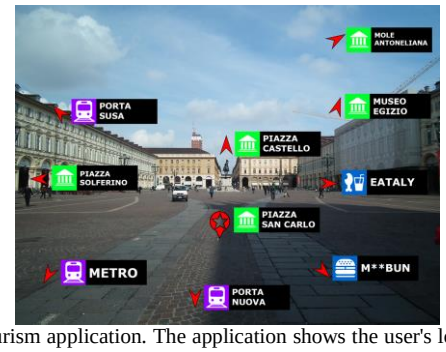


Fig. 6 Tourism application. The application shows the user's location and a selection of points of interest in the nearby area, such as: museums, restaurants and transport's lines.

The third type of AR application for tourism is related to the concept of fictionalization. In this case, tourist experience is augmented referring to very special places such as film sets or locations described in literature: AR can improve and enhance the fictionalized landscape visits.
A main issue affects the spread of AR for tourism: the lack of interoperability among applications; this impacts both on developers and content aggregators. If this problem is shared by all application domains of augmented reality, tourism is the one more affected because of the huge amount of information gatherable.

3.6 Architecture and construction

AR reality technologies find an important application field in architecture and construction tasks, which are usually classified under the category "architectural design and urban planning". Architects and designers deeply stress issues related to spatial communication: they would intuitively convey information about appearance, scale and features of a proposed project. Unfortunately, neither scale models nor virtual models completely tackle all the needs of spatial communication.
AR can be a valid support and several types of application are well known:

1. the first and simplest type aims to display buildings, from their early designs to final constructions, by representing them as virtual objects into the real world. These applications allow architects to plan and evaluate in advance the impact of a new construction (urban planning); moreover, this approach enables the so called walking tours, thus allowing users to move in a real scenario and observing a virtual building

as if it were real. Figure 7 shows an example of this type of application, with a 3D model overlapped to its blueprint through AR.

2. AR applications can also enable a sort of x-ray view. Hidden features such as structural supports, pipes and electric cables can be overlapped to a building. In this case, it is very important the integration of the AR application with BIM (Building Information Modeling) data.
3. Collaborative design. The final goal of architects and designers is to obtain a seamless transition between individual work with CAD tools and collaborative work. Collaborative work can strongly reduce design times and improve the interaction between designers and customers. Magic meetings refer to augmented/mixed spaces where people can cooperate to model, analyze and assess CAD projects. An example of magic meeting is MR^2, a mixed reality meeting room [22].
4. As presented in section "Assembly, maintenance and repair", AR can be used for maintenance tasks; in this case, special purpose applications are tailored to support building maintenance.

At the present time, a very few architects and designers use AR, and in the most of the time AR is only used to enhance presentations and marketing. This is due to a lack of integration of AR technologies in the design workflow. Although this problem has been investigated since the last decade [23], a lack of standards prevents from a real integration of AR in CAD software.

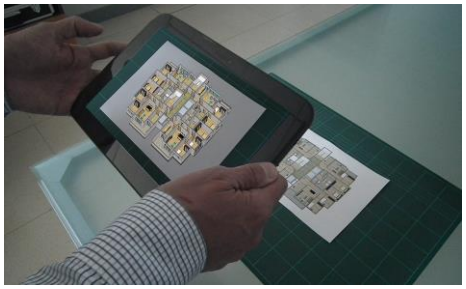


Fig. 7 An architecture application, which allows the user to visualize a 3D model of the apartment's blueprint framed by the camera.

3.7 Cultural heritage and museum visits

AR applications for tourism and cultural heritage share a lot of requirements and characteristics; on the other hand, some distinguishing features can be identified.

First of all, it is important to define the concept of cultural heritage attraction. It can be a monument, a ruin, a battlefield and everything can assume a cultural value. At the present time, videos and virtual tours are the main supports to promote cultural heritage; printed info and scale models can be also available for on site visitors. Augmented reality can strongly enhance this support by providing the users with multimedia contents ranging from the temporal reconstruction of a site to a virtual reconstruction of a battle. This kind of applications can be categorized as an augmented guide (see for instance [24]). Fig. 8 shows an example of an augmented guide that provides textual, audio and video information to the user, depending on the work of art framed by the camera.

When cultural heritage gives rise to cultural tourism, AR provides a set of undeniable advantages [25]:

1. printed info and other unnatural objects placed in a cultural site involve costs and, moreover, alter the site itself; AR applications do not need information boards or other "artifacts" and can be modified/updated more efficiently than traditional info.
2. AR does not limit the amount (and the type) of information for users, whereas info boards have a limited size, thus reducing the information potential.
3. AR applications, generally, enhance the user experience and this can trigger a virtuous circle where costs due to the development of an AR application are balanced by the positive publicity of satisfied users that visited the site.
4. Cultural heritage AR applications are often social, thus allowing users to share their experiences by social networks.

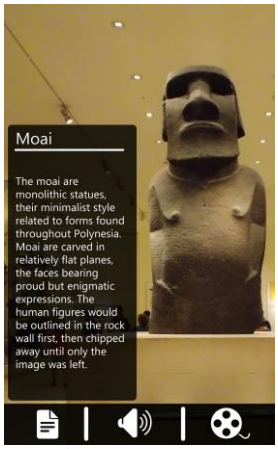


Fig. 8 A multimedia guide for museums that provides textual, audio and video information for the work of art framed by the camera.

A characteristic feature has to be mentioned for AR applications designed for museum visits. Every type of multimedia guide has to be supported by a navigation mechanism able to determine position and orientation of the user. Basically, a guide has to answer to the following two questions: 1) Where can I find the object for which I see multimedia content is available? 2) Where are the information related to a particular artwork that I can see?

Navigation systems based on Wi-Fi, Bluetooth, RFID and infrared technologies can either fail (or they can be

strongly limited) when the user is moving in crowed environments or provide a not sufficient precision in the computation of the user's orientation. AR tracking systems can mitigate these problems, thus providing developers an efficient solution for localization in GPS-denied environments.
Issues related to AR applications targeted for cultural heritage and art are almost the same as AR applications for tourism.

3.8 Teaching, education and training

Teachers, educators and trainers are always searching for new technologies able to enhance the learning experience of their students. AR proved to be an effective and efficient tool to improve traditional learning and training paths. AR changes the way users and machines interacts and this can stimulate students to approach the study of course material in a different and more pro-active way. Moreover, AR applications tailored to support teaching and learning can incentivize cooperative and collaborative learning, thus improving the teacher-student and student-student collaboration. In the same way as virtual reality, AR helps instructors to simulate and visualize microscopic or macroscopic scale systems; moreover, dangerous and/or destructive events can be represented. A lot of examples of projects and publications are known in the literature (readers can examine in depth this topic by two survey papers [26][27]) that provides a clear vision of how AR might impact on teaching and training methods. Medicine, Engineering, Architecture, Chemistry, Mathematics and Geometry, Physics, Geography, Astronomy, History, Archeology, Music and Art are just some examples of disciplines that might be taught in a different and more exciting way by AR.
All domains previously presented as application fields can take advantage from AR technologies for education purposes. Maintenance AR applications can be used to support expert technicians as well as to train beginners and AR games could have an educative value. AR is often used in the *discovery-based learning* where the recognition of a place (geo-referenced AR applications) or of a person is the starting point to present, for instance, historical events or personalities. On the other hand, the most common use of AR in education is related to interactive course material (often providing 3D visualizations), which allows educators to reduce the gap between real and virtual. Magic books (also called AR books) are the best example of interactive material: some pages present animations, audio contents and 3D objects, students can interact with. The magic book allows teachers to implement the so called *blended education*, that is a hybrid approach that uses two (or more) different teaching technologies (e.g. the traditional book and the augmented reality). Figure 9 provides an example of Magic book application: a 3D model of planet Earth is visualized on top of a 2D image of Earth itself depicted on the physical book; the application allows to rotate the model and to change its size to provide further interaction.

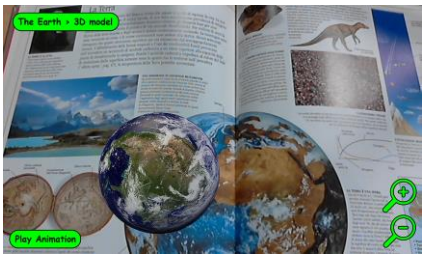


Fig. 9 A magic book that provides multimedia content related to the current page framed by the camera, e.g. a 3D model of planet Earth the user can interact with.

Despite of all these encouraging aspects, a main issue has constrained and limited the spread of AR in education so far: the difficulty for educators to make quickly deployable AR contents. AR applications and their contents are, at the moment, quite rigid, therefore it is very difficult for teachers to change and adapt them to the students' needs.

3.9 Military applications

AR applications for military purposes share a lot of issues with AR-based games and AR-based training systems. The term often used is BARS: Battlefield Augmented Reality System [28]; in other words, AR is used to synthetically create battlefields within the real world, which can be navigated without the limitation common to virtual environments. Usually, training scenarios for soldiers are made by projecting a virtual environment and virtual actors on walls within training facilities. This needs significant infrastructures and it is limited to indoor contexts. AR allows to overcome these constraints, thus allowing soldiers outdoor training activities where users can physically move through real environments.
Another use of AR is ordinary in military applications: the overlapping of information with the environment. For instance, soldiers can receive on their AR vision systems (mainly AR glasses) information about objects in the battlefield and the threat level related to each element of the environment. Figure 10 shows an example of military application: different kind of information is overlapped to the real scene to provide additional data on the surrounding area, such as target position, distance from the base camp and previous events in the area.
Some characteristic issues are related to the use of AR in military applications. First of all, if AR applications basically aim to track the position of a camera with respect to the real world, also weapons have to be often tracked in AR-based military systems. In first person shooting systems (see for instance [29]), the soldiers' heads have to be tracked as well as the positions and orientations of their

weapons. This is necessary to exactly determine what targets have been shot.

Fig. 10 A military application that provides soldiers additional information on the surrounding area, such as: target position, distance from the base camp and previous events in the area.

Moreover, information conveyed to the user has to be carefully managed; in fact, a huge amount of data might be available and an indiscriminate visualization of them will have a negative effect in terms of cognitive overload. This last issue can be partially mitigated by implementing a natural input interface based on speech and/or gesture recognition that enables soldiers to intuitively select and configure information to be overlaid to the battlefield.
Another important characteristic of AR-based systems deployed for military applications is the collaboration function. Soldiers have to be able to exchange information (e.g. each soldier should know the positions of the other ones, thus discriminating between them and potential foes), therefore affecting what assets are displayed on the interfaces of the other collaborating users.
AR-based systems for military applications should be also able to monitor each soldier's stress level, thus adapting the output of the interface in order to provide the best support.

4. Conclusion and future trends

The evolution of AR technologies is strictly related to component miniaturization. For instance, the availability of extremely small cameras now allows designers to provide users AR glasses almost unnoticeable from usual glasses. In a near future, contact lenses will be able to incorporate all functionalities now provided by AR glasses and this clearly introduces some privacy and security issues.
When users are in a public place, they are surrounded by a number of bystanders: they could be friends, acquaintances or strangers. These bystanders could be uncomfortable with the user recording them. One of the main concern of bystanders is to be located and tracked on the Internet through AR devices, similarly to what happens with video camera surveillance in public places (e.g. inside and outside banks). If using contact lenses is the case, bystanders could even not be aware of the AR device: their behavior would probably be different if they were aware of the AR device. Moreover, the user could cheat bystanders extorting and recording information (e.g. audio and video) without their consensus: this use will make the AR device similar to an advanced spy cam, with more data on the target (e.g. the location). The same concerns that affect bystanders could be applied to public and private places. A malicious application could be used to get information on a location to plan a robbery. For example, a user could enter a bank and get data about employees, security exits, alarm positions and much more. The capability to record huge amount of data unseen could even be used to plan terrorist attacks.
Beyond a technological development, the spread of AR depends on another key factor: the content availability. At the present time, AR applications are quite rigid and it is very difficult to change/adapt augmented content to users' needs. This lack of flexibility is a constraint for AR diffusion as only software developers are able to generate augmented contents; authoring tools are still very limited and tailored for a few application fields. For instance, teachers, educators and trainers could take advantage of AR by providing their students more effective and interesting course material; unfortunately, it is not easy neither to write a magic book nor to integrate AR tools in the traditional publishing pipeline. The lack of integration is another key issue and CAD tools should be modified to simplify asset generation. Maintenance, repair and assembly tasks would be strongly simplified by AR procedures but automatic (or semi-automatic) generation processes should be available. Unfortunately, a standard to describe AR contents still misses and this limits the integration of AR technologies with the existing tools. The Augmented Reality Markup Language (ARML) is an attempt to overcome this issue but it has not been still accepted by AR technology makers.
On the other hand, AR has got the potential to become a pervasive technology and a lot of everyday activities could be efficiently supported by AR. Ordinary car maintenance, furniture assembly as well as appliance installation are just three examples of routine activities that could be simplified, thus reducing efforts, costs and time spent.

References

[1] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," IEICE Transactions on Information Systems 1994, 1321–1329.
[2] I. Sutherland, "A head-mounted three dimensional display," In Proceedings of the Fall Joint Computer Conference, San Francisco, USA, Dec. 9-11, 1968, ACM New York, 757–764.
[3] R.T. Azuma, "A survey of augmented reality," Presence: Teleoperators and Virtual Environments 1997, vol. 6, no. 4, 355-385.
[4] O. Hugues, P. Fuchs, and O. Nannipieri, "New augmented reality taxonomy: Technologies and features of augmented environment," In Handbook of augmented reality, 47-63, 2011, Springer.
[5] J-M. Normand, M. Servières, and G. Moreau, "A new typology of augmented reality applications," In Proceedings of the 3rd Augmented Human International Conference, 2012, page 18, ACM.

[6] J.M. Braz and J.M. Pereira, "Tarcast: taxonomy for augmented reality casting with web support," The International Journal of Virtual Reality, vol. 7, no. 4, 47-56, 2008.
[7] R.T. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," IEEE Computer Graphics and Applications, vol. 21, no. 6, 34-47, 2001.
[8] F. Zhou, H. Been-Lirn Duh, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: a review of ten years of ismar," In Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 193-202, 2008. IEEE Computer Society.
[9] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo E. Damiani, and M. Ivković, "Augmented reality technologies, systems and applications," Multimedia Tools and Applications, vol. 51, no. 1, 341-377, 2011.
[10] E. Foxlin, "Motion tracking requirements and technologies;" Handbook of Virtual Environments: Design, Implementation, and Applications, Lawrence Erlbaum Associates, Hillsdale, NJ 2002, 163-210.
[11] H. Steinhaus, "Sur la localisation au moyen des rayons x," Comptes Rendus de L'Académie des Sciences 1938, 206, 1473–1475.
[12] T. Sielhorst, M. Feuerstein, and N. Navab, "Advanced medical displays: a literature review of augmented reality," Journal of Display Technologies 2008,vol. 4, no. 4, 451-467.
[13] S. Henderson and S.; Feiner, "Exploring the benefits of augmented reality documentation for maintenance and repair," IEEE Transactions on Visualization and Computer Graphics 2011, vol. 17, no. 10, 1355–1368.
[14] F. Lamberti, F. Manuri, A. Sanna, G. Paravati, P. Pezzolla, and P. Montuschi, "Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance" IEEE Transactions on Emerging Topics in Computing 2014, vol. 2, no. 4, 411-421.
[15] W. Piekarski and B. Thomas, "ARQuake: the outdoor augmented reality gaming system," Communications of the ACM 2002, vol. 45, no. 1, 36-38.
[16] R. Cavallaro, M. Hybinette, M. White, and T. Balch,. "Augmenting live broadcast sports with 3D tracking information," IEEE Multimedia 2011, vol. 18, no. 4, 38-47.
[17] J. Bule and P. Peer, "Interactive augmented reality marketing system," In: World Usability Day, Paper ID 2505, 2013.
[18] M. Billinghurst and H. Kato, "collaborative augmented reality," Communications of the ACM, vol. 45, no. 7, 2002, 64-70.
[19] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz, "Studierstube: an environment for collaboration in augmented reality," Virtual Reality 1998, vol. 3, no. 1, 37-48.
[20] K. Hannam, G. Butler, and C. Paris, "Developments and key issues in tourism mobilities," Annals of Tourism Research 2014, vol. 44, 171-185.
[21] C.D. Kounavis, A.E. Kasimati, and E.D. Zamani, "Enhancing the tourism experience through mobile augmented reality: challenges and prospects," International Journal of Engineering Business Management, 2012, vol. 4, no. 10, 1–6.
[22] K. Kiyokawa, M. Niimi, T. Ebina, and H. Ohno, "MR2 (MR Square): a mixed-reality meeting room," In Proceedings of the IEEE and ACM International Symposium on Augmented Reality, New York, USA, Oct. 29-30, 2001, IEEE, 169-170.
[23] W. Broll, I. Lindt, J. Ohlenburg, M. Wittkämper, C. Yuan, T. Novotny, A. Fatah gen. Schieck, C. Mottram, and A. Strothmann, "ARTHUR: a collaborative augmented environment for architectural design and urban planning," Journal of Virtual Reality and Broadcasting 2004, vol. 1, no. 1. 1-9.
[24] V. Vlahakis, N. Ioannidis, J. Karigiannis, M. Tsotros, M.; Gounaris, D. Stricker, T. Gleue, P. Daehne, and L. Almeida, "Archeoguide: an augmented reality guide for archaeological sites," IEEE Computer Graphics and Applications 2002, vol. 22, no. 5, 52-60.
[25] K. Attila and B. Edit, "Beyond reality – the possibilities of augmented reality in cultural and heritage tourism," In Proceedings of the 2nd International Tourism and Sport Management Conference, Debrecen, Hungary, Sep. 5-6, 2012.
[26] S.C.-Y. Yuen, G. Yaoyuneyong, and F. Johnson, "Augmented reality: an overview and five directions for AR in education," Journal of Educational Technology Development and Exchange 2011, vol. 4, no. 1, 119-140.
[27] H.-K. Wu, S.W.-Y. Lee, H.-Y. Chang, and J.-C. Liang, "Current status, opportunities and challenges of augmented reality in education," Computers & Education 2013, vol. 62, 41–49.
[28] M.A. Livingston, L.J. Rosenblum, S.J. Julier, D. Brown, Y. Baillot, J.E. Swan II, J.L. Gabbard, and D. Hix, "An augmented reality system for military operations in urban terrain," In Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, Orlando, Florida, USA, Dec. 2-5, 2002, 868-875.
[29] Z. Zhu, V. Branzoi, M. Sizintsev, N. Vitovitch, T. Oskiper, R. Villamil, A. Chaudhry, S. Samarasekera, and R. Kumar, "AR-Weapon: live augmented reality based first-person shooting system," In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Jan. 5-9, Waikoloa, HI, 2015, IEEE, 618 – 625.

Federico Manuri received the B.Sc. and M.Sc. degrees in computer engineering from Politecnico di Torino university in Italy in 2008 and 2011, respectively. He currently is a Ph.D. student at the Dipartimento di Automatica e Informatica at Politecnico di Torino university in Italy. His research interests include human machine interaction, computer graphics and augmented reality.

Andrea Sanna received the M.Sc. degree in electronic engineering and the Ph.D. degree in computer engineering from the Politecnico di Torino, Turin, Italy, in 1993 and 1997, respectively, where he is currently an Associate Professor with the Dipartimento di Automatica e Informatica. He has authored or co-authored several papers in the areas of computer graphics, virtual reality, distributed computing, and computational geometry. He is a Senior Member of the Association for Computing Machinery.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Provable Data Possession Scheme based on Homomorphic Hash Function in Cloud Storage

Li Yu[1] , Junyao Ye[1, 2], Kening Liu[1]

[1] School of Information Engineering, Jingdezhen Ceramic Institute, Jingdezhen 333403,China
lailul@163.com

2 Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200240, China
sdyejunyao@sjtu.edu.cn

Abstract

Cloud storage can satisfy the demand of accessing data at anytime, anyplace. In cloud storage, only when the users can verify that the cloud storage server possesses the data correctly, users shall feel relax to use cloud storage. Provable data possession(PDP) makes it easy for a third party to verify whether the data is integrity in the cloud storage server. We analyze the existing PDP schemes, find that these schemes have some drawbacks, such as computationally expensive, only performing a limited number provable data possession. This paper proposes a provable data possession scheme based on homomorphic hash function according to the problems exist in the existing algorithms. The advantage of homomorphic hash function is that it provides provable data possession and data integrity protection. The scheme is a good way to ensure the integrity of remote data and reduce redundant storage space and bandwidth consumption on the premise that users do not retrieve data. The main cost of the scheme is in the server side, it is suitable for mobile devices in the cloud storage environments. We prove that the scheme is feasible by analyzing the security and performance of the scheme.

Keywords: *Cloud Storage, Provable Data Possession, Homomorphic Hash Function, Data Possession Checking.*

1. Introduction

Cloud storage[1] can satisfy the demand of accessing data at anytime, anyplace. For the users who need inexpensive storage and unpredictable storage capacity, compared with purchasing the entire storage system, purchasing cloud storage capacity needed will obviously bring more convenience and efficiency. Cloud storage not only saves investment for the users but also saves resources and energy of society. However, there are still many problems to be solved such as security, reliability and service level of cloud storage, so it has not been widely used. When the users stores data in the cloud server, they are most concerned about whether the data is integrity. In cloud storage, only when the users can verify that the cloud storage server possesses the data correctly, they shall feel relax to use cloud storage.

Remote data verification allows the client to test the integrity of outsourcing data on an untrusted server. Proof of retrievability(POR) is the integrity verification algorithm proposed by Juels[2], the key method of the POR is to add some random data blocks to the stored data, its insertion position is determined by a pseudo-random sequence, and uses error-correcting codes. These data blocks and the stored data itself does not have any relationship, which are called sentinels, and these sentinels play in the role of tag. The tag is used to test the integrity of data. Provable data possession(PDP) is proposed by Ateniese[3], which has two notable characteristics, one is supporting public verification, another is using homomorphic signature algorithm[4]. These two characteristics make it easy for a third party to verify whether the data is integrity on the server.

Data possession checking(DPC) is proposed by Da Xiao[5], the basic idea is that the verifier randomly assigns several data blocks and its corresponding key, the server computes the hash value and returns to the verifier, then the verifier compares the hash value is consistent with the check block, thereby the verifier can determine whether the data is correctly held. The literature [6] proposed integrity checking for remote data based on RSA hash function. Let N be moduli of RSA, F be big integer representing file, the verifier saves $k = F \bmod \varphi(N)$. During the challenge, the verifier sends $g \in Z_N$, then the server returns $s = g^F mod\ N$, the verifier checks whether the equation $g^k mod\ N = s$ is satisfied.

Scalable provable data possession(SPDP) is proposed by Ateniese[7], the difference between SPDP algorithm and PDP algorithm is that the SPDP algorithm is supporting dynamic data. SPDP algorithm adopts label organization as the agreed number of tag, then encrypts the generating tag with a symmetric key and stores it in the server or local. Each challenge uses a tag to verify, but the scheme also restrict the number of verification. Dynamic provable data possession(DPDP) is proposed by Erway[8], the DPDP algorithm uses skip list which is like tree structure to generate tag, compared with SPDP algorithm, the DPDP algorithm is supporting dynamic data integrity verification. Chen proposed another algorithm[9] based DPDP algorithm, which uses RS code and Cauchy matrix

to intensify the robust and dynamic updates of the initial algorithm. In addition, there are integrity verification scheme IPDP in the private cloud and integrity verification CPDP[10] in the hybrid cloud.

In summary, the existing schemes have some drawbacks as follows: (1) most schemes based on public key cryptography are computationally expensive, especially when dealing with large volume of data. (2) Can only perform a limited number provable data possession. (3)Some schemes do not apply to the cloud storage service environment.

This paper proposes a provable data possession scheme based on homomorphic hash function according to the problems exist in the above algorithms. This paper uses the homomorphic hash function in literature [11], proposes a data integrity verification scheme based on homomorphic hash function supporting dynamic data and unlimited challenges. The advantage of homomorphic hash function is that it provides provable data possession and data integrity protection. The scheme is a good way to ensure the integrity of remote data and reduce redundant storage space and bandwidth consumption on the premise that users do not retrieve data. The main cost of the scheme is on the server side, it is suitable for mobile devices in the cloud storage environments. We prove the scheme is feasible by analyzing the security and performance of the scheme.

The remainder of this paper is organized as follows. Section 2 discusses theoretical preliminaries for the presentation. Section 3 describes provable data possession scheme based on homomorphic hash function. Section 4 analyzes the security of provable data possession scheme. Section 5 analyzes the performance of provable data possession scheme. We conclude in Section 6.

2. Preliminaries

We now recapitulate some essential concepts from homomorphic hash function.

2.1 Symbol of Homomorphic Hash Function

Firstly we explain some parameters of homomorphic hash function, all the parameters are generated in the setup stage, as is shown in Table 1.

TABLE 1
SYSTEM PARAMETERS AND PROPERTIES

Name	Description	e.g.
λ_p	discrete log security parameter	1024 bit
λ_q	discrete log security parameter	257 bit
p	random prime,$\lvert p\rvert=\lambda_p$	
q	random prime,$q \mid p-1, \lvert q\rvert=\lambda_q$	
β	block size in bits	16 KB
m	number of "sub-blocks" per block	512 bit
	$m=\lceil \beta/(\lambda_q-1) \rceil$	
g	$1 \times m$ row vector of order q elts in Z_p	
G	hash parameters, given by (p, q, g)	
n	number of file blocks	
seed	seed of key stream generator	
MAXR	maximum possible output of R(•)	
MINR	minimum possible output of R(•)	

In this paper, we uses the following two cryptographic primitives:

$H_G(\bullet):\{0,1\}^k \times \{0,1\}^\beta \rightarrow \{0,1\}^{\lambda_p}$

$R(\bullet):\{0,1\}^k \times \{0,1\}^l \rightarrow \{0,1\}^l$

Among which, k is the key length, $H_G(\bullet)$ is homomorphic hash function, $R(\bullet)$ is pseudo-random function, is used as pseudo-random generator.

2.2 Homomorphic Hash Algorithm

In algebra, homomorphism is the constant mapping between two algebraic structures, such groups, rings, fields or vector space. That is to say there exist mapping $\Phi: X \rightarrow Y$, satisfying $\Phi(x+y)=\Phi(x) \times \Phi(y)$, + is the operator of set X, and × is the operator of the set Y.

Generally, public key cryptography algorithm has the characteristic of homomorphism, for example, RSA algorithm has homomorphism for multiplication. Let k be public key, n be moduli. The encryption algorithm is $E_k(\cdot)$, for message x and y, we have the following:

$$E_k(x \times y)=(x \times y)^k \, mod \, n=(x^k mod \, n) \times (y^k mod \, n)=E_k(x) \times E_k(y) \quad (1)$$

If some algorithms satisfy the homomorphism for addition, then we have the equation:

$$E_k(x+y)=E_k(x)+E_k(y) \quad (2)$$

Some algorithms satisfy the homomorphism for multiplication, such as RSA. Some algorithms satisfy the homomorphism for addition, such as Paillier. If some algorithms satisfy homomorphism for addition and multiplication, then they are called full homomorphism algorithms. There is no genuine full homomorphic encryption algorithms available at present[12]. Homomorphic hash function means that the hash function has the characteristic of homomorphism. The homomorphic hash function used in this paper is based on literature[11].

Files F represented by $m \times n$ matrices, all the elements in the matrices belong to Z_q, because of $m=\lceil \beta/(\lambda_q-1) \rceil$, so every element less than $2^{\lambda_q}-1$, therefor less than q.

$$F=(b_1, b_2, \cdots, b_n)=\begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{pmatrix} \quad (3)$$

We add two blocks by adding their corresponding column-vectors. That is, to combine the i^{th} and j^{th} blocks of the file, we simply compute:

$$b_i + b_j = (b_{1,i} + b_{1,j}, \cdots, b_{m,i} + b_{m,j}) mod\ q \quad (4)$$

For file F, the computation of hash value is as following, firstly computes the hash value of each data block:

$$H_G(b_j) = \prod_{t=1}^{m} {g_t}^{b_{t,j}}\ mod\ p \quad (5)$$

The hash value of file F is $1 \times n$ row vector, every element in the row vector is the hash value of each block of the file:

$$H_G(F) = (H_G(b_1), H_G(b_2), \cdots, H_G(b_n)) \quad (6)$$

From the calculation process of each block, we can obtain the homomorphism of hash function:

$$\begin{aligned} H_G(b_i + b_j) &= \prod_{t=1}^{m} {g_t}^{b_{t,i}+b_{t,j}}\ mod\ p \\ &= \prod_{t=1}^{m} {g_t}^{b_{t,i}}\ {g_t}^{b_{t,j}}\ mod\ p \\ &= \prod_{t=1}^{m} {g_t}^{b_{t,i}}\ mod\ p \times \prod_{t=1}^{m} {g_t}^{b_{t,j}}\ mod\ p \\ &= H_G(b_i) \times H_G(b_j) \end{aligned} \quad (7)$$

Among which, each block is β bit, hash length is λ_p bit, if we choose β = 16 KB, $\lambda_p = 1024$ bit, then after the hash value of the file is calculated, the expansion of the file is $\frac{\lambda_p}{\beta} = \frac{1024}{16 \times 1024 \times 8} \approx 0.0078$.

2.3 Per-Publisher Homomorphic Hashing

The per-publisher hashing scheme is an optimization of the global hashing scheme just described. In the per-publisher hashing scheme, a given publisher picks group parameters G so that a logarithmic relation among the generators g isknown. The publisher picks q and p as above, but generates g by picking a random $g \in Z_p$ of order q, generating a random vector r whose elements are in Z_q and then computing $\boldsymbol{g} = g^r$.

Given the parameters g and r, the publisher can compute file hashes with many fewer modular exponentiations:

$$H_G(F) = g^{rF} \quad (8)$$

The publisher computes the product rF first, and then performs only one modular exponentiation per file block to obtain the full file hash. The hasher must be careful to never reveal g and r; doing so allows an adversary to compute arbitrary collisions for H_G.

3. Provable Data Possession based on Homomorphic Hash Function

3.1 Scheme Description

The purpose of provable data possession is to allow the user to verify whether the untrusted storage server holds data correctly. Generally, there are two parties: client and storage server. The scheme of provable data possession base on homomorphic hash function is composed of five phases: (1)Setup; (2)TagBlock; (3)Challenge; (4) ProofGen; (5) ProofVerify.

Firstly, we need to divide the file F into n blocks. In the following phases such as TagBlock phase and ProofVerify phase, all the calculations are based on the file blocks.

1. Setup Phase:

In the *Setup* phase, the input value is $(\lambda_p, \lambda_q, m, s)$, the output value is $G = (p, q, g)$. G is hash parameters, used in homomorphic hash function to produce hash value. *Setup* phase is described as Fig 1.

$Setup(\lambda_p, \lambda_q, m, s) \rightarrow G = (p, q, g)$
Seed PRNG R with s.
do
  q ← qGen(λ_q)
  p ← pGen(q, λ_p)
while p=0 **done**
for(i=1 **to** m) **do**
  do
    x ← R(p − 1) + 1
    $g_i \leftarrow x^{(p-1)/q}\ (mod\ p)$
  while $g_i = 1$ **done**
done
return (p,q,g)

$qGen(q, \lambda_q)$
do
  q ← R(2^{λ_q})
while q is not prime **done**
return q

$pGen(q, \lambda_p)$
for(i=1 to $4\lambda_p$) **do**
  X ← R(2^{λ_p})
  c ← X(mod 2q)
  p ← X − c + 1
  if p is prime **then return** p
done
return 0

Fig. 1Setup phase

2. TagBlock Phase:

In the TagBlock phase, the client uses pseudo-random generator to generate a series of pseudo-random numbers, then multiply each block of the file F with the corresponding pseudo-random number, and obtain the tag t_i of each block b_i. The client sends the b_i, tag t_i ,p, q to the server, the client saves the hash parameters G and the seed of pseudo-random generator. The detail is described in Fig 2.

TagBlock(G,F)
G=(p,q,g),F=$(b_1, b_2, \cdots, b_n)$
for(j=1 to n) **do**

$$x_j = H_G(b_j) = \prod_{t=1}^{m} g_t{}^{b_{t,j}} \ mod\ p$$

$$r_j = R(seed)$$

done
then Tag=$[x_1 \cdot r_1, x_2 \cdot r_2, \cdots, x_n \cdot r_n]$
return Tag=$[t_1, t_2, \cdots, t_n]$,$t_j = x_j \cdot r_j$
the client sends F,Tag,p,q to the server
saves G and seed

Fig. 2 TagBlock phase

3. Challenge Phase:

In the Challenge phase, the client uses pseudo-random generator to generate k challenge blocks to the server. The detail is described in Fig 3.

Challenge()
the client select k blocks to challenge randomly:
for (j=1 **to** k) **do**

$$r_j' = \lceil rand(time())/(MAXR - MINR) \times n \rceil$$

done
then the client sends<$r_1', \cdots, r_k', k$>to the server

Fig. 3 Challenge phase

4. ProofGen Phase:

In the ProofGen phase, the server calculates b_c and t_c using each block and its corresponding tag, then returns the b_c and t_c to the client. The detail is described in Fig 4.

ProofGen($r_1', \cdots, r_k', k$)

$$b_c = \left(\sum_{i=r_1'}^{r_k'} b_i \right) mod\ q$$

$$t_c = \left(\prod_{i=r_1'}^{r_k'} t_i \right) mod\ p$$

return(b_c, t_c)

Fig. 4 ProofGen phase

5. ProofVerify Phase:

In the ProofVerify phase, the client uses seed to reproduce the corresponding pseudo-random numbers, then verify whether the t_c is exactly the t_c that the client specified. Also verify that the t_c is corresponding to the b_c. The detail is described in Fig 5.

ProofVerify(b_c, t_c)
The client verifies b_c, recalls R(seed) to produce

$$(r_1, r_2, \cdots, r_n)$$

verify:

$$t_c \stackrel{?}{=} H_G(b_c \times r_{r_1'} \times \cdots \times r_{r_k'})$$

If the equation holds, it indicates that the file is intact.

Fig. 5 ProofVerify phase

There are two approaches for provable data possession, one is verified by data owner, and another is verified by a trusted third party. To delegate the work of provable data possession to a trusted third party has the following advantage, when a dispute is emerging, such as the service provider believes that it stores the data, but the data may be placed on secondary storage or offline storage. But the user demands that the server should provide online access, claiming that the performance does not meet the requirements. So it can be arbitrated by a trust third party.

When we use a third party to audit, we should provide privacy protection technology, that is to say, don't disclose the data to a third party. We can use the following privacy protection approaches :(1) First encrypt data and then calculate the relevant verification information, we use the encrypted data during verification, so won't disclose data; (2) Because we use sampling, the response of sampling is not continuous data, not returning the original data, but returning the verification information of original data. (3) Using a general method of privacy protection, add some random data in the data, this method will add extra cost. We will research on the privacy protection technology when a third party audit in the future.

3.2 Supporting Dynamic Data

Supporting dynamic data mainly consists of two operations: insert and delete.

1. Insert Data Block

Assume that insert data block b_s. The client sends to server for insert request, then the client receives from server: (F, Tag). The client will calculate the tag of the insert block b_s, and the tags of the s-th block after. Then send the updated F′ and Tag′ to the server. Then issue immediately the verification of the data block in order to ensure that the data uploaded is correct. The detail process is described in Fig 6.

Insert(b_s)
for(j=1 **to** n+1) **do**
$r_j = R(seed)$
if(j≥ s)
$x_j = H_G(b_j) = \prod_{t=1}^{m} g_t{}^{b_{t,j}} \ mod\ p$
$t_j = x_j \times r_j$
done
return $Tag' = [t_1, t_2, \cdots, t_{n+1}]$
client sends to server:F',Tag'

Fig. 6 Insert data block

2. Delete Data Block

Assume that delete data block b_k. The client sends to server for delete request, then the client receives from server: (F, Tag). The client will calculate the tags of the k-th block after. Then send the updated F' and Tag' to the server. Then issue immediately the verification of the data block in order to ensure that the data uploaded is correct. The detail process is described in Fig 7.

Delete(b_k)
for(j=1 **to** n-1) **do**
$r_j = R(seed)$
if(j≥ k)
$x_j = H_G(b_j) = \prod_{t=1}^{m} g_t{}^{b_{t,j}} \ mod\ p$
$t_j = x_j \times r_j$
done
return $Tag' = [t_1, t_2, \cdots, t_{n+1}]$
client sends to server:F',Tag'

Fig. 7 Delete data block

4. Security Analysis

4.1 Security of Homomorphic Hash Function

The homomorphic hash function used in this paper is based on discrete logarithm assumption. We analyze whether a probabilistic polynomial time(PPT) adversary can find a pair collision in probabilistic polynomial time. We use the method in literature [13] to define homomorphic hash function. A hash function family is defined by PPT algorithm $Q = (Hgen, H)$. Hgen represents a hash generator, input security parameters $(\lambda_p, \lambda_q, m)$, output a member G of hash function family. For hash function H_G, input the data of length $m\lambda_q$, output the hash value of length λ_p. $\mathcal{A}$ is a PPT adversary trying to a pair collision of the given hash function family.

Definition 1 For any hash function family Q, any PPT adversary $\mathcal{A}$, security parameters $\lambda = (\lambda_p, \lambda_q, m)$, and $\lambda_q < \lambda_p$, $m \le poly(\lambda_p)$, s.t. $Adv_{Q,\lambda} = Pr[G \leftarrow Hgen(\lambda); (x_1, x_2) \leftarrow \mathcal{A}(G): H_G(x_1) = H_G(x_2) \wedge x_1 \neq x_2]$

If PPT adversary $\mathcal{A}$'s time complexity is $\tau(\lambda)$, $Adv_{Q,\lambda}(\mathcal{A}) < \varepsilon(\lambda)$, $\varepsilon(\lambda)$ is a neglect function, $\tau(\lambda)$ is the polynomial of λ, then Q is a secure hash function.

The homomorphic hash function H_G in this paper is satisfying definition 1. If to construct discrete logarithm problem in finite field through parameters (λ_p, λ_q) is hard, then H_G is a collision-resist hash function. The detail provable process refer to literature [13].

4.2 Security of Pseudo-random Generator

In our scheme, the data blocks and their corresponding tags, p, q are saved in the server. The seed for generating pseudo-random numbers, p, q, g are saved in the client. The tags are verifiable, the construction of tags is using homomorphic hash algorithm and a series of pseudo-random numbers.

In the Challenge phase, the challenger generates k challenge blocks randomly, then send the k blocks to adversary $\mathcal{A}$, $\mathcal{A}$ generates integrity verification P, if P passes the validation, then $\mathcal{A}$ performs a successful deception. If $\mathcal{A}$ deletes the challenge blocks, then sends arbitrary data blocks and its corresponding tags to challenger. At this point, though the return value b_c can be verified is correct corresponding to t_c, $\mathcal{A}$ doesn't know the random numbers r_i used in constructing tags, so the challenger hash the data blocks he has received, using the same seed to generating random numbers, then computes the tags, compared with the tags $\mathcal{A}$ has returned, then you can verify whether the data blocks and tags $\mathcal{A}$ has returned are designated by challenger.

5. Performance Analysis

First, we analyze the performance cost of each phase in provable data possession. We convert all the exponentiations into multiplications. We denote the multiplication cost in Z_p^* as MultCost(p). For calculatingy^x, we need $1.5|x|$ times multiplications using Iterative Square method. First calculate ${y_i}^{2^z}$, build a list for ${y_i}^{2^z} (1 \le z \le \lambda_p)$, need $|x|$times multiplications, then looking for a list needs $|x|/2$ multiplications. During the process of calculating hash value, we all need to look for a list ${y_i}^{2^z}$. So the list is built in the Setup phase. To simplify the performance analysis, we will ignore the computation cost during the following analysis.

In the Setup phase, generating the key G of homomorphic hash function, relating to the random number generation and modulus exponentiation, for

parameters p and q, mainly uses a random number generator and a simple prime testing, for parameter g, needs $m(p-1)/2q$ mod p multiplications, its cost is $m(\lambda_p - 1)\text{MultCost}(p)/2\lambda_q$. However, these parameters are only generated once. For any approach of provable data possession, these parameters are indispensable and their cost is almost the same. In the TagBlock phase, the size of the data block β is 16KB, the output of homomorphic hash function is 1024 bit, so the hash function reduces the storage space of file to its original $\frac{\lambda_p}{\beta} = \frac{1024}{16\times1024\times8} = \frac{1}{128}$. This method of tag organization is very helpful in reducing storage redundancy. We need to compute the hash value of each data block, relating to $nm|p|/2$ mod p multiplications, its cost is $nm\lambda_p MultCost(p)/2$. In the Challenge phase, has the cost of generating two random numbers. In the ProofGen phase, has k times mod q additions, also has k times mod p multiplications, here the cost of multiplication is large, is $cMultCost(p)/2$. In the ProofVerify phase, has one time homomorphic hash calculation, relating to m times mod p multiplications, its cost is $m\lambda_p MultCost(p)/2$.

In practical use of cloud storage service, performance is always limited by the network bandwidth. Modular multiplication algorithm can be optimized, the optimization method refer to literature [11]. Optimized performance can improve more than 4 times. Zhao et al. [14] proposed the use of the graphics processing unit[15,16] to accelerate the performance of homomorphic hash function. We can also use this method to improve the performance in our scheme.

6. Conclusions

This paper proposes a provable data possession scheme based on homomorphic hash function according to the problems exist in the above algorithms. This method allows users to verify data integrity on the server for unlimited number of times.It also provides provable data possession on the server and data integrity protection. Users only need to save key G, transmission information is little during the verification process, and the verification of provable data possession is just one time homomorphic hash calculation. Through security analysis and performance analysis shows that the method is feasible. The scheme can achieve data recovery. We can use error-correcting code or erasure code to encode data before calculating hash value.

Acknowledgments

We are grateful to the anonymous referees for their invaluable suggestions. This work is supported by the National Natural Science Foundation of China (Nos. 61472472). This work is also supported by JiangXi Education Department (Nos. GJJ14650 and GJJ14642).

References

[1] Feng Dengguo, Zhang Min, Zhang Yan, et al. Study on cloud computing security. Journal of Software, 2011, 22(1): 71-83

[2] Juels A，Kaliski B S，Por S. Proof of rerievability for large files. Proc of the 14th ACM Conf on Computer and Communications Security. New York: ACM， 2007: 584-597.

[3] Ateniese G，Burns R，Curtmola R， et al. Provable data possession at untrusted stores. Proc of the 14th ACM Conf on Computer and Communications Security. New York: ACM， 2007: 598-609.

[4] Johnson R, Molnar D, Song D, et al. Homomorphic signature schemes. Proc of CT-RSA. New York: Springer, 2002:244-262.

[5] Da Xiao，Jiwu Shu，Kang Chen,et al. A practical data possession checking scheme for networked archival storage. Journal of Computer Research and Development, 2009, 46(10):1660-1668.

[6] Deswarte Y, Quisquater J J, and Saidane A. Remote integrity checking. Proc of IICIS'03, Switzerland, Nov.13-14, 2003: 1-11.

[7] Ateniese G, Dipr, Mangini L V, et al. Scalable and efficient provable data possession. Proc of the 4th International Confon Security and Privacy in Communication Netowrks (SecureComm 2008). New York: ACM, 2008:1-10.

[8] Erway C，Kupcu A，Papamanthou C，et al. Dynamic provable data possession. Proc of the 16th ACM Conf on Computer and Communications Security (CCS 2009). New York: ACM,2009: 213-222.

[9] Chen B， Curtmola R. Robust dynamic provable data possession. Distributed Computing Systems Workshops (ICDCSW), 32nd International Conference on . Macau: IEEE, 2012: 515-525.

[10] Zhu Y, Wang H, Hu Z, et al. Cooperative provable data possession. Beijing: Peking University and Arizona University，2010.

[11] Krohn M, Freedman M J， Mazieres D. On-the-fly verification of rateless erasure codes for efficient content distribution. Proc of IEEE Symposium on Security and Privacy. Lee Badger: IEEE, 2004: 226-240.

[12] Bruce Schneier. Schneier on Security. http://www.schneier.com/blog/archives/2009/07/homomorphic_enc.html.

[13] Bellare M, Goldreich O, and Goldwasser S. Incremental cryptography: the case of hashing and signing. Advances in Cryptology-CRYPTO'94, Santa Barbara, CA, Aug. 1994: 216-233.

[14] Zhao Kaiyong, Chu Xiaowen, Wang Mea. Speeding up homomorpic Hashing using GPUs. The 2009 (44th) IEEE Conference on Communication (ICC 2009), Dresden, Germany, June 14-18, 2009: 1-5.

[15]Bowers K D, Juels A, and Oprea A. HAIL: a high-vailability and integrity layer for cloud storage. Proceedings of ACMCCS'09, Chicago, Illinois, USA, Nov. 9-13, 2009: 187-198.

[16] Shacham H and Waters B. Compact proofs of retrievability. Proceedings of ASIACRYPT '08, Melbourne, Australia, Dec.7-11, 2008: 90-107.

Li Yu received her M.S. degree from JXAU University, Nanchang, China, in 2004. His research interests include Cryptography, Information Security, Space Information Networks and Internet of Things.

Junyao Ye is a Ph.D. student in Department of Computer Science and Engineering, Shanghai JiaoTong University, China. His research interests include information security and code-based cryptography.

Kening Liu received his M.S. degree from Minzu University of China, Beijing, China, in 2001. His research interests include subliminal channel, LFSR, code-based systems and other new cryptographic technology.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Summarization of Various Security Aspects and Attacks in Distributed Systems: A Review

Istam Shadmanov [1], Kamola Shadmanova [2]

[1,2] Information Technology, Bukhara State University, Uzbekistan
istam.shadmanov89@gmail.com, kamola.shadmanova94@gmail.com

Abstract

The modern world is filled with a huge data. As data is distributed across different networks by the distributed systems so security becomes the most important issues in these systems. Data is distributed via public networks so the data and other resources can be attacked by hackers. In this paper we define different security aspects including on the basis of authorization, authentication, encryption and access control for distributed systems.

Keywords: *distributed system, security issues, security threats, integrity, security services.*

1. Introduction

Nowadays, security threats are a growing concern since the complexity of the collaborative processes increases. Another way of looking at security in systems is that we attempt to protect the services and data from various threats and attacks, which are discussed in a section 2. The objective of this paper is to understanding of the various threats and security aspects associated with distributed system.

Security issues in modern distributed systems can roughly be divided into two parts. [1] The first part is the communication between users or processes, possibly residing on different machines. The principal mechanism for ensuring secure communication is a secure channel and more specifically, authentication, message integrity, and confidentiality. Confidentiality refers to the property of a computer system whereby its information is disclosed only to authorized parties. Integrity is the characteristic that alterations to a system's assets can be made only in an authorized way. In other words, improper alterations in a secure computer system should be detectable and recoverable. The second part is the authorization, which deals with ensuring that a process gets only those access rights to the resources in a distributed system it is entitled to. Authorization is covered in a separate section 3, dealing with access control. Secure channels and access control require mechanisms to distribute cryptographic keys, but also mechanisms to add and remove users from a system. Before starting our description of security in distributed systems, it is necessary to define what a secure system is. Security system, that enforces boundaries between computer networks, consisting of a combination of hardware and software that limits the exposure of a computer or computer network from attack.

The systems security requires the existence of the followings:

Confidentiality: Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information;

Integrity: that supposes avoiding data corruption and keeping data integrity;

Availability: which means ensuring that data and applications can always be accessed, regardless of any interferences, to authorized entities. [5]

These three concepts embody the fundamental security objectives for both data and for information and computing services.

2. Threats and attacks in distributed system

A threat is a potential for violation of security, which exists when there is a circumstance, capability, action, or event that could breach security and cause harm. That is, a threat is a possible danger that might exploit vulnerability. A threat can be either "intentional" (i.e., intelligent; e.g., an individual cracker or a criminal organization) or "accidental" (e.g., the possibility of a computer malfunctioning).

There are some types of security threats to consider:

-interception, the concept of interception refers to the situation that an unauthorized party has gained access to a service or data. Interception happens when data are illegally copied, for example, after breaking into a person's private directory in a file system;

-interruption, refers to the situation in which services or data become unavailable, unusable, destroyed, and so on;

-modification, include intercepting and subsequently changing transmitted data, tampering with database entries and changing a program;

-fabrication, refers to the situation in which additional data or activity that would normally not exist are generated.

An attack on system security can drives from an intelligent threat; that is, an intelligent act that is a deliberate attempt to evade security services and violate the security policy of a system.

2.1 Classes of attacks in distributed system

Classes of attack might include passive monitoring of communications, active network attacks, close-in attacks, exploitation by insiders, and attacks through the service provider. The following are common types of network attacks in distributed system as, A Passive Attack, An Active Attack, Syn Flood Attack, Password Attack, Distributed Attack, An Insider Attack and Phishing Attack:

A passive attack usually monitors unencrypted traffic and looks for clear-text passwords and sensitive information that can be used in other types of attacks. Passive attacks include traffic analysis, monitoring of unprotected communications, decrypting weakly encrypted traffic, and capturing authentication information such as passwords. Passive interception of network operations enables adversaries to see upcoming actions.

In an active attack, the attacker tries to bypass or break into secured systems. This can be done through stealth, viruses, worms or Trojan horses. Active attacks include attempts to circumvent or break protection features, to introduce malicious code, and to steal or modify information. These attacks are mounted against a network backbone, exploit information in transit, electronically penetrate an enclave, or attack an authorized remote user during an attempt to connect to an enclave. Active attacks result in the disclosure or dissemination of data files or modification of data.

A distributed denial of service (DDoS) attack attempts to consume the target's resources so that it cannot provide service. One way to classify DDoS attacks is in terms of the type of resource that is consumed. Broadly speaking, the resource consumed is either an internal host resource on the target system or data transmission capacity in the local network to which the target is attacked. A simple example of an internal resource attack is the SYN flood attack. [6]

Figure 2.1 shows the steps involved:

1. The attacker takes control of multiple hosts over the Internet, instructing them to contact the target Web server.
2. The slave hosts begin sending packets, with erroneous return IP address information, to the target.
3. Each packet is a request to open a TCP connection. For each such packet, the Web server responds with a SYN/ACK (synchronize/acknowledge) packet, trying to establish a TCP connection with a TCP entity at a spurious IP address. The Web server maintains a data structure for each SYN request waiting for a response back and becomes bogged down as more traffic floods in. The result is that legitimate connections are denied while the victim machine is waiting to complete bogus "half-open" connections.

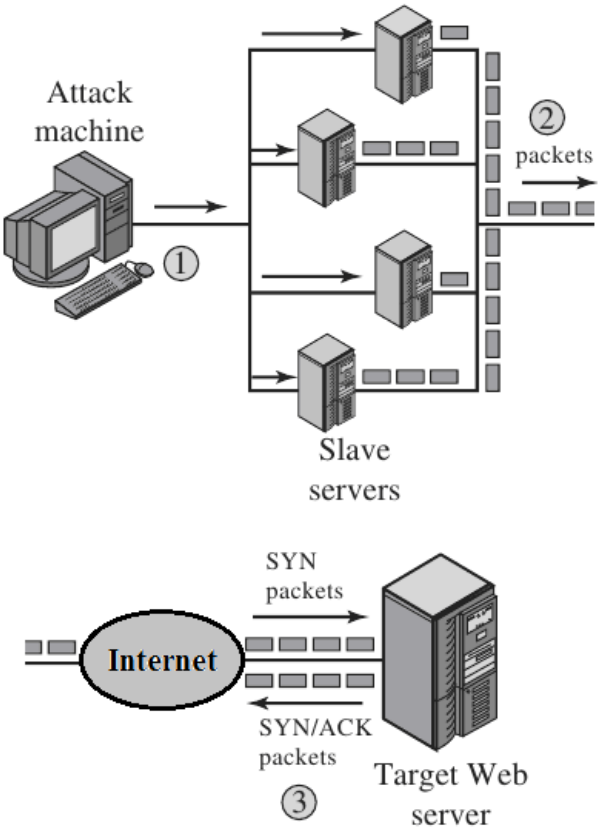


Figure 2.1 Distributed SYN flood attack

Password attack, an attacker tries to crack the passwords stored in a network account database or a password-protected file. There are three major types of password attacks: a dictionary attack, a brute-force attack, and a hybrid attack. A dictionary attack uses a word list file, which is a list of potential passwords. In brute-force attack the attacker tries every possible combination of characters.

A distributed attack requires that the adversary introduce code, such as a Trojan horse or back-door program, to a "trusted" component or software that will later be distributed to many other companies and users Distribution attacks focus on the malicious modification of hardware or software at the factory or during distribution. These attacks introduce malicious code such as a back door to a product to gain unauthorized access to information or to a system function at a later date.

An insider attack involves someone from the inside, such as a disgruntled employee, attacking the network Insider attacks can be malicious or no malicious. Malicious insiders intentionally eavesdrop, steal, or damage information; use information in a fraudulent manner; or deny access to other authorized users. No malicious attacks typically result from carelessness, lack of knowledge, or intentional circumvention of security for such reasons as performing a task.

In phishing attack, the hacker creates a fake web site that looks exactly like a popular web site companies such as Facebook, Hotmail, Yahoo or consumer products companies. The phishing part of the attack is that the hacker then sends an e-mail message trying to trick the user into clicking a link that leads to the fake site. [2] When the user attempts to log on with their account information, the hacker records the username and password and then tries that information on the real site. An example of a phishing email posted as a warning on Microsoft's web site is shown in Figure 2.2.

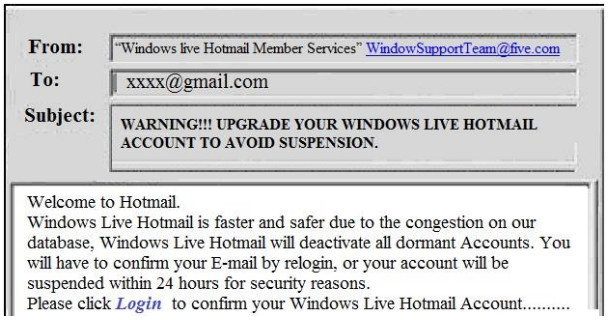


Figure 2.2 Example Phishing Email Message

Sometimes the phishing email advises the recipient of an error, and the message includes a link to click to enter data about an account. The link, of course, is not genuine, its only purpose is to solicit account names, numbers and authenticators.

Apart from attacks originated from external parties, many break-ins occurs due to poor information security policies and procedures, or internal misuse of information systems. Also, new security risks could arise from evolving attack methods or newly detected holes in existing software and hardware. But a system must be able to limit damage and recover rapidly when attacks occur.

2.2 Security Issues in Distributed system

The informatics security is an important issue that must be analyzed in order to identify security requests, to discover possible vulnerabilities or threats and to avoid loss of information [4].

For enforcement of security, the distributed system must have the following additional requirements:

- It should be possible for the sender of a message to know that the message was received by the intended receiver;
- It should be possible for the receiver of a message to know that the message was sent by the genuine sender;
- It should be possible for both the sender and receiver of a message to be guaranteed that the contents of the message were not changed while it was in transfer.

The secured implementation of distributed systems has been generated lot of critical issues. Some of these are as follows:

- Identification of methodology which access the security level in any system;
- Application of middleware in distributed system security;
- Application of web services in security purposes;
- Monitoring of the system security;
- Development of security metrics;
- Integration of techniques, like Cryptography etc., for secure distributed data communication.

3. Methods of the solution on security issues

In the context of distributed systems security is oriented on collaborative side, which means that security components cooperate to achieve a common goal, represented by vulnerabilities elimination. There are some broad areas of security in distributed systems:

- Authentication
- Access Control
- Data Confidentiality
- Data Integrity
- Encryption
- Digital Signature
- Nonrepudiation

Authentication. A fundamental concern in building a secure distributed system is a authentication of local and remote entities in the system. The authentication service is concerned with assuring that a communication is authentic. In the case of a single message, such as a warning or alarm signal, the function of the authentication service is to assure the recipient that the message is from the source that it claims to be from. In the case of an ongoing interaction, such as the connection of a terminal to a host, two aspects are involved. First, at the time of connection initiation, the service assures that the two entities are authentic, that is, that each is the entity that it claims to be. Second, the service must assure that the connection is not interfered with in such a way that a third party can masquerade as one of the two legitimate parties for the purposes of unauthorized transmission or reception.

Access Control. In the context of network security, access control is the ability to limit and control the access to host systems and applications via communications links and the prevention of unauthorized use of a resource. To achieve this, each entity trying to gain access must first be identified, or authenticated, so that access rights can be tailored to the individual.

Data Confidentiality. Confidentiality is the protection of transmitted data from passive attacks and unauthorized disclosure. With respect to the content of a data transmission, several levels of protection can be identified. The broadest service protects all user data transmitted between two users over a period of time. The other aspect of confidentiality is the protection of traffic flow from analysis. This requires that an attacker not be able to observe the source and destination, frequency, length, or other characteristics of the traffic on a communications facility.

Data Integrity. As with confidentiality, integrity can apply to a stream of messages, a single message, or selected fields within a message. Again, the most useful and straightforward approach is total stream protection. A connection - oriented integrity service, one that deals with a stream of messages, assures that messages are received as sent with no duplication, insertion, modification, reordering, or replays. The destruction of data is also covered under this service. On the other hand, a connectionless integrity service, one that deals with individual messages without regard to any larger context, generally provides protection against message modification only. The integrity service relates to active attacks, we are concerned with detection rather than prevention. If a violation of integrity is detected, then the service may simply report this violation.

Encryption. The use of mathematical algorithms to transform data into a form that is not readily intelligible. The transformation and subsequent recovery of the data depend on an algorithms and encryption keys. A security related transformation on the information to be sent. Examples include the encryption of the message, which scrambles the message so that it is unreadable by the opponent, and the addition of a code based on the contents of the message, which can be used to verify the identity of the sender. (Figure 4.1)

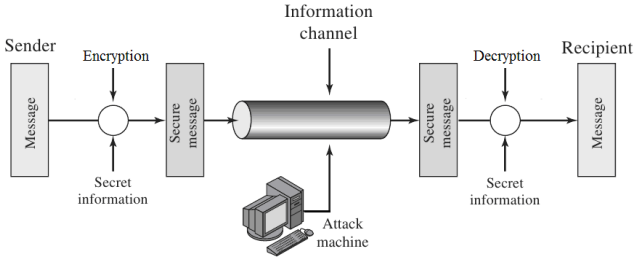


Figure 3.1 Example For Message Security

Digital Signature. Data appended to, or a cryptographic transformation of, a data unit that allows a recipient of the data unit to prove the source and integrity of the data unit and protect against forgery e.g., by the recipient. Digital information can be signed, producing digital certificates. Certificates enable trust to be established among users and organizations. [3]

Nonrepudiation. Nonrepudiation prevents either sender or receiver from denying a transmitted message. Thus, when a message is sent, the receiver can prove that the alleged sender in fact sent the message. Similarly, when a message is received, the sender can prove that the alleged receiver in fact received the message. In such cases a notary is used to register messages, that neither of the participants can not back out of a transaction and disputes can be resolved by presenting relevant signatures or encrypted text.

Security aspects come into play when it is necessary or desirable to protect the information transmission from an opponent who may present a threat to confidentiality, authenticity, and so on.

4. Conclusion

In order that the users can trust the system and rely on it, the various resources of a computer system must be protected against destruction and unauthorized access. Enforcing security in a distributed system is more difficult than in a centralized system because of the lack of a single point of control and the use of insecure networks for data communication. Authentication, access control, notarisation, data confidentiality, data integrity, digital signature etc. are well-studied and used methods of secure distributed systems.

References

[1] Andrew S. Tanenbaum, Maarten Van Steen, “Distributed systems Principles and Paradigms”, 2nd ed., Upper Saddle River, NJ, USA: Pearson Higher Education, 2007, pp. 377-389.

[2] Charles P. Pfleeger, Shari Lawrence Pfleeger, Jonathan Margulies, “Security in computing”, 5th ed., USA, Pearson Education, Inc., Massachusetts, January 2015, pp. 300-304.

[3] George Coulouris, Jean Dollimore and Tim Kindberg, “Distributed System- Concepts and design”, 5th ed., USA, Addison- Wesley, Massachusetts, 2012, pp. 481-483.

[4] I. Ivan, C. Ciurea, “Security of Collaborative Banking Systems”, Proceedings of the 4th International Conference on Security for Information Technology and Communications, November 17-18, 2011, Bucharest, Romania.

[5] Manoj Kumar, Nikhil Agrawal, “Analysis of Different Security Issues and Attacks in Distributed System A-Review”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, April 2013.

[6] William Stallings, “Cryptography And Network Security Principles And Practice”, 5th ed., NY, Pearson Education, Inc., 2011, pp. 809-812.

First Author -ISTAM UKTAMOVICH SHADMANOV

1. 2013-UP TO NOW - TEACHER OF THE DEPARTMENT OF INFORMATION TECHNOLOGY
2. 2011-2013 MASTER`S DIPLOMA ON APPLIED MATHEMATICS AND INFORMATION TECHNOLOGY
3. 2007-2011 BACHELOR`S DIPLOMA ON APPLIED MATHEMATICS AND INFORMATICS
4. NUMBER OF PAPERS- 9.

Second Author-KAMOLA SHADMANOVA UMED QIZI

1. 2012-UP TO NOW – STUDENT OF THE DEPARTMENT OF INFORMATION TECHNOLOGY
2. NUMBER OF PAPERS- 2.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Persons' Personality Traits Recognition using Machine Learning Algorithms and Image Processing Techniques

Kalani Ilmini[1] and TGI Fernando[2]

[1] Department of Computer Science, University of Sri Jayewardenepura, Nugegoda, Sri Lanka
ilmini@dscs.sjp.ac.lk

[2] Department of Computer Science, University of Sri Jayewardenepura, Nugegoda, Sri Lanka
gishantha@dscs.sjp.ac.lk

Abstract

The context of this work is the development of persons' personality recognition system using machine learning techniques. Identifying the personality traits from a face image are helpful in many situations, such as identification of criminal behavior in criminology, students' learning attitudes in education sector and recruiting employees.

Identifying the personality traits from a face image has rarely been studied. In this research identifying the personality traits from a face image includes three separate methods; ANN, SVM and deep learning. Face area of an image is identified by a color segmentation algorithm. Then that extracted image is input to personality recognition process. Features of the face are identified manually and input them to ANN and SVM. Each personality trait is valued from 1 to 9. In the second attempt m-SVM is used because outputs are multi-valued. ANN gave better results than m-SVM. In the third attempt we propose a methodology to identify personality traits using deep learning.

Keywords: *Artificial Neural Networks (ANN); Support Vector Machine (SVM); Multi Valued Multi Class Support Vector Machine (m-SVM); Deep learning; Personality trait*

1. Introduction

Recognition of a person's personality using the face images is one of the main interesting areas in Psychology. Some people believe that it is not possible to deduce a person's characteristics using the features of the face, but there are some researches which have proved that it is possible to deduce a person's characteristics using the face images. Such characteristics are known as psychological characteristics. Some psychological characteristics which have identified and studied by psychologists are emotional stability, dominance, sensitivity, intelligence, confidence, trustworthy, responsibility. Physiognomy is a broadly used approach for psychological characteristics recognition using the face. There are two other approaches also; they are phase facial portrait and ophthalmogeometry [1]. Identifying psychological characteristics using a face is helpful in many situations. In day-to-day social affairs it is very useful and easy to deal if we can find personality type of each person. There are some areas which we can apply automatic personality traits identification and get benefits. Such as when recruiting employees for a job it is very useful if we can identify the applicant's personality type because the development of a company depends on the employees who are working in the company. Another example is identifying criminals' behavior in criminology. Also automatic personality recognition system can help teachers to identify students' characteristics and change the teaching methods accordingly. The main reason of this research is to seek better solution for recognizing psychological characteristics using a face image by applying machine learning algorithms and image processing techniques.

2. Related works

The article Recognition of Psychological Characteristics from Face [1] describes that personality is a complex combination of traits and characteristics which determines our expectations, self-perceptions, values and attitudes, and predicts our reactions to people, subjects and events. Traits and characteristics are same things, but trait is distinguishing from characteristic since traits consider about feature or quality. So when considering traits theorists assume that traits are relatively stable over time, traits differ among individuals, and traits have influence behavior. To measure personality characteristics psychological researches apply psychometrics. Psychometrics is the construction of instrument and procedure for measurement and development of theoretical approaches to measurement. When applying psychology for personality traits recognition not only face, it can be applied to body, skull, fingers, hand, leg etc. This article describes three main approaches to psychological recognition from a face; they are physiognomy, phase facial portrait and ophthalmogeometry. Physiognomy is originally interprets different facial features as this method is based upon the idea that the assessment of the person's outer appearance, primarily the face, facial features, skin texture and quality, may give insights into one's character or personality. Phase facial portrait is mainly based on

calculating angles of facial feature lines directions. Ophthalmogeometry is based on the idea that person's emotional, physical and psychological states can be recognized by 22 parameters of eye part of the face.
There is less number of researches regarding to the current research area but facial expression recognition and age estimation using ANN and SVM have been done in several researches [2], [3], [4], [5].
To detect face area from an image, different techniques such as integral projection method [3] and color based segmentation algorithm [6] can be used. Color based segmentation is mostly used approach in face detection than other methods because it is the almost invariant against the changes of the size of the face, orientation of the face.
There is an existing application used to identify the personality traits from a face image [7]. This software uses a neural-network to identify corrections between facial features and psychological characteristics using photo identification techniques recognized by law enforcement professionals.

3. Technologies

3.1 Artificial neural networks

An Artificial Neural Network (ANN) is an information processing system which imitates the process of human brain. Human brain composed with neurons which are responsible for processing information and produce an output. An ANN also likes brain, it processes information and solves problems and also learns through examples.
Since ANN has the ability to derive information from complex and noisy data, ANN can be used in specific applications such as pattern recognition, face recognition, age estimation, different objects identification and emotion identification. A well trained ANN can predict the category of a new input by extracting its features. The optimum neural network for a specific problem is difficult to obtain and the solution is depend on the training dataset.

3.2 Backpropagation algorithm

The backpropagation algorithm is an algorithm used to train the multilayer neural networks. Backpropagation is a supervised learning method for multilayer neural networks and also known as the generalized delta rule. This method is used by different research communities in different contexts. First algorithm to train multilayer neural networks was introduced in the thesis of Paul Werbos in 1974 [8], but it was not disseminated in the neural network community. And rediscovered by David Rumelhart, Geoffrey Hinton, and Ronald Williams in 1986 [9]; David Parker in 1985 published it at M.I.T. [10]; Yann Le Cun in 1985 [11] independently.
Since backpropagation is a variation of gradient search algorithm, it is converged to a local minimum while search space has many local minimums and one global minimum. The converged point depends on the initial weights and biases. The network learning time becomes large when sample size is large. The network can be over learning if training does not stop at right time.

3.3 Leave some out multi-fold cross validation method

The Marsland's book [12] describes neural network training ratios depending on the dataset size. There is a common ratio which is used when dividing the dataset into training, testing and validation portions in network design. That is if it is plenty of data applicable then use 50:25:25 otherwise 60:20:20 ratios for training, testing and validation respectively. These datasets should be selected randomly so that they represent all classes.
If the training dataset is really small and if there are separate testing and validation datasets then one cannot guarantee that the network will be sufficiently trained. Then it is possible to perform leave-some-out, multi-fold cross-validation. In this method dataset is randomly partitioned into K subsets, and one subset is used as a validation set, while the neural network is trained on all the others. A different subset is then left out and a new network is trained on that subset, repeating the same process for all of the different subsets. Finally, the network that produces the lowest validation error is tested and used. In the most extreme case of this is leave-one-out cross-validation, where network is validated on just one piece of data, training on all of the rest.

3.4 Support vector machine

Support Vector Machine (SVM) was first introduced by Boser, Guyon, and Vapnik in COLT-92 in 1993 [13]. SVM provides significant classification than the other machine learning algorithms. SVM modifies the data to solve the problem by changing the data so that it uses more dimensions than the original data. Its implementation includes matrix calculation so works on reasonably sized data sets. SVM finds maximum margin which separate classes. That is SVM gives optimal separation between classes.

3.5 Multi class classification SVM

If there are more than two classes in the classification problem, one can use m-class SVM. But m-class SVM is in under research. Common approaches used in m-class SVM are one against all method and one against one strategy. Out of these two strategies one against all is the most common approach when considering about m-class SVM. The one against all strategy defines M number of

classifiers for m-class problem and in the ith iteration, ith class is assigned as positive and all others are assigned as negative. Finally all classifiers are combined to get the final solution.

3.6 Physiognomy

Backbone of physiognomy dates back to ancient Greece, and is still very popular. Physiognomy refers to identify person's psychological characteristics or personality using outer appearance, mainly using the face. So physiognomy also refers the relatively unchanging facial features which might use to interpret inner or hidden aspects of a person. These features include details of the forehead, eyebrows, eyes, nose and mouth.
Some people believe that it's not possible to deduce person's characteristics using the features of the face, but there are research studies proved that the possibility of deduce person's characteristics using face images. In 1772 J. C. Lavater said that physiognomy is a science and he proved the truth of the physiognomy on his book Essays on Physiognomy [14]. This is a historical article about physiognomy and it is still accepted. Since thousands of years researchers try to study the relationship between psychological characteristics and facial features. After doing such researches they have reached to judge person's personality via face features with a high degree of confidence. Simply we can say that physiognomy is true because people who look alike tend to behave in the same way or at least have some common behaviors. We got our outer appearance from genes that are from the parents and ancestors. Because of that we have same appearance to our parents. This implies that we should have approximately same personality with our parents. In day-to-day life we heard something like "She is smart just like her mother" which prove the truth of the physiognomy.

3.7 Deep learning

Deep learning is a machine learning technique which comes with ANN. It is emerged to machine learning research from 2006. Deep learning is also called as deep structured learning or hierarchical learning. Deep learning is mostly used with signal and machine learning researches. Deep Learning is a set of algorithms in machine learning that attempts to model high level abstractions in data by using architectures composed of multiple non-linear transformations [15].
When consider ANN, SVM, logistic regression and hidden markov model they are known as shallow architectures. For example SVM uses a shallow linear pattern separation model with one or zero feature transformation layer. Shallow architectures have been showed good results when dealing with real world problems but in some context we faced difficulties with those shallow architectures. For example when choosing input to the ANN, it is depending on the problem. So suggest the need of deep architectures for extracting complex structure and building internal representation from rich inputs. For an example the image classification system gets inputs as images. The Foundation to deep learning is artificial neural network and feed-forward neural network or MLPs with many hidden layers which are often referred to as deep neural networks. Deep learning is categorized in to three classes such as deep networks for unsupervised or generative learning, deep networks for supervised learning and hybrid deep networks. Deep networks for unsupervised or generative learning are used to capture unknown pattern in the data and categorized data depending on those patterns. Deep networks for supervised learning used for pattern classification purposes. Target label data are always available.

4. Methodology

The proposed system applies ANN, SVM and image processing techniques. The development stage is divided into three stages depending on the machine learning techniques used, first stage of implementation includes ANN; second stage of implementation used SVM and in the third stage proposed a methodology to identify personality traits using deep learning. The data set used in this research is obtained from social cognition and social neuroscience lab Alexander Todorov, Department of Psychology, Princeton University [16]. This database include several databases created for psychological research purposes out of them chose mean ratings on 66 faces from the Karolinska data set on 14 trait dimensions [17].

4.1 First stage of implementation

The main goal of this research was to identify the personality traits of a person using the face image by applying the technique neural networks. First put same number of landmarks on each face image as shown in Fig.1. The structure of the landmarks is defined after studying on physiognomy.

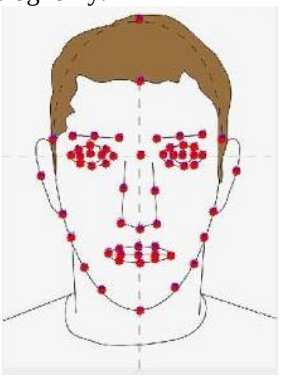


Fig.1 Image after put landmarks.

After the process of putting landmarks on each image, the coordinates of each landmark point and the expected trait values are saved in a text file.
The input to the neural network is corresponding landmark point coordinate values of each image which are stored in the text files. The text file includes 60 point coordinates and expected traits values. Since the dataset is small, the leave some out multi-fold cross validation method has been used by randomizing the dataset and ran several times to obtain better training, testing and validation results.

4.2 Second stage of implementation

In this stage SVM is used as the technology to identify personality traits. Input to the algorithm is coordinate values of landmark points stored in text files as first stage. In SVM expected values are considered as classes. Since our problem has multiple classes, an m-class SVM classifier has been used.

4.3 Third stage of implementation

In this stage color segmentation algorithm is used to detect face automatically. As next phase a deep neural network is proposed to design in recognizing personality traits from face image. Input to the network is face area of the image which is the output of automatic face detection algorithm. Many research have proved that Deep Neural Networks (DNN) is extremely good with performing recognition and classification tasks.
A DNN extract features from images automatically during the training process. Training of DNNs is computationally expensive and they need large datasets. Training of DNNs is computationally expensive and they need large datasets.
It is proposed to use a supervised DNN (e.g. Convolution Neural Network) to train the system to identify personality traits from a face image.

5. Results

5.1 First stage results

The neural network is tested using portion of the dataset and the best validation performance obtained in this process is 0.80706. Fig. 2 shows the performance plot of the network.

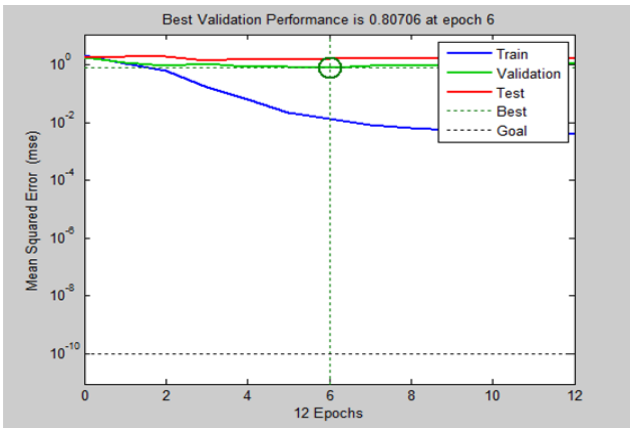


Fig.2 Performance plot of neural network.

5.2 Second stage results

Separate SVM models have been designed for each personality trait value, for instance if a trait has three unique values then built three models. Then each model is tested with four kernel functions. They are linear, quadratic, polynomial and RBF. Some personality traits were not classified and some traits were successfully classified with these four kernels. Table 1 shows the results obtained in this stage.

Table 1: SVM results with different kernels

Trait	*Linear*	*Quadratic*	*Polynomial*	*RBF*
Attractive	√	×	×	×
Caring	√	×	×	×
Aggressive	×	×	×	×
Mean	×	√	×	×
Intelligent	√	√	√	×
Confident	×	×	×	×
Emotionally stable	√	×	×	×
Trustworthy	×	×	×	×
Responsible	√	×	×	×
Weird	×	×	×	×
Unhappy	√	×	√	×
Dominant	×	√	√	×

5.3 Third stage results

The color segmentation algorithm used in this stage obtained 98.5% accuracy of detecting face area.

6. Conclusions

The first stage of implementation (using a neural network) gave better results than the second stage of implementation (using a support vector machine). In this research study, m-class SVM was used because each personality trait is multi valued. The dataset used in this research is small. The size of the dataset may affect to the results obtained. So accuracy of the results may increase by using a large dataset.

The m-SVM strategy used in this research is one against all. There is an another strategy called one against one. One can develop the one against one strategy.

In third stage of implementation involved automatic face detection algorithm to detect face area from face image which includes non-face areas. To detect the face color based segmentation method was used. This algorithm gave 98.5% accuracy without any preprocessing on images and with preprocessing on images may obtain higher accuracy in automatic face detection phase.

There may be hidden factors other than introduced in this study in feature extraction phase in ANN and SVM stages. So study more on psychology and improve feature extraction phase may improve the final results.

The third stage of implementation proposed to use supervised deep neural network. By implementing a deep neural network with a large dataset can improve the accuracy of the classification.

References

[1] Ekaterina Kamenskaya and Georgy Kukharev,"Recognition of psychological characteristics from face", Metody Informatyki Stosowanej, Issue 1, 59–73, 2008.

[2] Jawad Nagi, Syed Khaleel Ahmed and Farrukh Nagi, "A MATLAB based face recognition system using image processing and neural networks", Proc. of the 4th International Colloquium on Signal Processing and its Applications (CSPA), March 7-9, 2008.

[3] Zaenal Abidina and Agus Harjoko, "A Neural Network based Facial Expression Recognition using Fisherface", International Journal of Computer Application, Vol. 59– No.3, pp 0975 – 8887, 2012.

[4] Christine L. Lisetti and David E. Rumelhart, "Facial Expression Recognition Using a Neural Network", FLAIRS Conference, pp328–332, 1998.

[5] Sangeeta Agrawal, Rohit Raja and Sonu Agrawal,"Support Vector Machine for age classification", International Journal of Emerging Technology and Advanced Engineering ISSN 2250-2459, Volume 2, Issue 5, 2012.

[6] H C Vijay Lakshmi and S. PatilKulakarni, "Segmentation algorithm for multiple face detection in color images with skin tone regions using color spaces and edge detection techniques", International journal of computer theory and engineering, Volume 2, Issue 4, 1793-8201, 2010.

[7] UNIPHIZ Lab, *Digital Physiognomy Software*, Computer Sowftware. UNIPHIZ Lab, 2001 – 2014.

[8] Paul J. Werbos. "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences". PhD thesis, Harvard University, 1974.

[9] Rumelhart, David E.; Hinton, Geoffrey E., Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature* 323 (6088): 533–536.

[10] D.Parker, "*Learning Logic*", Technical Report TR-47, Center for Computational Research in Economics and Management Scince, M.I.T., Apr.1985.

[11] Y. LeCun, "Une procédure d'apprentissage pour réseau a seuil asymmetrique (a Learning Scheme for Asymmetric Threshold Networks)", Proceedings of Cognitiva 85, 599–604, Paris, France, 1985.

[12] Stephen Marsland, *MACHINE LEARNING: An Algorithm perspective*. NewYork: CRC Press Taylor & Francis Group Boca Raton, 2009.

[13] BOSER, Bernhard E., Isabelle M. GUYON, and Vladimir N. VAPNIK, "*A training algorithm for optimal margin classifiers*", COLT '92: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. New York, NY, USA: ACM Press, pp. 144–152, 1992.

[14] Lavater, J. C., *Essays on physiognomy* (C. Moore, Trans.). London: H. D. Symonds. Original work ad.), 1797.

[15] Y. Bengio, A. Courville, and P. Vincent., "Representation Learning: A Review and New Perspectives," IEEE Trans. PAMI, special issue Learning Deep Architectures, 2013.

[16] Alexander Todorov, Department of Psychology, Princeton University, "Social cognition and social neuroscience", http://tlab.princeton.edu/databases/.

[17] Lundqvist, Flykt, and Ohman 1998, and Oosterhof and Todorov 2008. The exact references are: Lundqvist, D., Flykt, A., and Ohman, A. (1998) Karolinska Directed Emotional Faces [Database of standardized facial images]. Psychology Section, Department of Clinical Neuroscience, Karolinska Hospital, S-171 76 Stockholm, Sweden; Oosterhof, N. N., and Todorov, A. (2008) .

Kalani Ilmini obtained her B.Sc. (Special) degree in computer science from the University of Sri Jayewardenepura in 2014. She is currently working as a temporary lecturer at the Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura, Sri Lanka. Her domains of interest are: Machine Learning, Image Processing and Data Mining.

T. G. I. Fernando is a senior lecturer/researcher in Computer Science at the Department of Computer Science, University of Sri Jayewardenepura, Sri Lanka. He obtained his B.Sc. degree in Mathematics from the University of Sri Jayewardenepura in 1993. He holds two M.Sc. degrees; In Industrial Mathematics from the University of Sri Jayewardenepura, Sri Lanka in 1998 and in Computer Science from the Asian Institute of Technology, Thailand in 2002. He holds a Ph.D. degree in Intelligent Systems from the Brunel University, UK in 2009.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Ranking user's comments by use of proposed weighting method

Zahra Hariri[1]

[1] **Student in Faculty of Graduate Studies, Safahan Institute of Higher Education, Isfahan, Iran, P.O.Box 817474319**
Zahara.hariri2003@gmail.com

Abstract

The main challenges which are posed in opinion mining is information retrieval of large volumes of ideas and categorize and classify them for use in related fields. The ranking can help the users to make better choices and manufacturers in order to help improve the quality. As one of the pre-processing techniques in the field of classification, weighting methods have a crucial role in ranking ideas and comments. So, we decided to offer a new weighting method to improve some other similar methods, especially Dirichlet weighting method. In this paper, the proposed method will be described in detail, and the comparison with the three weighting methods: Dirichlet, Pivoted and Okapi also described. The proposed weighting method has higher accuracy and efficiency in comparison to similar methods. In the following, user comments of online newspapers are ranked and classified by use of proposed method. The purpose is to provide more efficient and more accurate weighting method, therefor the results of ranking will be more reliable and acceptable to users.

Keywords: *Opinion mining, Information retrieval, ranking comments, weighting methods, weighting methods constraints.*

1. Introduction

Word Wide Web can be considered as a repository of ideas from users. The challenge that the manufacturers and web administrators are faced by is to analyze and organize their ideas.

Analysis of emotions in online publications is a way of organizing user's ideas, which requires weighting of the words in comments. The weighting methods include genetic algorithms, artificial neural networks, regression equations, TF-IDF, Pivoted, Okapi, Dirichlet. In this article we propose a new weighting method which is improved of Dirichlet weighting method, also it satisfy all 7 constraints.

The paper is organized as follows. Section 2 state some weighting methods and, also proposed weighting method. All the constraints are checked for the proposed method in section 3. In section 4, accuracy and performance of the proposed method is compared with methods such as: Pivoted, Okapi and Dirichlet. Section 5 is about implementation of ranking comments by use of proposed weighting method. Dataset is discussed in section 6. Conclusion and future works are expressed in section 7. Finally, section 8 states references.

This document is set in 10-point Times New Roman. If absolutely necessary, we suggest the use of condensed line spacing rather than smaller point sizes. Some technical formatting software print mathematical formulas in italic type, with subscripts and superscripts in a slightly smaller font size. This is acceptable.

2. Weighting methods

There are a lot of weighting methods that are used. But there are different in 2 aspects: satisfying constraints and the value of parameters like efficiency and accuracy. In the following some weighting methods are mentioned.

2.1 Pivoted method

Vector space model is displayed as a vector of words. Documents are ranked based on the similarity between query and document vector. Pivoted retrieval method is one of the best retrieval formula which is expressed in equation 1 as bellow [1].

$$S(Q,D) = \sum_{t\in Q\cap D} \frac{1+\ln(1+\ln(c(t,D)))}{(1-S)+S\frac{|D|}{awdl}} . c(t,Q). ln \frac{N+1}{df(t)} \quad (1)$$

Where S is retrieval parameter, c(t,D) is The number of repetitions of word t in document D. |D| is the length of document D. c(t,Q) is The number of repetitions of word t in query Q. N is the number of documents and df(t) is the number of documents including word t.

2.2 Okapi method

This formula is an effective retrieval formula that uses classical probabilistic model. It is expressed in equation 2[1].

$$S(Q,D) = \sum_{t\in Q\cap D} ln\frac{N-df(t)+0.5}{df(t)+0.5} \times \frac{(k_1+1)\times c(t,D)}{k_1((1-b)+b\frac{|D|}{awdl})+c(t,D)} \times \frac{(k_3+1)\times c(t,Q)}{k_3+c(t,Q)} \quad (2)$$

K_1 is between 1 and 2 , b is equal to 0.75 , K_3 is between 0 and 1000. df(t) is the number of documents including word t, c(t,D) is The number of repetitions of word t in document D and awdl is the average of document's length.

2.3 Dirichlet prioir method

This method is one of the best methods of language modeling approach which working based on similarity between query and document. It is expressed in equation 3[1].

$$s(Q,D) = \sum_{t \in Q \cap D} c(t,Q).\ln(1+\frac{c(t,D)}{\mu.p(t|C)} + |Q|.ln\frac{\mu}{|D|+\mu} \quad (3)$$

Where μ is retrieval parameter, c(t,D) is The number of repetitions of word t in document D. |D| is the length of document D and |Q| is the length of query Q. $p(t|C)$ is possibility of existence of word t in the collection.

2.4 Proposed weighting method

In this study a new weighting method is proposed in equation 4 which aim is to satisfy all the constraints and, also improve accuracy and efficiency of previous methods such as Pivoted, Okapi and Dirichlet.

$$S(Q,D) = \sum_{t \in Q \cap D} (C(t,Q).C(t,D).\ln\left(1+\frac{C(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D|} \qquad (4)$$

Where μ is retrieval parameter, c(t,D) is The number of repetitions of word t in document D. |D| is the length of document D and |Q| is the length of query Q. c(t,Q) is The number of repetitions of word t in query Q.
Dirichlet method don't satisfy the LNC2 constraint but proposed method satisfy all the weighting method constraints which will be explain in the next section.

3.CHEKING WEIGHTING METHOD CONSTRAINTS FOR PROPOSED METHOD

There are 7 constrains which are good to be satisfied by weighting methods. In the following all 7 constraints are checked for proposed method.[1,2]

3.1. TFC1

In equation 5, it is shown that proposed method satisfies TFC1.

$$s(Q,D_2) = c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_2|}$$

$$s(Q,D_1) = c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_1|}$$

C(t,D_1) > C(t,D_2) then s(Q,D_1) > s(Q,D_2). (5)

3.2. TFC2

In equation 6, it is shown that proposed method satisfies TFC2.

$$\begin{aligned}
s(Q,D_2) &= c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) \\
&\quad + \frac{|Q|}{|D_2|}s(Q,D_3) \\
&= c(t,Q).c(t,D_3).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) \\
&\quad + \frac{|Q|}{|D_3|}s(Q,D_1) \\
&= c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) \\
&\quad + \frac{|Q|}{|D_1|}c(q,D_2) = 1 + c(q,D_1) \\
&\rightarrow \ c(q,D_2) > c(q,D_1) \qquad 1c(q,D_3) \\
&= 1 + c(q,D_2) \ \rightarrow \ c(q,D_3) \\
&> c(q,D_2)
\end{aligned}$$

$$c(q,D_1) < c(q,D_2) < c(q,D_3) \ \rightarrow \ 2c(q,D_1) < c(q,D_2) < c(q,D_3)$$

$$s(Q,D_2) - s(Q,D_1) > s(Q,D_3) - s(Q,D_2)$$

$$\left[c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D_2)}{\mu.df(t)}\right) + \frac{|Q|}{|D_2|}\right] - \left[c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_1|}\right] > \left[c(t,Q).c(t,D_3).\ln\left(1+\frac{c(t,D_3)}{\mu.df(t)}\right) + \frac{|Q|}{|D_3|}\right] - \left[c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_2|}\right]$$

$$c(t,Q).c(t,D_2) - c(t,Q).c(t,D_1) > c(t,Q).c(t,D_3) - c(t,Q).c(t,D_2)$$

$$\rightarrow \ c(t,D_2) - c(t,D_1) > c(t,D_3) - c(t,D_2) \ \rightarrow \ 2c(t,D_2) > c(t,D_3) + c(t,D_1) \qquad (6)$$

3.3. TFC3

In equation 7, it is shown that proposed method satisfies TFC3.

$$s(Q,D_1) = c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_1|}s(Q,D_2) = c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_2|}s(Q,D_1) > s(Q,D_2) \ \rightarrow$$

$$s(Q,D_1) = c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_1|} > s(Q,D_2) = c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_2|}td(q_1) = td(q_2)c(t,Q).c(t,D_1) > c(t,Q).c(t,D_2) \ \rightarrow \ c(t,D_1) > c(t,D_2) \qquad (7)$$

3.4. TD

In equation 8, it is shown that proposed method satisfies TDC.

$$\begin{cases} t \notin Q \rightarrow c(t,D_2) = c(t,D_1)+1 \rightarrow s(Q,D_1) > s(Q,D_2) \\ t \in Q \rightarrow c(t,D_2) = c(t,D_1) \rightarrow s(Q,D_1) = s(Q,D_2) \end{cases}$$
$$\rightarrow s(Q,D_1) \geq s(Q,D_2) s(Q,D_1)$$
$$= c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right)+\frac{|Q|}{|D_1|}s(Q,D_2)$$
$$= c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right)+\frac{|Q|}{|D_2|}$$
$$c(t,D_2) = c(t,D_1) \text{ then } s(Q,D_1) = s(Q,D_2). \quad (8)$$

3.5. LNC1

In equation 9, it is shown that proposed method satisfies LNC1.

$$\begin{cases} t \notin Q \rightarrow c(t,D_2) = c(t,D_1)+1 \rightarrow s(Q,D_1) > s(Q,D_2) \\ t \in Q \rightarrow c(t,D_2) = c(t,D_1) \rightarrow s(Q,D_1) = s(Q,D_2) \end{cases} \rightarrow$$
$$s(Q,D_1) \geq s(Q,D_2) s(Q,D_1) = c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right)+\frac{|Q|}{|D_1|}s(Q,D_2) = c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right) + \frac{|Q|}{|D_2|} \quad (9)$$

3.6. LNC2

In equation 10, it is shown that proposed method satisfies LNC2.

$$s(Q,D_2) = c(t,Q).c(t,D_2).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right)+$$
$$\frac{|Q|}{|D_2|}s(Q,D_1) = c(t,Q).c(t,D_1).\ln\left(1+\frac{c(t,D)}{\mu.df(t)}\right)+$$
$$\frac{|Q|}{|D_1|}c(w,D_1) = K.c(w,D_2) \rightarrow c(w,D_1) > c(w,D_2) \quad 1 \quad (10)$$

Due to equation 1 of 10, $s(Q,D_1)>s(Q,D_2)$, $\frac{|Q|}{|D|}$ will not grow as much as $c(w,D)$, then $s(Q,D_1) \geq s(Q,D_2)$.

3.7. TF-LNC

In equation 11, it is shown that proposed method satisfies TF-LNC.

$c(q,D_1) > c(q,D_2)$ then $c(q,D_1) - c(q,D_2)>0$ and, also $|D_1| = |D_2| + c(q,D_1) - c(q,D_2)$ then $|D_1| > |D_2|$ (11)

Table 1: Weighting method constraints

Formula	*TFC1*	*TFC2*	*TFC3*	*TDC*	*LNC1*	*LNC2*	*TF-LNC*
Pivoted	Y	Y	Y	Y	Y	C	C
Dirichlet	Y	Y	Y	Y	Y	C	Y
BM25	C	C	C	Y	C	C	C
PL2	C	C	C	C	C	C	C
Proposed Method	Y	Y	Y	Y	Y	Y	Y

4. Comparison of Accuracy and Efficiency

To compare the proposed method with three methods which were mentioned before, we need confusion matrix that are explained in next parts. Also we need TN, TP, FN, FP parameters owing to calculating accuracy and efficiency.[3,4]

TN: Are correct but have been misdiagnosed by machine.
TP: Are correct and have been diagnosed correctly by machine.
FN: Are false and have been misdiagnosed by machine.
FP: Are false but have been diagnosed correctly by machine.

Accuracy and efficiency can be calculated by use of equations 12, 13 and 14.[5,6,7]

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \quad (12)$$

$$efficiency = \frac{2*Recall*Accuracy}{Recall+Accuracy} \quad (13)$$

$$Recall = \frac{TP}{TP+FP} \quad (14)$$

Calculating parameters which are necessary from confusion matrix, is shown in tables 2, 3, 4 and 5.

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 25 & 234 & 96 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Okapi confusion matrix

Table 2: Okapi parameters from confusion matrix

	First row	*Second row*	*Third row*	*Forth row*
TN	234	0	234	234
TP	0	234	0	0
FN	0	124	0	0
FP	25	0	96	3

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 275 & 0 & 0 \\ 0 & 0 & 3 & 17 \end{bmatrix}$$

Dirichlet confusion matrix

Table 3: Dirichlet parameters from confusion matrix

	First row	*Second row*	*Third row*	*Forth row*
TN	27	17	27	17
TP	0	10	0	3
FN	1	0	275	0
FP	0	275	1	0

$$\begin{bmatrix} 391 & 1 & 11 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivoted confusion matrix

Table 4: Pivoted parameters from confusion matrix

	First row	*Second row*	*Third row*	*Forth row*
TN	0	391	391	391
TP	391	0	0	0
FN	17	0	0	0
FP	0	1	11	5

$$\begin{bmatrix} 72 & 0 & 0 & 0 \\ 0 & 173 & 0 & 0 \\ 0 & 0 & 37 & 0 \\ 0 & 0 & 13 & 11 \end{bmatrix}$$

Proposed method confusion matrix

Table 5: Proposed method parameters from confusion matrix

	First row	*Second row*	*Third row*	*Forth row*
TN	221	120	256	282
TP	72	173	37	11
FN	0	0	0	13
FP	0	0	13	0

For computing confusion matrix, we need a matrix which its rows are documents and columns are words that remain after preprocessing. Equation 15 shows an example.

$$\begin{bmatrix} -1.6708 & -1.6708 & -1.6708 & -1.6707 \\ -1.6708 & -1.6707 & -1.6708 & -1.6708 \\ -1.6708 & -1.6708 & -1.6708 & -1.6707 \\ -1.6708 & -1.6708 & -1.6708 & -1.6707 \end{bmatrix} \quad (15)$$

Then this matrix will be an input for matlab for computing confusion matrix. After that accuracy and efficiency are computable. Table 6 shows the results of accuracy and efficiency comparison.

Table 6: Comparison of accuracy and efficiency

Formula	Confusion Matrix	Accuracy	Efficiency
Okapi	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 25 & 234 & 96 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	80%	31%
Pivoted	$\begin{bmatrix} 391 & 1 & 11 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	97%	36%
Dirichlet	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 275 & 0 & 0 \\ 0 & 0 & 3 & 17 \end{bmatrix}$	50%	47%
Proposed Method	$\begin{bmatrix} 72 & 0 & 0 & 0 \\ 0 & 173 & 0 & 0 \\ 0 & 0 & 37 & 0 \\ 0 & 0 & 13 & 11 \end{bmatrix}$	97%	90%

A code that give us an input matrix for computing confusion matrix has some steps as below:

1. A query is written by user.
2. Eliminating stop words.
3. Allocating weight to words by use of one of weighting methods.
4. Creating output matrix(which is input in matlab to calculate confusion matrix)

In figure 1 a schema of an output matrix based on proposed weighting method is shown.

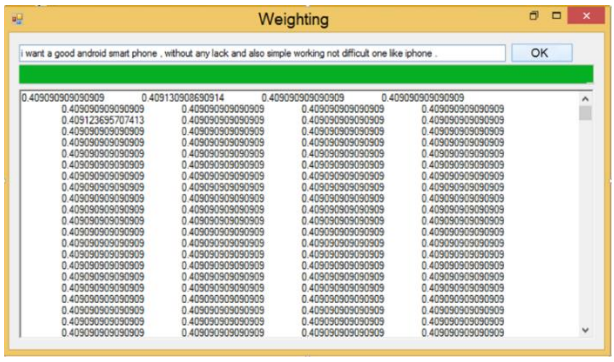


Fig. 1 weighting matrix.

To calculate confusion matrix in matlab some steps has been taken:

1. Test set is created.
2. A random order is created.
3. Sorting input matrix and test matrix based on random order which was created before.
4. Learning set is created.
5. Sample set is classified based on test and learning sets.
6. Confusion matrix is created.

Figure 2 shows a confusion matrix based on proposed method in matlab.

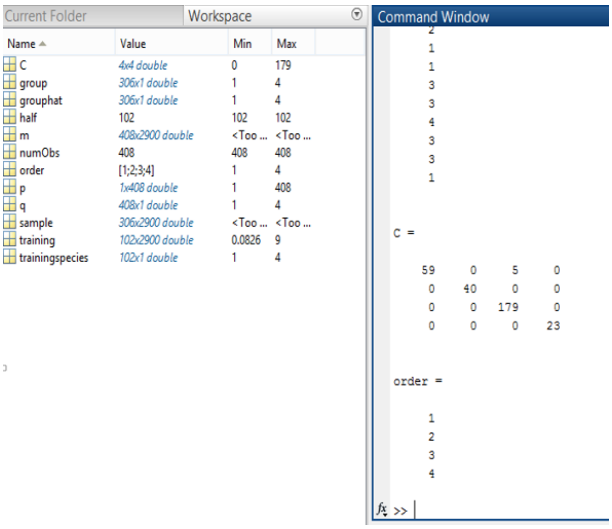


Figure. 2 Confusion matrix in Matlab.

5. Implementation

Text mining and sentiment analysis such as analyzing user's comments can be implemented by using c#.net programming framework. In figure 3 shows a summarized flowchart of an implemented code for ranking user's comments based on proposed method.

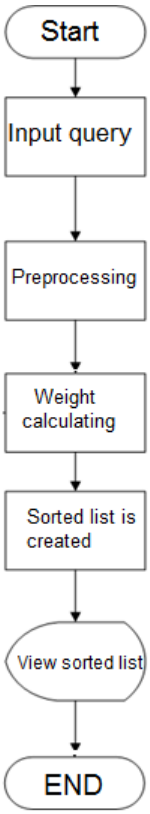


Figure. 3 Implemented code flowchart.

Implemented code is consist of 2 parts, dataset preparation and ranking comments. As is shown in figure 3, first of all, user insert a query in the weighting part and query is sent to preprocessing part. After that, words are weighted based on proposed method. Then a sorted list is created by use of cosine similarity.

Overally, all the steps that has been taken due to ranking comments are as follow:

1. A query is written by user.
2. Eliminating stop words.
3. Price and property of a product is inserted by user.
4. Words are weighted based on proposed method.
5. Documents are ranked based on cosine similarity.
6. Ranked comments are shown.

Implemented code is in c#.net framework for ranking user's comments.

First user insert a query, price and property after that the ok key is selected all of the calculations and computations are done. Then, cosine similarity is calculated by use of equation 16.[8,9]

$$\text{Similarity}=\frac{\vec{d}j \cdot \vec{q}}{|\vec{d}j|.|\vec{q}\ |} = \frac{\sum_{i=1}^{t}(w_{ij}.w_{iq})}{\sqrt{\sum_{i=1}^{t} w_{ij}^2 . \sum_{i=1}^{t} w_{iq}^2}} \quad (16)$$

Where $\vec{d}j$ is document vector, $\vec{q}$ is query vector, $|\vec{d}j|$ is size of document vector, $|\vec{q}\ |$ is size of query vector. w_{ij} is a weight of word I in document j and w_{iq} is a weight of word I in query q.

Figure 4 is an example of Proposed product to user based on below query and ranked comments by use of proposed weighting method.

Query: I want a good android smart phone, without any lack and also simple working not difficult one like iPhone.

Samsung Galaxy S6 Dual-SIM
Samsung Galaxy Note Edge
Samsung Galaxy S4 SGH
Huawei Ascend Y530

Figure. 4 Sample output.

6. **Dataset**

We couldn't find profitable dataset, so we collect a dataset from Amazon.com in the period of times about 2 months. This dataset is about cellphones by 2 property (price and operating system type). Figure 5 shows dataset program's flowchart.

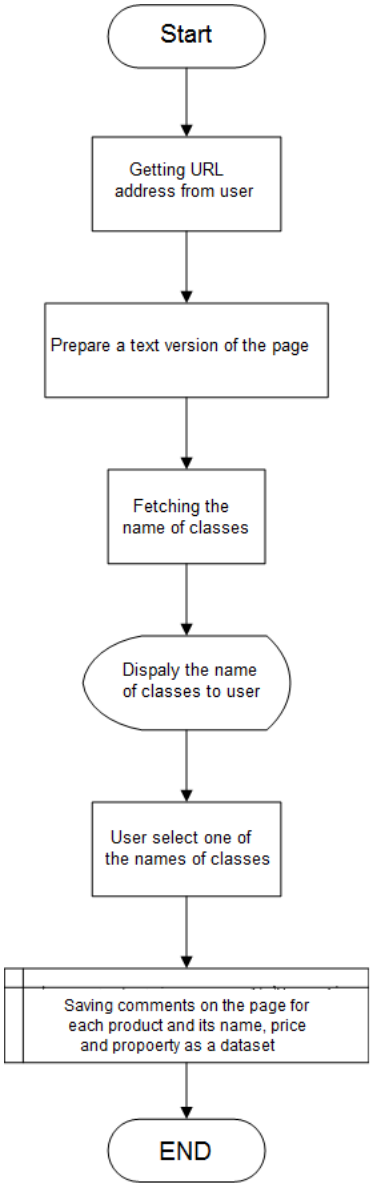


Figure. 5 Gathering Dataset flowchart.

As is shown in figures 6 and 7, all the steps are such as figure 6.

Figure. 6 Gathering Dataset program (steps 1 and 2).

Figure. 7 Gathering Dataset program (step 3)

7. Conclusion

One of the crucial usage of weighting methods is in text mining and information retrieval. Nowadays, ranking user's comments plays a vital role owing to its important help to users for selecting the best product and also help the producer to know best about their products in user's point of view.

Using one of the weighting methods is so important due to ranking comments. Weighting methods are comparable in 2 aspects. Proposed method in this research can satisfy these 2 aspect as well. It can satisfy all the 7 constraints and also it has better accuracy and efficiency in comparison to similar methods such as Okapi, Pivoted and Dirichlet. Another advantage of this research is its recommendation to user about a product he needs.

Figure 8 shows a comparison between proposed method and similar ones.

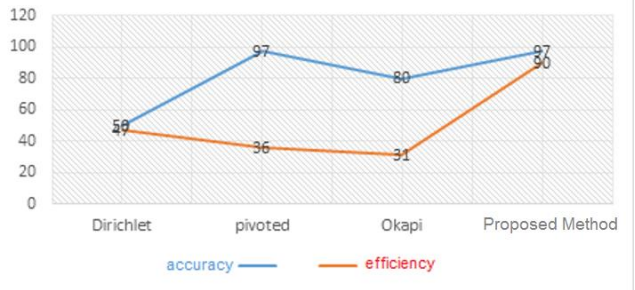


Figure. 8 Comparison of accuracy and efficiency.

Generally, this research's benefits are:

1. User can specify the category of the product.
2. User can determine 2 important property about product.
3. Others comments can be used in field of each product.
4. Weighting methods which was proposed has a better accuracy and efficiency in comparison to others.

Although obtained results show a good performance of the proposed method, we can't claim that it is the best method. The aim of this study was to use the results to provide useful suggestions to the user, but the results can be used for other purposes, too.

Some future works are:

- Improve executive order to enhance the speed of ranking.
- Use proposed method for showing results to the owners of online communities.
- Integrate database issues and user personal profile's information, in order to omitting the stage of sending gathered information from user.

Acknowledgments

"This article is fetched from the thesis in Faculty of Graduate Studies, Safahan Institute of Higher Education."

References

[1] HUI FANG, TAO TAO, CHENGXIANG ZHAI, "Diagnosti c Evaluation of Information Retrieval Models",ACM Transaction on Information Systems, Vol 29, No 2, Article 7,April 2011.

[2] Alexandru Tatar, Panayotis Antoniadis, Marcelo Dias de Amorim, and Serge Fdida," Ranking news articles based on popularity Prediction", International Conference on Advances in Social Networks Analysis and Mining, IEEE/ACM ,2012, pp 106 – 110, ISBN/ISSN: 978-1-4673-2497-7 .

[3] Chiao-Fang Hsu, Elham Khabiri, James Caverlee," Ranking Comments on the Social Web",Proceedings of the 2009 International Conference on Computational Science and Engineering , Vol 04, ISBN: 978-0-7695-3823-5 , pp 90-97 .

[4] Robertson, A.M. and Willett, P., "An Upperbound to the Performance of Ranked-Output Searching: Optimal Weighting of Query Terms Using a Genetic Algorithm", Journal of Documentation, Vol. 52, pp. 405–420, 1996.

[5] Ghose, A. and P. Ipeirotis. Designing novel review ranking systems: predicting the usefulness and impact of reviews. In Proceedings of the International Conference on Electronic Commerce, 2007.

[6] Giorgos Giannopoulos, Ingmar Weber, Alejandro Jaimes, Timos Sellis" Diversifying User Comments on News Articles",Lecture Notes in Computer Science Vol 7651, 2012, pp 100-113.

[7] Stefan Siersdorfer, Sergiu Chelaru, Jose San Pedro," How Useful are Your Comments?- Analyzing and Predicting YouTube Comments and Comment Ratings",WWW '10 Proceedings of the 19th international conference on World wide web ,2010,ISBN: 978-1-60558-799-8 , pp 891-900 .

[8] Pooja Kherwa, Arjit Sachdeva, Dhruv Mahajan Nishtha Pande, Prashast Kumar, "An approach towards comprehensive sentimental data analysis and opinion mining", Advance Computing Conference (IACC), 2014 IEEE International, 21-22 Feb. 2014,ISBN: 978-1-4799-2571-1,pp:606-612.

[9] Liu, Bing, "Sentiment analysis: A multi-faceted problem". IEEE Intelligent Systems, 2010.25,3: 76-80.

First Author bachelor degree was achieved in 2012 from TABARESTAN Chalus , Iran in computer software engineering , Student of master degree in SAFAHAN , Isfahan , Iran. Working on Text mining, Sentiment analysis, Public Opinion Research.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

A Simple Study on Search Engine Text Classification for Retails Store

Renien Joseph[1], Samith Sandanayake[2] and Thanuja Perera[3]

[1] Zone24x7 (Private) Limited
460, Nawala Road Koswatte, Sri Lanka
renien.john@gmail.com

[2] Zone24x7 (Private) Limited
460, Nawala Road Koswatte, Sri Lanka
samithdisal@gmail.com

[3] Zone24x7 (Private) Limited
460, Nawala Road Koswatte, Sri Lanka
tm.thanuja.perera@gmail.com

Abstract

It is obvious, the continuing growth of textual content rapidly increasing within the Word Wide Web (WWW). So certainly with the combination of sophisticated text processing and classification techniques it leads to produce high accurate search results. Even though a large body of research has delved into these problems; each has their theories and different approaches according to their data collection. This has been very challenging task continuously and this paper converges solutions, comprehensive comparisons that leads to different approaches. Therefore it will help to implement a robust search engine. The research proves probability text classification models classify documents robustly. But to improve the search result that involves short texts, we should certainly go through a hybrid approach including rules and statistical neural network models. As a pruning components the pre-processing and post-processing modules should adapted. And also due to the dynamic data the process pipeline should be frequently update.

Keywords: Search queries, *Text Classification, Rules, Machine Learning, and Information Retrieval*

1. Introduction

Nowadays the Internet usage is very common. With the rapid growing technologies world is shrinking too small. People do their activities on World Wide Web and they are now getting comfortable with online shopping experience. Hence, it is important the consumers should be provided with concrete online-based search solutions [12][10].

Catering a solution to all kind of consumers is a challenging task. There are consumers use search engine with different behaviorism. In the context of search engine many researchers still have higher involvement to provide a robust solution. Queries can be categorized in to two [9] and they are,

1) Key word/ Short word search queries
2) Long tail search queries

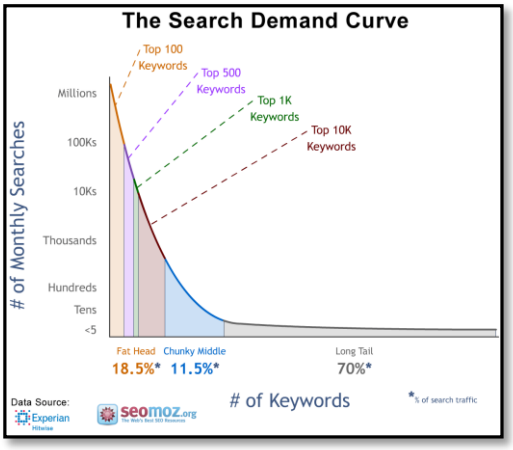


Fig. 1 Search query demand [11]

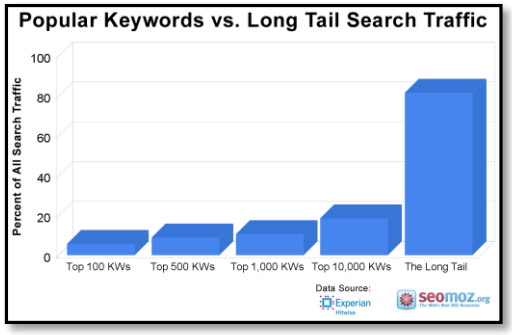


Fig. 2 Benchmarking of search traffic [11]

The graphs (Figure 1,Figure 2) explain the variation of the search queries and its demand. It clearly shows consumers try mostly with short end queries more than long tail queries. Therefore, this research approach can be catered to short and long tail queries appropriately.

During the ionic stages researchers were dealing with handful amount of data and their approach to the search engine was with relational databases [16][21]. But now the evolution has begun in technology and Big Data is extraordinarily growing up in full space [19]. So, with the help of Big Data Analytics [13] this research focuses to provide accurate solution not only to the retail domain but it serve the solution to the various domains precisely.

2. Approach

The data that contains within the retail company will not be enough to implement a robust search engine. Web is increasingly becoming the dominant information seeking method. Every search engine has its root information retrieval (IR) module. IR is all about data retrieval, analysis and emphasizing data as the basic unit (Figure 3). Figure 3 exposes the modules that involves in a search engine component. The indexing component mainly deals with the data mining [2], web mining and social network analysis [4] job that keeps the data collection up to date. The intelligent component tries to understand data collection, user input queries and stay has a decision-making module.

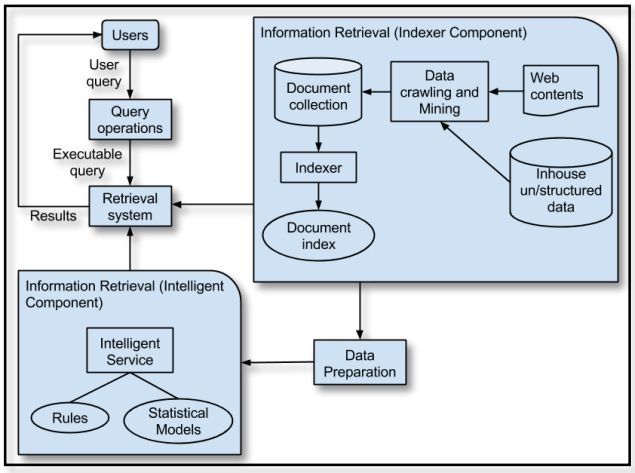


Fig. 3 IR system overall architecture

The main important data process layer is known as Natural Language Processing (NLP) [7]. During this process it will help the engineers to find the information that matching to their domain that helps to boost and improve the search result. In NLP, pre-processing is an inaugural step to processes the query [15]. During this process it helps to clean up and throw away the unwanted words from the query [6]. Figure 4 shows the step-by-step approach for text pre-processing. Due to the pre-processing pipeline, it significantly allows to improve the accuracy in the stage of text classification.

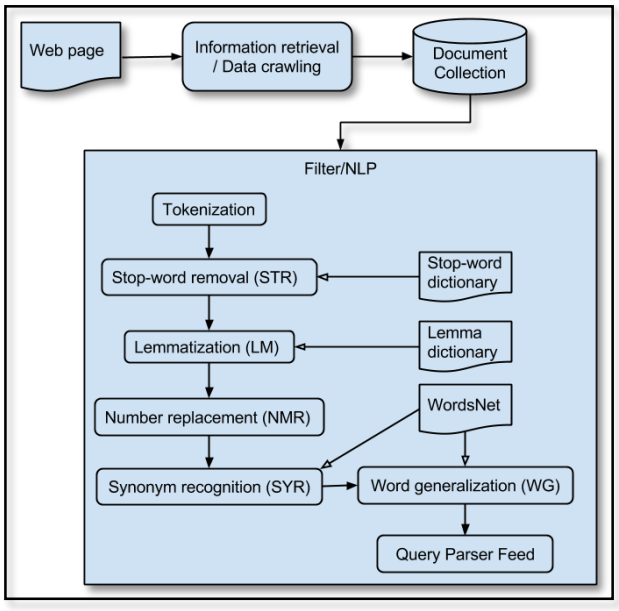


Fig. 4 Text pre-processing schema

In depth, many different components come together according to the users input query. As mention in introduction section the main categories of search queries can be classified in to the following types [15],

- Boolean queries
- Phrase queries
- Proximity queries
- Full document queries
- Natural Language queries

According to the users query usage type, it is important to change the query operational module. In the below sections it covers mostly to tackle the Boolean, Phrase and Natural Language queries.

2.1 Rules Engine

Before the prediction models were invented the developers/researchers were using Rules based search engine. In technology era too still people try to depend on the classical approach to identify the user queries. Rules based engines were implemented focusing into the business domains [1].

The Boolean search queries can be easily handled by rules. The Boolean operators, AND, OR and with negativity words the queries can be constructed. The Boolean operators are filtered during the rules processing stage and certain partial queries are filtered after the Part of Speech tagging (POS) [7]. POS tagging plays an important role in speech and natural language parsing stage. Many researchers attempted to this problem and implemented solution with different approaches [18][17]. Hence, it's very vital to choose the correct approach based on data collection structure and the query patterns also need to be considered during the decision.

2.2 Multinomial Naive Bayes (MNB)

To measure the similarity of two documents, researchers mainly focus on word frequency, which is the most traditional technique. It provides enough word co-occurrences or Shared context for good similarity measures. The Naive Bayes classifier is a simple probabilistic classifier, which is based on Bayes theorem with strong, and naive independence assumptions [14].

This Supervised Learning model falls into Bag-of-Words category and most text classification methods use the bag-of-words representation because of its simplicity for classification purposes [3]. It is one of the most basic text classification techniques with various applications in email spam detection, personal email sorting, document categorization, sexually explicit content detection, language detection and sentiment detection. Despite the naive design and oversimplified assumptions that this technique uses, Naive Bayes performs well in many complex real-world problems [3].

In the ambience of retail domain the data type and structure is specific. An abstract data collection is shown in the Table1. After the pre-processing stage the model should identify the feature classes. But due to the sparseness of short text data set list (Table 1), state-of-the-art techniques failed to achieve the desired accuracy [20]

Table 1 Sample data set

Color	Brand	Patterns	Product
Red	Nike	Stripe	Nike Shoes
Yellow	Jennifer Lopez	Stars	Red Jewelry
Green	Hot Wheels	Animal Print	Watches
Purple	American Flyer	Black Squares	Electronics
Black	Nike	-	Jumping Beans
Silver	NORDSTROM	Plane	Sterling Silver Chain
White	Croft & Barrow	Box	Damask Sheet Set
White Lemon	Sleepy Days	Stars	Pajama Set
Khaki	Cuddl Duds	-	Cozy Soft Bed
Black	Wool rich	Circle & Dots	Toys

2.3 Artificial Neural Network

Due to the complex data structure (Table 1) probability theories failed to provide sufficient accurate results. Therefore to improve the result and to deal complex data the focus went towards Artificial Neural Network (ANN). Humans perceive everything as a pattern, whereas for a machine everything is data. So, to think in the direction of human, machines should be trained to understand the data patterns.

Data's are the very important asset for machine and more the data processed the accuracy will gradually will increase or decrease [5]. Most of the prediction problems absolutely focused on pattern matching and recognition. Over the several decades in the filed of pattern recognition, neural network can be regarded as an extension of the many standard techniques. The common approach is to make use of feed-forward network architectures such as perceptron.

2.4 Single Layer Perceptron Model

In its simplest form Single Layer Perceptron is network that classify linearly separable pattern. Various techniques exist for determining the weigh (W) values in single-layer networks. It is widely studied in the 1960's and Widrow and Lehr [22] reviewed and converged for linearly separable pattern.

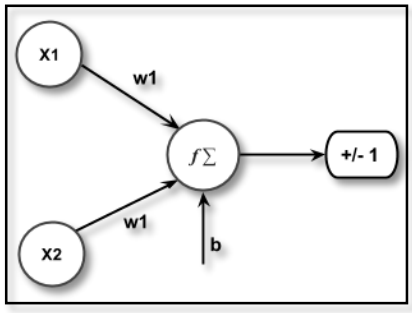


Fig. 5 Perceptron Neural Network

In this basic Perceptron schema (Figure 5) two inputs are the vector elements (x1, x2). The vectors are multiplied with respective weight element (w1, w2) and then summed. This produces a single value that is passed to a threshold function that has only two possible values (1 and -1). 'b' denotes the bias element to prune the model.

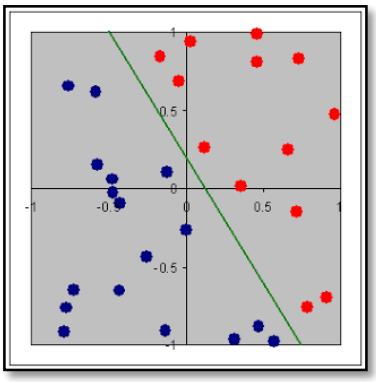


Fig. 6 Linearly separated dataset with Single Layer Perceptron

After Single layer perceptron training; the Figure 6 shows a linearly separated two classes by the green line. The blue dots belong to the first class and the red one belongs to the second. The Figure 6 shows 2 dimension case and by extending to 3 dimension by parse plane that should separate the patterns and dimension that are greater than 3 then it becomes hyper plane that will separate -1 patterns from the 1 patterns.

Therefore according to the retail domain data set (Table 1), it will only produce the correct prediction for data's that converges with AND function (Figure 7) and OR Function (Figure 8).

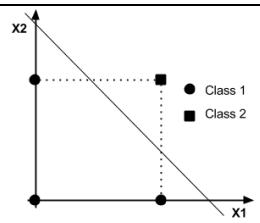


Fig. 7 AND Function

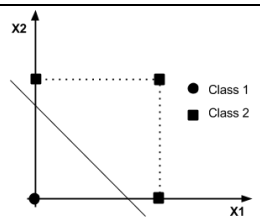


Fig. 8 OR Function

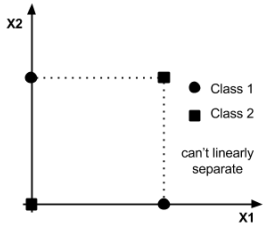


Fig. 9 EX-OR Function

The non-linear separable (Table 1) data that converges with Ex-OR function (Figure 9) cannot be solved with one layer of neural network. So, to improve the classification result need to consider multiple layers and it was proved by Minsky and Papert [5].

2.5 Multi Layer Perceptron Model

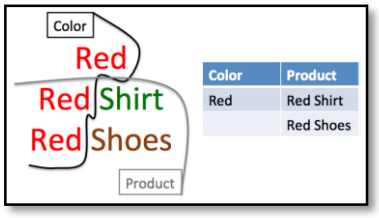


Fig.10 Retail complex data

The above Figure 10 obviously shows the complexity of the data and it's non-linearly spreadable data collection. Therefore MNB and single-layer perceptron models failed to produce a robust prediction.

To allow for more general mapping and to properly classify the data collection need to consider successive transformations corresponding to networks having several layers of adaptive weights. Neurons that are connected to each other's and that send each signal. Commonly neurons can be connected between any neurons and even themselves, but in that cases it gets difficult to train the data set.

Muliti-Layer Perceptron, neurons are arranged into layers (Figure 11). It consists of input layer, one or more hidden layers and an output layer. The signals are passed from one layer to another layer unlit it reaches to the threshold function. The training of the network is done by the highly popular algorithms know as the error back-propagation. This algorithm is based on the error correcting learning rules. Basically there are two passes through the different layers of the networks; forward pass and backward pass.

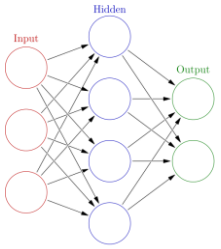


Fig. 11 Neurons mapping

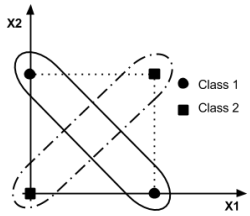


Fig. 12 Separated EX-OR Function

Therefore, with the shortlist of data collection with retail domain Multilayer Perceptron model allowed to classify non-linearly separable data collection (Figure 12).

3. Comparison Of Text Classification Approaches

The test result clearly proves that ANNs are superior over the traditional statistical model when relationship between output and input variables is implicit, complex and nonlinear.

The main strength of each of these algorithms is depended on certain criteria and depends on the context of the domain data. Table 2 presets a concise comparison of all the statistical approaches covered in the above analysis.

Corrêa and Ludermir [8] have done an experimental on ANNs model and the result were extraordinary. The below result (Table 2) expose clearly in general, the MLP Networks distinguished as best classifiers nonlinear data collection [8] and for linear separable data collection the Single Layer perceptron model has much more accurate than Multinomial Navie Base model which is know for probability.

Table 2 Comparison of text classification

Algorithm	Basic techniques	Strengths	Weakness	Highest accuracy reported (%)
Rules	1. Defined rules 2. Defined pattern matching	1. Robust approached for small data set 2. Easy to implement 3. For static data recommended 4. Best for Boolean search queries	1. Dynamic data it fails 2. Difficult to add rule for big data set 3. Can not guarantee	48.6
Navie Bayes	1. Supervised learning 2. Probability based classifier	1. Simple and robust algorithm. □ 2. Independence assumption minimizes computational complexity. 3. Wide applicability 4. Best for document classification	1. Susceptible to Bayesian Poisoning and unrelated video insertion 2. Failed to short text	72.5
Single Layer Perceptron	1. Supervised learning 2. One layer neural network	1. Classify linearly separable data collection 2. Difficult to solve complex problems	1. Non-linear separable data collection can not be classified 2. Lots of training data needed	85.3
Multi Layer Perceptron	1. Supervised learning 2. One or more hidden layer	1. Classify linearly and non-linearly separable data collection 2. Complex problems can be solved 3. Multiple neurons	1. Time consuming to train all the hidden layers	96.7

4. Conclusion

This paper presented a quantitative as well as qualitative comprehensive study of search engine text processing pipeline, text classification techniques and the approaches with the focus of retail domain. This paper includes the accuracy according the dynamic data set that frequently change and the approaches that were taken to make it more accurate. It also assessed the strength and the weakness of the models and the approaches that have taken to improve the search results.

From the research, the observation clearly explicit that significant work has to be done to improve the text class classification models and it obviously going to improve the returning search result which will help the consumers to purchase their needed items within short time period.

References

[1] AGARAM, M. K. & LIU, C. 2011. An Engine-independent Framework for Business Rules Development. Enterprise Distributed Object Computing Conference (EDOC). Helsinki: IEEE.

[2] AGARWAL, S. 2013. Data Mining: Data Mining Concepts and Techniques. Machine Intelligence and Research Advancement (ICMIRA), 2013 International Conference on. Katra: IEEE.

[3] AGGARWAL, C. C. & ZHAI, C. 2012. A survey of text classification algorithms. In: Mining Text Data. Springer.

[4] ALAVIJEH, A. Z. 2015. The Application of Link Mining in Social Network Analysis. Advances in Computer Science : an International Journal.

[5] BISHOP, C. M. 1996. Neural Networks for Pattern Recognition, Clarendon Press; 1 edition

[6] CESKA, Z. & FOX, C. 2009. The Influence of Text Pre-processing on Plagiarism Detection. International Conference RANLP. Borovets, Bulgaria.

[7] COLLOBERT, R., WESTON, J., BOTTOU, L., KARLEN, M., KAVUKCUOGLU, K. & KUKSA, P. 2011. Natural Language Processing (Almost) from Scratch. Journal of Machine Learning Research 12.

[8] CORRÊA, R. F. & LUDERMIR, T. B. 2002. Automatic Text Categorization: Case Study. IEEE Conference Publications. Neural Networks, 2002. SBRN 2002. Proceedings. VII Brazilian Symposium on.

[9] DEMERS, T. 2010. Short Vs. Long Tail: Which Search Queries Perform Best? [Online]. Search Engine Land. Available: http://searchengineland.com/short-vs-long-tail-which-search-queries-perform-best-36762 [Accessed July, 04th 2015].

[10] DREYER, K. 2014. Study: Consumers Demand More Flexibility When Shopping Online.

[11] FISHKIN, R. 2009. Illustrating the Long Tail [Online]. Moz. Available: https://moz.com/blog/illustrating-the-long-tail [Accessed July 04th 2015].

[12] GOOGLE, N. R. S. H. D. C. S. T. L. S. T. W. 2015. New Research Shows How Digital Connects Shoppers to Local Stores – Think with Google [Online]. Available: https://http://www.thinkwithgoogle.com/articles/how-digital-connects-shoppers-to-local-stores.html.

[13] HU, H., WEN, Y., CHUA, T.-S. & LI, X. 2014. Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. Access, IEEE. IEEE.

[14] KIBRIYA, A. M., FRANK, E., PFAHRINGER, B. & HOLMES, G. 2005. Multinomial Naive Bayes for Text Categorization Revisited. AI'04 Proceedings of the 17th Australian joint conference on Advances in Artificial Intelligence, 3339.

[15] LIU, B. 2011. Web Data Mining, SIGKDD Explorations, springer

[16] LUO, Y., WANG, W. & LIN, X. 2008. SPARK: A Keyword Search Engine on Relational Databases. Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on. IEEE.

[17] MAHAR, J. A. & MEMON, G. Q. 2010. Rule Based Part of Speech Tagging of Sindhi Language. Signal Acquisition and Processing, 2010. ICSAP '10. International Conference Bangalore.

[18] P.J, A., MOHAN, S. P. & K.P., S. 2010. SVM Based Part of Speech Tagger for Malayalam. Recent Trends in Information, Telecommunication and Computing (ITC), 2010 International Conference Kochi, Kerala.

[19] PARK, H. 2014. Bigger, Better, Faster, Stronger: The Future of Big Data [Online]. cmswire. Available: http://www.cmswire.com/cms/big-data/bigger-better-faster-stronger-the-future-of-big-data-027026.php 06 March 2015].

[20] QUAN, X., LIU, G., LU, Z., NI, X. & WENYIN, L. 2010. Short text similarity based on probabilistic topics. Knowledge and Information Systems, 25.

[21] WANG, W., LIN, X. & LUO, Y. 2007. Keyword Search on Relational Databases. Network and Parallel Computing Workshops, 2007. NPC Workshops. IFIP International Conference IEEE.

[22] WIDROW, B. & LEHR, M. A. 1990. 30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation. IEEE.

First Author Graduated from University of Westminster and holds BEng (Hons) in Software Engineering, Computer Science. Currently works as a Senior Software Engineer at Zone24x7 (Private) Limited. Have a quite a good experience in Research and Development side. An active developer with very broad knowledge and keen interests towards Big Data, NLP, 3D visualization, Computer vision technologies, modern web and mobile technologies.

Second Author Bachelor in Computing Science degree holder from the Staffordshire University, UK and currently a Senior Software Engineer at Zone24x7 (Private) Limited, who is passionate in data science and actively researching and developing in projects related to NLP distributed systems.

Third Author A Software Engineer currently works at Zone24x7 (Private) Limited and obtained B.Sc. (Hons) special degree in Computer Science from University of Sri Jayewardenepura. Interested key areas are NLP, Big data and web technologies.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

A Novel Performance Evaluation Approach for the College Teachers Based on Individual Contribution

XING Lining

College of Information System and Management, National University of Defense Technology
Changsha 410073, P.R. China
xinglining@gmail.com

Abstract

The performance evaluation to college teachers has very important theoretical significance and practical value. Therefore, a novel performance evaluation approach is proposed to the college teachers based on individual contribution. Experimental results suggest that this approach is feasible and efficacious.

Keywords: *Performance Evaluation, College Teacher, Individual Contribution*

1. Introduction

At present, there are two main kinds of teacher evaluation system, they are reward and punishment evaluation system and the developing evaluation system [1].

Reward and Punishment Evaluation System. It is aimed at strengthening the performance management of teachers, giving corresponding reward or punishment according to their performance [2]. This kind of evaluation system can mobilize teachers' enthusiasm and creativity only by external rewards, and it can punish incompetent teachers to make them improve deficiencies constantly and make progress. Teachers can develop constantly through awarding part of excellent teachers and punish part of incompetent teachers, and then promoting the increase of educational level [3]. There are some following disadvantages of reward and punishment teacher evaluation system: teachers concern much about the evaluation results, which is easy to cause that teachers are not willing to exchange their own advanced information with other teachers with the purpose of competing for rewards or promotion; this kind of system is a superincumbent teacher evaluation system, which brings teachers with great psychological stress, affecting the relationship between teachers and leaders, between teachers and teachers, between teachers and administrators; this kind of system overly concerns with short-term results, which makes the evaluation indicators overly quantization and unification, neglecting the individual differences of teachers and the exchange between estimators and the evaluated, so that teachers cannot preserve their own interests during the evaluation process [4-5].

Developing Evaluation System. It is aimed at promoting teachers' development, promoting teachers' professional development and promoting schooling quality by teacher evaluation in a relaxing and democratic atmosphere, thereby realizing a win-win situation among the college and teachers [6-8]. The evaluation system of developing teachers believes that intrinsic motivation is much more motivate than external motivation because teachers have got high-level education. It believes that self-motivation should be the primary because external pressure can only make them reach the minimum requirement while intrinsic motivation can make them develop great enthusiasm and mobilize their initiative [9]. If giving them necessary working conditions, they can make their works excellent. The evaluation system of developing teachers provides teachers with necessary working conditions in a relaxing and democratic atmosphere, cultivating their professional ethics, mobilizing their working enthusiasm, stimulating their working enthusiasm, and then to realize the management and development objectives while in meeting teachers' self-value needs; the evaluation system of developing teachers lays emphasis on the subjectivity and difference of teachers [10]. The subjectivity mainly reflects in the evaluation process that it pays great attention to the dominant role of teachers, trusting and respecting them, attaching importance to the equal dialogue, exchange and communication between estimators and the evaluated. The difference mainly reflects in adopting different evaluation criterion to implement discrepancy evaluation according to teachers' different backgrounds, personalities, teaching styles and the current stage of their career [11]. Teachers participate in the evaluation process actively, estimators gather information from multiple channels and then evaluate teachers, feedback evaluation results in time and apply the evaluation results scientifically and reasonably.

2. Individual Contribution

In 1929, people held a tug-of-war test in Rangeland Germany. When a person participated in the tug-of-war, the power he or she contributed was 63 kg; while eight person participated in the tug-of-war, the power they contributed

was 63×8=504 kg in theory, however, the actual power was just 248 kg, which occupied 49% of the theoretical value; each person only contributed half of his or her whole power [12]. In 1979, Latanne etc. also did a similar test: when a person did his or her best to scream and clap, the voice he or she contributed was 100%, while six person screamed and clapped together, each person only contributed 40% of their voice in average. We are wondering how an administrator should motivate and stimulate employees to contribute more of their power after seeing the abovementioned two tests [13].

Why does it happen that the contribution of the same person will decrease in a team? It is called 'Social Loafing' in academe. In the actual life with increasingly fierce competition, the occurrence probability of 'Social Loafing' is higher in greater-scale and better-brand organization. The half-heartedness resulted by 'Social Loafing' can be summarized as 'Loss of coordination' and 'Reducing of sense of responsibility' from an administrative perspective. 'Reducing of sense of responsibility' is first because the personal contribution degree is not specific and then shuffling off responsibility onto others; the second is because the administrators do not implement organization objectives to each team member definitely. It cannot be prevented of the decreasing of individual initiative without specific measurement of personal contribution degree and impartial evaluation [14-15].

The paper will define personal contribution degree as: 'the increased degree of comprehensive competitiveness index in a team when adds someone' or 'the decreased degree of comprehensive competitiveness index in a team when reduces someone'. Obviously, the larger the range caused by adding (or reducing) someone to increase (or decrease) the team comprehensive competitiveness index, the more personal contribution the person made to the team, who should be one of the members to be taken seriously or protected; otherwise, the less personal contribution the person made to the team, who should be dispensable role in the team [16].

In the performance appraisal process of teachers in universities and colleges, the essay will measure the comprehensive competitiveness index of a team from five aspects: teaching (having lessons and counselling students), scientific research (application and implementation of scientific research projects), academic (basic research and composing academic paper), engineering development (design and development of engineering projects) and laboratory management (daily management of laboratory). In order to simplify the treatment of problems, the essay supposes that the works teachers have done in these five aspects all can be quantized according to specific standard; in a team, unilateral workload can be accumulated [17].

3. Performance Evaluation Approach

The paper applies TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method to synthesize the workloads in teaching, scientific research, academic, engineering development and laboratory management, and then get the comprehensive competitiveness index of multiple teams with different individuals. TOPSIS method is ordering according to the limited evaluation objects and the approaching degree of ideal solution, and making the evaluation of relative superior or inferior among existing objects. Ideal solution has positive ideal solution and negative ideal solution, the best evaluation object should be closest to the positive ideal solution and be farthest to the negative ideal solution. The basic principles of TOPSIS method: ordering by testing the distance of evaluation object to the positive ideal solution and the negative ideal solution, if the evaluation object is the closest to the positive ideal solution and the farthest to the negative ideal solution, then he or she is the best; if not, he or she is the worst.

Let $M=\{1,2,\ldots,m\}$ and $N=\{1,2,\ldots,n\}$, suppose that the decision scheme set is $U=\{u_i\}$, the attribute set $V=\{v_j\}$, the index weight set $W=\{w_j\}$, the decision matrix $A=(a_{ij})_{m\times n}$ ($i\in M$, $j\in N$), where a_{ij} is the value obtained from the measure by scheme u_i according to index v_j, w_j is the index weight to be determined and $\sum_{j=1}^{n} w_j = 1$, then the quadruple $\langle U, V, W, A\rangle$ constitutes the mathematical model.

The physical dimensions of various indexes in the attribute set may be different, thus the decision matrix should be normalized by following some rules before making a decision. There are several attribute types, including the benefit-type, cost-type, fixed-type and interval-type, of which the most commonly used types are the benefit-type and the cost-type. Suppose that I_1 and I_2 are respectively the subscript sets of the benefit-type and cost-type attributes, and the normalized decision matrix can be written as $B=(b_{ij})_{m\times n}$, then the normalized formulas for the benefit-type and cost-type indexes are respectively

$$b_{ij} = \begin{cases} \dfrac{a_{ij} - a_{i\min}}{a_{i\max} - a_{i\min}}, & j \in I_1 \\ \dfrac{a_{i\max} - a_{ij}}{a_{i\max} - a_{i\min}}, & j \in I_2 \end{cases} \quad (1)$$

In which, $a_{i\max} = \max_{i\in M}\{a_{ij}\}$ and $a_{i\min} = \min_{i\in M}\{a_{ij}\}$.

(1) Determination of index weights by entropy method. The entropy H_j for the j^{th} index v_j calculated by the normalized decision matrix $B=(b_{ij})_{m\times n}$ is

$$H_j = -k\sum_{i=1}^{m}(\overline{b_{ij}}\ln\overline{b_{ij}}),\ \ j \in N \tag{2}$$

In here, $k=(\ln\ m)^{-1}$, $\overline{b_{ij}} = \dfrac{b_{ij}}{\sum_{i=1}^{m} b_{ij}}$,and suppose that when $\overline{b_{ij}} = 0$, $\overline{b_{ij}}\ln\overline{b_{ij}} = 0$. The entropy weight of the j^{th} index v_j calculated by H_j is

$$w_j = \frac{1-H_j}{n-\sum_{j=1}^{n} H_j},\ \ j \in N \tag{3}$$

After the index weight is determined, the weighted normalized decision matrix can be written as $C=(c_{ij})_{m\times n}$, its computational formula is

$$c_{ij} = w_j \times b_{ij},\ i \in M;\ \ j \in N \tag{4}$$

(2) Nearness Degree Computing. In this section, suppose $\Phi^+ = (c_j^+)$ and $\Phi^- = (c_j^-)$ ($j \in N$) are respectively the positive and negative ideal points, in which

$$c_j^+ = \max_{i\in M} c_{ij} \tag{5}$$

$$c_j^- = \min_{i\in M} c_{ij} \tag{6}$$

Written as $\Psi^+ = (d_i^+)$, $\Psi^- = (d_i^-)$ ($i \in M$), where

$$d_i^+ = \sqrt{\sum_{j=1}^{n}(c_{ij} - c_j^+)^2} \tag{7}$$

$$d_i^- = \sqrt{\sum_{j=1}^{n}(c_{ij} - c_j^-)^2} \tag{8}$$

d_i^+ and d_i^- are respectively the nearness degrees of scheme u_i to the positive ideal point Φ^+ and the negative ideal point Φ^- . Their physical meaning is that: the smaller d_i^+ and d_i^- , the larger the degrees of similarity between scheme u_i and the positive and negative ideal points respectively.

(3) Calculation for comprehensive index values. Suppose that the vector of the comprehensive ranking index value for scheme u_i is $Z=(z_i)(i \in M)$, in which

$$z_i = \frac{d_i^-}{d_i^- + d_i^+} \tag{9}$$

The schemes are sorted according to the comprehensive index values, and the larger the comprehensive index values, the better the schemes.

(4) General steps of general TOPSIS. Based on the above analysis, the solving steps of TOPSIS ranking model are listed as follows.

Step 1: Suppose that there is a MADM problem, and its decision matrix is $A=(a_{ij})_{m\times n}$, then the normalized decision matrix $B=(b_{ij})_{m\times n}$ is obtained by (1);

Step 2: The index weights w_j are calculated by (2) and (3), and the weighted normalized decision matrix $C=(c_{ij})_{m\times n}$;

Step 3: The positive ideal point $\Phi^+ = (c_j^+)$ and the negative one $\Phi^- = (c_j^-)$ are solved by (5) and (6), and the nearness degrees of scheme u_i to c^+ and c^- by (7) and (8);

Step 4: The comprehensive ranking index value z_i of scheme u_i is solved by (9), and determine the relative merits of the schemes using the values of z_i.

4. Experimental Results

The paper supposes that there are 10 teachers in the team, their workloads in teaching, scientific research, academic, engineering development and laboratory management are as Table 1. From Table 1, we can see that the first teacher is outstanding in all aspects; the second to fifth teachers are in the medium in all aspects; the sixth to tenth teachers are outstanding in a single aspect of teaching, scientific research, academic, engineering development and laboratory management, and a slightly short in other aspects. Then, we will use a method based on individual contribution degree to evaluate the performance of each teacher.

Table 1: The Workload of Each Teacher in the Team

No.	Teaching (class hours)	Scientific research (thousand yuan)	Academic (scores)	Engineering development (hours)	Laboratory management (hours)
1	100	40	30	150	150
2	50	25	15	100	110
3	55	21	16	110	105
4	50	30	15	120	100
5	60	20	15	110	110
6	800	10	10	50	50
7	30	320	10	50	50
8	30	10	240	50	50
9	30	10	10	1600	50
10	30	10	10	50	1600

Table 2 enumerates that the total team workloads in teaching, scientific research, academic, engineering development and laboratory management after missing any a teacher. Table 2 also provides the positive ideal solution and the negative ideal solution under ten situations.

In order to simplify the treatment of problems, we synthesize the workloads of teaching, scientific research, academic, engineering development and laboratory management and set all their weight coefficient as 0.2. Table 3 provides the comprehensive competitiveness index of each team under different situations. We can calculate the individual contribution degree of single teacher by comparing the comprehensive competitiveness index of a team with missing a teacher and the comprehensive competitiveness index of a team without missing a teacher (see Table 3).

Table 2: The Total Team Workloads under Different Situations

No.	Group status	Teaching (class hours)	Scientific Research (thousand yuan)	Academic (scores)	Engineering Development (hours)	Laboratory Management (hours)
1	The first teacher is missing	1135	456	341	2240	2225
2	The second teacher is missing	1185	471	356	2290	2265
3	The third teacher is missing	1180	475	355	2280	2270
4	The fourth teacher is missing	1185	466	356	2270	2275
5	The fifth teacher is missing	1175	476	356	2280	2265
6	The sixth teacher is missing	435	486	361	2340	2325
7	The seventh teacher is missing	1205	176	361	2340	2325
8	The eighth teacher is missing	1205	486	131	2340	2325
9	The ninth teacher is missing	1205	486	361	790	2325
10	The tenth teacher is missing	1205	486	361	2340	775
11	No one missing	1235	496	371	2390	2375
12	Positive ideal solution	1235	496	371	2390	2375
13	Negative ideal solution	435	176	131	790	775

Table 3: Experimental Results of Simulation Example

No.	Group status	Comprehensive Competitiveness Index of Team	Individual Contribution of Single Teacher
1	The first teacher is missing	0.8872	0.1128
2	The second teacher is missing	0.9329	0.0671
3	The third teacher is missing	0.9329	0.0671
4	The fourth teacher is missing	0.9279	0.0721
5	The fifth teacher is missing	0.9323	0.0677
6	The sixth teacher is missing	0.6623	0.3377
7	The seventh teacher is missing	0.6630	0.3370
8	The eighth teacher is missing	0.6629	0.3371
9	The ninth teacher is missing	0.6526	0.3474
10	The tenth teacher is missing	0.6508	0.3492
11	No one missing	1	—

It can be seen from the results of Table 3: (1) the first teacher is outstanding in all aspects and his or her individual contribution degree to the team is higher; (2) the second to fifth teachers are in the medium in all aspects, their individual contribution degrees to the team are lower; the sixth to tenth teachers are extremely outstanding in a single aspect of teaching, scientific research, academic, engineering development and laboratory management although they are a slightly short in other aspects, they have higher individual contribution degrees to the team.

5. Conclusions

In conclusion, if a teacher in universities and colleges can be outstanding in many aspects like teaching, scientific research, academic, engineering development and laboratory management, then his or her individual contribution degree to a team is higher, which is an ideal development pattern; at the same time, if he or she is extremely outstanding in a single aspect of teaching, scientific research, academic, engineering development and laboratory management, he or she can also make great contribution to the team, and this part of person are deserved to be provided with good development opportunities and wide development platform.

Acknowledgment

The authors thank the reviewers for valuable comments and constructive suggestions. We also thank the editor-in-chief, associate editor and editors for helpful suggestions. This research work was supported by the major project of undergraduate teaching and education research among Twelfth-Five-Year-Plan at National University of Defense Technology (U2012006), and the major project of Education Science Planning among Twelfth-Five-Year-Plan at Hunan province (XJK014AGD001).

References

[1] Chen CT, Lin CT, Huang SF (2006). A fuzzy approach for supplier evaluation and selection in supply chain management. International Journal of Production Economics, 102(2), 289-301

[2] Gan Q, Zhang FC, Zhang ZY. Multi-objective optimal allocation for regional water resources based on ant colony optimization algorithm. International Journal of Smart Home, 2015, 9(5): 103-110

[3] Firouzabadi SMAK, Henson B, Barnes C (2008). A multiple stakeholders' approach to strategic selection decisions. Computers and Industrial Engineering, 54(4), 851-865

[4] Yang YS, Xu BW, Li JJ. An improved genetic algorithm for fast configuration design of large-scale container cranes. Journal of Information and Computational Science, 2015, 12(11): 4319-4330

[5] Gumus AT (2009). Evaluation of hazardous waste transportation firms by using a two-step fuzzy-AHP and TOPSIS methodology. Expert Systems with Applications, 36(2), 4067-4074

[6] Wang FJ, Li JL, Liu SW, et al. An improved adaptive genetic algorithm for image segmentation and vision alignment used in microelectronic bonding. IEEE / ASME Transactions on Mechatronics, 2014, 19(3): 916-923

[7] Abo-Sinna MA, Amer AH, Ibrahim AS (2008). Extensions of TOPSIS for large scale multi-objective non-linear programming problems with block angular structure. Applied Mathematical Modelling, 32(3), 292-302

[8] Chen TY, Tsao CY (2008). The interval-valued fuzzy TOPSIS method and experimental analysis. Fuzzy Sets and Systems, 159(11), 1410-1428

[9] Wang HF. Computer networks reliability using improved genetic algorithm optimization. Energy Education Science and Technology Part A: Energy Science and Research, 2014, 32(6): 8737-8742

[10] Lin MC, Wang CC, Chen MS, Chang CA (2008). Using AHP and TOPSIS approaches in customer-driven product design process. Computers in Industry, 59(1), 17-31

[11] Opricovic S, Tzeng GH (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. European Journal of Operational Research, 156(2), 445-455

[12] Changdar C, Mahapatra GS, Pal RK. An improved genetic algorithm based approach to solve constrained knapsack problem in fuzzy environment. Expert Systems with Applications, 2015, 42(4): 2276-2286

[13] Wang TC, Chang TH (2007). Application of TOPSIS in evaluating initial training aircraft under a fuzzy environment. Expert Systems with Applications, 33(4), 870-880

[14] Zuo XQ, Chen C, Tan W, et al. Vehicle Scheduling of an Urban Bus Line via an Improved Multiobjective Genetic Algorithm. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 1030-1041

[15] Zou YY, Zhang YD, Li QH, et al. Improved multi-objective genetic algorithm based on parallel hybrid evolutionary theory. International Journal of Hybrid Information Technology, 2015, 8(1): 133-140

[16] Wang YM, Elhag TMS (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. Expert Systems with Applications, 31(2), 309-319

[17] Wang YJ, Lee HS (2007). Generalizing TOPSIS for fuzzy multiple-criteria group decision-making. Computers and Mathematics with Applications, 53(11), 1762-1772

Lining Xing was born in 1980. He received his Ph.D at the National University of Defense Technology, P.R. China in 2009. Now he is one research fellow at National University of Defense Technology. His research interests include mission planning, intelligent optimization and resource scheduling.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Text Anomalies Detection Using Histograms of Words

Abdulwahed Almarimi[1] and Gabriela Andrejková[2]

[1,2]Institute of Computer Science, Faculty of Science, P. J. Šafárik University in Košice
04001 Košice, Slovakia
abdoalmarimi@gmail.com
gabriela.andrejkova@upjs.sk

Abstract

Authors of written texts mainly can be characterized by some collection of attributes obtained from texts. Texts of the same author are very similar from the style point of view. We can consider that attributes of a full text are very similar to attributes of parts in the same text. In the same thoughts can be compared different parts of the same text. In the paper, we describe an algorithm based on histograms of a mapped text to interval $\langle 0,1 \rangle$. In the mapping, it is kipped the word order as in the text. Histograms are analyzed from a cluster point of view. If a cluster dispersion is not large, the text is probably written by the same author. If the cluster dispersion is large, the text will be split in two or more parts and the same analysis will be done for the text parts. The experiments were done on English and Arabic texts. For combined English texts our algorithm covers that texts were not written by one author. We have got the similar results for combined Arabic texts. Our algorithm can be used to basic text analysis if the text was written by one author.

Keywords: *Authorship attribution, stylometry*, anomaly detection, histogram

1. Introduction

In the text processing, there are solved many problems connected to an authorship of texts, for example external plagiarism, internal plagiarism, authorship verification, document verification, authorship attribution. The problems are followed in a PAN competition [http://pan.webis.de/], benchmarked texts are on web page [1], [2] and [3]. The Authorship Attribution (AA) problem is formulated as a problem to identify the author of the given text from the group of potential candidate authors. The solutions of AA problem are based on basic features extracted from the text (statistics information, stylistics information, syntax information and semantics information). Some interesting approaches to solving of the problem can be found in [4], [5], [6],[7],[8] and [9].
Some different situation is in the Text Anomaly Detection (TAD) problem. The problem is formulated as an identification of text parts they have unseen behavior in a comparison to full text or to the other parts of the text. The result on anomalies could answer to the question "if the given text was written by one author or it has some parts probably written by the other authors". In the paper we developed an algorithm to cover some anomalies in the texts using histograms of mapped texts to time sequences and modified by a kernel smoothing function. The motivation to the algorithm we found in [6].

The paper is written in the following structure: The second section describes some background and statistics of some analyzed Arabic and English texts. In the third section, it is described our developed method, the algorithm TAD_Histo. The fourth section contains results for analyzed texts. In the conclusion, we formulate summary of results and the plan of the following research.

2. Basic Background and Text Statistics

In the text anomaly detections we used English texts from benchmark [1] and Arabic texts from [2], [3].

2.1 English Texts

We illustrate some statistical analysis of 5 English analyzed texts in Table 1. We show the number of words, the number of letters and the number of different words, the number of words by the length 3 and 4 with percentage. More statistics we described in the paper [10].

Table 1: Statistics of 5 English Texts,

The number	E1	E2	E3	E4	E5
of words	176598	132020	125487	106359	93085
of letters	874761	607783	498696	471566	417899
of diff. words	22954	19268	15853	15321	12929
words by the length 3	**38733** 21.93%	**27599** 20.90%	**23780** 18.95%	**26497** 24.91%	**19875** 21.35%
words by the length 4	26321 14.90%	20909 15.83%	18429 14.68%	19224 18.07%	18520 19.89%

In the Fig. 1 we show graphs of frequencies for five Arabic and English texts. In the left panel there are Arabic texts and in the right panel English texts. In the

Arabic texts the highest percentage of occurrences is for words of the length 3 and 4. In the English texts the highest percentage of occurrences is for words of the length 3. In English texts the majority of words has the length 1-15, bigger than Arabic texts where the majority words has the length 1-10.

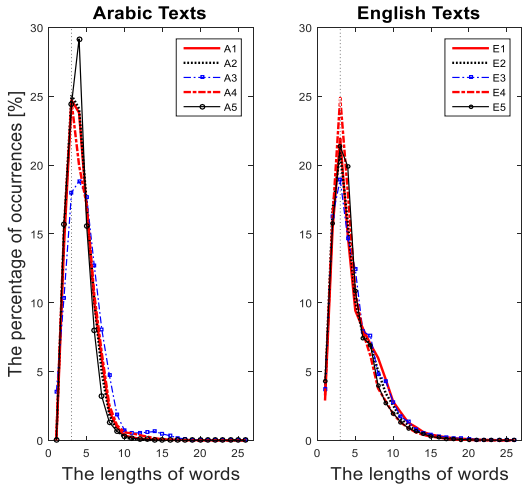


Fig. 1. The percentage of occurrences according to the lengths of words in 5 Arabic and 5 English texts. We prepared the figure in [11].

2.2 Arabic Texts

In the Table 2, we illustrate the number of words, letters and different words, we describe the number of words by the length 3 and 4 with percentage of five Arabic texts. More statistics we described in the paper [10].

Table 2: Statistics of 5 Arabic Texts,

The number	A1	A2	A3	A4	A5
of words	94197	48358	51938	31656	39340
of letters	395065	198019	247448	135573	152905
of diff. words	14110	9061	25755	10098	7036
Words by the length 3	**23287** 24.72%	**12130** 25.08%	9353 18.00%	**7795** 24.62%	9619 24.45%
Words by the length 4	22426 23.80%	11653 24.09%	**9779** 18.82%	6324 19.97%	**11476** 29.17%

2.3 Basic Background

We will use the following symbols:

- Γ - a finite alphabet of letters; $|\Gamma|$ is the number of letters in Γ ; in our texts: Γ_A is Arabic alphabet and Γ_E is English alphabet.
- V - a finite vocabulary of words in the alphabet Γ presented in the alphabetic order; $|V|$ - the numbers of words in the vocabulary V ;
- D - text; a finite sequence of words, D = $\langle w_1,...,w_n\rangle; w_i \in V; N -$ the number of words in the texts.
- $D = \langle d_1 d_2 ... d_{|D|}\rangle; |D|$ - the number of symbols in the text D;

3. Histograms Method

The analysis using n-gram profiles method [11] do not keep the sequence of the words, it follows occurrences of the n-grams (sequences of letters) in texts. If the vocabulary V is alphabetically ordered, then it is possible to do its mapping to integer numbers 1…|V|. The texts should be considered as integer valued time series, the sequences of the numbers of words in the vocabulary V. According to [5] and [12] the sequences of words can be followed in time using weighted bag of words (lowbow). Lowbows can follow a track of changes in histograms connected to words through all text. Histograms will be done in the interval $\langle 0,1\rangle$ and it is necessary to map texts into the interval $\langle 0,1\rangle$. A length normalized text x_D is a function $x_D : \langle 0,1\rangle \times V \to \langle 0,1\rangle$ such that

$$\sum_{j \in V} x_D(t, j) = 1, \forall t \in \langle 0,1\rangle.$$

If f_D^j is the frequency of $j \in V$ in the text D then $x_D(t, j) = f_D^j / N$ of word *j* in the mapping position *t*. The mapping to the interval is important because of the different lengths of texts.

The main idea behind the locally weighted bag of words framework [12] is to use a local smoothing kernel to smooth the original word sequence temporally. The first version of our modified algorithm was published in [5]. We have developed a new modification in the last 3 steps and it can be formulated in the following steps.

The algorithm TAD_Histo:

Step 1:

To map a text *D* to the interval $\langle 0,1\rangle$. Let $t = (t_1, t_2, ..., t_N) = (1/N, 2/N, ...N/N)$ be the vector of values from $\langle 0,1\rangle$, $\sum_{j=1}^{N} t_j = (N+1)/2$. Each t_j should be associated to the word in the position *j* in the text.

Mapping of the text is

$$MD(t) = \langle md_{t1}, md_{t2}, ...md_{tN}\rangle,$$

where $md_{t_i} = f_D^j / N$, *i* is a word in the text *D* and *j* is index of word I in the vocabulary *V* .

Step 2:

Let $K^{s}_{\mu,\sigma}(x):\langle 0,1\rangle \to R$ be some kernel smoothing function with location parameter $\mu, \mu \in \langle 0,1\rangle$ and a scale parameter σ. We take k positions of parameter $\mu, (\mu_1, \mu_2, ...\mu_k)$, such that $\sum_{j=1}^{k} k^{s}_{\mu,\sigma}(t_j) = 1.$

It is possible to use Gaussian Probability Density Function (PDF) (7) restricted to the interval $\langle 0,1\rangle$ and renormalized

$$N(x;\mu;\sigma) = \frac{1}{\sigma\sqrt{(2\pi)}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad (1)$$

We will use a modification of the function N, the function (2)

$$K^{s}_{\mu,\sigma}(x) = \begin{cases} \frac{N(x,\mu,\sigma)}{\phi(1,\mu,\sigma)-\phi(0,\mu,\sigma)}, & if\ x \in \langle 0,1\rangle, \\ 0, & otherwise. \end{cases} \quad (2)$$

where $\phi(x)$ is a Cumulative Distribution Function (CDF)

$$\phi(x,\mu,\sigma) = (1 + er\ f(\frac{x-\mu}{\sigma\sqrt{(2)}})), \quad (3)$$

where $er\ f(x) = \frac{1}{\sqrt{(pi)}} \int_{-x}^{x} \exp(-t^2) dt.$

Compute vectors $K^{s}_{\mu,\sigma}(t)$ for each positions $\mu_j, j = 1,...,k$ and chosen σ.

Step 3:

Compute local modified vectors LH^{i}_{D} for each position $\mu_j, j = 1,...,k$ as follows:

$$LH^{j}_{D}(t) = MD(t) \times K^{s}_{\mu_j,\sigma}(t) \quad (4)$$

LH^{i}_{D} present the sequences of vectors for a computation of histograms usable for some possible analysis of the texts.

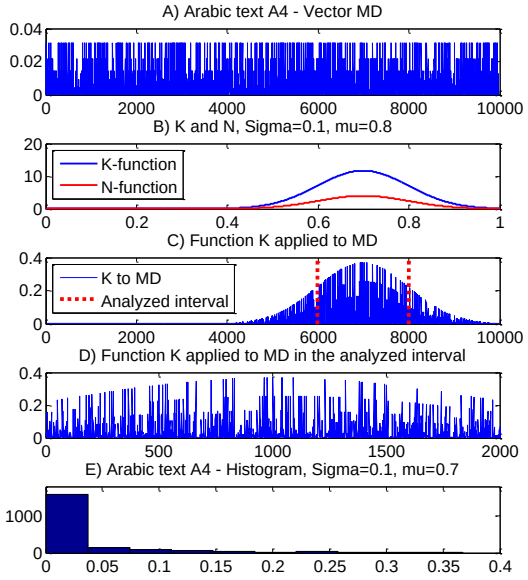
Step 4:

a) Reduce LH^{i}_{D} vectors to analyzed intervals $\langle ibeg, iend\rangle$. In the intervals the values of the function K are higher than some constant C_K. The value of C_K is developed in experiments.
Let the reduced LH^{i}_{D} be $LH^{i}_{D}r$.

b) Compute histograms to $LH^{i}_{D}r$ in *q* equidistant intervals. We will get *k* histogram points in *q*-dimensional space.

Step 5:

Cluster analysis will be applied to cover if all histogram points belong to the same cluster. If all histogram points are in one compact cluster then the analyzed text is probably written by one author if the points belong to more clusters then some anomalies were found in the text and analyzes will continue by splitting the text into two or more parts.
Compute the center C of the cluster and analyze distances histogram points from the center.



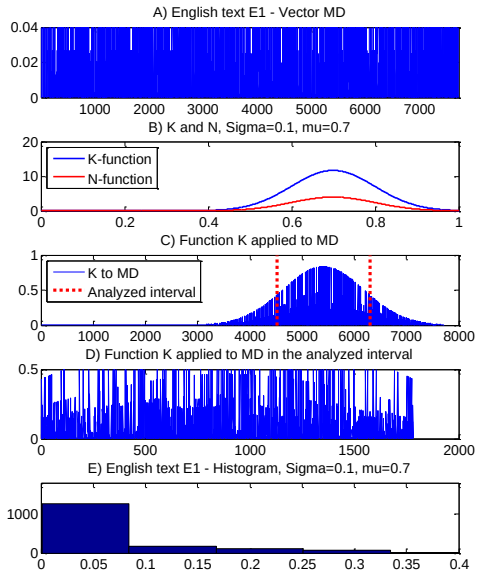


Fig. 2. The illustration of the algorithm on Arabic text A4 and English text E1. In the panels A), the vectors MD are constructed for text words. The prepared smooth functions $K^{s}_{\mu,\sigma}$ and N, where sigma=0.1 and mu=0.7, are shown in the panels B). The panels C) show the application of function K to vectors MD. In the panels D) there are shown values of the analyzed intervals and the down panels E) represent the histograms of the result in the panels C).

The steps 1-3, 4a) of the algorithm are shown in Fig. 2 using Arabic text A4 and English text E1.
The step 4b) is illustrated in Fig. 3 using Arabic text A4. In Fig. 3, there are shown some applications of Kμ function to MD vector for different values μ in the first and the third columns. Histograms are plotted in the second and the fourth columns.

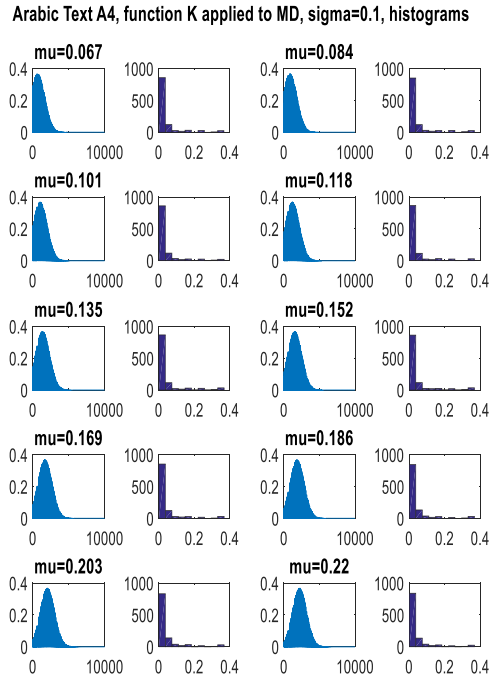


Fig. 3. The panels in the first and third column illustrate application of the function K with different values μ to MD of text A4. The second and fourth column show histograms of the analyzed histograms.

The step 5 of the algorithm is illustrated in Fig. 4 using Arabic text A4 and English text E1. To a visualization of the histogram points, *q* – dimensional space was reduced to 3 dimensional using dimensions 2, 3 and 4. The first interval of $LH_D^i r$ histogram contains all values closed to 0, it means we did not use them in the visualization but in the evaluation complete histogram points were used.

4. Evaluation

The analysis of Arabic text A4 and English text E1 is shown in Figs. 2-4. According to the presented visualization of results it is visible that the points are in one cluster.

In our experiments we used the following values of constants: C_K=10 for Arabic texts and C_K=8, 9 for English texts, q=10, σ = 0.1, μ_1 = 0.05, μ_{i+1} = μ_i+0.17, i=1, 2,… 49. The number of prepared histograms for each text was 50.
The results of 6 Arabic and 6 English texts are shown in the Tables 3 and 4. Five English texts were chosen from [9] and five Arabic texts from [7], [8]. In the Tables, there are coordinates of the centers, the distance and the index of the histogram point with the maximal distance d_{max} from the center and the number of points in the bigger distance than 50% from d_{max}.

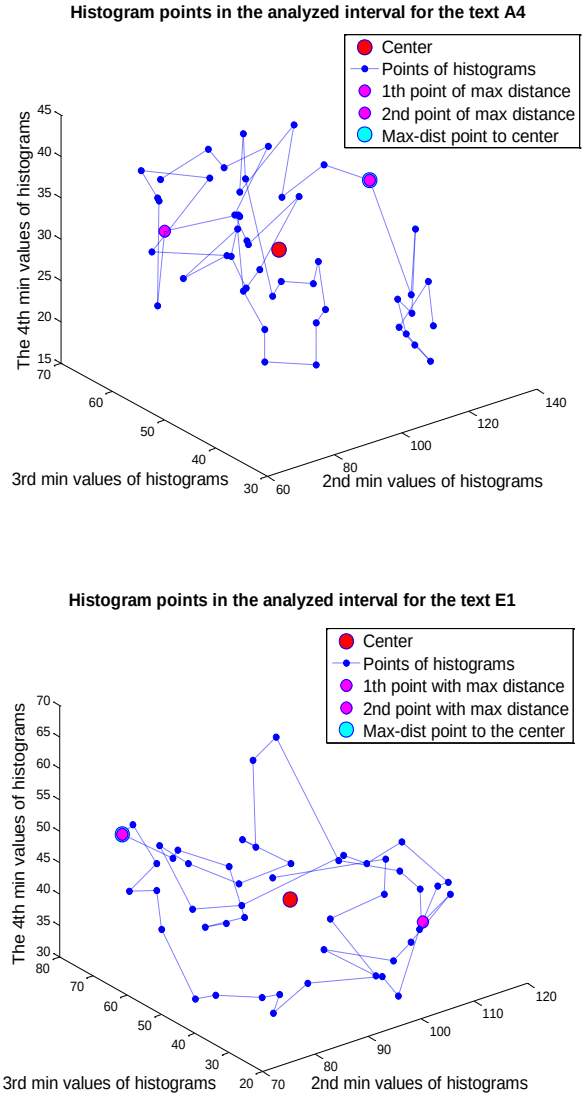


Fig. 4. The first panel illustrates histogram points of Arabic text A4, the second panel shows histogram points of English text E1. Except histogram points, we show the centers of the clusters, the points with the maximal distance from the center and two pints with the maximal distance.

Our first idea is to compare the numbers of histogram point in the distance less than 50% of d_{max} – d50l to the number of histogram points in the bigger distance than d_{max} – d50b.

Let p50 be

$$p50 = \frac{d50l}{d50b} \qquad (5)$$

If p50>1.00 then the cluster is quite compact else it is necessary to do some new analysis. In the text, there are some anomalies.

Table 3: The results of 5 English original and 1 combined texts.

Texts	Coordinates of the histogram center	d_{max} - Maximal distance between the center and histogram point	50% of distance / p50
E1	[536.2200, 70.3400, 34.8600, 33.0800, 17.4800, 30.2800, 16.3800, 30.5600, 0, 0]	94.5163 H-Point: 1	47.2581 36/14 = 2.5714
E2	[9010.2, 1218.6, 972.5, 283.9, 134.0, 173.9, 206.6, 584.1, 601.0, 0]	498.4786 H-Point: 1	249.2393 46/4= 11.5000
E3	[8773.7, 767.7, 980.5, 416.1, 226.6, 110.9, 564.1, 285.9, 84.3, 0]	428.9243 H-Point: 50	214.4621 39/11= 3.5454
E4	[6962.7, 1061.3, 722.7, 819.8, 302.5, 558.6, 181.6, 291.8, 387.9, 0]	383.1302 H-Point: 32	191.5651 32/18= 1.2777
E5	[6434.2, 954.7, 381.2, 716.9, 294.2, 293.9, 264.2, 143.9, 400.9, 0]	214.7325 H-Point: 1	107.3663 35/15= 2.3333
E1-E5	[1188.7, 163.6, 97.3, 93.4, 44.8,1.5, 90.6, 94.1, 0, 0]	113.7178 H-Point: 26	56.8589 24/26= 0.9230

Table 4: The results of 5 Arabic original and 1 combined texts.

Texts	Coordinates of the histogram center	d_{max} - Maximal distance between the center and histogram point	50% of distance / p50
A1	[7674.2, 1111.5, 565.9, 571.3, 164.4, 183.7, 203.8, 0, 0, 0]	508.2270 H-Point: 49	254.1135 30/20= 1.5
A2	[3864.5, 670.4, 291.8, 160.6, 268.6, 0, 49.3, 87.2]	280.4267 H-Point: 1	140.2133 32/18= 1.7777
A3	[5089.7, 286.9, 97.6, 96.7, 187.8, 38.1, 0, 0, 0, 0]	105.4364 H-Point: 29	52.7182 33/17= 1.9411
A4	[862.0, 97.62, 52.22, 29.9, 16.48, 8.06, 14.88, 1.94, 19.86, 0]	48.9918 H-Point: 11	24.4959 27/23= 1.1739
A5	[2574.0, 229.0, 229.5, 114.8, 114.6, 143.2, 170.9, 96.4, 87.7, 0]	201.8314 H-Point: 1	100.9157 39/11= 3.5454
A1-A3	[2394.5, 230.6, 192.2, 69.9, 1.5, 79.9, 42.8, 0, 0, 0]	341.5122 H-Point: 50	170.7561 24/26= 0.9230

Tested texts not written by one author were prepared as a combination of 2 texts. One part was taken from the first text and the second part was taken from the second text. It was clear that for us that the combined text was not written by one author. The Arabic combined text A1-A3 and the English combined text E1-E5 are analyzed in Fig. 4. It is possible to see different structure of vector MD in both parts and using function K to construct histograms does important some parts according to μ.

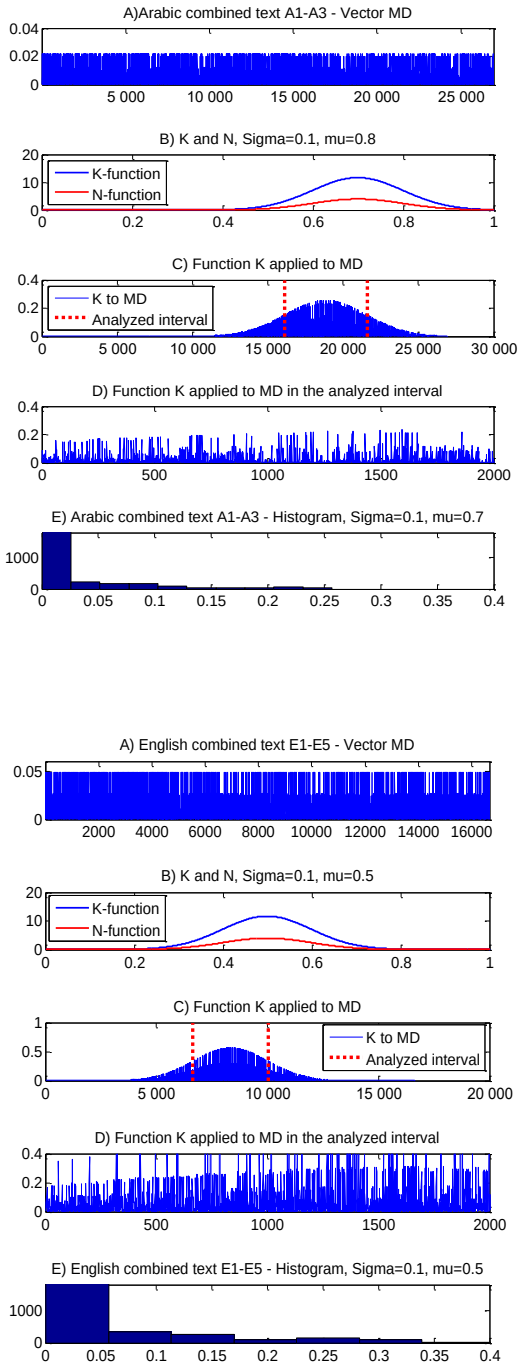


Fig. 5. The illustration of the algorithm on Arabic combined text A1-A3 and English combined text E1-E5. In the panels A), the vectors MD are constructed for text words. The prepared smooth functions $K^{s}_{\mu,\sigma}$ and N, where sigma=0.1 and mu=0.7, are shown in the panels B). The panels C) show the application of function K to vectors MD. In the panels D) there are shown values of the analyzed intervals and the down panels E) represent the histograms of the result in the panels C).

The analysis of histogram points is plotted in Fig. 5. It is visible that the histogram points are arrange in two clusters. The value p50 for Arabic combined text is 0.9230 and the English combined text is 0.9230. It means our method covers in both texts some anomalies.

In the experiments, we found that the constant C_K is different for Arabic and English texts.

5. Conclusion

In the paper, we developed the modified algorithm to cover anomalies in the paper. It is based on histograms computed on mapped texts but it is the order of word in the text is not disturb. We illustrate results for 6 Arabic and 6 English texts. Our method finds anomalies in artificial combined texts. Our next plan is to do some statistics on benchmarked texts.

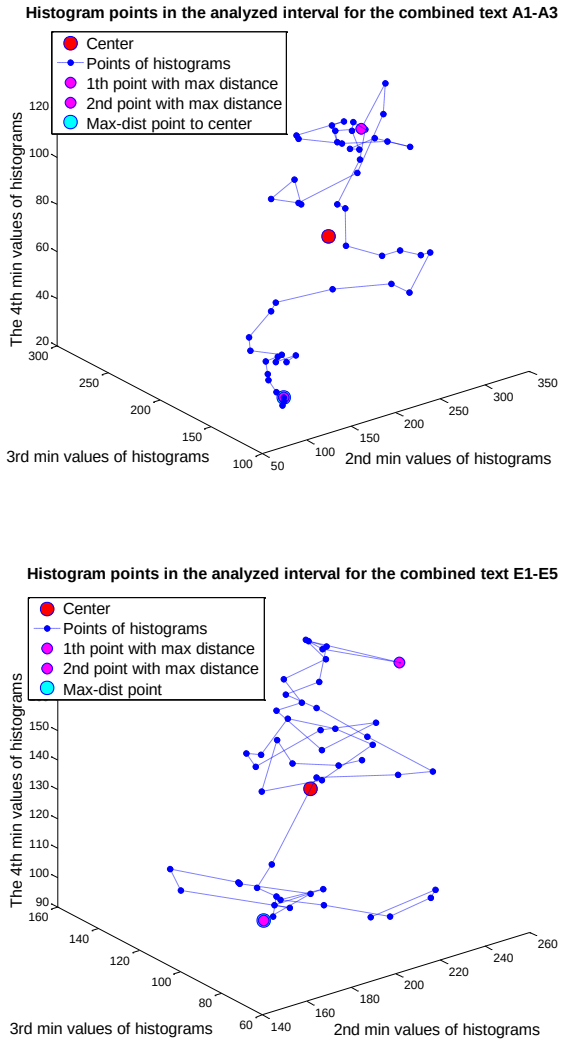


Fig. 6. The top panel illustrates histogram points of Arabic combined text A1-A3, the down panel shows histogram points of English combined text E1-E5. Except histogram points, we show the centers of the clusters and the points with the maximal distance from the center.

Acknowledgements

The research is supported by the Slovak Scientific Grant Agency VEGA, Grant No. 1/0142/15.

References

[1] pan-plagiarism-corpus-2011.part1.rar, ttp://www.uniweimar.de/en/media/chairs/webis/corpora/pan-pc-11/ .

[2] King Saud University Corpus of Classical Arabic. http://ksucorpus.ksu.edu.sa .

[3] I. Bensalem, P. Rosso, S. Chikhi: A New Corpus for the Evaluation of Arabic Intrinsic Plagiarism Detection. CLEF 2013, LNCS 8138, pp. 53-58, 2013.

[4] A. Neme, J.R.G. Pulido, A. Muñoz, S. Hernández, T. Dey: Stylistics analysis and authorship attribution algorithms based on self-organizing maps**,** Neurocomputing, 147 5 January 2015, pp. 147–159.

[5] E. Stamatatos: Authorship attribution based on feature set subspacing ensembles. International Journal on Artificial Intelligence Tools, 15.5, pp. 823-838.

[6] H. J. Escalante, T. Solorio, M. Montes-y-Gomez: Local Histograms of Character N-grams for Authorship Attribution. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, pp. 288-298, Portland, Oregon, June 19-24, 2011. c 2011 Association for Computational Linguistics.

[7] G. Oberreuter, G. LHuillier, S. A. R´ıos, and J. D. Velasquez: Approaches for Intrinsic and External Plagiarism Detection. Notebook for PAN at CLEF 2011 .

[8] F. I. Haj Hassan, M. A. Chaurasia: N-Gram Based Text Author Verification. IPCSI vol. 36 (2012), IACSIT press, Singapore.

[9] E. Stamatatos, A survey of modern authorship attribution methods, J. Am. Soc. Inf. Sci. Technol. 60(3) (2010), pp. 538-556.

[10] A. Almarimi, G. Andrejková: Discrepancies Detection in Arabic and English Documents, ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 5, No.17 , September 2015. ISSN : 2322-5157, pp. 69-75.

[11] A. Almarimi, G. Andrejková: Document Verification using N-grams and Histograms of Words. IEEE 13th International Scientific Conference on Informatics, November 18-20 Poprad Slovakia, pp. 21-26.

[12] G. Lebanon, Y. Mao, J. Dillon: The locally weighted bag of words framework for document representation. Journal of Machine Learning Research 8 (2007), pp. 2405-2441.

First Author: Master of Science from P. J. Šafárik University in Košice, Institute of Computer Science, Faculty of Science, now PhD. Student.

Second Author: Associate Professor, Deputy director for academic affairs in of Computer Science, Faculty of Science, P. J. Šafárik University in Košice. She is interested in artificial neural networks and in stringology.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Software-As-A-Service for improving Drug Research towards a standardized approach

Keis Husein[1], Prof. Ezendu Ariwa[2] and Paul Bocij[3]

[1] London School of Commerce, Associate College of Cardiff Metropolitan University,
London SE1 1NX, United Kingdom
l0234szsz1112@student.lsclondon.co.uk

[2] Department of Computer Science and Technology, University of Bedfordshire
Luton, LU1 3JU, United Kingdom
ezendu.ariwa@lsclondon.co.uk

[3] Aston Business School, Aston University
Birmingham, B4 7ET, United Kingdom
P.BOCIJ@aston.ac.uk

Abstract

Software as a Service (SaaS) is one of the most interesting applications of a service-oriented architecture. SaaS spares users the cost of acquiring and maintaining hardware and software. The SaaS business model is very widely appreciated in the economy, because it brings not only financial benefits but also process-optimizing benefits. Drug research is a lengthy, complex and costly process that involves a number of disciplines, from medicine through economics to natural science. Until recently, the standard programs and infrastructure used for data analysis were almost exclusively commercial, proprietary, closed source and expensive. One of the major attractions of the open-source model is to customize the platform to suite the requirements and react faster on changes. There are different proprietary approaches. However there is a gap in knowledge regarding using closed and proprietary infrastructure. The aim of this research is to investigate the impact of an open source SaaS approach on the drug research process.

Keywords: *SaaS; OpenStack; Drug Research; Virtual Screening*

1. Introduction

Drug research and development (R & D) is a complex, time intensive, expensive task and involve many risks. According to general estimates[2], it is assumed that the conception of the "drug to market" has duration of 12 years and cost an average of more than $ 800 million. From this fact technologies developed to reduce the research time and cost. Computer-aided drug design (CADD) is such a modern development[3].

1.1. Software as a Service

Cloud computing is defined by Feuerlicht[4]: It includes computing, data storage and software services, which are accessible through the Internet. Coombe[5] defines cloud computing as a change of paradigm, which enables scalable execution and storage across distributed and networked systems. Cisco[6] believes that cloud computing has the potential to provide on-demand services available and at lower cost than current options, with less complexity, greater scalability and greater range. Catteddu and Hogben[7] describing that cloud computing provides IT-resources in a new way. Cloud services e.g. software or data storage, processing can immediately make on demand available. Especially in times of rising costs, the growth of cloud services is a ray of hope. Another definition [1] describes SaaS as software that is used as a hosted service and accessed over the Internet via standard web browser. This may include IIS or Apache software, but typically SaaS means the provision of business applications, including collaboration software and industry-specific applications, to companies that need these applications to run their business. Fig. 1. Illustrates how the distribution and operation of SaaS compared to in-house looks like.

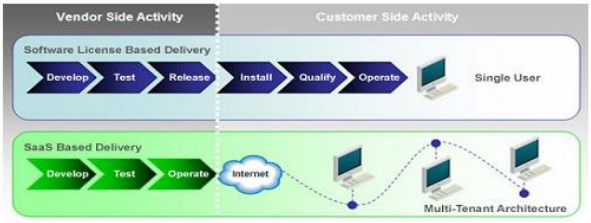


Fig. 1 SaaS model vs. traditional Software[8]

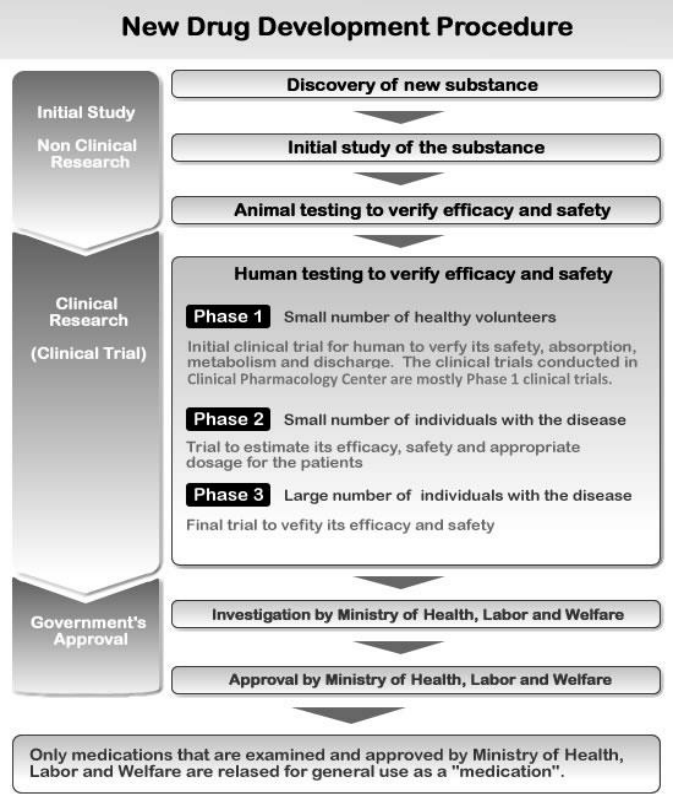


Fig. 2 Procedure for developing new drugs [9].

1.2. Drug research and development (R & D)

From 10,000 compounds, which were under investigation for a drug, only one ever come to market. Even if a compound reaches the market, only one of three bring their development costs back. Therefore, the development is associated with high risks! Drug development is a scientific challenge that is strictly regulated, because of the legitimate concerns of public health. Fig. 2 shows the procedure for development of a new drug.

Drug development phases: The drug development process can be split up into three phases as seen in Fig. 2. (1) Pre-clinical research and development. (2) Clinical research and development and (3) Government's approval.

2. DRUG DISCCOVERY METHODS

One of the most successful ways to find promising drug candidates is to examine how the target protein interacts with randomly selected compounds that usually are part of libraries. The test is common in so-called high-throughput screening (HTS) facilities. Compound libraries are commercially available with a size of millions of compounds. The most promising compounds from the screening will come forth, called hits, which are the compounds that exhibit binding activity at the target.

2.1. Virtual Screening

VS is based on the basis of the derived mathematical or simulated real screening. Computational methods can be used to predict or simulate how a particular connection with a specific target protein interacts. These findings are used to help build a hypothesis over chemical properties and in the design of active substances and they can also be used to refine and modify drug candidates. The following three virtual screening or computational methods are used in modern drug development; (A) Molecular Docking, (B) Quantitive Structure-Activity Relationships (QSAR) and (C) Pharmacopeia Mapping.

A. Molecular Docking: If the structure of the target is available, usually from x-ray crystallography, molecular docking is the most used VS-method. Molecular docking may also be used to test if the hypothesis is right, before carrying out the expensive laboratory tests. This software can predict how a drug binds to a target protein.

B. QSAR: As mentioned above, it is necessary to know the geometrical structure of both the ligand and the target protein in order to use molecular docking. QSAR is an example of a process regardless of whether the structure is known or cannot be applied. QSAR models are used for VS of compounds for investigation to find suitable drug candidates for the target.

C. **Pharmacopeia Mapping:** While QSAR concentrates on a number of descriptors, such as electrostatic and thermodynamic properties. Pharmacopoeia is a geometric mapping approach. A pharmacophore can be thought of a 3D model of the characteristic features of the binding site of the protein under examination (target).

2.2. High-throughput screening

HTS is a method for scientific experimentation especially used in drug discovery and relevant for the areas of biology and chemistry. Robotics, data processing and control software, liquid handling devices and detectors makes HTS available. HTS researcher can quickly perform millions of biochemical, genetic or pharmacological tests. Through this process, one can quickly identify active substances, antibodies or genes, which modulate a particular bimolecular path. The results of these experiments provide starting points for drug design and for understanding the interaction or role of a particular biochemical process in biology.

3. Virtual Screening Software

In this section, AutoDock, VS software will be specifically highlighted. AutoDock is docking software also specifically found in cloud application. AutoDock is a pioneer among the docking programs to model ligand confrontational with full flexibility. The Institute Prof. Arthur J. Olsen Laboratory, the first version (1989-1990) was written in FORTRAN 77 by Dr. David S. Goodsell[10]. (It was almost AutoDoq called because it uses quaternions for rotations, quaternions do not come from the so-called Gymbal lock problem associated with the Euler angles.) AutoDock base algorithm is the so-called Monte Carlo method in combination with genetic algorithm for giving conformations.

3.1. Inhibox

InhibOx[11] focuses on the development and provision of services and technology in computer-aided drug discovery (CADD). Inhibox infrastructure is based upon the Amazon Web Services. InhibOx delivers: Lead identification capabilities, lead optimization and VS capability with a compound database of 110 million synthesized compounds.

3.2. Accelrys HEOS

HEOS[12] was a development and a starting point for neglected diseases collaborations. HEOS was able to manage information, which was generated by more than 400 users across the world. HEOS is hosting over 90 drug research projects and has a library with over 600,000 chemical compounds.

3.3. eScience Central

e-Science Central[13] is another SaaS provider on the VS market. It is a solution working 100% in the cloud and needs only a web browser for running studies. They offer working with total privacy or in collaboration with trusted partners. One of the most important key features of e-science central is to analyze data and not only share them in the cloud like other collaboration solutions does. The infrastructure is Windows Azure. The difference is that e-science central is Platform as a Service provider and deliver a complete platform to run specific applications like AutoDock or similar.

4. Conclusions and Future Work

Virtual screening services are currently offered as web services. They are accessible for interactive use or batch processing of entire library of compounds with transparent

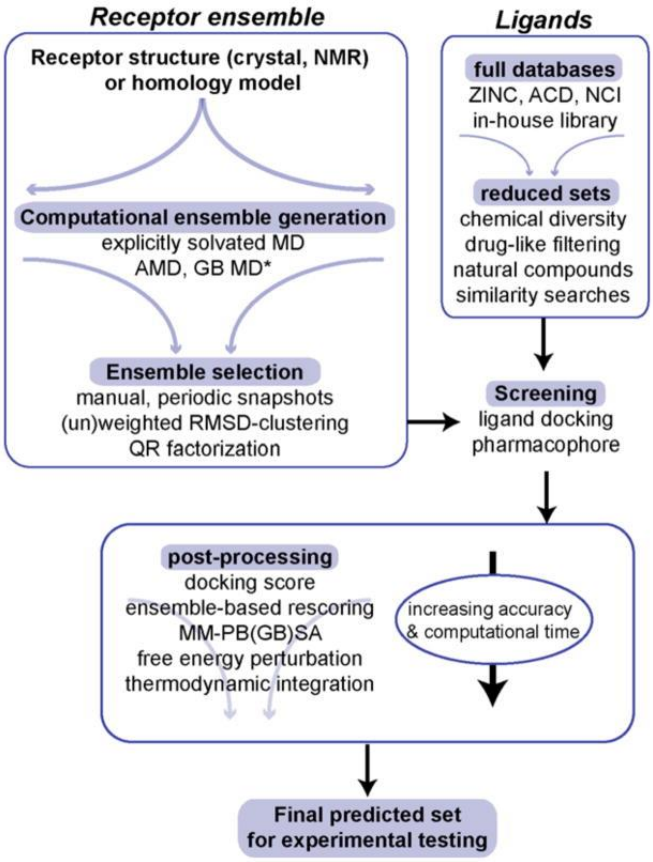


Fig. 3 General workflow for ensemble-based virtual screen experiment [15]

access to cloud or cluster resources[14]. The systems will be available as a computational workflow (see Fig.3) and be easily accessible to a much wider audience. Commercial offers, like Accelrys Pipeline Pilot provides a complete solution for academic research in computer-aided drug design. Increasingly, both academic and commercial enterprises rely on cloud-based virtual screening services that are completely transparent to the end user. Research, such as the World Community Grid use the idle time of computers and provide another venue for virtual screening represents. Better theories and efficient numerical Procedure to allow selection of the most relevant conformations of ligand binding* in a predictive manner are still needed. New algorithms, like Iterated Local Search global optimizer[16] and multi threading optimized algorithms, used in AutoDock Vina present with 10 to 100 times faster compared to AutoDock with equal or better performance. In summary, make the further acceleration by algorithmic and hardware improvements, utility computing and cloud-based virtual screening make it accessible to a wider audience, and its further validation and optimization. Algorithmic acceleration is an important point, because results

delivered faster without increasing hardware power. Further research will be done in running AutoDock Vienna in a standardized environment using OpenStack. Openstack[17] is a cloud operating system based on standardized programs. Results will be compared with proprietary alternatives like Amazon Web Services or Microsoft Azure.

References

[1] K. Husein, SaaS als Alternative zu traditionellen IS in Unternehmen, Freie Universität Berlin, Germany: Diploma thesis, 2009.

[2] Y. Tang, W. Zhu, K. Chen, H. Jiang, "New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery", Drug Discovery Today: Technologies vol. 3, no. 3, pp. 307-313, 2006.

[3] W.L. Jorgensen, "The many roles of computation in drug discovery" Science, vol. 303, pp. 1813-1818, 2004.

[4] G. Feuerlicht, "Next generation SOA: Can SOA survive cloud computing?" in Advances in Intelligent Web Mastering -2, AISC 67, pp. 19-29, 2010.

[5] B. Coombe, "Cloud computing-overview, advantages, and challenges for enterprise deployment", Bechtel Technology Journal, 2(1), pp. 1-11, 2009.

[6] Cisco, "The Cisco powered network cloud: An exciting managed services opportunity." Whitepaper, 2009.

[7] D. Catteddu & G. Hogben, "Cloud computing: Benefits, risks and recommendations for information security", European Network and Information Security Agency (ENISA), pp. 1-125, 2009.

[8] T. Rapczynski, R. Siripanya, "Software as a Service für Unternehmensanwendungen", Hft Stuttgart – University of Applied Science, Stuttgart – Germany, Research report, p. 2, 2008.

[9] Medical Co. LTA Clinical Center, "How new drugs are developed", www.rinsho.or.jp/en/01clinical_02.html, call on 16.12.2010.

[10] AutoDock, "AutoDock history", autodock.scripps.edu/resources/science/AutoDock/, call on 14.07.2013.

[11] InhibOx, "InhibOx: Cloud-Enabled Drug Discovery", http://www.inhibox.com/inhibox, called on 14.07.2013

[12] F. Bost, R. T. Jacobs & P. Kowalczyk, "Informatics for neglected diseases collaborations", Current Opinion in Drug Discovery & Development, vol. 13, no. 3, pp.286-296, 2010.

[13] E-Science Central, "About: e-science central", http://www.esciencecentral.co.uk/?page_id=10, called on 14.07.2013.

[14] R. E. Amaro & W. W. Li, "Emerging Methods for Ensamble-Based Virtual Screening", Current Topics in Medical Chemistry, Vol. 10, Issue 1, pp. 3-10, 2010.

[15] Chemgapedia, "Liganden", http://www.chemgapedia.de/ called on 22.03.2014

[16] Baxter J. Local optima avoidance in depot location. Journal of the Operational Research Society. 1981; vol. 32(no. 9):815–819.

[17] OpenStack,"What is OpenStack", http://www.openstack.org/, called on 22.03.2014

Keis Husein earned his Master degree in 2009. He is a PhD student and his passion and specialization area is SaaS. He is actually working at HQ plus providing Software as a Service.

Ezendu Ariwa is Professor for Information Systems, Computing in Social Science at University of Bedfordshire

Paul Bocij is BAM Course Director & Lecturer at Aston Business School. His research interests are Educational Technology & computer-based assessment, online deviant behavior (cyber stalking) and professionalism & ethics in Information Systems.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Investigation of Coherent Multicarrier Code Division Multiple Access for Optical Access Networks

Ali Tamini[1], Mohammad Molavi Kakhki[2]

[1] Department of Electrical Engineering, Ferdowsi University of Mashhad
Mashhad, Iran
ali_tamini@stu.um.ac.ir

[2] Department of Electrical Engineering, Ferdowsi University of Mashhad
Mashhad, Iran
molavi@um.ac.ir

Abstract

Orthogonal frequency division multiplexing (OFDM) has proved to be a promising technique to increase the reach and bit rate both in long-haul communications and in passive optical networks. This paper, for the first time, investigates the use of OFDM combined with electrical CDMA in presence of coherent detection as a multiple access scheme. The proposed multicarrier-CDMA system is simulated using Walsh-Hadamard codes and its performance is compared to that of coherent WDM-OFDM system in terms of bit-error-rate and bandwidth efficiency. It is shown that MC-CDMA benefits from better spectral efficiency while its performance slightly deteriorates in comparison to WDM-OFDM when the number of users is increased.

Keywords: *Fiber Optic Communication; Optical MC-CDMA; Coherent Detection; Optical Access Network; OptiSystem*

1. Introduction

The increase in internet traffic has led to increase in demands for optical access networks with higher bit rates that can be used in longer distances. Therefore, exploitation of advance techniques is necessary for future passive optical networks (PONs). Optical orthogonal frequency division multiplexing is a promising technique that has been the subject of interest in recent years. Not only this modulation type benefits from good spectral efficiency due to the overlap of the frequency sub-channels, but it also has the ability to fight chromatic dispersion (CD) in fiber optic channel, improving system performance [1]. OFDM has been used in optical communications either with direct detection [2] or with coherent detection [3] and has shown better receiver sensitivity, spectral efficiency and robustness against chromatic dispersion when used with coherent scheme, but it has a more complex transceiver design [4].

In passive optical networks, OFDM has been studied either as an access scheme (OFDMA) where each subcarrier is subscribed to a certain user, or simply as a modulation where the subcarriers are all dedicated to a single user [5]. Sometimes it has been combined with access schemes such as wavelength division multiplexing (WDM) to meet the requirements of future optical access networks [6-8]. In [9], A comparative study between different hybrid WDM passive optical network (XDMA-WDM-PON) architectures, namely TDMA WDM PON, OCDM-WDM-PON and OFDM-WDM-PON, demonstrated superiority of OFDM-WDM-PON in terms of spectral efficiency, transmission length, compensation for chromatic dispersion, high transmission speed, high optical network units (ONUs) capacity and compatibility with RF over the other two schemes. Nevertheless, in WDM the necessity for guard band decreases spectral efficiency and capacity of the system [10]. Therefore, the use of a hybrid access scheme which makes better use of the bandwidth allocation than WDM can further improve performance of OFDM-based passive optical networks. In this paper, the use of OFDM combined with electrical CDMA in presence of coherent optical receivers is investigated for the first time as a hybrid optical access scheme and is compared with coherent WDM-OFDM regarding spectral efficiency and BER performance for different number of users in different transmission distances and different launch powers.

In section 2, some background information about OFDM and MC-CDMA is provided. Section 3 describes the designed system and the relating simulations parameters. The simulations have been done by co-simulation of OptiSystem 13 software with Matlab. The results of these simulations and their analysis are presented in section 4. Finally, section 5 is dedicated to the conclusion and possible future works.

2. Technical Background

In this section, optical OFDM and multicarrier-CDMA structures are briefly introduced.

2.1 Orthogonal Frequency Division Multiplexing

In OFDM, the spectrum is divided into overlapping subcarriers. Each subcarrier is transmitted with a slower rate in parallel with other subcarriers. Therefore, OFDM has the ability to reduce the effect of chromatic dispersion in optical fiber communications by extending subcarriers in time domain.

Figure 1 shows the block diagram of OFDM modulator. The input data is the output of an M-ary modulator. Serial input data is converted into parallel by serial to parallel converter. Then IFFT is applied to the parallel data and cyclic prefix is added to each stream. Finally, the digital data is converted to analog using an interpolation technique and the parallel data is converted back to serial. In demodulator, reverse operations are applied. Schematic of OFDM demodulator is shown in figure 2.

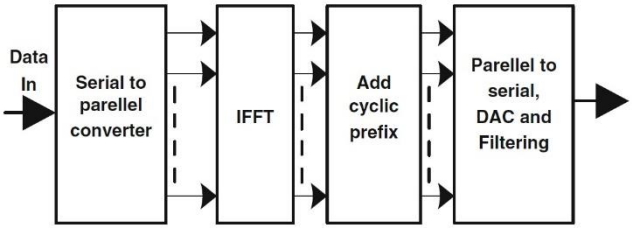


Fig. 1 OFDM modulator [8].

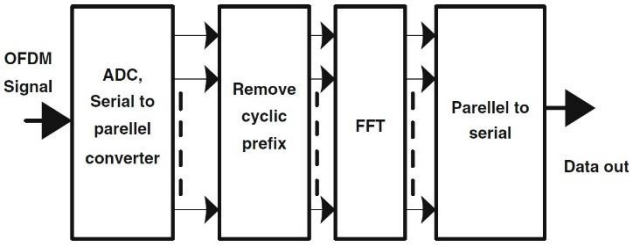


Fig. 2 OFDM demodulator [8]

2.2 Coherent Optical OFDM

A generic coherent optical OFDM (CO-OFDM) system is shown in figure 3. The system can be divided into five parts [4]: (1) electrical baseband OFDM transmitter, (2) electrical to optical up-converter consisting of a pair of Mach-Zehnder modulators (MZM) and a continuous-wave (CW) laser, (3) optical channel, (4) optical to electrical down-converter by a pair of balanced photodetectors and a continuous-wave laser, and (5) electrical baseband OFDM receiver.

Due to differential structure of balanced photodetectors, the common mode noise is reduced and as a result, CO-OFDM has a better sensitivity than direct detection optical OFDM (DDO-OFDM) [4]. Moreover, with appropriate settings CO-OFDM benefits from linear transmission of data between electrical and optical domains which makes OFDM perform better in alleviating channel dispersion [3]. However, DDO-OFDM is less susceptible to phase noise. Besides, it has a lower complexity and therefore, it has been investigated extensively in PONs [11].

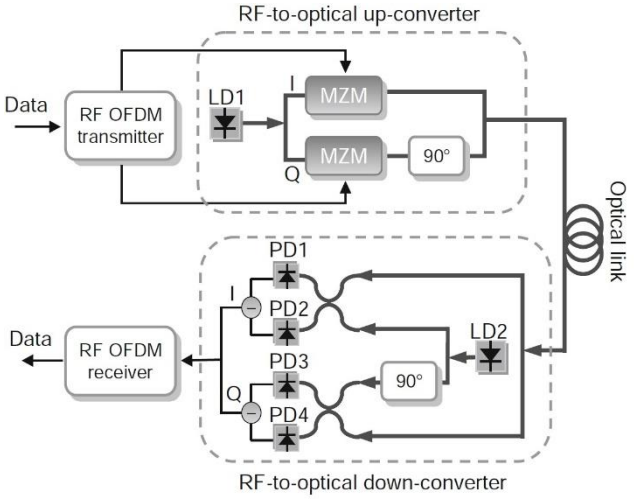


Fig. 3 A coherent optical OFDM system [4].

2.3 Multicarrier-CDMA

Multicarrier-CDMA (MC-CDMA) access scheme is a combination of CDMA and OFDM in which a CDMA signature code is assigned to each OFDM user. As depicted is figure 4, the signature code is multiplied by each of the input symbols at the transmitter before the IFFT block. At the receiver side, the signature code is multiplied by the signal to regenerate the input symbols after the FFT operation.

MC-CDMA has been widely studied in wireless communications. However, the first demonstration of MC-CDMA in fiber optic communications was not until 2009. In [12], direct detection scheme was used to transmit data with an overall bit rate of 15 Gbps over a 70 km single mode fiber (SMF).

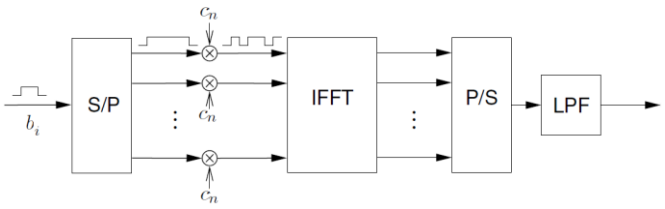


Fig. 4 Schematic of a MC-CDMA transmitter [13].

3. Design Parameters

Our simulations have been performed by co-simulation of OptiSystem 13 and Matlab, where Matlab is used only for simulation of MC-CDMA modulator and demodulator blocks. Figure 4 shows the block diagram of the simulation MC-CDMA system with N users. Each of the transceivers has a structure as shown in figures 3. Walsh-Hadamard codes are employed as signature codes. For each simulation, the length of the code equals the number of users. For each user, a 4-QAM coder is used before MC-CDMA modulator with 512 subcarriers and 1024 FFT points with no cyclic prefix. At the output of the MC-CDMA modulator, the in-phase and quadrature signals are converted into an optical signal and up-converted by a pair of Lithium Liobate Mach-Zehnder modulators and a continuous-wave laser of 193.05 THz with a linewidth of 100 kHz. Here all users have the same laser carrier frequency unlike WDM-OFDM where each user has a specific carrier frequency. Then the signals from all users are combined by an optical power combiner and sent into the channel.

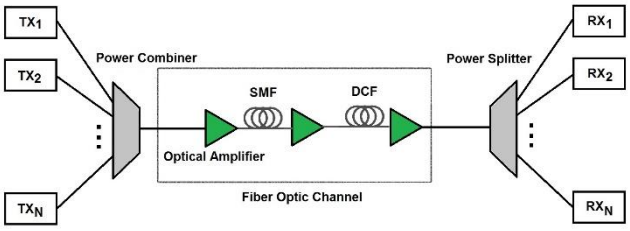


Fig. 5 Block diagram of the MC-CDMA system.

The transmitted signal passes through the channel shown in figure 5. The first optical amplifier acts as a power booster with a gain of 15 dB, while the other two are used for compensation of fiber power loss. The amplifiers have a noise figure of 4 dB. The parameters of optical fibers are shown in table 1.

Table 1: Characteristics of optical fibers in simulated by OptiSystem

Parameter	SMF	DCF
Dispersion (ps/nm/km)	16	-80
Dispersion Slope (ps/nm^2/km)	0.08	-0.4
Attenuation (dB/km)	0.2	0.4
Effective Area (um^2)	80	30
Nonlinear Index of Refraction (m^2/W)	2.6e-20	2.6e-20

At the receiver, the signal is split by a power splitter, the output of which is converted to electrical form by a pair of balanced photo-detectors. A CW laser with the same carrier frequency and linewidth as the transmitter one is used as the local oscillator. Each of the PIN photo-detectors has a responsivity of 1 A/W, thermal noise of 100e-24 and dark current of 10 nA. The resulting electrical signal passes through the MC-CDMA demodulator followed by a 4-QAM decoder to regenerate the binary data for each user.

A coherent WDM-OFDM system is also simulated for the purpose of comparison with MC-CDMA. It makes use of the same components with the same parameters as those used in MC-CDMA. Here, the guard band between WDM channels is considered 0.25 of the channel OFDM signal bandwidth. The overall word length is 524288 in all simulations.

4. Results and Discussion

In figure 6, a comparison has been made between BER performance of MC-CDMA and WDM-OFDM systems versus the launch power per user for various number of users. The systems are simulated with a bit rate of 50 Gbps for each user and the channel length is 240 km. It can be seen that although the performance of both systems deteriorates when the number of users increases, MC-CDMA is more susceptible to degeneration. Figure 7 shows the changes in BER for both systems with various number of users through different values of optical fiber channel length. The system has a bit rate of 50 Gbps for each user and a launch power of -20dBm per user. Figures 6 and 7 are quite consistent with each other in showing the relation between BER performance and the number of users.

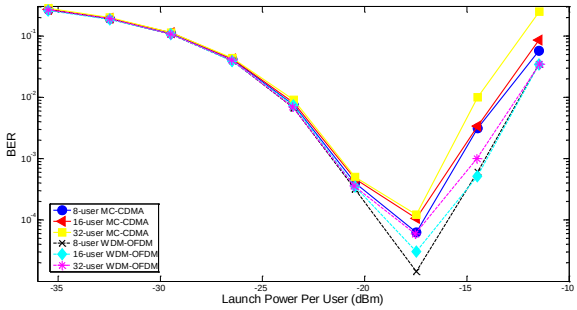


Fig. 6 Comparison of BER vs launch power per user for MC-CDMA and WDM-OFDM systems with various number of users.

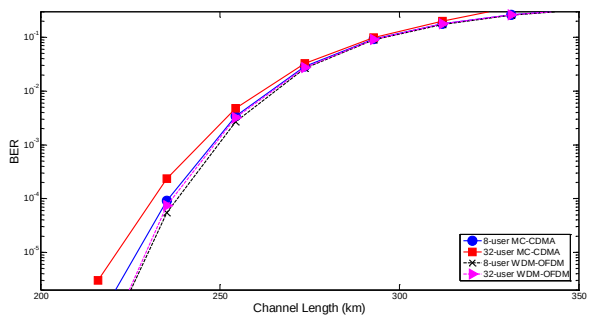


Fig. 7 Comparison of BER vs channel length for MC-CDMA and WDM-OFDM systems with various number of users.

Figure 8 shows spectral efficiency of the simulated systems with regard to the number of users. This figure shows the superiority of MC-CDMA scheme when spectral efficiency is concerned. As seen in figure 8, the spectral efficiency of MC-CDMA increases when the number of users increases, while the spectral efficiency of WDM-OFDM is constant and always less than that of MC-CDMA. This implies a superior spectral efficiency of MC-CDMA over WDM-OFDM especially for greater number of users. The guard band between the WDM channels results in degradation of spectral efficiency for WDM-OFDM. However, the spectral efficiency for the simulated WDM-OFDM would not exceed 1.42 bits/s/Hz even if no guard band is assigned between the channels. Figure 9 shows the overall bandwidth allocation in MC-CDMA and WDM-OFDM systems with 8 users.

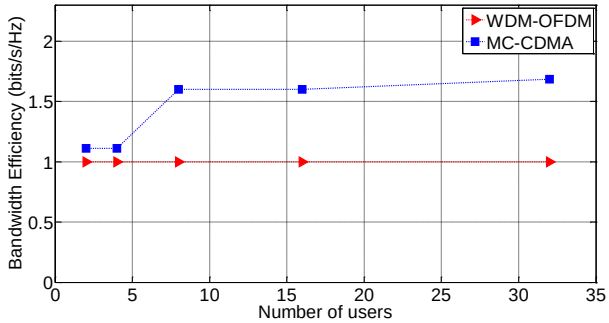


Fig. 8 Comparison of bandwidth efficiency between MC-CDMA and WDM-OFDM for systems with 2, 4, 8, 16 and 32 users.

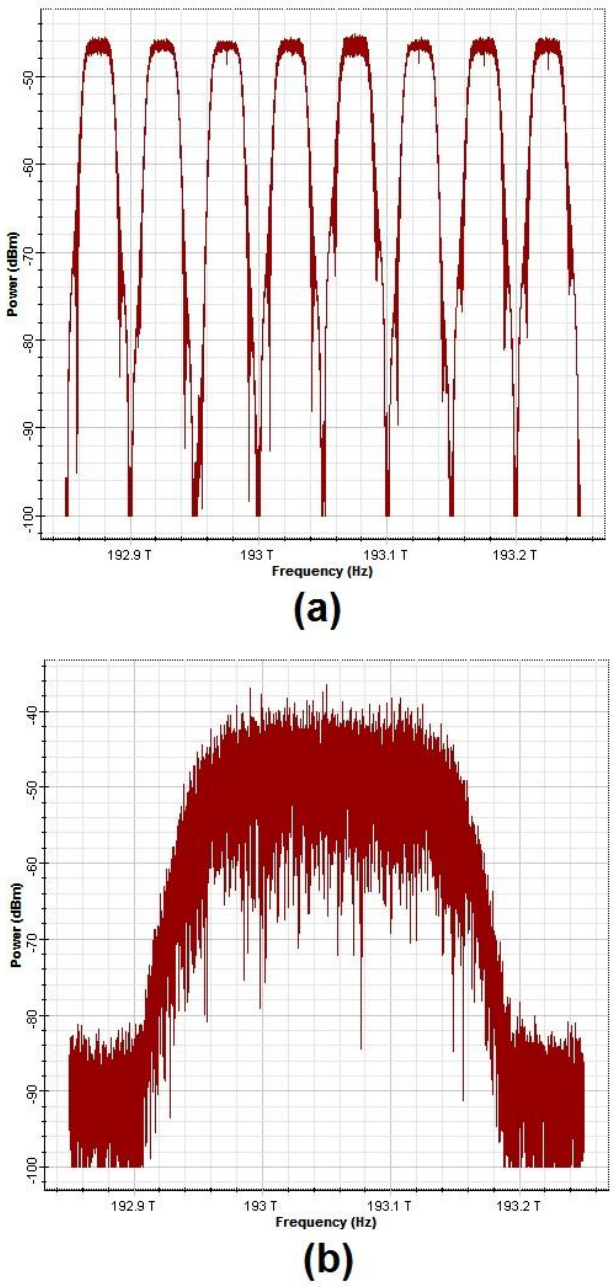


Fig. 9 Optical spectrum of (a) WDM-OFDM, and (b) MC-CDMA signals before entering the channel.

5. Conclusion

In this work, we investigated the use of coherent MC-CDMA in optical fiber communications as an access scheme for the first time. The proposed scheme showed a better spectral efficiency compared to WDM-OFDM especially when the number of users increases, but at the same time its BER performance declines more rapidly than that of WDM-OFDM. The results suggest that even

though MC-CDMA makes better use of bandwidth, it would have a better performance in PONs with smaller number of subscribers at ONUs. For future work, coherent MC-CDMA can be analyzed in PON structure where asynchronous upstream data transmission is an issue and therefore, non-orthogonal codes such as M-sequence should be used as signature codes. It is noted that performance evaluation of asynchronous MC-CDMA for direct detection systems has been reported by [14].

References

[1] B.J. Dixon, R.D. Pollard, and S. Iezekeil, "Orthogonal frequency-division multiplexing in wireless communication systems with multimode fiber feeds", IEEE Trans Microwave Theory Techniques, Vol. 49, No. 9, 2001, pp. 1404-9.

[2] I. Djordjevic, and B. Vasic, "Orthogonal frequency division multiplexing for high-speed optical transmission", Opt. Express, Vol. 14, 2006, pp. 3767–75.

[3] W. Shieh, and C. Athaudage, "Coherent optical orthogonal frequency division multiplexing", Electronic Letters, Vol. 42, No. 10, 2006, pp. 587-9.

[4] W. Shieh, and I. Djordjevic, "OFDM for Optical Communications", Academic Press, 2009.

[5] N. Cvijetic, "OFDM for Next-Generation Optical Access Networks", Lightwave Technology, vol. 30, no. 4, 2012, pp. 384-398.

[6] H. Miaoqing, W. Shibao, M. Fang, and L. Tang, "A wavelength reuse OFDM-WDM-PON architecture with downstream OFDM and upstream OOK modulations", in 2011 IEEE 3rd International Conference on Communication Software and Networks (ICCSN), 2011, pp. 389-392.

[7] N. Cvijetic, et al., "Terabit Optical Access Networks Based on WDM-OFDMA-PON," Journal of Lightwave Technology, vol. 30, 2012, pp. 493-503.

[8] G. Pandey and A. Goel, "Long reach colorless WDM OFDM-PON using direct detection OFDM transmission for downstream and OOK for upstream," Optical and Quantum Electronics, vol. 46, no. 12, 2014, pp. 1509–1518.

[9] M. Hernandez, A. Arcia, R. Alvizu, M. Huerta, "A review of XDMA-WDM-PON for Next Generation Optical Access Networks," in Global Information Infrastructure and Networking Symposium (GIIS), 2012, pp. 1-6, 17-19.

[10] A. Banerjee, et al., "Wavelength-division-multiplexed passive optical network (WDM-PON) technologies for broadband access: a review [Invited]", J. Opt. Netw., vol. 4, 2005, pp. 737-758.

[11] K. Alatawi and M. Matin, "High Data Rate Coherent Optical OFDM Systems for Long-Haul Transmission", M.S. Thesis, University of Denver, Denver, Colorado, 2013.

[12] V.R. Arbab, W. Peng; X. Wu; A.E. Willner, "Experimental demonstration of multicarrier-CDMA for passive optical networks," in 34th European Conference on Optical Communication, 2008, pp. 1-2, 21-25.

[13] L. Hanzo, M. Munster, et al., "OFDM and MC-CDMA for Broadband Multiuser Communications, WLANs and Broadcasting", IEEE Press and John Wiley & Sons, England, 2003.

[14] M.H. Shoreh, H. Beyranvand, and J.A. Salehi, "Performance evaluation of asynchronous multi-carrier code division multiple access for next-generation long-reach fibre optic access networks", IET Optoelectronics, vol. 9, no. 6, 2015, pp. 325-332.

Ali Tamini received his B.S. degree in electrical engineering in 2012. He is currently working towards his M.S. degree in telecommunications engineering at Ferdowsi University of Mashhad. His major research interests include orthogonal frequency division multiplexing passive optical networks, space division multiplexing and optical wireless systems.

Dr. Mohammad Molavi Kakhki (S. M. IEEE 1993) received the B.S. degree in physical electronics in 1971 from university of Tabriz, Iran, and M.S. and Ph.D. degrees in electronic communications from University of Kent at Canterbury, UK. In 1977 and 1981 respectively. From 1981 to 1984 he worked as a research fellow at Kent University. In 1985 he joined the department of electrical engineering at Ferdowsi University of Mashhad (FUM), Iran, and he is currently a professor at FUM. His research interests are optical communications, digital communications and power line communications.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Digital Image Encryption Based On Multiple Chaotic Maps

Amirhoushang Arab Avval[1], Jila ayubi[2] and Farangiz Arab[3]

[1] Department of Research and Development Ports and Maritime Organization
Chabahar, Sistan and Baluchestan, Iran
amirhoushang.arab@gmail.com

[2] Department of Electrical engineering Meraaj Higher Education Institue
Salmas, West Azarbayjan, Iran
Jila.ayubi@gmail.com

[3] Department of Gaedi Health Center, Isfahan Medical Sciences University
Isfahan, Isfahan, Iran
Dr.f.arab@gmail.com

Abstract

A novel and robust chaos-based digital image encryption is proposed. The present paper presents a cipher block image encryption using multiple chaotic maps to lead increased security. An image block is encrypted by the block-based permutation process and cipher block encryption process. In the proposed scheme, secret key includes nineteen control and initial conditions parameter of the four chaotic maps and the calculated key space is 2^{883}. The effectiveness and security of the proposed encryption scheme has been performed using the histograms, correlation coefficients, information entropy, differential analysis, key space analysis, etc. It can be concluded that the proposed image encryption technique is a suitable choice for practical applications.

Keywords: *Image Encryption; Chaos; Chaotic Maps; High Security.*

1. Introduction

With the rapid growth in digital image processing and network technology, more information and multimedia files has been transmitted over the computer networks and internet. Protection of digital information against illegal access and distribution has become extremely important. Therefore, techniques are required to provide security functionalities like privacy, integrity, or authentication especially suited for these data types of multimedia. A few applications of these techniques for providing privacy and confidentiality of visual data are in the areas of telemedicine, videoconferencing, military surveillance, and video-on-demand, pay TV etc. However, visual data such as image and video is different from text, conventional algorithms such as DES, IDEA, AES and most other methods are not suitable for image and video encryption.

Chaos theory has been established since 1970s in many practical applications to the real world, including synchronization, control, neural network, communication, etc [1-4]. Many researchers have noticed that there exists a close relationship between chaos and cryptography [5], [6]; many properties of chaotic systems have their corresponding counterparts in traditional cryptosystems.

Chaotic systems have several significant advantage in establish secure communications, such as ergodicity, sensitivity to initial condition, control parameters and random like behavior [7], [8]. A lot of image encryption schemes based on chaotic map have been already presented [9-12]. In the digital world nowadays, the security of digital images and performance speed has become more important since the communications of digital information over network occur more and more frequently.

In this paper, a new design of a class of chaotic cryptosystems is suggested to overcome the aforementioned drawbacks. Experimental results and security analysis indicate that the encryption algorithm based on multiple chaotic maps is advantageous from the point of view of large key space and high security.

The rest of the Letter is organized as follows. Section 2 describes used chaotic maps in proposed algorithm. An image encryption and decryption procedure is shown in Section 3. Also, the selected example and simulation results are discussed in Section 4. Section 5 is the conclusion.

2. Chaotic maps

2.1 Jacobian Elliptic Maps

In the past twenty year's dynamical systems, particularly one-dimensional iterative maps have attracted much attention and have become an important area of research activity [13]. This is also true in the case of so-called elliptic maps [14,15]. One-parameter families of jacobian elliptic rational maps [16] of the interval [0,1] with an invariant measure can be defined as:

$$X_{n+1} = \frac{4\alpha^2 X_n(1-k^2X_n)(1-X_n)}{(1-k^2X_n^2)^2 + 4(\alpha^2-1)X_n(1-k^2X_n)(1-X_n)} \quad (1)$$

Where $X_0 \in [0,1]$, $\alpha \in [0,4]$ and $k \in [0,1]$, k (modulus) represent the parameter of the elliptic functions. Bifurcation diagram of jacobian elliptic map is shown in Fig.1.

2.2 Chaotic Coupled Map

CML (coupled map lattices) based spatiotemporal chaotic systems have drawn initial attention in chaotic cryptography in recent years due to their excellent chaotic dynamical properties [17], [18] and [19].

Coupled map lattices are arrays of states whose values are continuous, usually within the unit interval, or discrete space and time [20]. The pair-coupled map with ergodic behavior can be considered as a one-dimensional dynamical map defined as:

$$X_{n+1} = [(1-\varepsilon) f_1(X_n)^P + \varepsilon f_2(X_n)^P]^{\frac{1}{P}} \quad (2)$$

Where, in general, P is an arbitrary parameter, ε the strength of the coupling, and the functions $f_1(X_n)$, $f_1(X_n)$ are two arbitrary one-dimensional maps. Obviously, by choosing $P = 1$, we get ordinary linearly coupled maps. $f_1(X_n)$, $f_1(X_n)$ defined as:

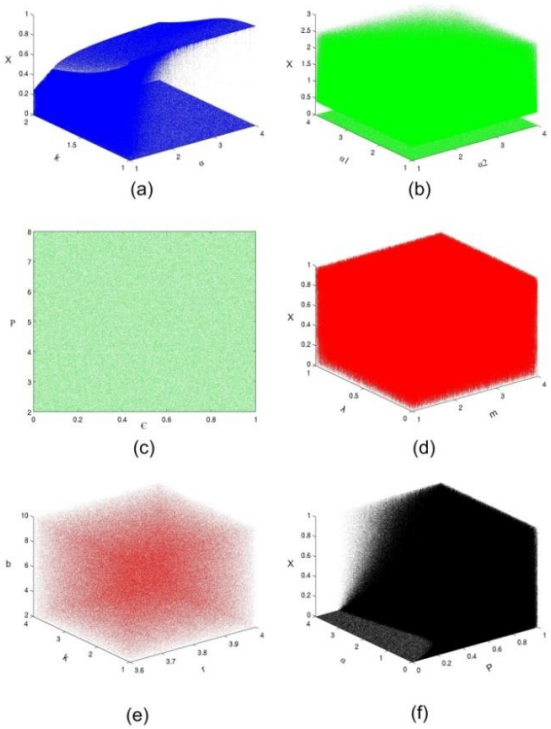


Fig. 1 Bifurcation diagram of (a) Jacobian elliptic map (b-c) Chaotic coupled map (d-e) Quantum map (f) piecewise nonlinear map.

$$f_1(X_n) = \frac{1}{\alpha_1^2} \tan(|\ N \times \arctan(|\ X_n\ |))$$
$$f_2(X_n) = \frac{1}{\alpha_2^2} \cot(|\ N \times \arctan(|\ X_n\ |)) \quad (3)$$

Where $X_0 \in [0,1]$, $\alpha_1 \in [0,4]$, $\alpha_2 \in [0,4]$, $\varepsilon \in [0,1]$ and $P \in [2,10]$.Fig.1(b-c) show the bifurcation plot of chaotic coupled map.

2.3 Quantum Map

The quantum rotators model has been widely used to study the dynamics of classically chaotic quantum systems and is specified in a simple form by:

$$X_{n+1} = r(X_n - X_n^2)\cos^k(-\lambda \frac{e^{-mb}}{b}) \quad (4)$$

Where $X_0 \in [0,1]$, $r \in [3.6,4]$, $\lambda \in [0,1]$, $m \in [1,4]$, $b \in [1,4]$ and $k \in [2,10]$.Bifurcation diagram of quantum map are shown in Fig.1(d-e).

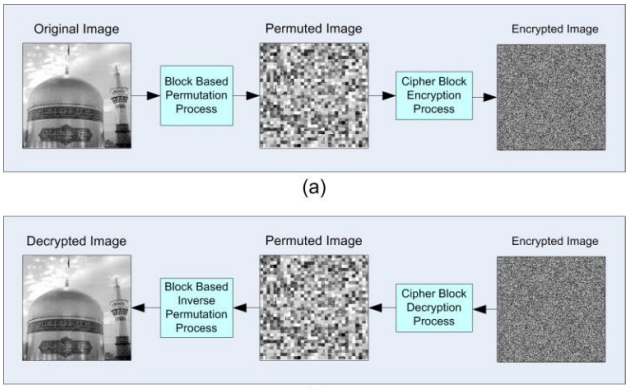


Fig. 2 Block Diagram Of (a) Encryption Process (b) Decryption Process

2.4 Piecewise nonlinear chaotic Map

A brief review of one-parameter families of piecewise nonlinear chaotic maps with an invariant measure is presented in [7]. These maps can be defined as:

$$X_{n+1} = \frac{\alpha^2 F}{1 + (\alpha^2 - 1)F} \quad (5)$$

Where

$$F = \begin{cases} \dfrac{X_n}{P} & 0 \le X \le P \\ \dfrac{X_n - P}{1 - P} & P < X \le 1 \end{cases} \quad (6)$$

Then, the probability parameter of the piecewise nonlinear chaotic maps p is generated by using the results of iteration of the trigonometric map can be defined as:

$$Y_{n+1} = \frac{1}{\beta^2} \tan^2(N \times \arctan(\sqrt{X_n}) \quad (7)$$

Therefore

$$P=\begin{Bmatrix} Y_{n+1} & 0\le Y_{n+1}\le 1 \\ \dfrac{1}{Y_{n+1}} & Y_{n+1}>1 \end{Bmatrix} \quad (8)$$

Where $X_0 \in [0,1]$, $\alpha \in [0,4]$, $\beta \in [0,4]$, $Y_0 \in [0,1]$, $b \in [1,4]$ and $P \in [0,1]$. Bifurcation diagram of quantum map is shown in Fig.1(f).

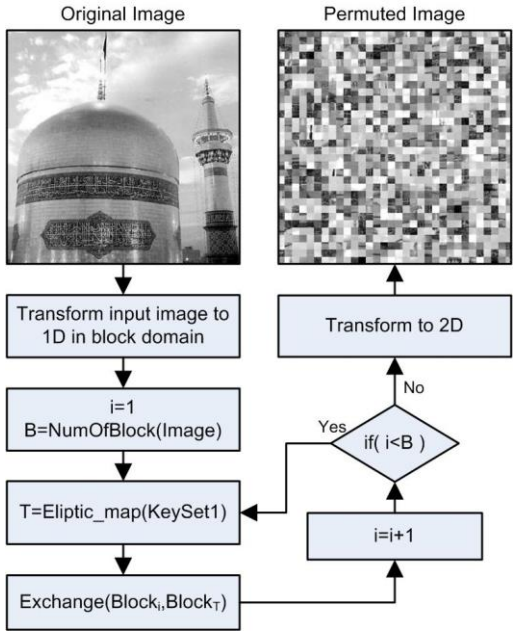


Fig. 3 Flowchart of Permutation Process.

3. THE ENCRYPTION AND DECRYPTION PROCEDURES

The proposed cryptosystem is a stream cipher algorithm based on multiple chaotic Maps. The block diagram of the proposed algorithm is presented in Fig.2 .This algorithm consists of the following major parts:

3.1 Permutation Process

Permutation procedure algorithm on each block is as follows:

- Step 1: Transform the input image from 1D to 2D in block domain.
- Step 2: Let size of image blocks in B and initialize $i=1$.
- Step 3: Generate chaotic pseudo-random number by jacobian elliptic map and set in T. ($T \in [1, B]$)
- Step 4: Exchange i-th block of image with T-th block.
- Step 5: Let $i=i+1$.
- Step 6: Repeat steps 3 to 5 until you reach the last block.
- Step 7: To display obtained image, transform the blocks from 1D to 2D.

3.2 Cipher Block Encryption Process

Cipher block encryption procedure is as follows:

- Step 1: Initialize $i=1$.
- Step 2: Generate chaotic pseudo-random number by jacobian elliptic map and set in T_1 . $T_1 \in \{1,2,3\}$)
- Step 3: Apply bit XOR operator :
- $Block_i = Block_i \oplus Block_{i-1}$

Flowchart of permutation process is shown in Fig.3.

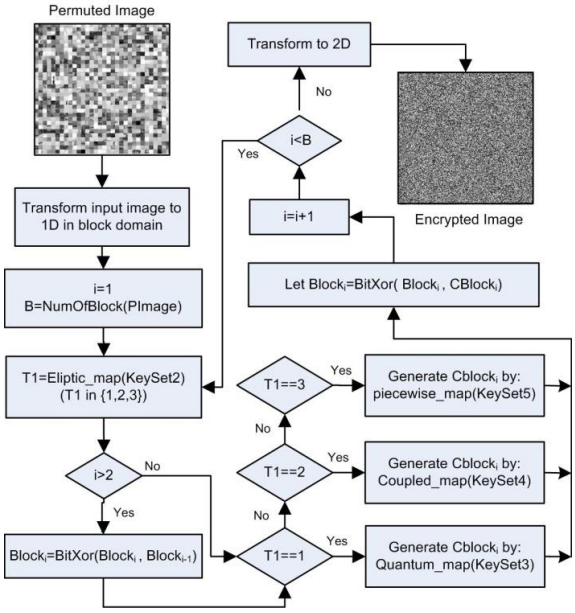


Fig. 4 Flowchart of Cipher Block Encryption Process.

- Step 4: if $T_1 = 1$ then, generate chaotic pseudo-random number by chaotic coupled map and set in $CBlock_i$. (Each elements in $CBlock_i \in [0,255]$).
- Step 5: if $T_1 = 1$ then, Generate chaotic pseudo-random number by quantum map and set in $CBlock_i$. (Each elements in $CBlock_i \in [0,255]$).
- Step 6: if $T_1 = 3$ then, generate chaotic pseudo-random number by piecewise map and set in $CBlock_i$. (Each elements in $CBlock_i \in [0,255]$).
- Step 7: Apply bit XOR operator :

$$Block_i = Block_i \oplus CBlock_i$$

- Step 8: Let $i = i + 1$.
- Step 9: Repeat steps 2 to 8 until you reach the last block.
- Step 10: To display obtained image, transform the blocks from 1D to 2D.

Flowchart of Cipher Blocked Encryption process is shown in Fig. 4. Since both decryption and encryption procedures have similar structure, they essentially have the same algorithmic complexity and time consumption.

4. EXPERIMENTAL RESULTS

In order to test the efficiency of the proposed chaotic cryptographic scheme a gray scale image "Mashhad" with the size 512×512 pixels is used (Fig. 5(a)). The results of the encryption are presented in Fig. 5(b-c). As can be seen from the figures there is no patterns or shadows visible in the corresponding cipher text. The test has been carried out in other familiar images as well (see Fig. 5).

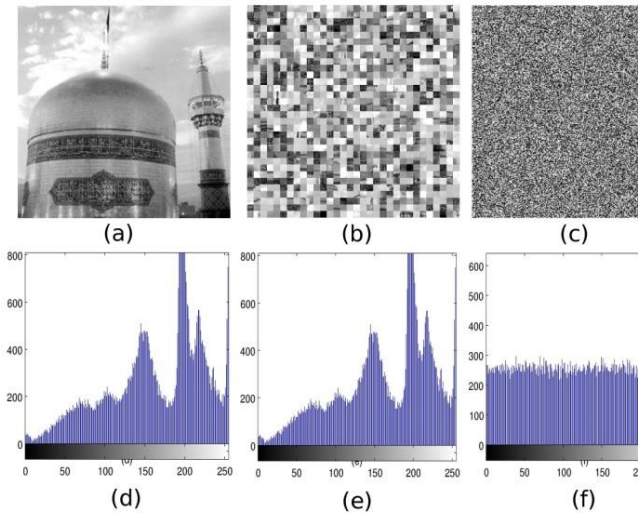


Fig. 5 Encryption and Decryption of Mashhad image. (a) Original Image (b) Permuted Image (c) Cipher text Image (d) Histogram of Plaintext (e) Histogram of Permuted Image Cipher text.

4.1 SECURITY ANALYSIS

When a new cryptosystem is proposed, it should always be accompanied by some security analysis. A good encryption procedure should be robust against all kinds of cryptanalytic, statistical and brute-force attacks. Here, some security analysis has been performed on the proposed scheme like key space analysis, statistical analysis, etc. The security analysis demonstrated a high security level of the new scheme as demonstrated through following test methods.

4.1.1 Histogram analysis

An image histogram illustrates that how pixels in an image are distributed by plotting the number of pixels. By taking a (512×512) sized "Mashhad" image as a plaintext, the histogram of the plaintext and permutated image and corresponding cipher text are shown in Fig. 6(d-f). As it was shown, the histograms of the original image and hence it does not provide any clue to employ any statistical analysis attack on the encryption image [7], [21].

4.1.2 Information Entropy

The entropy (such as KS-entropy, information entropy,) is the most outstanding feature of the randomness [22]. Information theory is a mathematical theory of data communication and storage founded in 1949 by Claude E. Shannon. To calculate the entropy $H(s)$ of a source s, we have:

$$H(s) = \sum_{i=0}^{255} P(s_i) \log_2 \frac{1}{P(s_i)} \qquad (9)$$

Where $P(s_i)$ represents the probability of symbol s_i. Actually, given that a real information source seldom transmits random messages, in general, the entropy value of the source is smaller than the ideal one. However, when these messages are encrypted, their entropy should ideally be 8.If the output of such a cipher emits symbols with entropy of less than 8, Then there exists a predictability which threatens its security. For the introduced encrypted image (Fig. 5(c)) and standard test images (Fig. 7-9), we have calculated the information entropy and the result have been presented in Table 1. The obtained values are very close to the theoretical value 8.

Apparently, comparing it with the other existing algorithms, the proposed algorithm is much closer to the ideal situation. That is, the information leakage in the encryption process is negligible, and so the encryption system is secure against the entropy attack.

4.2 Correlation of two adjacent pixels:

We have also analyzed the correlation between two vertical, two horizontal, and two diagonally adjacent pixels in "Mashhad" cipher image. To analyze the correlation of the adjacent pixels the following relation has been used [9]:

$$C_r = \frac{(N\sum_{j=1}^{N} x_j y_j - \sum_{j=1}^{N} x_j \sum_{j=1}^{N} y_j)}{(N\sum_{j=1}^{N} x_j^2 - (\sum_{j=1}^{N} x_i)^2)(N\sum_{j=1}^{N} y_j^2 - \sum_{j=1}^{N} y_i)^2)} \qquad (10)$$

Where x_i and y_j are the values of the adjacent pixels in the image and N is the total number of pixels selected from the image for the calculation. We have chosen randomly5000 image pixels in the plain image and the ciphered image respectively to calculate the correlation coefficients of the adjacent pixels in diagonal, horizontal

and vertical direction (See Table 2). The same result for ciphered image presented in Table 3. It demonstrates that the encryption algorithm has covered up all the characters of the plain image showing a good performance of balanced 0 - 1 ratio. The correlation of the plaintext and cipher text is shown in Fig. (6).

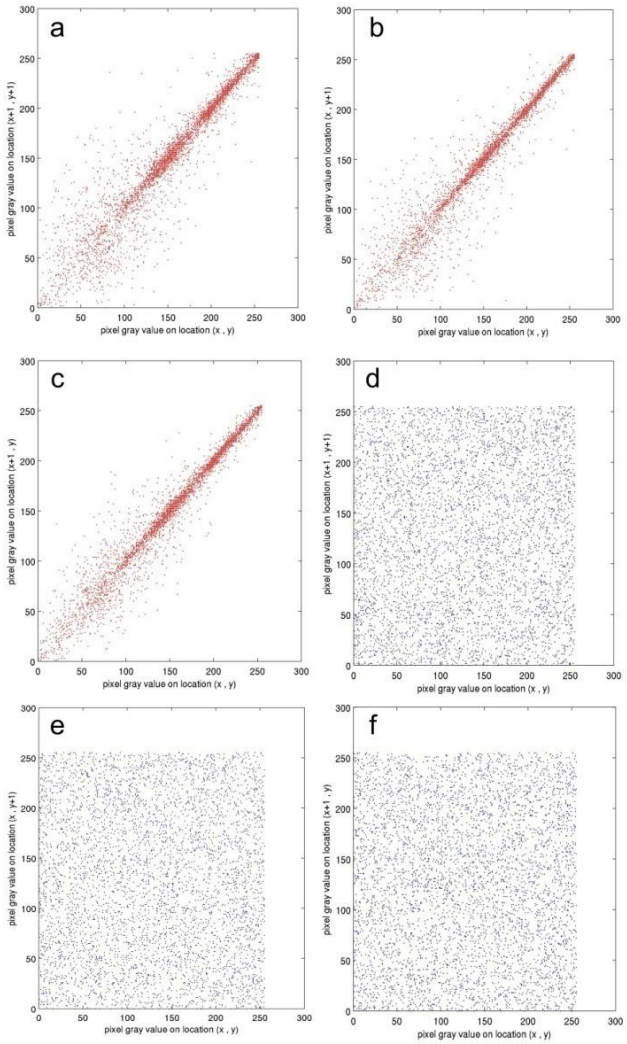


Fig. 6 Correlations of two diagonal, horizontal and vertical adjacent pixels in the plain-image and in the cipher-image: (a-c) Correlation analysis of plain-image. (d-f) Correlation analysis of cipher-image.

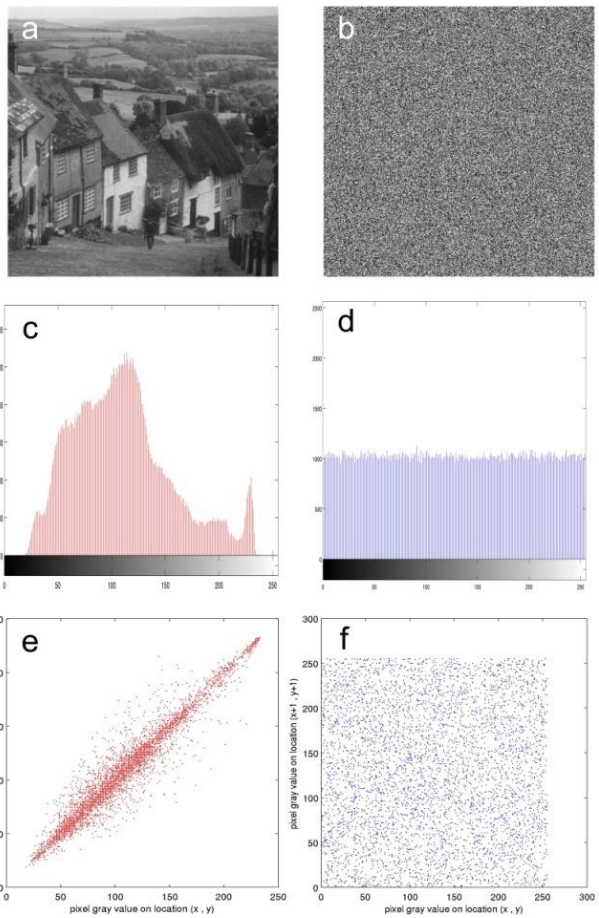


Fig. 7 (a) Original Hill image (b) encrypted image (c) histogram of original image (d) histogram of encrypted image (e-f) Correlations of two diagonal adjacent pixels in original and encrypted image.

4.3 Plaintext sensitivity analysis (Differential analysis):

In order to resist differential attack, a minor alternation in the plain-image should cause a substantial change in the cipher-image. To test the influence of one-pixel change on the whole image encrypted by the proposed algorithm, two common measures were used: NPCR and UACI [23]. NPCR represents the change rate of the ciphered image provided that only one pixel of plain image changed. UACI which is the unified average changing intensity, measures the average intensity of the differences between the plain-image and ciphered image. For calculation of NPCR and UACI, let us assume two ciphered images C_1 and C_2 whose corresponding plain images have only one-pixel difference. Label the grey-scale values of the pixels at grid (i, j) of C_1 and C_2 by $C_1(i, j)$ and $C_2(i, j)$, respectively. Define a bipolar array, D, with the same size as image C_1 or C_2. Then, $D(i, j)$ is determined by $C_1(i, j)$ and $C_2(i, j)$, namely, if $C_1(i, j) = C_2(i, j)$ then

$D(i,j)=1$; otherwise, $D(i,j)=0$. NPCR and UACI are defined by the following formulas [24]:

$$UACI=\frac{1}{W\times H}[\frac{|C_1(i,j)-C_2(i,j)|}{255}]\times 100\% \quad (11)$$

$$NPCR=\frac{\sum_{i,j}D(i,j)}{W\times H}\times 100\% \quad (12)$$

Where W and H are the width and height of C_1 or C_2. Tests have been performed on the proposed scheme by considering the one-pixel change influence on a 256 grayscale image of size 512×512. The obtained result presented in Table 4. The calculated value of UACI and NPCR for the proposed algorithm is compared with other chaos-based image encryption algorithm and comparison results is shown in Table 5. The table shows that the proposed algorithm is so sensitive to the plaintext.

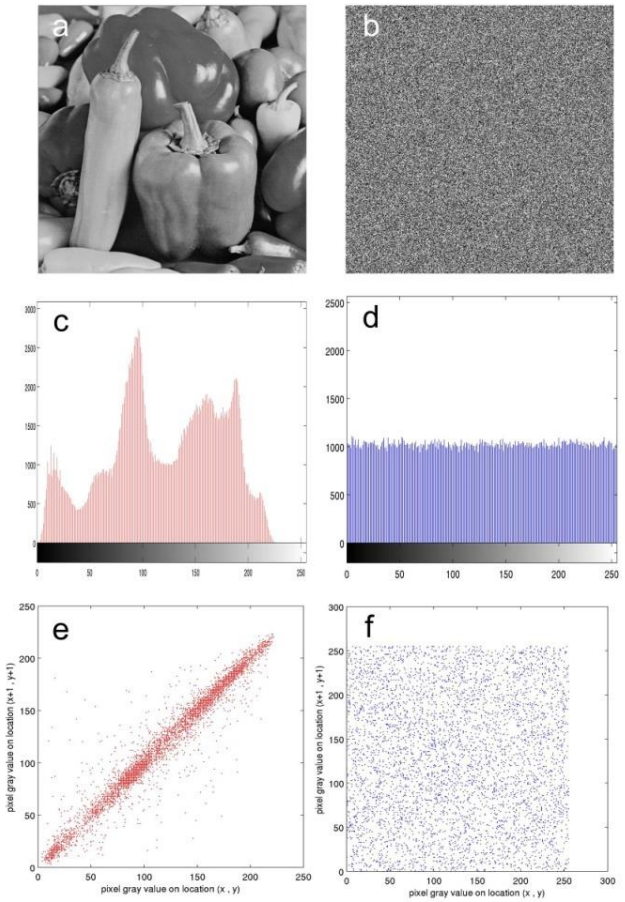


Fig. 8 (a) Original peppers image (b) encrypted image (c) histogram of original image (d) histogram of encrypted image (e-f) Correlations of two diagonal adjacent pixels in original and encrypted image.

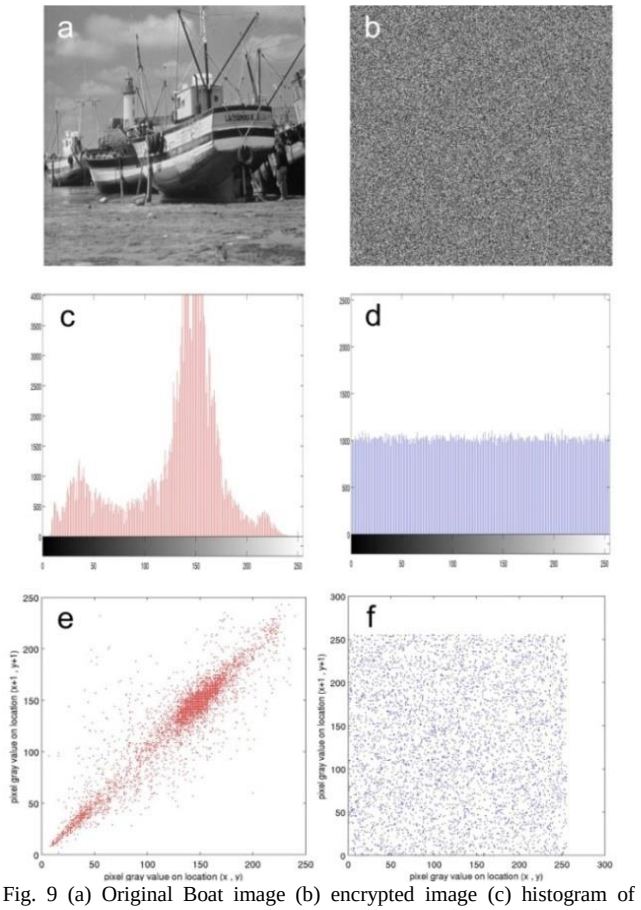


Fig. 9 (a) Original Boat image (b) encrypted image (c) histogram of original image (d) histogram of encrypted image (e-f) Correlations of two diagonal adjacent pixels in original and encrypted image.

5. Key Space

The key is the fundamental aspect of every cryptosystem. An algorithm is as secure as its key. No matter how strong and well designed the algorithm might be, if the key is poorly chosen or the key space is small enough, the cryptosystem will be broken. The size of the key space is the number of encryption/decryption key pairs that are available in the cipher system.

In the proposed scheme, the secret key includes nineteen control and initial conditions parameter of the four chaotic maps. The sensitivity to these initial parameters is shown as follows:

- Jacobian Elliptic Maps: ($X_0\in[0,1]$, $\alpha\in[0,4]$ and $k\in[0,1]$).
- Chaotic Coupled Map: ($X_0\in[0,1]$, $\alpha_1\in[0,4]$, $\alpha_2\in[0,4]$, $\varepsilon\in[0,1]$ and $P\in[2,10]$).
- Quantum Map: ($X_0\in[0,1]$, $r\in[3.6,4]$, $\lambda\in[0,1]$, $m\in[1,4]$, $b\in[1,4]$ and

$k \in [2,10]$).

- Piecewise Nonlinear Chaotic Map:

($X_0 \in [0,1]$, $\alpha \in [0,4]$, $\beta \in [0,4]$, $Y_0 \in [0,1]$, $b \in [1,4]$ and $P \in [0,1]$).

If the precision is 10^{-14} for each of nineteen parameters, the size of key space for initial conditions and control parameters is $2^{883}((10^{-14})^{19})$ .The key space is large enough to resist all kinds of brute-force attacks [24].

6. CONCLUSION

Cryptography is the art of achieving security by encoding messages to make them non- readable. We propose a novel encryption scheme for color image based on multiple chaotic maps. This algorithm tries to address the shortcoming of encryption such as small key space and level of security. Secret key includes nineteen control and initial conditions parameter of the four chaotic maps and the calculated key space is 2^{883} and the key space is large enough to resist all kinds of brute-force attacks. Therefore, it is an effective technique for image encryption. The goal is to realize an encryption method with a private code. Further studies must be started to develop encryption methods with a public key.

TABLE 1: INFORMATION ENTROPY

Image	*Entropy*
Mashhad	7.9994
Hill	7.9994
Peppers	7.9993

TABLE 2: SIMULATION RESULT OF CORRELATION OF TWO ADJACENT PIXELS IN ORIGINAL IMAGE

Image	*Diagonal*	*Horizontal*	*Vertical*
Mashhad	0.9493	0.9725	0.9762
Hill	0.9655	0.9818	0.9820
Peppers	0.9715	0.9822	0.9812
Boat	0.9259	0.9358	0.9728

TABLE 3: SIMULATION RESULT OF CORREATION OF TWO ADJACENT PIXELS IN CIPHER IMAGE

Image	*Diagonal*	*Horizontal*	*Vertical*
Mashhad	-0.0188	0.0015	-0.0038
Hill	0.0029	0.0211	0.0014
Peppers	0.0072	0.0051	0.0092
Boat	0.0027	0.0419	-0.0059

TABLE 4: PLAINTEXT SENSITIVITY ANALYSIS OF PROPOSED ALGORITHM

Image	*UACI*	*NPCR*
Mashhad	0.3345	0.3841
Hill	0.3352	0.3925
Peppers	0.3353	0.3937
Boat	0.3352	0.3979

TABLE 5: COMPARATION RESULT OF PREPOSED ALGORITHM WITH OTHER RELATION CHAOTIC METHOD

Algorithm	*UACI*	*NPCR*
Behnia et al.[7]	0.39	0.46
Akhavan et al.[28]	0.39	0.39
Sun et al.[29]	0.3192	0.40
Rhouma et al.[30]	0.3346	0.389
Tong et al. [31]	0.3356	0.39453
Proposed Algorithm	0.3345	0.3841

References

[1] Ira Aviram, AvinoamRabinovitch, Bifurcation analysis of bacteria and bacteriophage coexistence in the presence of bacterial debris, ommunications in Nonlinear Science and Numerical Simulation, Volume 17, Issue 1, January 2012, Pages 242-254, ISSN 1007-5704, DOI: 10.1016/j.cnsns.2011.04.031.

[2] RongweiGuo, Finite-time stabilization of a class of chaotic systems via adaptive control method, Communications in Nonlinear Science and numerical Simulation, Volume 17, Issue 1, January 2012, Pages 255-262, ISSN 1007-5704, DOI: 10.1016/j.cnsns.2011.05.001.

[3] Li-Guo Yuan, Qi-Gui Yang, Parameter identification and synchronization of fractional-order chaotic systems, Communications in Nonlinear Science and Numerical Simulation, Volume 17, Issue 1, January 2012, Pages 305-316, ISSN 1007-5704, DOI: 10.1016/j.cnsns.2011.04.005.

[4] Kun Zhang, Hua Wang, Hui Fang, Feedback control and hybrid projective synchronization of a fractional-order Newton-Leipnik system, Communications in Nonlinear Science and Numerical Simulation, Volume 17, Issue 1, January 2012, Pages 317-328, ISSN 1007-5704, DOI: 10.1016/j.cnsns.2011.04.003.

[5] Wei J, Liao XF, Wong KW, Xiang T. "A new chaotic cryptosystem", Chaos, Solitons& Fractals, vol. 30, pp.1143-52, 2006.

[6] Lian S, Sun J, Wang J, Wang Z. "A chaotic stream cipher and the usage in video protection", Chaos, Solitons&Fractals , vol.34, pp.851-59, 2007.

[7] S. Behnia, A. Akhshani, S. Ahadpour, H. Mahmodi, and A. Akhavan, "A fast chaotic encryption scheme based on piecewise nonlinear chaotic maps," Phys. Lett. A vol. 366, pp.391-396, 2007.

[8] S. Behnia, A. Akhshani, H. Mahmodi, and A. Akhavan, A. [2008] "A novel algorithm for image encryption based on mixture of chaotic maps," Chaos, Solitons& Fractals 35, pp. 408-19.

[9] VinodPatidar, N.K. Pareek, G. Purohit, K.K. Sud, A robust and se-cure haotic standard map based pseudorandom permutation-substitution scheme for image encryption, Optics Communications, In Press, Cor-rected Proof, Available online 26 May 2011, ISSN 0030-4018, DOI: 10.1016/j.optcom.2011.05.028.

[10] Zhi-liang Zhu, Wei Zhang, Kwok-wo Wong, Hai Yu, A chaos-based symmetric image encryption scheme using a bit-level permutation,Information Sciences, Volume 181, Issue 6, 15 March 2011, Pages 1171-186, ISSN 0020-0255, DOI: 10.1016/j.ins.2010.11.009.
[11] Yong Wang, Kwok-Wo Wong, Xiaofeng Liao, Guanrong Chen, A new chaos-based fast image encryption algorithm, Applied Soft Computing, Volume 11, Issue 1, January 2011, Pages 514-522, ISSN 1568-4946, DOI: 10.1016/j.asoc.2009.12.011.
[12] Di Xiao, Frank Y. Shih, Using the self-synchronizing method to improve security of the multi chaotic systems-based image encryption, Optics Communications, Volume 283, Issue 15, 1 August 2010, Pages 3030-036, ISSN 0030-4018, DOI: 10.1016/j.optcom.2010.03.063.
[13] K. Umeno, Phys. Rev. E 58 (1998) 2644.
[14] R. Chacon, A.M. Garcia-Hoz, Europhys. Lett. 57 (1) (2002) 7.
[15] K. Umeno, RIMS Kokyuroku 1098 (1999) 104.
[16] R.L. Devancy, An Introduction to Chaotic Dynamical Systems. Addison Wesley, 1982.
[17] T. Xiang, K. Wong and X. Liao, Selective image encryption using a spatiotemporal chaotic system, Chaos 17 (2007), p. 023115.
[18] P. Li, Z. Li, W. Halang and G. Chen, A stream cipher based on a spatiotemporal chaotic system, Chaos SolitonsFract 32 (2007), pp. 1867–1876.
[19] S. Wang, J. Kuang, J. Li, Y. Luo, H. Lu and G. Hu, Chaos-based secure communications in a large community, Phys Rev E 66 (2002), p. 065202.
[20] S. Behnia, M. Teshnehlab, P. Ayubi, Multiple-watermarking scheme based on improved chaotic maps, Communications in Nonlinear Science and Numerical Simulation, Volume 15, Issue 9, September 2010, Pages 2469-2478, ISSN 1007-5704, DOI: 10.1016/j.cnsns.2009.09.042.
[21] M. A. Jafarizadeh and S. Behnia, “Hierarchy of one- and many-parameter families of elliptic chaotic maps of cn and sn types”, Physics Letters A vol. 310, pp.168-176, 2003.
[22] R.L. Devancy, An Introduction to Chaotic Dynamical Systems. Addison Wesley, 1982, pp. - .
[23] E. Ott. Chaos in dynamical system .Cambidge university pess, 2002, pp. .
[24] A. Akhavan, H. Mahmodi, A. Akhshani, A new image encryption algorithm based on one-dimensional polynomial chaotic maps, Lect. Notes Comput. Sci. 4263 (2006) 963971.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Using k-means clustering algorithm for common lecturers timetabling among departments

Hamed Babaei[1], Jaber Karimpour [2] and Sajjad Mavizi [3]

[1] Department of Computer Engineering
Islamic Azad University, Ahar Branch
Ahar, Iran
hamedbabaei63@gmail.com

[2] Department of Computer Sciences
University of Tabriz
Tabriz, Iran
karimpour@tabrizu.ac.ir

[3] Department of Computer Engineering
Islamic Azad University, Shabestar Branch
Shabestar, Iran
sajjad.mavizy@gmail.com

Abstract

University course timetabling problem is one of the hard problems and it must be done for each term frequently which is an exhausting and time consuming task. The main technique in the presented approach is focused on developing and making the process of timetabling common lecturers among different departments of a university scalable. The aim of this paper is to improve the satisfaction of common lecturers among departments and then minimize the loss of resources within departments. The applied method is to use a collaborative search approach. In this method, at first all departments perform their scheduling process locally; then two clustering and traversing agents are used where the former is to cluster common lecturers among departments and the latter is to find unused resources among departments. After performing the clustering and traversing processes, the mapping operation in done based on principles of common lecturers constraint in redundant resources in order to gain the objectives of the problem. The problem's evaluation metric is evaluated via using clustering algorithm k-means on common lecturer constraints within a multi agent system. An applied dataset is based on meeting the requirements of scheduling in real world among various departments of Islamic Azad University, Ahar Branch and the success of results would be in respect of satisfying uniform distribution and allocation of common lecturers on redundant resources among different departments.

Keywords: *University Course TimeTabling Problem (UCTTP), Common Lecturer TimeTabling Problem (CLTTP), Multi-Agent Systems, K-mean Clustering Algorithms.*

1. Introduction

The goal of the university course timetabling problem (UCTTP[1]) is to find a method to allocate whole events to fix predefined timeslots and rooms, where all constraints within the problem must be satisfied. Events include students, teachers and courses where resources encompasses the facilities and equipment's of classrooms such as theoretical and practical rooms. Also timeslots include two main components, namely daily and weekly timeslots which it varies from one institution to another. However, each classroom also has its own components allocated to those classrooms (the capacity of theory and practical rooms), number of blackboards and whiteboards related to each theory and practice classroom and etc. [1, 2, and 3].

1.1 Description of the Problem

UCTTP is a hybrid optimization problem in the class of NP-hard problems occur at the beginning of each semester of universities and includes the allocation of events (courses, teachers and students) to a number of fixed timeslots and rooms. This problem must satisfy both hard and soft constraints during allocation of events to resources, so that the possible timetables are obtained after full satisfaction of whole hard constraints and also soft constraints to increase and promote the quality of possible generated timetables as necessary. There are some problems and complexities in UCTTP process; firstly, the scheduling process is an NP-complete problem, then it could not be solved in the polynomial time classes because of the exponential growth of this problem and the existence of some variations in the fast growth of students' numbers in this problem, so we must seek heuristic approaches. Secondly, the number of constraints (hard and soft) in this problem differs from one institution to another. Therefore, the main aim of all of the mentioned algorithms is to maximize the number of soft constraints satisfied in the final timetables [1, 2, 3, and 4].

1.2 The basic definitions of the problem

- Event: a scheduled activity, like: teacher, course, and student.
- Timeslot: a time interval in which each event is scheduled, like: weekly timeslot such as Tuesday and daily timeslot such as 8 a.m. to 9 a.m. and etc.
- Resource: resources are used by events, like: equipment's, rooms, timeslots and etc.
- Constraint: a constraint is a restriction in scheduling of events, categorized into two types of hard and soft constraints, like the capacity of classrooms, given timeslot and etc.
- People: People include lecturers and students and are a part of events.

[1] University Course TimeTabling Problem

- Conflict: the confliction of two events with each other, like: scheduling of more than one teacher for one classroom at the same time.

1.3 Different types of constraints in the problem

Constraints in UCTTP problem are classified into two classes of hard and soft constraints. Hard constraints must be satisfied in the problem completely so that the generated solution would be possible and without conflict; no violation is allowed in these constraints. Soft constraints are related to objective function; objective function is to maximize the number of satisfied soft constraints. Unlike hard constraints, soft constraints are not necessarily required to satisfy; but as the number of these satisfied constraints increases, the quality of solutions of objective function increases. In the following, a list of hard and soft constraints presented which are taken from literature [1, 2, 3, 4, 5, 6, and 7].

1.3.1 Hard Constraints

- A teacher could not attend two classes at the same time.
- A course could not be taught in two different classes at the same time.
- A teacher teaches only one course in one room at each timeslot.
- At each daily timeslot in one room only one group of students and one teacher could attend.
- A teacher teaches for only one group of students at each daily timeslot.
- There are some predefined courses which are scheduled in a given timeslots.
- The capacity of the classrooms should be proportional to the number of students of the given course.

1.3.2 Soft Constraints

- The teacher can have the choice to suggest priority certain timeslots for her/his courses either public or private times.
- A teacher may request a special classroom for a given course.
- The courses should be scheduled in a way that the empty timeslots of both teacher and student to be minimized.
- Timetabling of the courses should be conducted in a way that the courses not scheduled at evening timeslots, as it is possible; unless an evening timeslot has been requested by a particular teacher.
- The lunch break is either 12 p.m. to 13 p.m. or 13 p.m. to 14 p.m., usually.
- The start time of classes may be 8 a.m. and the ending time may be 20:30 p.m. (evening), usually.
- The maximum teaching hours for teachers in a classroom are 4 hours.
- The maximum learning hours for students is 4 hours.
- Scheduling should be conducted in a way that one or a group of students not attend university for one timeslot in a day.

1.4 Mathematical formulation of the problem

Formal definition of UCTTP problem includes n: the number of events $E=\{e_1, e_2, \ldots, e_n\}$, k: the number of timeslots $T=\{t_1, t_2, \ldots, t_k\}$, m: the number of rooms $R=\{r_1, r_2, \ldots, r_m\}$, L: the number of rooms' features $F=\{f_1, f_2, \ldots, f_l\}$ and s: the set of students $S=\{s_1, s_2, \ldots, s_s\}$. For example, if the number of daily timeslots is 9 and the number of weekly timeslots is 5, then the total timeslots will be $T= 9 \times 5 =45$.
The input data for each sample problem (data sets) include the size and features of each room, the number of students in an event and information about conflicting events. So, we should know the procedure of measuring violation and non-violation of hard and soft constraints in order to have the ability to replace events within matrixes. At first the penalty function per violation from soft constraint must be calculated for each solution which is corresponding to a timetable, as bellow [3, 5, 6, and 7]:

$$PF(S) = \sum_{j=1}^{SC} W_j \times (-1) \quad (1)$$

In Eq. (1), S is the solution, W_j is the weight of each soft constraint (value 0 means non-violation, value 1 means violation and -1 shows the cost of each violation per soft constraint) and SC is the number of soft constraints. However, PF represents the penalty function. Value of objective function per solution considering hard constraints can be calculated as:

$$OF(S) = \sum_{i=1}^{HC} W_i \times (-1) + PF(S) \quad (2)$$

In Eq. (2), W_i is the weight of each hard constraint where value 0 means non-violation, value 1 means violation and -1 shows the cost of each violation per hard constraint. Also HC and OF are the number of hard constraints, and the objective function, respectively. Always the value of first term of right hand side of the Eq. (2) is equal to zero ($\sum_{i=1}^{HC} W_i \times (-1) = 0$), this means that the violation of hard constraints is not feasible. So OF (S) = 0 + PF (S), consequently OF (S) = PF (S).
In order to determine the violation of solutions, from hard and soft constraints, results of sample problems are stored in 5 matrixes namely STUDENT-EVENT, EVENT-CONFLICT, ROOM-FEATURES, EVENT-FEATURES and EVENT-ROOM which is introduced in the following.
Each event is met by each student which is stored in the matrix STUDENT-EVENT. This matrix called matrix A is a $k \times n$ matrix. If the value of $U_{i,j}$ in the matrix $A_{k,n}$ be 1, then student $i \epsilon S$ must attend event $j \epsilon s$, otherwise, its value will be 0. The matrix size is $k \times n = |S| \times n$. The EVENT-CONFLICT matrix is an $n \times n$ matrix with two arbitrary events which could be scheduled in the same timeslots. This matrix called matrix B is used to quickly identify events which potentially allocated to same timeslots. ROOM-FEATURES matrix is a $m \times l$ matrix which shows the features of each room; this matrix called matrix C. If the value of $C_{i,j}$ be 1, then each $i \epsilon R$ has a feature of $j \epsilon F$, and otherwise its value will be 0. The matrix size is $m \times l = m \times |F|$. The EVENT - FEATURE matrix also called matrix D is a $n \times l$ matrix and represents the features required by each event. Namely, event $i \epsilon E$ requires features of $j \epsilon F$, if and only if $d_{i,j}$=1. The matrix size is $n \times l = n \times |F|$. Finally the EVENT-ROOM matrix called G

matrix is an n × m matrix which represents the list of possible rooms so that each event could be allocated in those rooms. This matrix represents the quick identification of all rooms in terms of their size and features for each appropriate event. The matrix size is n × m [1, 3, 5, 6, and 7].

1.5 The approaches used in the study of UCTTP

The first definition of timetabling has been presented as three sets of: 1) teachers, 2) classrooms and 3) timeslots (Gotlib, 1963). Approaches used to solving the UCTTP problem up to now are as follows: 1) Operational Researches (OR) based techniques including graph coloring theory based technique, IP/LP method and Constraint Based Satisfaction(s) technique (CPSs); 2) Metaheuristic approaches also including Case Base Reasoning method (CBR), population based approaches and single solution based approaches where the population based approaches includes Genetic Algorithms (GAs), Ant Colony Optimization (ACO), Memetic Algorithm (MA), Harmonic Search Algorithm (HAS) and single solution algorithms also includes Tabu Search Algorithm (TS), Variable Neighborhood Search (VNS), Randomized Iterative Improvement with Composite Neighboring algorithm (RIICN), Simulated Annealing (SA) and Great Deluge Algorithm (GD); 3) multi criteria and multi objective approaches; 4) intelligent novel approaches such as hybrid approaches, artificial intelligence based approaches, fuzzy theory based approaches and 5) distributed multi agent systems approach [1, 2, 3, 5, 6, and 7].

1.6 Motivation and historical perspective of the problem

Agents are technologies inspired from global environment to develop initial instances of systems. Whenever a distributed multi agent system is considered, it means that there is a network of agents cooperates with each other to solve problems which are out of capability of each single agent [8]. Recently, using distributed multi agent systems based approach to solve UCTTP problem has been applied by [9] where in the this method, a solution is used to deal with UCTTP problem using distributed environment and an interface agent -which is responsible to cooperate different timetabling agents- collaborate with each other to improve the solution of common goal. The initial timetables are generated for multi agent systems by using multiple hybrid metaheuristics which are a combination of graph coloring metaheuristics and local search in different methods. The hybrid metaheuristics provide the capability to generate possible solutions for all samples of both Socha et al. (2002) and international competitions timetabling 2002 datasets. However, recently, [10] has used distributed agents to create UCTTP by considering hard (necessary) and soft (desirable) constraints. Also, he presented fairly meeting of distribution in allocating resources in his Ph.D. thesis. There are two types of agents in that model which are year- programmer agent and rooms' agent. However, there are four principles to efficiently organize agents, including: 1) queue and the sequential queue algorithm, 2) queue and interleaved queue algorithm, 3) round robin and sequential round robin algorithm and 4) round robin and interleaved round robin algorithm. The problem formulation and dataset have been adopted from the third section of ICT-2007. The obtained result ensures the consistency of interleaved round robin principle for year-programmer agents in the system and the fairest chance in obtaining the required resources. However, recently [11] has used a novel clustering technique based on FP-Tree to solve UCTTP where the given technique is done to classify students based on their selective courses who submitted for the next semester. The aim of this clustering is to solve scheduling of courses where in the previous semesters the submission of students in some courses due to simultaneous scheduling has been prevented, while in this technique no conflict would happen over scheduling of exams since no two exams at the same time would be taken for courses by two identical groups of students.

1.7 Claim

In this article our main goal is to schedule common lecturers (ˀCLTTP) among different departments based on redundant resources among departments. Clustering algorithms have been used to schedule common lecturers within a distributed system based approach. Since the system uses a distributed multi agent architecture so in order to reach the goal of CLTTP problem, two agents, clustering and traverser, are considered, respectively. The clustering agent performs the act of clustering common lecturers among departments within clusters according to the common, semi-common and uncommon priorities, constraints and features of lecturers so that lecturers who are similar and closer to each other in terms of selecting priorities and constraints are places within high value clusters (primary and more dense clusters) in order to be allocated to their demanded and prioritized resources. After clustering process, the mapping of these clusters is done due to the clusters of common lecturers among departments in to traversed groups of redundant resources among departments collected by traverser agent. The research performed in this article is to present a new and different approach of timetabling problem to develop and make the process of timetabling common lecturers among departments over existing (redundant) resources in departments of a university scalable. The contributions presented in this article to solve the CLTTP problem include: 1) descending satisfaction (from desirable to undesirable priorities) of constraints and priorities of common lecturers among departments and 2) minimizing the loss of redundant resources among departments. Of course, these goals are evaluated by using clustering common lecturers among departments and grouping the redundant resources among departments.

2. Related works

Those approaches solved UCTTP problem by now include the mentioned methods in section 1.5.

2.1 Operational research approaches

Graph coloring approach is on how to model a UCTTP problem by using a non-directional graph where [12] has used vertices as events, colors as time slots and edges as constraints in a graph to solve timetabling problem where no two adjacent vertices have co-colors; since a sign of conflict has been authenticated in the time table. Another hybrid approach has also been proposed to solve UCTTP problem using genetic algorithm by [13] which reduces the cost of finding the number of minimum required colors to color a graph with this hybrid method. In [14], IP method (integer programming) has

ˀ Common Lecturer TimeTabling Problem

been presented to solve UCTTP problem where the goal is to allocate a set of courses among lecturers and groups of students and also a set of weekly and daily time period pairs. Again, [15] has presented an IP-based two-step simplification method where during step 1, the classes require sequence are scheduled by allocating courses to given days and times and during step 2, ensuring the sequence of those courses requiring more than one time period for the same student groups is also done.

2.2 Meta heuristic approaches

In [16], a genetic algorithm has been used in respect of ordering a university timetabling where the intersection rate was 70% and no hard constraint was violated and the applied constraints were almost on room's occupation and capacity. However, [17] has proposed a new GA technique to solve UCTTP problem which uses a learner machine. The results of this technique include minimization of the number of violated soft constraints, high usage of available rooms and reduction of lecturers' workload. Of course applying ant colony optimization algorithm by [18] to UCTTP problem after submission has been done according to ITC-2007 dataset where ants allocate events to rooms and time slots based on two types of pheromone T_{ij}^{s} and T_{jk}^{y} . This algorithm has performed well on timetabling and generated good results during longer. Applying a hybrid ant colony system has been proposed to solve UCTTP problem in [19], where two types of hybrid ant systems including combination of SA with AC and combination of TS with AC have been presented. A number of ants perform entire allocation of courses to time slots based on a predefined list. Selection of time slots' probabilities is done by ants to allocation courses using heuristic information and an indirect coordinator mechanist among agents (Stigmergic) and existing activities within an environment. The memetic algorithm has been done using [20] to solve UCTTP problem via combination of local search method in genetic algorithm. One of the local searches is done on events and the other one is performed on time slots.

The Tabu search algorithm has been applied by [21] for the first time to allocate students to courses and also balance the number of students within whole submitted group where the first phase is: generating a set of solutions for a student, and the second phase is: combining a set of solutions and applying Tabu search with local strategies and the third phase is also: allocating room and improving allocation, while without changing the initial allocation of courses to timeslots. In [22], the influence of neighborhood structures has been presented on Tabu search algorithm to solve UCTTP problem where the effect of simple and swap transitions has been tested on Tabu search operations based on neighborhood structures. Here, four new neighborhood structures have been used and compared. To solve UCTTP problem, the combination of kempe neighborhood chain has been presented in simulated annealing algorithm by [23] where one of the hard constraints of reformulation is done by relaxation and then this constraint is created in the form of relaxed soft constraint. However, the relaxation problem is analyzed in two steps: 1- to create a feasible solution, a heuristic based graph is used and 2- a simulated annealing algorithm has been used to minimize the violations of soft constraints (in the second phase, a kempe neighborhood chain based heuristic has been used).

[24] Also has used directed local search strategy in genetic algorithm to solve UCTTP problem where the directed search strategy uses a data structure to create offspring that stores the extracted information of good individuals of previous generations in itself. The results are satisfactory with this local search combined in the genetic algorithm. The aim is to maximize allocations and minimize the violations from soft constraints. The variable neighborhood search algorithm (VNS) has been presented by [25] to solve UCTTP problem which proposes the base VNS and then states some modifications to each solution which apply an exponential Monte Carlo acceptance criterion. However, the main idea of applying Monte Carlo acceptance criterion was to improve the heuristics by admitting the best solution with given probability so that the number of promised neighbors would be found.

2.3 Modern intelligent approaches

A hybrid algorithm has been presented by [26] which is the combination of sequential heuristic and simulated annealing to solve UCTTP problem on ITC-2002 dataset. This method includes three phases: Phase 1: using a sequential heuristic to generate feasible time tables; phase 2: applying simulated annealing to minimize the number of soft constraints' violations and phase 3: uses simulated annealing to increase the improvement of the generated time tables' quality. Recently, a multi population hybrid genetic algorithm has been proposed by [27] to solve UCTTP problem based on three genetic algorithms FGARI, FGASA and FGATS. In this algorithm, fuzzy logic is used to evaluate the number of violations from soft constraints in fitness function to deal with real worlds data which are ambiguous and non-deterministic and random methods, local search, simulated annealing and Tabu search would also be beneficial in addition to fuzzy method to improve inductive search in order to meet the need of search ability.

To solve UCTTP problem, [28] has presented a fuzzy multi criteria heuristic ordering method where the ordering of events has been done according to three independent heuristics simultaneously using fuzzy methods. The sequential combination of three heuristics is ordered as follows: 1- the highest degree, 2- saturation degree and 3- enrollments degree and the fuzzy weight of an event is also used to represent what problem the event has to be scheduled. The ordered events are allocated to the last time slot with the least value of penalty cost as a descending manner while the feasibility is maintained throughout whole process. A fuzzy solution has been presented by [29] based on memetic approach to solve university timetabling where a time table has been compared with both genetic and memetic algorithm and its results may satisfy the existing constraints simultaneously in a shorter time interval. The aim was to use fuzzy logic as a tool to local search in memetic algorithm. [30] Has proposed the fuzzy genetic heuristic idea to solve UCTTP problem where the genetic algorithm has been applied by using indirect representation based on the features of integrating events and modeling the fuzzy set to evaluate the violation from soft constraints in the objective function according to uncertainty of real world data. Here, a degree of uncertainty which is in an objective function is considered for each soft constraint and this uncertainty is evaluated by formulation of soft constraint violation parameter in objective function by using fuzzy membership functions.

3. The proposed method

In [8], an agent could observe and receive everything through sensors from its environment and then perform within environment via the stimulus. Agents are classified into various classes based on their applications including the following agents: 1- autonomous, 2- intelligent, 3- reactive, 4- pro-active, 5- learner, 6- mobile, 7- collaborative/communicative. So, agents must have a common language and a communicative media to communicate and cooperate with each other where these two components are vital among agents.

3.1 Common lecturers timetabling problem among departments

Common lecturers' timetabling problem among departments is one of the challenges among university departments where in this article it has been tried to perform this scheduling based on regarding priorities and requirements of common lecturers to allocate redundant resources among departments. Since the common lecturers among departments always deal with facilitating their timetabling, then a new idea and solution must be researched to facilitate the timetabling of common lecturers so that some challenges such as collision among lecturers and other common events in departments and not promoting the satisfaction of common lecturers based on their desirable choices would be avoided. However, solving CLTTP problem has led to a developed and scalable scheduling process where in this research we have considered this by performing scheduling and distribution of common lecturers over redundant resources among departments. Therefore, to solve a CLTTP problem, the solution in the form of distributed multi agent system accompanied with applying clustering algorithms must follow the process of minimizing the collision of common lecturers among departments. Fig. 1 represents a holistic view on CLTTP problem in a tree structure.

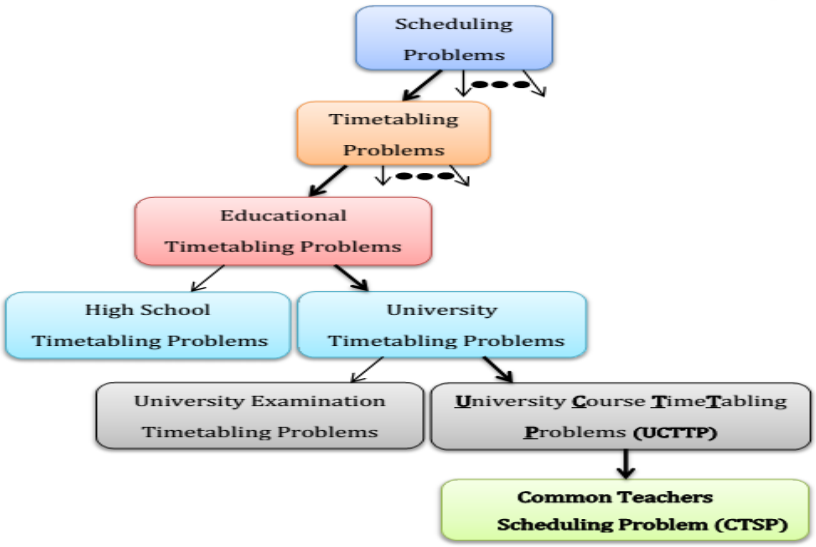


Fig. 1: The tree structure of common lecturers' timetabling problem among departments

3.2 Frameworks and infrastructures of the proposed algorithm

The proposed algorithm consists of four agents: 1- time table (each ith department or agent, $TA_i; i = 1,2,\ldots$), 2- mediator agent (MA), 3- clustering agent (CA) and 4- traverser agent (TraA) which have been shown in Fig. 2, with their relations in three phases. The first phase includes steps 1 and 2 which are planned by the timetabling agent to produce feasible with no conflict time tables. Of course, in this phase, the identification and collection of common lecturers among departments is done by the mediator agent in step 3, the second phase includes steps 4, 5 and 6 which performs the process of clustering common lecturers among departments within the clustering agent to make uniform distribution on the traversed redundant resources of each department by the traverser agent and the third phase consists of steps 7 and 8 where the process of mapping the common lecturers' clusters is done in redundant resources based on the constraints of common lecturers and send the time tables with the capability of planning to each department for a semester.

3.2.1 The first phase

The first phase includes the hard constraints related to lecturers of each department satisfied by TA_i agent and contains the following constraints: 1- a lecturer could not teach more than 6 hours per day, 2- a lecturer could not be in more than one department at the same time slot, simultaneously, 3- a lecturer could not be in two classes at one or more departments in one day or at the same time slot, 4- a class is allocated to one lecturer at one time slot, and 5- two lecturers could not be in the same class of a department at the same time. Fig. 3 show the lecturers timetabling algorithms on the resources related to each department by TA_i agent. Between the first and the second phases, the mediator agent (MA) studies the operation of extracting common lecturers among departments accompanied with their features to cluster in the next step without any conflict based on the aim of the problem which is to time table the common lecturers among departments and sends them to their related departments (TA_i) in order to modify the conflicts when it discovers a conflict and inconsistency in the time tables of common lecturers among departments. And then the time tables of common lecturers of

each department fixed in the respect of the problem aim by the mediator agent are sent during step 3 to the clustering agent (CA).

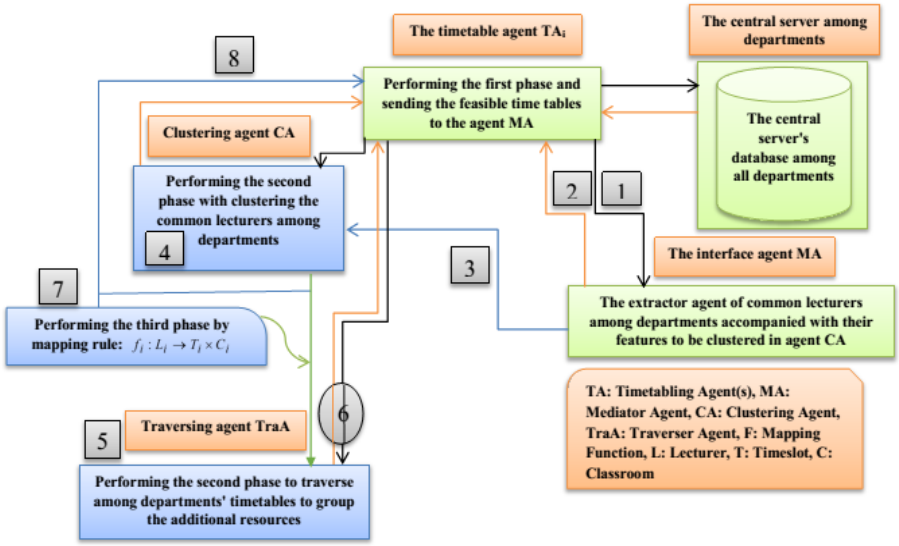


Fig. 2: The general view of CLTTP problem's schematic

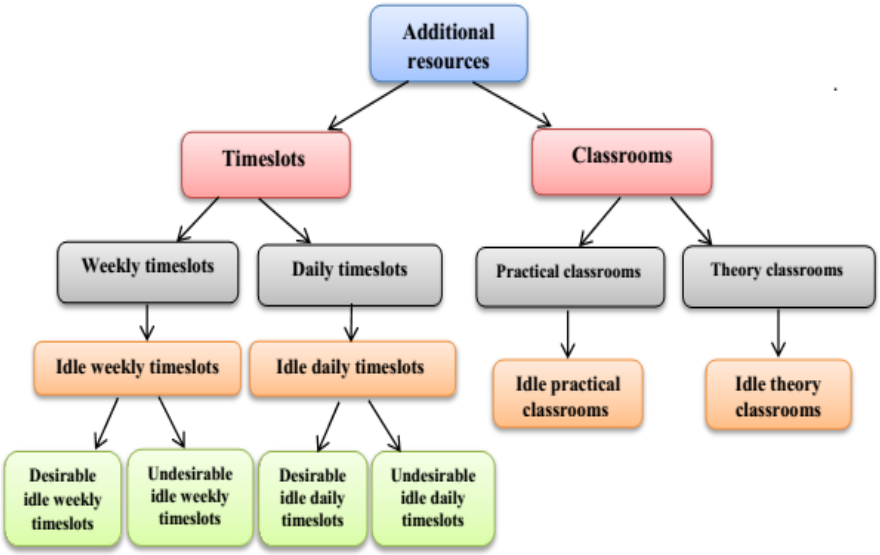


Fig. 3: The structure of grouping redundant resources among departments in TraA

3.3 The second phase

In the second phase, CA clusters common lecturers among departments based on their constraints (step 4) and TraA agent is applied through traversing and grouping the redundant resources among departments (among TA_i agents) (step 5). Of course, before entering step 5, all busy and redundant resources have been determined entirely through time tables of each department (TA_i) in step 6 and sent to step 5 by TraA agent to perform traversing and grouping. In the second phase, two ideas have been proposed where the former is to consider two new agents of CA and TraA in the architecture of multi agent system and perform the mapping process by CA in TraA and the latter is to state a clustering method coinciding the type of problem called k-means clustering to perform the process of clustering common lecturers among departments applied within CA agent. The algorithms of two CA and TraA agents have been shown in figures 4 and 5. The third phase

In the third phase, the process of mapping priorities and requirements of common lecturers is presented to uniformly distribute and allocate redundant resources among departments. In the last step of the third phase (step 8) the final solution (timetabling of common lecturers among departments for one semester) is sent to all the departments based on each department's (TAi) identification codes after the process of mapping clustering agents in the traversed redundant resources in TraA agent.

3.3.1 Clustering and traversing in the second phase

In the second phase, the clustering of common lecturers among departments is performed in the clustering agent (CA) by two algorithms of k-means, fuzzy c- means clustering and the proposed funnel-shape clustering where the clustering process is described through four features of each common

lecturer as: desirable daily time slots, undesirable daily time slots, desirable weekly time slots and undesirable weekly time slots. Following the process of clustering common lecturers among departments based on their features, traversing and grouping of the redundant resources is done by TraA agent as Fig. 3.

3.4 The complete description of adapted k-means clustering algorithm's details

After stating the priorities and soft constraints of each common lecturer among departments based on (3), now in (3) let consider L_k as the kth common lecturer, $WeeklyTimeslots_k$ as the kth weekly time slot, $DailyTimeslots_k$ as the kth daily time slot, $Departement_k$ as kth department and χ_k as the membership degree of each common lecturer.

$$L_k = \left\{ WeeklyTimeslot_k \Big/ Departement_k = \left\{ \chi_k \Big/ DailyTimeslot_k \right\} \right\}; \chi_k \in [0,1] \tag{3}$$

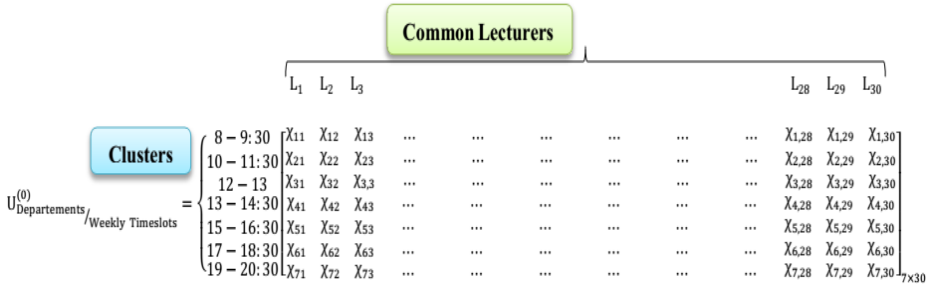
In (4), the default pattern of primary matrixes is represented as $U^{(0)}_{Dep/WeeklyTimeslots}$ related to each department and each weekly timeslot and the values of membership degree of each common lecturers is denoted by χ_{ik} per row or daily timeslot per department are represented as following: each daily time slot $DailyTimeslots_{(1-7)}$ from 8-9:30$_1$ to 19-20:30$_7$ as one cluster which would be 7 clusters and weekly timeslots $WeeklyTimeslots_{(1-7)}$ from Saturday (1) to Friday (7) and Dep as five departments $Departement_{(1-5)}$ in (4). Finally we would reach to the final matrix of $U^{(0)}$ consisting of 7 rows (clusters) and 30 columns (common lecturers). The resulted matrix is represented as (5). In k-means clustering, the membership degree and no membership of each kth common lecturer in the ith cluster would be computable based on (5).

$$U^{(0)}_{Dep_{(1-5)}/WeeklyTimeslots_{(1-7)}} = \begin{Bmatrix} 8-9:30_1 \\ 10-11:30_2 \\ 12-13_3 \\ 13-14:30_4 \\ 15-16:30_5 \\ 17-18:30_6 \\ 19-20:30_7 \end{Bmatrix} \begin{bmatrix} \chi_{11} & \cdots & \chi_{1,30} \\ \cdots & & \cdots \\ & \cdots & \\ \cdots & & \cdots \\ & \cdots & \\ \cdots & & \cdots \\ \chi_{71} & \cdots & \chi_{7,30} \end{bmatrix}_{7,30} \tag{4}$$

$$\chi_{A_i}(x_k) = \begin{cases} 1; x_k \in A_i \\ 0; x_k \notin A_i \end{cases}; A_i is ith cluster and x_k is kth common lecturer \tag{5}$$

In (5), parameter $\chi_{A_i}(x_k)$ meaning the membership degree of x_k (kth common lecturer) in cluster A_i (ith cluster) is represented with two binary values of 0 or 1 to no-membership and membership of kth common lecturer in ith cluster. However, in order to use (5) within the membership matrix, we could review (6). In (6), parameter χ_{ij} is extended as parameter $\chi_{A_i}(x_j)$ where i represents the number of clusters and j is considered as the number of common lecturers. After stating all features and priorities of each common lecturer among departments, we would reach a final matrix of $U^{(0)}$ based on matrixes of (4) in terms of department, daily time slots and weekly time slots parameters which includes 7 rows (clusters) and 30 columns (common lecturers). The obtained matrix is represented as (7).

$$\chi_{ij} = \chi_{A_i}(x_j); \quad i = 1,\ldots,c;\ j = 1,\ldots,n, and\ x_j is jth common lecturer sin\ ith cluster \tag{6}$$



(7)

3.4.1 The steps of adapted k-means clustering for CLTTP problem

By given initial matrix of $U^{(0)}$ for each common lecturer among departments, we have the following steps:
Finding the centers of each i cluster according to j feature of each common lecturer among departments would be calculated through (8). In (8). We could find the center of each i cluster per each j feature and the priorities of common lecturers among departments. In (8) parameters $k=1,...,n$, X_{kj} and χ_{ik}, represent the number of common lecturers among departments, the membership degree of each common lecturer due to i[th] cluster and the contribution amount of each k[th] common lecturer to j[th] feature. Equation (9) is the extension of variable X_{kj} of each common lecturer over three parameters or features (priority) X_{kj_1} of departments, X_{kj_2} is the daily time slot and X_{kj_3} is the weekly time slot. After describing the structure of each i feature of common lecturers in (9), now the rule of finding the center of cluster must be presented in terms of (10) which is consistent with the common lecturers' timetabling problem. It must be noted that since the common lecturers have been distributed among departments, then the (8) must be cycled among all five departments ($Dep=1,...,5$) based on three features and priorities of each common lecturer determined by parameter v_i per given common lecturer.

$$V_{ij}=\frac{\sum_{k=1}^{n}\chi_{ik}\times X_{kj}}{\sum_{k=1}^{n}\chi_{ik}} \tag{8}$$

$$\sum_{k=1,\,j=1}^{k=n,\,j=3} X_{kj}=\sum_{k=1}^{n}\left(X_{kj_1}+X_{kj_2}+X_{kj_3}\right) \xrightarrow{j_1,j_2,j_3}(j_1 isDepartements,\ j_2 isWeeklyTimeslots,\ j_3 isDailyTimeslots) \tag{9}$$

$$V_{ij}=\sum_{Dep=1}^{5}\left\{\frac{\sum_{k=1}^{n}\left(\chi_{ik}\times\left(\sum_{j=1}^{3}X_{kj}\right)\right)}{\sum_{k=1}^{n}\chi_{ik}}\right\};v_i=\left\{v_{i_1},v_{i_2},\ldots,v_{i_j}\right\};i=1,\ldots,7;\ j=1,\ldots,3 \tag{10}$$

Obtaining the distance of each k common lecturer out of cluster i over the lecturer placed in the center of cluster i and extending the distance of k- means clustering would be used to be compatible with common lecturers timetabling problem based on (11). In (11), let d_{ik} be the distance parameter of k[th] common lecturer over i[th] cluster and two parameters X_{kj} and V_{ij} represent the ratio of k[th] common lecturer over each feature j (department, weekly time slot and daily time slot), respectively and the other parameter would be the variable of finding the center of i[th] cluster over feature j of each k[th] common lecturer.
Now, we must obtain the updating process of initial matrix's elements $U^{(0)}$ called the values of membership degree in order to reach matrix $U^{(1)}$. Equation (12) would be the main rule to update the elements of matrix $U^{(0)}$ namely χ_{ik} s in k- means clustering.

$$d_{ik}=d(X_k-V_i)=\sqrt{\sum_{j=1}^{3}\left(X_{kj}-V_{ij}\right)^2}\xrightarrow{k=1,\ldots,30;i=1,\ldots,7}\sqrt{(X_{k1}-V_{i1})^2+(X_{k2}-V_{i2})^2+(X_{k3}-V_{i3})^2} \tag{11}$$

$$Update:U^{(r)};\forall i,k:\chi_{ik}^{(r+1)}=\begin{cases}1;d_{ik}^{(r)}=\min\left\{d_{ik}^{(r)}\right\};\forall i\in c,\\ k=1,\ldots,n;i=1,\ldots,c;r=0,1,\ldots\\ 0;the\,Otherwise\end{cases} \tag{12}$$

Equation (12) performs the updating process of the initial matrix's elements $U^{(0)}$. The parameters of (12) are as follows: r is the counter and repeater of updating, i as i[th] cluster ($i=1,...,c$), k means the k[th] common lecturer ($k=1,...,n$), $d_{ik}^{(r)}$ as r[th] iteration of the rule of finding the distance of k[th] common lecturer over i[th] cluster and $\chi_{ik}^{(r+1)}$ represents the membership degree of k[th] common lecturer over i[th] cluster after the first iteration and upon the initial matrix $U^{(0)}$ (described either randomly or based on the requirements of each common lecturer). In (13), the process of extending the rule of updating the (12) is presented based on the common lecturer timetabling problem. In (12), that element of initial matrix $U^{(0)}$ with the minimum value in terms of distance of k[th] lecturer over i[th] cluster, $\min_{i=1,\ldots,c;k=1,\ldots,n;r=0,1,\ldots}\left\{d_{ik}^{(r)}\right\}$, equals to $\chi_{ik}=1$ and otherwise it would be $\chi_{ik}=0$. In (12) the value of χ_{ik} would be 0 or 1. In fact, according to (13), $d_{ik}^{(r)}$ with the minimum value is replaced with 1 and others with 0. Now, after computing each updated value of χ_{ik} based on (12), matrix $U^{(1)}$ is formed as (14).

$$k=1 \Rightarrow \min(d_{11},\ldots,d_{71}) = \chi_{11} or \ldots or \chi_{71}$$
$$\ldots\ldots\ldots\ldots$$
$$k=30 \Rightarrow \min(d_{1,30},\ldots,d_{7,30}) = \chi_{1,30} or \ldots or \chi_{7,30} \quad (13)$$

$$U^{(1)}_{Departements/Weekly\ Timeslots} = \begin{array}{c} \\ 8-9{:}30 \\ 10-11{:}30 \\ 12-13 \\ 13-14{:}30 \\ 15-16{:}30 \\ 17-18{:}30 \\ 19-20{:}30 \end{array} \begin{array}{c} \overbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{\text{Common Lecturers}} \\ L_1 \quad L_2 \quad L_3 \quad \ldots \quad L_{28} \quad L_{29} \quad L_{30} \\ \begin{bmatrix} \chi_{11} & \chi_{12} & \chi_{13} & \ldots & \chi_{1,28} & \chi_{1,29} & \chi_{1,30} \\ \chi_{21} & \chi_{22} & \chi_{23} & \ldots & \chi_{2,28} & \chi_{2,29} & \chi_{2,30} \\ \chi_{31} & \chi_{32} & \chi_{3,3} & \ldots & \chi_{3,28} & \chi_{3,29} & \chi_{3,30} \\ \chi_{41} & \chi_{42} & \chi_{43} & \ldots & \chi_{4,28} & \chi_{4,29} & \chi_{4,30} \\ \chi_{51} & \chi_{52} & \chi_{53} & \ldots & \chi_{5,28} & \chi_{5,29} & \chi_{5,30} \\ \chi_{61} & \chi_{62} & \chi_{63} & \ldots & \chi_{6,28} & \chi_{6,29} & \chi_{6,30} \\ \chi_{71} & \chi_{72} & \chi_{73} & \ldots & \chi_{7,28} & \chi_{7,29} & \chi_{7,30} \end{bmatrix}_{7\times 30} \end{array} \quad (14)$$

At each step, in order to terminate the updating process of matrix $U^{(r)}$ to $U^{(r+1)}$, (15), the matrix norm rule, must be used to terminate the execution of k- means clustering algorithm. However, it must be said that the iteration process of (15) is in a way that we would reach to an optimal solution matrix $U^{(*)}$ and this procedure follows the $U^{(r=0)} \to U^{(r=1)} \to \ldots \to U^{(*)}$ rule. Equation (16) represents how to apply (15) for common lecturer timetabling problem. Step 4 is the final phase of k-means algorithm for membership of each common lecturer where the (14) fails, the restoration would continue from the step 1 with recently created matrix $U^{(1)}$ so that we reach a new matrix $U^{(2)}$ which is the updated matrix $U^{(1)}$ and so on.

After terminating each membership matrix U 's updating, the value of objective function must be obtained in terms of two parameters χ_{ik} and d_{ik} based on (17), where k is the number of common lecturers, c is the number of clusters, χ_{ik} is the membership degree of each k[th] common lecturer in i[th] cluster and d_{ik} is also the distance of k[th] common lecturer over the common lecturers within the center of i[th] cluster in c cluster.

$$\left\| U^{(r+1)} - U^{(r)} \right\| = \max_{i,k} \left| \chi_{ik}^{(1)} - \chi_{ik}^{(0)} \right| \le \varepsilon_L; \quad (15)$$
$$\varepsilon_L = 0.1 or 0.01, r = 0,1,2,\ldots$$

$$J(U,v) = \sum_{k=1}^{n=30} \sum_{i=1}^{c=7} \chi_{ik} \left(d_{ik}^2 \right) Objective function \quad (16)$$

$$\left\| U^{(r=1)} - U^{(r=0)} \right\| =$$
$$\max_{i=1,\ldots,7;k=1,\ldots 30} \begin{vmatrix} \chi_{11}^{(1)} - \chi_{11}^{(0)} & \cdots & \cdots & \chi_{1,30}^{(1)} - \chi_{1,30}^{(0)} \\ \chi_{21}^{(1)} - \chi_{21}^{(0)} & \cdots & \cdots & \chi_{2,30}^{(1)} - \chi_{2,30}^{(0)} \\ \chi_{31}^{(1)} - \chi_{31}^{(0)} & \cdots & \cdots & \chi_{3,30}^{(1)} - \chi_{3,30}^{(0)} \\ \chi_{41}^{(1)} - \chi_{41}^{(0)} & \cdots & \cdots & \chi_{4,30}^{(1)} - \chi_{4,30}^{(0)} \\ \chi_{51}^{(1)} - \chi_{51}^{(0)} & \cdots & \cdots & \chi_{5,30}^{(1)} - \chi_{5,30}^{(0)} \\ \chi_{61}^{(1)} - \chi_{61}^{(0)} & \cdots & \cdots & \chi_{6,30}^{(1)} - \chi_{6,30}^{(0)} \\ \chi_{71}^{(1)} - \chi_{71}^{(0)} & \cdots & \cdots & \chi_{7,30}^{(1)} - \chi_{7,30}^{(0)} \end{vmatrix} \le \varepsilon_L \quad (17)$$

Before mapping these functions, the way of independently mapping of function g has been shown in Fig. 4 for the resources of each department and the function f within Fig. 5 has been represented to map the common lecturers among departments in additional resources.

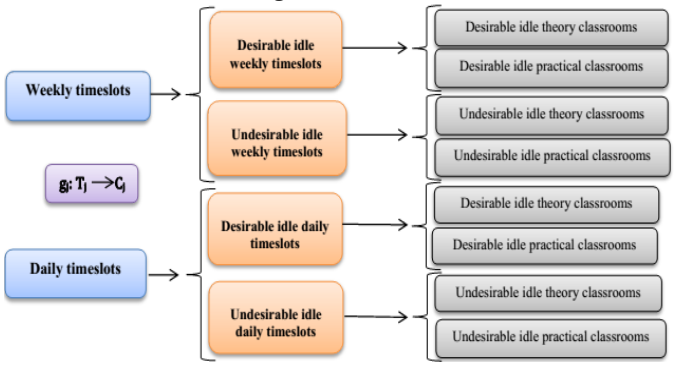


Fig. 4: Mapping time slots in classes

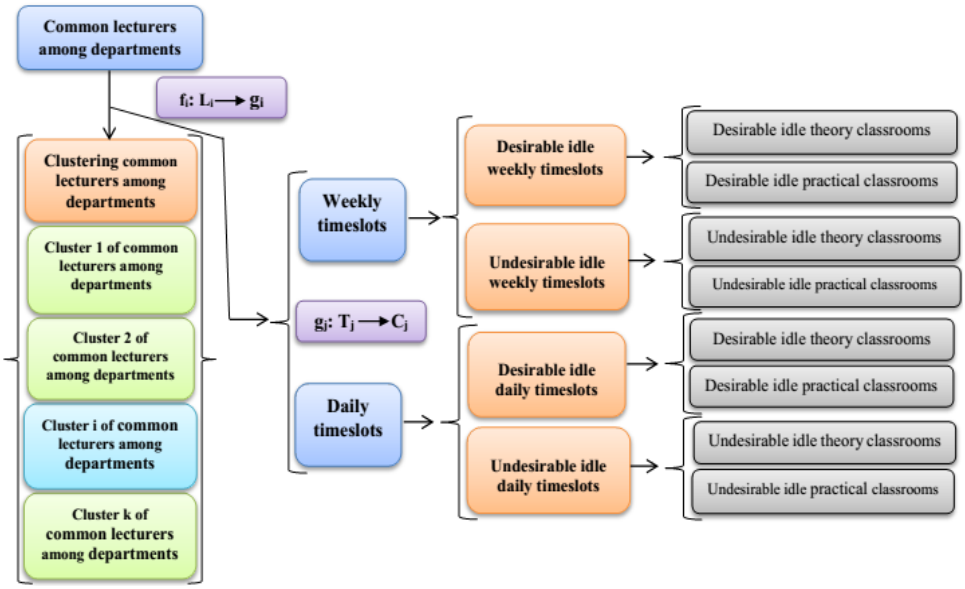


Fig. 5: Mapping the clusters of common lecturers in additional resources among departments

3.5 Mapping

In the third phase the mapping function is as $f_i : L_i \to T_i \times C_i$, where f_i is the mapping function of priorities and requirements of common lecturers (soft constraints of common lecturers), L_i s are the representative clusters of common lecturers among departments, T_i s represent additional time slots among departments and C_i s also represent additional classes among departments. However, before mapping function f, the mapping of function f must also been performed by agent $TraA$ for the resources among departments as $g_j : T_j \to C_j$. Fig. 6 presents the way of mapping two functions f and g for the common lecturers to the additional resources among departments.

4. Results and experiments

To test the structure of the proposed algorithm, we consider a data set including 30 lecturers, 5 departments (computer engineering, electronic engineering, civil engineering, humanity science and mathematics), 7 weekly time slots (Saturday, Sunday, Monday, Tuesday, Wednesday, Thursday, Friday), 7 daily time slots (8-9:30, 10-11:30, 12-13, 13-14:30, 15-16:30, 17-18:30 and 19-20:30) and 13 classrooms per department (3 practical classes an 10 theoretical classes). The properties of the system to implement include a CPU with 2.13 GHZ speed, 3GB RAM and Win7 operating system and the implementation tools also include 1) C#.net 2010 programming language, 2) using SQL server 2008 software for querying from the databases and 3) reporting by Crystal Report v.13. Total number of resources in the university equals to $7\times7\times5\times(10+3)$ and if we want to calculate the separate resources of each department we would have $(7\times7\times5\times(10+3))\div5$ and the total number of the remained additional resources is obtained as $[(7\times7\times5\times(10+3))-(7\times7\times(10+3))]$. The k-means clustering algorithm must be performed to find the loss percent of additional resources per department so that the minimized percent of additional resources per department, $\frac{Dep_D}{(7\times7\times(10+3))}\times100, D=1,\ldots,5$, is obtained as the dedicated resources of each department divided by whole resources of departments, therefore, each Dth department minimizes the loss percent of its additional resources.The criteria of evaluating the CLTTP problem's purposes After using the k-means clustering algorithm and allocating to (additional) resources, the following relations are presented to evaluate the criteria of the paper. Equation (18), $CTDS_1^{(i)}$, computes the descending satisfaction percent of each common lecturer among departments' features at each cluster and (19), $CTDS_2^{(i_j)}$, also obtains the descending satisfaction percent of each common lecturer among departments' priorities and features among clusters and over each cluster.

(18) is calculated per cluster. The numerator of (18) means how many requirements and features of the kth common lecturer in ith cluster presented initially as a report (selections and requirements of each department also could be considered) have been satisfied and the denominator of (18) represents the total number of requests, priorities and requirements of kth common lecturer at that ith cluster which is the sum of satisfied priorities and requirements accompanied with the dissatisfied priorities at ith cluster for the kth common lecturer and the satisfaction percent of kth common lecturer's feature is obtained at ith cluster.

$$CTDS_1^{(i)} = \frac{\sum_{k=1}^{n} W_{ik}^{SC} \times X_{ik}}{\sum_{const=1}^{r} \left(Total_{const}^{SC}\right)} \times 100 = \frac{\sum_{Const=1}^{r} \sum_{k=1}^{n} \left(W_{ik}^{Const} \times L_{ik}\right)}{\sum_{Const=1}^{r} \sum_{k=1}^{n} Total_{k}^{Const}} \times 100$$
$$L_{ik} = \sum_{k=1}^{n=30} \sum_{j=1}^{m=3} X_{k_j} = \sum_{k=1}^{n} \left(X_{k_1} + X_{k_2} + X_{k_3}\right) = \left(X_{1_1} + X_{1_2} + X_{1_3}\right) + \ldots + \left(X_{30_1} + X_{30_2} + X_{30_3}\right) \quad (18)$$

In (18), the i[th] cluster with $i = 1,\ldots,c; c = 7$ shows the k number of common lecturers $k = 1,\ldots,n; n = 30$ and W_{ik}^{SC} constraints satisfied for X_{ik} common lecturer (k[th] lecturer at i[th] cluster). In (18), $Total_{const}^{SC}$ expresses all the constraints of common lecturers at each cluster per common lecturer. For example, $\left(X_{1_1} + X_{1_2} + X_{1_3}\right)$ means the feature of common lecturer 1 has been satisfied at cluster 1.

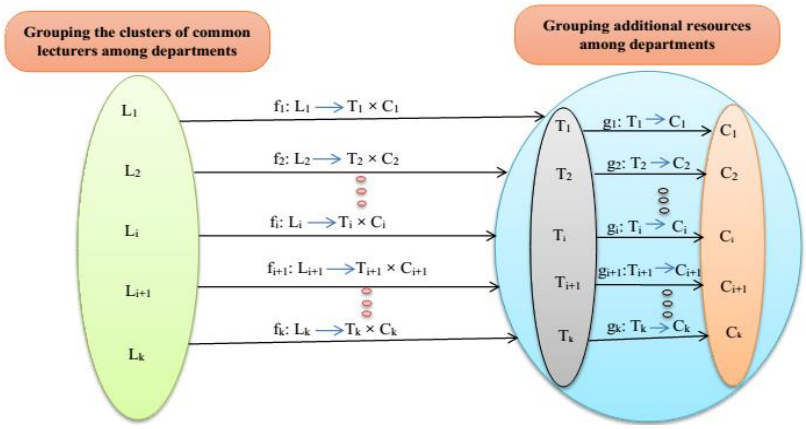


Fig. 6: Mapping the clusters of common lecturers in additional resources

Equation (19), represents the amount of competitiveness among clusters in terms of satisfaction percent of requirements, constraints and priorities of common lecturers among departments of each cluster, it means that we could find that at which i[th] cluster which k[th] common lecturer has more satisfied priorities and requirements over other common lecturers within each i[th] cluster and other j cluster with $j = i+1, i+2, \ldots$. The numerator of this fraction must compute the satisfaction percent of each k[th] common lecturer in terms of each i[th] cluster and the obtain that percent over other j cluster and the denominator of this fraction must find the sum of whole satisfactions of each common lecturer at i[th] cluster with whole dissatisfactions of each common lecturer at i[th] cluster and then this iterates per j remained clusters so that the percent of real satisfactions of each cluster with their common lecturers would be obtained over whole satisfactions and dissatisfaction of per cluster and then the satisfaction percent of each cluster could be obtained over common lecturers and their allocation priority to the additional resources by dividing and averaging the obtained values of each cluster.

$$CTDS_2^{(i_j)} = \frac{\sum_{i=1}^{c} \left(W_i^{SC} \times i\right)}{\sum_{j=i+1}^{c} \left(Total_{Const_j}^{SC} \times j\right)} \times 100 = \frac{\sum_{Const=1}^{r} \sum_{k=1}^{n} \sum_{i=1}^{c} W_{ik}^{Const} \times i}{\sum_{Const=1}^{r} \sum_{j=i+1}^{c} Total_{Const_j}^{Const} \times j} \times 100 \quad (19)$$

In (19), $i = 1,\ldots,c$ is the number of clusters, W_i^{SC} is the satisfaction percent of common lecturers' constraints of i[th] cluster and j also represents the number of other clusters in addition to i[th] cluster where $j = i+1,\ldots,c$ and $c = 7$. In (19), the value of W_i^{SC} must be calculated in terms of the number of satisfied constraints for k[th] common lecturer at i[th] cluster over total number of i[th] cluster's constraints for the common lecturers within this cluster. After obtaining a percent for each i[th] cluster and common lecturers of those clusters, we could observe that which clusters have maximum satisfaction degree or minimum violation, so at first that cluster would have the priority of allocation and after reaching for instance to i[th] cluster, now we must look for those common lecturer within i[th] cluster whose satisfaction percent is the highest or they have minimum violation percent over his/her features and requirements and this is done upon (18).

We could obtain the loss percent of additional resources among departments after clustering and mapping process per department based on (20).

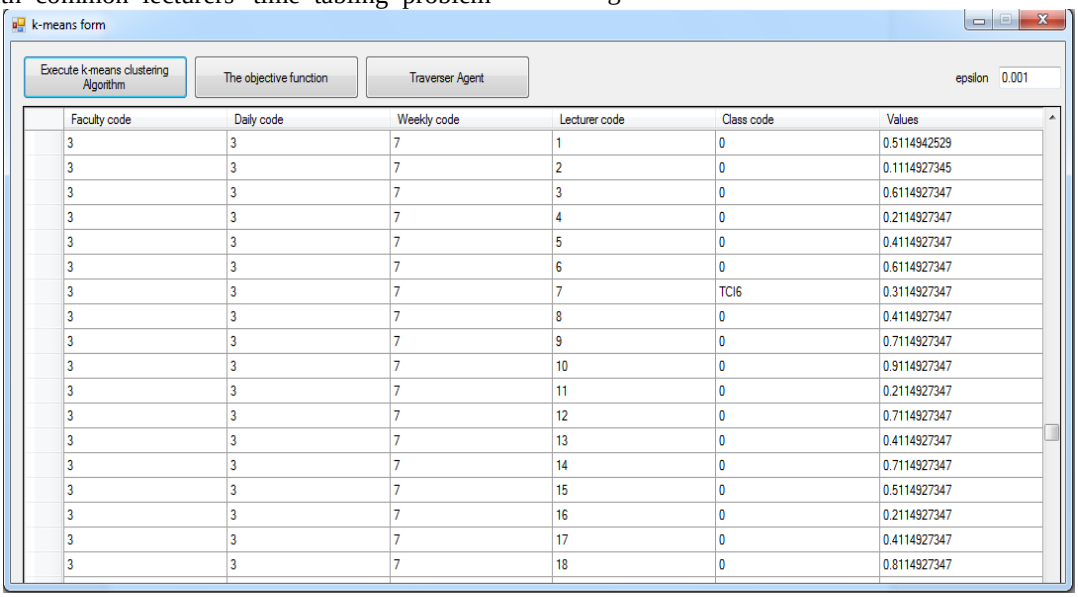
$$ERW_A = \frac{a}{b} \times 100 \quad (20)$$

In (20), $ERW_A : \left(Extra\,\mathrm{Re}\,sources Waste\right)_{After}$ means the loss of additional resources after clustering and mapping processes. Here, a represents the number of the remained additional resources of each department after allocation and b corresponds to the total number of existing resources at each department. To realize ERW_A equation, each

department must apply its resources' allocation process to each common lecturer selectively (from the common lecturer himself/herself) and mandatory (from each department). The remained additional resources among departments equals to the subtraction of total number of departments' resources to the allocated resources by common lecturers and trainings of each department.

4.1 The performance of k-means clustering algorithm over dataset

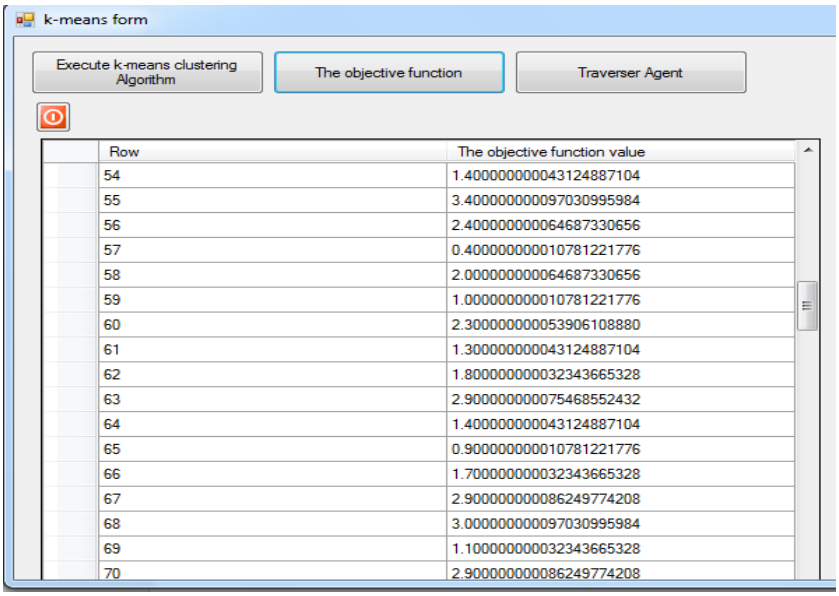
Based on the dataset presented in the first part of section 4, now we could test the k-means clustering algorithm on them. In Fig. 7, the k-means clustering algorithm based on the descriptions in sections 3.3.1 and following the sequence in applying the relations on k-means clustering algorithm compatible with common lecturers' time tabling problem have been shown. In Fig. 7, buttons Execute k-Means Algorithm, Traverser Agent and epsilon represent the performing of compatible k-means algorithm, the traversing agent and the value of parameter $\varepsilon = 0.001$, respectively. The six columns in Fig. 7, each one from left to right represent the faculty code, daily timeslot code, weekly timeslot code, common lecturer code, classroom code (theory-practical) and the computed values after pressing button Execute k-Means Algorithm with considering the value of epsilon=0.001. Button Traverser Agent in section 4.4 would present the way of traversing additional resources of faculties accompanied with mapping common lecturers to those additional resources. The column 6 which is the computed value is obtained after applying the rules of section 3.3.1 on the compatible k-means clustering algorithm.



Faculty code	Daily code	Weekly code	Lecturer code	Class code	Values
3	3	7	1	0	0.5114942529
3	3	7	2	0	0.1114927345
3	3	7	3	0	0.6114927347
3	3	7	4	0	0.2114927347
3	3	7	5	0	0.4114927347
3	3	7	6	0	0.6114927347
3	3	7	7	TCI6	0.3114927347
3	3	7	8	0	0.4114927347
3	3	7	9	0	0.7114927347
3	3	7	10	0	0.9114927347
3	3	7	11	0	0.2114927347
3	3	7	12	0	0.7114927347
3	3	7	13	0	0.4114927347
3	3	7	14	0	0.7114927347
3	3	7	15	0	0.5114927347
3	3	7	16	0	0.2114927347
3	3	7	17	0	0.4114927347
3	3	7	18	0	0.8114927347

Fig. 7: The result of applying k-means clustering algorithm

In Fig. 8, objective function in k-means clustering algorithm computed.



Row	The objective function value
54	1.400000000043124887104
55	3.400000000097030995984
56	2.400000000064687330656
57	0.400000000010781221776
58	2.000000000064687330656
59	1.000000000010781221776
60	2.300000000053906108880
61	1.300000000043124887104
62	1.800000000032343665328
63	2.900000000075468552432
64	1.400000000043124887104
65	0.900000000010781221776
66	1.700000000032343665328
67	2.900000000086249774208
68	3.000000000097030995984
69	1.100000000032343665328
70	2.900000000086249774208

Fig. 8: The result of objective function computed in k-means clustering algorithm

4.2 Comparison of adopted fuzzy c-means and proposed funnel-shape clustering algorithm with the k-means clustering algorithm adopted

In this section, we have shown the process of comparing k-means, fuzzy c-means and funnel-shape clustering algorithms in figures 9, 10 and 11, respectively and also provided a brief comparison as distinct for each faculty based on each 3 clustering algorithms in Fig. 12. In Fig. 12, we have shown a final pie chart in terms of common lecturers satisfaction percent based on each clustering algorithm.

In Fig. 9, the comparison result of k-means algorithm's satisfaction is shown for each 25 common lecturers among faculties as 3D (three dimensional). In this Fig., three length, width and height dimensions represent the common lecturer's code, the faculty code and the satisfaction percent of common lecturers, respectively.

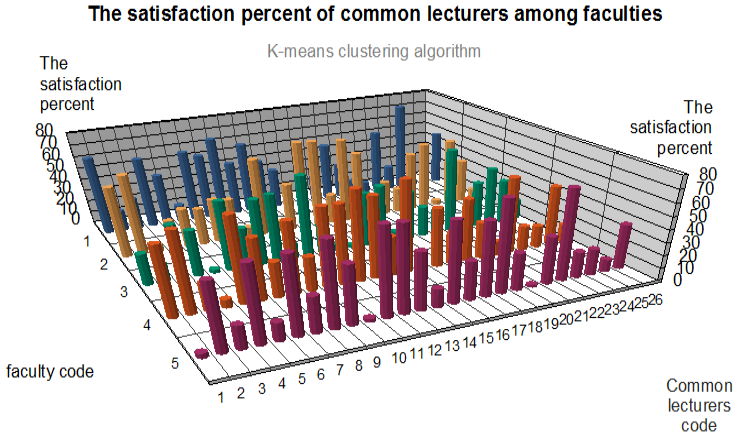


Fig. 9: The satisfaction percent of common lecturers based on k-means algorithm.

In Fig. 10, the comparison result of fuzzy c-means algorithm's satisfaction of each 25 common lecturers among faculties is shown as 3D (three dimensional). In this figure, three length, width and height dimensions represent the common lecturers' code, the faculty code and the satisfaction percent of common lecturers, respectively.

In Fig. 11, the comparison result of funnel-shape algorithm's satisfaction of each 25 common lecturers among faculties is shown as 3D (three dimensional). In this figure, three length, width and height dimensions represent the common lecturers' code, the faculty code and the satisfaction percent of common lecturers, respectively.

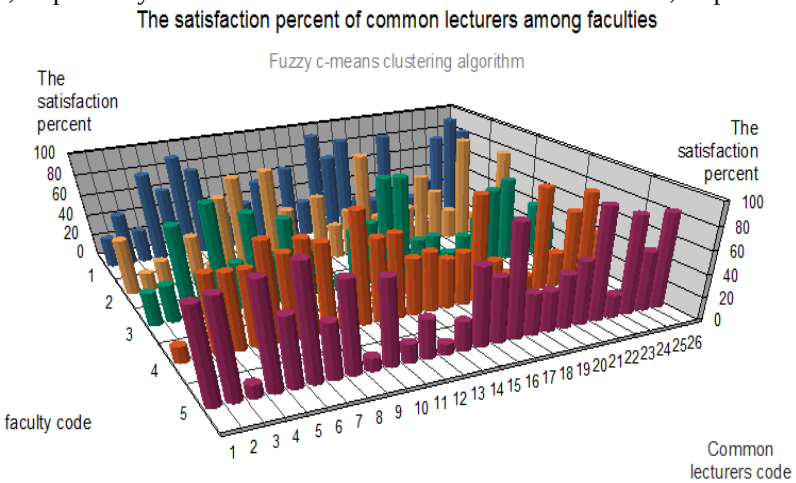


Fig. 10: The satisfaction percent of common lecturers based on fuzzy c-means algorithm

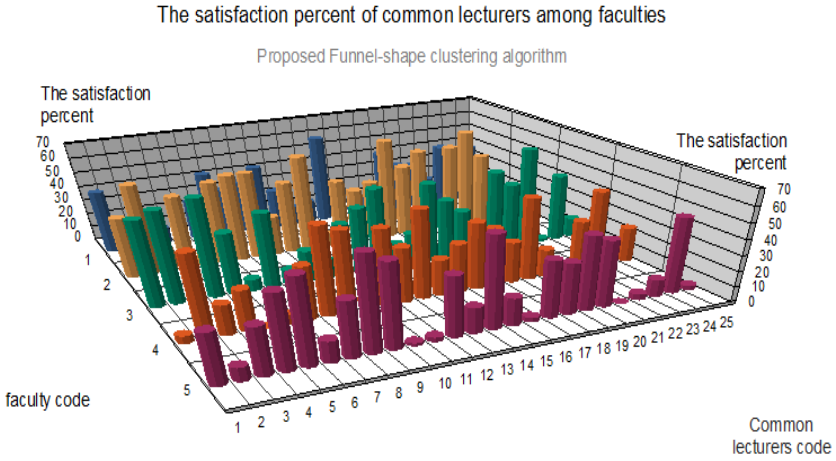


Fig. 11: The satisfaction percent of common lecturers based on funnel-shape algorithm

In Fig. 12, the minimum and maximum satisfaction percent of common lecturers among faculties have been shown for 5 faculties, 25 common lecturers corresponding to the dataset and 3 clustering algorithms. In the first five figures, the satisfaction percent of common lecturers is shown based on each clustering algorithm per faculty and finally the pie chart in the Fig. 12 shows the summary of satisfaction percent of common lecturers among faculties separately and in terms of clustering algorithms. The satisfaction percent of k-means, fuzzy c-means and funnel-shape clustering algorithms are as 28.19%, 38.6% and 33.2 %.

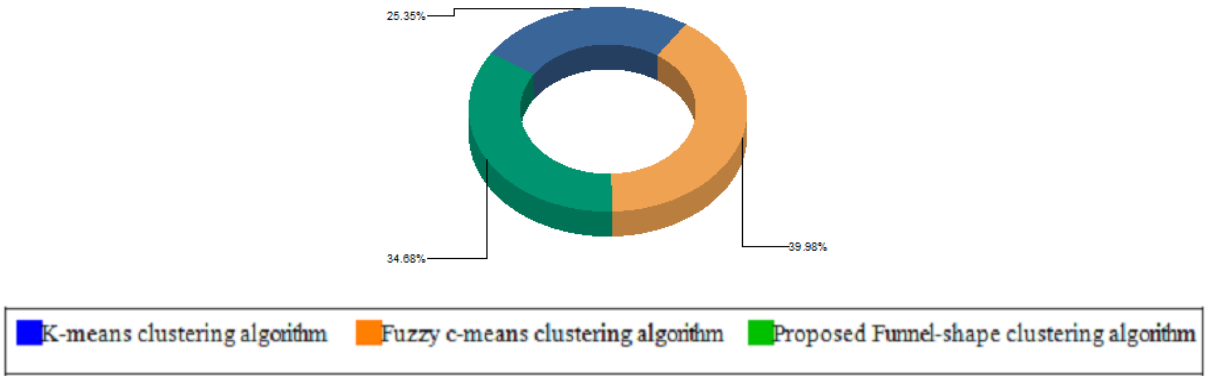


Fig. 12: The descending satisfaction percent of priorities of common lecturers among departments based on clustering algorithms

4.3 Traversing (additional) resources among departments and mapping the clusters of common lecturers by k-means clustering algorithm

Fig. 13 show the way of mapping the clusters of common lecturers to the additional resources of each 5 faculties by using k-means clustering algorithm.
In this shape, by clicking the button of deleting the allocated resources, all previously allocated resources per faculty are removed and by selecting the button of allocating the additional resources to the common lecturers, the act of emptying the stack of common lecturers' clusters list is done to map to the additional resources among faculties.
Since the assumptions related to the constraints and resources have been considered constant per faculty, so the allocation is done based on two selections where one is from the education (the related group) of each faculty and the other one is from the common lecturers among faculties.

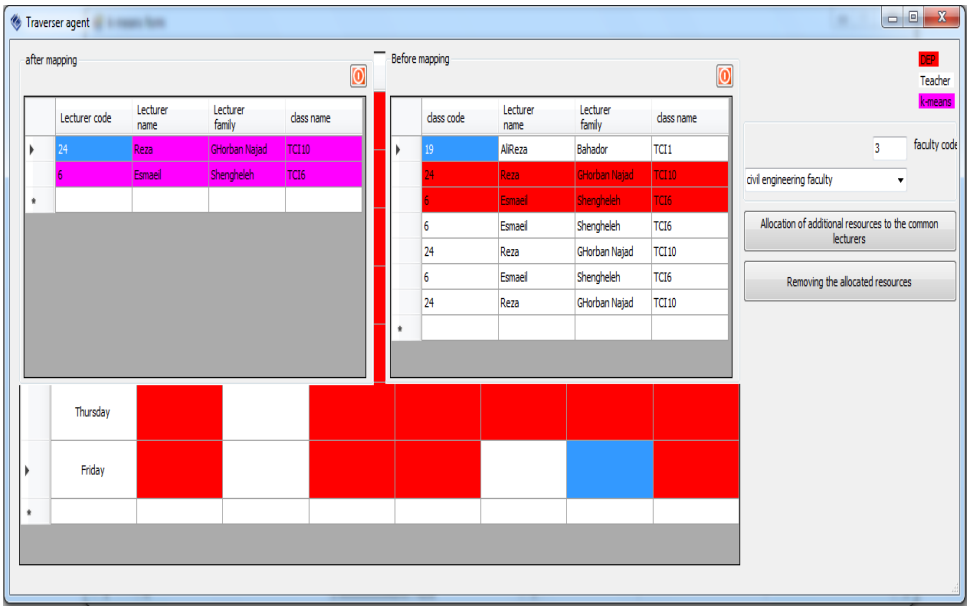


Fig. 13: Mapping the clusters of common lecturers in additional resources among departments with k-means algorithm

In Fig. 13, the red color shows the education (group) selections of each faculty, the white color represents the selections of each common lecturer, in Fig. 13 the purple color show the allocations of selections of each common lecturer to their constraints and priorities in k-means clustering algorithm after mapping process. Fig. 14 shows the additional resources loss percent per five faculties corresponding to each clustering algorithm. Table 1 shows the overall result of each three algorithms based on three clustering algorithms. However, here we could say that the first goal is to minimize the loss of additional resources of faculties for clustering algorithms from the maximum to minimum fuzzy c-means clustering (41.288%), funnel shape clustering (the proposed funnel) (32.55%) and k-means clustering (26.16%) and the second goal is to satisfy the priorities of common lecturers among faculties in a descending manner where for clustering algorithms from the maximum to minimum as fuzzy c-means clustering (38.6%), the proposed funnel clustering (33.2%) and k-means clustering (28.1%).

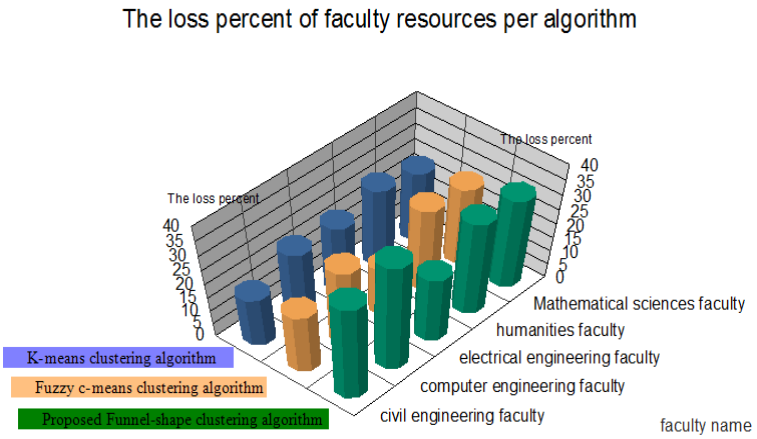


Fig. 14: Minimizing the loss of additional resources among departments through clustering algorithms

Table 1: Comparison of clustering algorithms based on research goals

Research goals	***Standard clustering***		***The proposed clustering***
	Fuzzy c- means clustering	***k-means clustering***	***The proposed funnel-shape clustering***
Loss minimization Faculties additional resources	41.288%	26.16%	32.55%
Descending satisfaction of common lecturers priorities	38.6%	28.1%	33.2%

5. Discussion

In this section, effects of the proposed method's advantages and disadvantages are discussed.

5.1 Disadvantages

1- Variability of lecturers' constraints and priorities in department where in the real context, it is not possible to satisfy all the requirements and priorities of involved events in a desirable extent and for this purpose a descending satisfaction is considered.
2- Limitation of appropriate and desirable resources in system to perform lecturers' timetabling process and traversing resources.
3- Not applying meta-heuristic and hybrid methods which leads to relative loss of efficiency of proposed algorithm in generating tables with primary timetabling ability within the existing agents in the system.

5.2 Advantages

1- Considering the priorities of lecturers specifically and their constraints in order to uniformly distribution over available resources.
2- In timetabling lecturers, most of their clear features are employed sufficiently.
3- Applying multi agent system based method to increase the autonomy of each department's timetabling where this autonomy prevents unplanned collisions and allocations among agents within distributed environment.

6. Conclusion

In this article, the obtained results from the CLTTP problem's purposes through the proposed approach include: 1- the proposed method results in a descending satisfaction from the priorities (soft constraints) of common lecturers among departments to allocate additional resources and 2- the loss of additional resources (unused) at each minimized department which represents the allocation of common lecturers to resources with an improving process. The future approach to solve UCTTP problem would be to work on multi agent based methods as a distributed architecture and apply modern syntactic and fuzzy meta-heuristic approaches where for example we can use meta-heuristic algorithms for two agents TA_i and MA in order to increase throughput in generating and improving time tables. In this problem we can use fuzzy c-means clustering algorithm by applying features weight learning (soft constraints of common lecturers) in generating more improved time tables based on common lecturers among departments where this algorithm could be executed after performing the process of mapping function f and transferring time tables to each agent (department). It must be noted that this method could be used to generate improved time tables in the first phase for each department locally. However, various types of events and resources' features within CLTTP problem could be considered in different kinds of clustering methods and various mapping methods could be used in such clustering approaches.

References

[1] Babaei, H., Karimpour, J., and Hadidi, A., "A survey of approaches for university course timetabling problem," Computers & Industrial Engineering 86, 2015, pp. 43–59.
[2] Feizi-Derakhshi, M. R., Babaei, H., and Heidarzadeh, J., "A Survey of Approaches for University Course TimeTabling Problem," Proceedings of 8th International Symposium on Intelligent and Manufacturing Systems, Sakarya University Department of Industrial Engineering, Adrasan, Antalya, Turkey, 2012, pp. 307-321.
[3] Obit, J. H., Developing Novel Meta-heuristic, Hyper-heuristic and Cooperative Search for Course Timetabling Problems, Ph.D. Thesis, School of Computer Science University of Nottingham, Nottingham, UK, 2010.
[4] Gotlib, C. C., "The Construction of Class-Teacher TimeTables,", Proc IFIP Congress, Vol. 62, 1963, pp. 73-77.
[5] Asmuni, H., Fuzzy Methodologies for Automated University Timetabling Solution Construction and Evaluation, Ph.D. Thesis, School of Computer Science University of Nottingham, Nottingham, UK, 2008.
[6] Lewis, M. R., Metaheuristics for University Course Timetabling, Ph.D. Thesis, Napier University, Napier, UK, 2006.
[7] Redl, T. A., A Study of University Timetabling that Blends Graph Coloring with the Satisfaction of Various Essential and Preferential Conditions, Ph.D. Thesis, Rice University, Houston, Texas, USA, 2004.
[8] Srinivasan, S., Singh, J., and Kumar, V., "Multi-Agent based Decision Support System Using Data Mining and Case Based Reasoning," IJCSI International Journal of Computer Science Issues, 2011, Vol. 8, Issue 4, No 2.
[9] Obit, J. H., Landa-Silva, D., Ouelhadj, D., Khan Vun, T., and Alfred, R., "Designing a Multi-Agent Approach System for Distributed Course TimeTabling," 2011, IEEE.
[10] Wangmaeteekul, P., Using Distributed Agents to Create University Course TimeTables Addressing Essential Desirable Constraints and Fair Allocation of Resources, Ph.D. Thesis, School of Engineering & Computing Sciences Durham University, Durham, UK, 2011.
[11] Shatnawi, S., Al -Rababah, K., and Bani-Ismail, B., "Applying a Novel Clustering Technique Based on FP- Tree to University Timetabling Problem: A Case Study," 2010, IEEE.
[12] DeWerra, D., "An Introduction to TimeTabling," European Journal of Operational Research 19, 1985, pp. 151-162.
[13] Asham, G.M., Soliman, M.M. and Ramadan, A.R., "Trans Genetic Coloring Approach for Timetabling Problem," Artificial Intelligence Techniques Novel Approaches & Practical Applications, IJCA, 2011, pp. 17-25.
[14] Daskalaki, S., Birbas, T., and Housos, E., "An integer programming formulation for a case study in university timetabling," European Journal of Operational Research 153, 2004, pp. 117–135.
[15] Daskalaki, S., and Birbas, T., "Efficient solutions for a university timetabling problem through integer programming," European Journal of Operational Research 160, 2005, pp. 106–120.
[16] Khonggamnerd, P., and Innet, S., "On Improvement of Effectiveness in Automatic University Timetabling Arrangement with Applied Genetic Algorithm," 2009, IEEE.
[17] Alsmadi, O. MK., Abo-Hammour, Z. S., Abu-Al-Nadi, D. I., and Algsoon, A. , "A Novel Genetic Algorithm Technique for Solving University Course Timetabling Problems," 2011, IEEE.
[18] Mayer, A., Nothegger, C., Chwatal, A., and Raidl, G., "Solving the Post Enrolment Course Timetabling Problem by Ant Colony Optimization," In Proceedings of the 7th International Conference on the Practice and Theory of Automated Timetabling, 2008.
[19] Ayob, M., and Jaradat, G., "Hybrid Ant Colony Systems For Course Timetabling Problems," 2nd Conference on Data Mining and Optimization 27-28 October 2009, Selangor, Malaysia, IEEE, 2009, pp. 120-126.

[20] Jat N. S., and Shengxiang, Y., "A Memetic Algorithm for the University Course Timetabling Problem," 20th IEEE International Conference on Tools with Artificial Intelligence, IEEE, 2008, pp. 427-433.
[21] Alvarez, R., Crespo, E., and Tamarit, J. M., "Design and Implementation of a Course Scheduling System Using Tabu Search," European Journal of Operational Research 137, 2002, pp. 512-523.
[22] Aladag, C. H., Hocaoglu, G., and Basaran, A. M., "The effect of neighborhood structures on tabu search algorithm in solving course timetabling problem," Expert Systems with Application 36, 2009, pp. 12349–12356.
[23] Tuga, M. Berretta, R. and Mendes, A., "A Hybrid Simulated Annealing with Kempe Chain Neighborhood for the university Timetabling Problem," 6th IEEE/ACIS International Conference on Computer and Information Science, 2007, ICIS.
[24] Shengxiang, Y., and Jat, N.S., "Genetic Algorithms with Guided and Local Search Strategies for University Course Timetabling," IEEE Transactions on Systems, MAN, and Cybernetics-PART C: Applications and Reviews, 2011, Vol. 41, No. 1.
[25] Abdullah, S., Burke, E.K., and McColloum, B., "An Investigation of Variable Neighborhood Search for University Course Timetabling," In The 2th Multidisciplinary Conference on Scheduling: Theory and Applications, NY, 2005, pp. 413-427.
[26] Kostuch, P., "The University Course Timetabling Problem with a Three-Phase Approach," In Lecture Notes in Computer science, pages 109-125, 2005, Springer-Berlin / Heidelberg.
[27] Shahvali Kohshori, M., and Saniee Abadeh, M., "Hybrid Genetic Algorithms for University Course Timetabling," IJCSI International Journal of Computer Science Issues, 2012, Vol. 9, Issue 2, No 2.
[28] Asmuni, H., Burke, E.K., and Garibaldi, J.M., "Fuzzy multiple heuristic ordering for course timetabling," The Proceedings of the 5th United Kingdom Workshop on Computational Intelligence (UKCI05), 2005b, pp 302-309.
[29] Golabpour, A., Mozdorani Shirazi, H., Farahi, A. kootiani, M., and H. beige, "A fuzzy solution based on Memetic algorithms for timetabling," International Conference on MultiMedia and Information Technology, IEEE, 2008, pp. 108-110.
[30] Chaudhuri, A., and Kajal, D., "Fuzzy Genetic Heuristic for University Course Timetable Problem," Int. J. Advance. Soft Comput. Appl., 2010, Vol. 2, No. 1, ISSN 2074-8523.

Hamed Babaei received the B.Sc. and M.Sc. degrees in computer engineering from Islamic Azad University (IAU), Iran, in 2007 and 2011, respectively. He is currently teaching in Department of Computer Engineering, Islamic Azad University of Ahar, Ahar, Iran. His current research interests include evolutionary algorithms, Meta-heuristic approaches, Distributed data base, Distributed systems, and their applications for timetabling problems.

Jaber Karimpour received the B.Sc, M.Sc. and Ph.D. degrees in computer systems from university of tabriz, Tabriz, Iran, in 1997, 2000 and 2009, respectively. He is currently teaching in Department of Computer Sciences, University of tabriz, Tabriz, Iran. His current research interests include theoretical Computer Science, Software Component Technology; Software and System specification, design and verification; Applied Mathematics in Software Engineering (Formal Methods); Temporal Logic; Modal Logic; Object Oriented Programming.

Sajjad Mavizi received the B.Sc. and M.Sc. degrees in computer engineering from Islamic Azad University (IAU), Iran, in 2008 and 2012, respectively. His current research interests include evolutionary algorithms, Meta-heuristic approaches, mobile ad hoc networks, vehicular ad hoc networks.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

New Media Display Technology and Exhibition Experience

A Case Study from the National Palace Museum

Chen-Wo Kuo[1], Chih-ta Lin[2] and Michelle C. Wang[3]

[1]National Palace Museum
Taipei City, 11143, Taiwan
paulkuo@npm.gov.tw

[2]National Taiwan University of Arts
New Taipei City, 22058, Taiwan
richarlin@gmail.com

[3] National Palace Museum
Taipei City, 11143, Taiwan
mwang@npm.gov.tw

Abstract

As the inheritor of Chinese civilization, the National Palace Museum (hereafter referred to as the NPM), houses a world-class collection of cultural art and artifacts. Since the NPM began promoting the National Digital Archives Project in 2002, its efforts have expanded to develop a digital museum and various e-learning programs. Extending the use of digital archives to its educational and cultural industrial endeavors, the NPM has maximized the value of its exhibitions, publications, and educational programs. In 2013, the NPM integrated creative thinking and interdisciplinary technologies, such as floating projection, augmented reality, and other sensory interactive media, to recreate the historical circumstance of 19th century East Asian maritime cultures in "Rebuilding the Tong-an Ships—New Media Art Exhibition," which opened at Huashan 1914 Creative Park and later won the Gold Award at the 2014 Digital Education Innovation Competition. Through a thorough exploration of the factors contributing to the success of "Rebuilding the Tong-an Ships," this study has isolated the two main factors of the exhibition's popularity, namely, the compactness of the metadata and the atmosphere created by the interactive display technology.

Keywords: *National Palace Museum; New Media Arts; Interactive; Exhibition*

1. Background and Motivation

1.1 Museum Function in the Age of Information

Exhibition is a very important part of museum operations. After all, all museum exhibitions are made for the public's education and the transmission of human knowledge. Mirroring our time and reflecting our current aesthetic tastes, exhibition is an altruistic cause done for public welfare. [1]

As times evolve, museum operation patterns have also reached a point of transition. A museum's value and function are both expanding. In recent years, the NPM has been promoting a "museum without walls" project by which it hopes to allow those who cannot visit the museum in person to become familiar with the collection. The concept of a "museum without walls" is enabled by the proliferation of photography and printing, which allow for the mass replication of collection images and change people's habits of appreciating art. In the present age, digitization, the internet, and multimedia technologies add to the trend of limitless exchange and transmission. Items in a museum's collection are no longer trapped beneath the dust in a museum's physical archives. From still life displays to dynamic interactive media, exhibition materials seek to shake up art and artifacts with new technology and new media. The digital archive is believed to be the foundation on which the NPM will develop into a world-class museum. The far-reaching effects of the internet will allow the essence of Chinese civilization to spread across the world and attract more visitors to the museum.

1.2 NPM Digital Archive Applications and Results

The NPM's digital archive started officially in 2001. In the following year, the NPM took part in the National Digital Archives Project. With as many as 650,000 items in its collection, preservation and management was extremely tasking prior to digitization. Thanks to the development of technology, the NPM now has cutting-edge methods to organize its collection. Seven departments--Antiquities, Painting and Calligraphy, Rare Books and Historical Documents, Publishing, Registration, Technology, and Information Center--initiated the NPM's Digital Archive Plan in 2001 with the aim of establishing a complete collection database for public appreciation and use.

The brunt of the plan lies in digitizing all the art and artifacts in the collection, a process which begins with

photographing, scanning, color correcting, watermarking, printing, composing item descriptions and extends to developing metadata, knowledge base, and applications. In other words, photographing or scanning art and artifacts to create digital files enables more efficient organization, research, and application. Digital files can be copied or edited to create useful backup images that will not suffer the ravages of time and can be easily preserved.

Digitizing the collection offers advantages not only in assisting the museum's exhibition planning, publishing, archival, educational, and research efforts but also in opening up the treasures of Chinese civilization to people all over the world. Image licensing services could also be offered in the future, allowing the digitized materials to achieve the goals of educational promotion, to stimulate the development of value-added applications and creative products, and to contribute to the knowledge based economy. [2]

After several years of implementing the 2002 National Digital Archives Project, the NPM advanced to establish a digital museum and other digital learning programs. Applying its digital archives to educational and cultural industrial endeavors, the NPM has not only maximized the value of its exhibitions, publications, and educational programs but also catalyzed the government's endeavors in the cultural and creative industries by injecting a new wave of inspiration and resources.

1.3 New Media Interactive Displays and the Museum Experience

Explorations of relevant fields in new media art, from early creative media discussions to recent speculative studies of aesthetics and museum industry archival mechanisms, have seen a considerable growth in the amount of data accumulated, demonstrating that this field is indeed developing vigorously. Concurrently, the government proposed the Digital Art Creation Project when implementing the 2008 National Development Challenge. This project gave rise to the international exhibition "Wanderer," curated by Junjie Wang and "Pleasure," curated by Shuming Lin for the 2005 Ars Electronica 25 Year Anniversary. Both exhibitions were held at the National Museum of Fine Arts and soon after ignited a wave in new media art. As a result, new media art became the mainstream in the Taiwanese art scene. [3]

In Pine II and Gilmore's concept of the experience economy (1998), they noted that enterprises should be thinking of how to enhance service to create a more attractive consumer experience. Pine II and Gilmore proposed to consider experience from two levels: (a) degree of participation--active versus passive--and (b) the relationship between experience and environment, which can be subdivided into two types: absorption, the method of manipulating experience to attract consumer attention, versus immersion, the method of including audience as part of the physical experience. There are four types of experiences: (1) Entertainment; (2) Education; (3) Escapist; and (4) Esthetic. So, applying Pine II and Gilmore concept of the experience economy, in order for museums to truly achieve long term competitive advantage, museum curators need to create an immersive, stimulating, and memorable experience.

In the past, exhibitions mostly relied on textual, auditory, and especially visual content to explain art and artifacts; therefore, audience members are traditionally the most familiar with looking. But, as the direction of museum exhibition and educational ideals shift, more and more emphasis is being placed upon audience participation, audience immersion, guided observation and reflection, and development of participant insight. Taking into account the ever-changing nature of the technology used in exhibitions, how then do we achieve an exhibition's original intention and faithfully apply new media art to exhibition content? We must transform the NPM's collection into a wellspring for the museum's creative industry by generating creative content from the research derived from the collection. Adhering to the core concept of using science and technology as a platform for marketing NPM's world-class collection to the world, the exhibition combines ancient artifacts with new technology to show the fruits of NPM's latest digitization efforts. [4]

Focusing on "Rebuilding the Tong-an Ships—New Media Art Exhibition" as the object of study, this paper explores the extent to which the new media installations interact with and generate feedback from the audience while showcasing the fruits of NPM's digitization overhaul and its application of new media art for the purpose of providing a point of reference for future exhibitions related to history.

2. "Rebuilding the Tong-an Ships" New Media Art Exhibition

2.1 Tong-an Ship History

Tong-an ships refer to the large sized ancient sailing vessels that saw rise during the middle of the Qing Dynasty. The ships were so named because they originated in Fujian Province, in the municipality of Tong-an. These ships were not only widely common for civil use, but they were also immensely popular with the pirates. Since the Age of Discovery ushered in a surge of maritime activities in the East Asian seas, the region, under the combined influences of harsh weather conditions and the demands of the imperial Chinese tributary system, experienced a new maritime era made possible by extensive economic contact. The comings

and goings of trading vessels, adventurers and navigators from various countries created business opportunities and promoted political and economic contact between China and other Asian countries, even with Western European countries. Situated on the western hub of the Pacific Ocean, Taiwan experienced a steady improvement in its political, military, and economic status as well.

Beginning in the early 19th century, increasing maritime activities led the Qing Empire to strengthen its military forces in an effort to preserve order and stability on the seas. But conflicts between the Vietnamese regimes caused the pirates, whose military might ran supreme in those times, to be courted for their prowess. The Tây Sơn dynasty granted the pirates many titles and favors, creating a hotbed for the rise of piracy. After repeated restoration efforts by the Ming loyalists in Taiwan, the Qing Empire kept a vigilant effort along the Southern Coast, building a naval fleet in Taiwan and Penghu to regulate the seas and stabilize relations in South East Asia.

Tong-an ships do not merely act as methods of transportation, but enable connections to form in the region's political, military, economic, and cultural activities. At last, they became the main naval force off coast and are the most representative of ancient Chinese ships before the appearance of steamboats. "Rebuilding the Tong-an Ships--New Media Art Exhibition," organized by the NPM, is the first exhibition with the combined theme of maritime history and Taiwanese history. The exhibition is based on two rare warship diagrams from the Qing Dynasty in the National Palace Museum collection, *Diagram of the Tong-an Ship Ji* and *Diagram of Tong-an Ship No. 1*,which were reckoned to be attachments to a memorandum. However, the original memorandum in question remained long hidden in the archives of NPM until a researcher discovered its appendix out of nearly 200,000 military official records in the digital archive and finally was able to reconstruct the conflict between two major players in Chinese naval history, Li Changgeng and Cai Qian.

2.2 Exhibition Overview

Devoting itself fully to incorporating new media into its exhibitions, the NPM, after planning *NPM Digital* and other special new media exhibitions, cooperated with Huashan 1914 Creative Park and launched *Rebuilding the Tong-an Ships--New Media Art Exhibition* as part of the museum without walls series. The Department of Education, Exhibition, and Information Services and the Department of Rare Books and Historical Documents are the central members of this exhibition's curatorial team. Using interactive installations to fulfill the exhibition's educational goals, The Department of Education, Exhibition, and Information Services developed new forms of display techniques to present the findings of researchers from the Department of Rare Books and Historical Documents.

This exhibition space is organized into three main sections--period background, main characters, and ship structure. Within these sections are seven main new media art pieces: *The "Bon Voyage!" Projection Wall*, *Cross-over dialogue: Holographic Projection*, *Cloud Gallery*, *Deconstructing the Tong-an Ship*, *Looking Through the Tong-an Ship*, *The Augmented Reality Clothes Changing System*, *Breaking Waves: Tongan Ship Interactive Game*, *Nautical Chart Interactive Tabletop*, *Rebuilding the Tong-an Ship: Theater.*

2.3 New Media Art in "Rebuilding the Tong-an Ships"

The exhibition uses holographic projection, naked-eye 3D, augmented reality, kinect sensory interactive devices, and other new media technologies to stimulate the five senses. Each piece reproduces an aspect of the prosperous 19th century East Asian maritime culture to help audiences understand the military, cultural and other historical background information related to the Tongan ship. The metaphorical use of *chao*, meaning tide, in the Chinese exhibition title originated in connection with the ocean tide, emigration tide, and economic tide brought in by the Tong-an ships. First, the actual ocean tide brought unlimited possibilities by opening up China to external contact; immigration brought about abundant manpower, leading to the development of Taiwan; the ensuing economic tide created a rich and diverse maritime culture, from which sprung figures such as the Great Seafarer Cai Qian. Finally, using new media technologies to reinterpret and present the museum collection is a new trend, or tide, in contemporary exhibition. Manifesting the different meanings of "tide," therefore, is one of the exhibition objectives. [5]

The seven installations integrate interactive technology with woodworking equipment, lighting, and fragrances to create an appropriate atmosphere for each piece. The documentary *Rebuilding the Tong-an Ships* provides a supporting overview of the ship's cultural and naval history and some insights into the historical documents left behind.

The themes of the exhibition are the Tong-an ship and the ocean, both of which are closely linked to the main visual design concept. Manifested in the typeface of *Rebuilding the Tong-an Ships* documentary, the overlapping red and blue in *chao* from the Chinese title represent the use of 3D and other new media technology. The blue-green and yellow gradients in the base layer symbolize youth, the ocean, and the ocean tides. The floor covering the exhibition entrance uses light brown tones to mimic a

ship's brown wooden color. Surrounding interactive installations and fences are also created in the same colors as those painted upon the Tong-an ships' hulls.

In the exhibition's audio design, the curatorial team installed sounds of seagulls and ocean sprays at the entrance to create a soothing ocean atmosphere. Music of Chinese string instruments play in the background. The hope is that audio settings can not only help relax the audience but also bring them into the exhibition environment.

Scent, which can have a direct impact on mood, is also an important factor in the creation of experience. This exhibition especially designed a fragrance exclusive to the Tongan ships. Inspired from blue amber, the exhibition fragrance is supposed to symbolize the ship's adventurous, resolute spirit, capable of forging ahead in the face of adverse sea weather. The reason for the emphasis on scent is that scent has the power to strengthen human memory and emotional experience. Olfactory information passes more easily to the limbic system in the brain and can be registered more firmly in human memory. Scent can facilitate the diffusion of information and allow audience members to have a more varied experience. Exhibition facilities use wood for the deck and side gate fixtures, improving tactile sensations. Combined with the virtual interaction system, these construction material choices enable audiences to have an entirely new tactile experience when visiting "Rebuilding the Tong-an Ships--New Media Art Exhibition."

3. Visitor Experience

Held in the Boiler Room at Huashan 1914 Creative Park, the exhibition ran from July 20 to September 22, 2013. The Park closes on Mondays. Weekends and holidays crowds peak (Table 1). Total attendance reached 54,067 visitors. The highest percentage visiting group falls in the 20-29 age range (Table 2), with more female than male visitors (Table 3); sex and age factor little into overall satisfaction, but familiarity with electronic equipment has a significant difference on visitor response (Table 4), indicating that this exhibition is immensely suitable for the internet generation and corresponds to curatorial objectives.

The overall satisfaction rating for this exhibition is 99.4% (Table 10). The most popular exhibition piece is *Breaking Waves: Tongan Ship Interactive Game*, with 35.1% average satisfaction rating (Table 11); The satisfaction rating for the hardware equipment is 87.9% （Table 12). A few visitors complained of devices freezing or that there were too many visitors and insufficient time to experiment with the devices.

Out of all the exhibition pieces, the *Cloud Gallery* is the extension of the National Palace Museum's Cloud ICT Platform Development Project. Since the execution of the Digital Archive Plan, the NPM gathered an enormous quantity of digital content not only for the purpose of developing its creative cultural industry but also for the integration and application of more effective business models. Particularly, now with the substantial improvements in screen resolution and network bandwidth, cloud galleries are bound to become the new mainstream. Due to budget constraints, certain exhibitions offer only a uni-directional presentation; however, starting in 2012, in "Four Seasons of the NPM," installations have managed to fulfill interactive goals in their designs. Compared to two-directional, interactive presentations, uni-directional media have a satisfaction rating of only 4.1%. From comparing this data, audience preference for interactive installations is evident.

The satisfaction rating for interactive installations is 95% (Table 13). 89.3% of audiences think that the installations adequately convey the exhibition theme. These results are enough to demonstrate audience recognition of this exhibition's curatorial efforts.

On the scheme of the exhibition, 96.2% (Table 15) agree that using games to help people understand the artifacts is an effective method, but there are some who think that not all artifacts are suited to this method. 89.2% (Table 16) agree that the spatial design creates the right atmosphere for a museum. These numbers demonstrate that, overall, the design of the exhibition space was a success.

On the level of education, audience interest rose from 58.9% to 81.6% after viewing the exhibition, demonstrating that this exhibition achieved its educational goals by a large margin (Table 17 and 18).

To sum up the statistical results, this exhibition takes advantage of the ambiance at Huashan 1914 Creative Park's and new media display methods to attract successfully more than fifty thousand visitors. The highest frequency visiting age group is ages 20-29. There were more female than male visitors. The display methods and new media interactive technologies cater to the younger generation's curiosity for new things. The familiarity of electronic equipment acquired by those living in the Information Age makes the exhibition installations easily accessible. The high satisfaction ratings gathered from the data are the best supporting evidence for proving that the NPM has fulfilled its purpose to create memorable and enjoyable educational experiences.

4. Tables

Table 1: Exhibition visitors by days of the week

		Frequency	*Percentage*	*Effective Percentage*	*Cumulative Percentage*
Efficient	Tuesday	207	12.9	13	13
	Wednesday	198	12.4	12.4	25.3
	Thursday	232	14.5	14.5	39.9
	Friday	284	17.8	17.8	57.6
	Saturday	381	23.8	23.8	81.5
	Sunday	296	18.5	18.5	100
	Total	1598	99.9	100	
Missing Values	System-Missing Values	2	0.1		
Total		1600	100		

Table 2: Exhibition visitors distributed by age

Age		*Frequency*	*Percentage*	*Effective Percentage*	*Cumulative Percentage*
Efficient	< 20	342	21.4	21.4	21.4
	20~29	572	35.8	35.9	57.3
	30~39	288	18	18.1	75.4
	40~49	298	18.6	18.7	94
	>50	95	5.9	6	100
	Total	1595	99.7	100	
Missing Values	System Missing Values	5	0.3		
Total		1600	100		

Table 3: Exhibition visitors distributed by gender

		Frequency	*Percentage*	*Effective Percentage*	*Cumulative Percentage*
Efficient	Female	1100	68.8	68.8	68.8
	Male	498	31.1	31.2	100
	Total	1598	99.9	100	
Missing Values	System Missing Values	2	0.1		
Total		1600	100		

Table 4: The effect of gender on overall satisfaction data

Group Statistics

	Gender	*Frequency*	*Mean*	*Standard Deviation*	*Standard Error of the Mean*
Total	Female	1078	4.1313	0.42458	0.01293
	Male	487	4.1471	0.41111	0.01863

Table 5: The effect of gender on overall satisfaction test

Independent Sample Test

		Equal mean t test				
		t	Degree of Freedom	Significant (two-tailed)	Mean Deviation	Standard Error Difference
Total	Equal Variances Assumed	-0.688	1563	0.492	-0.01579	0.02296
	Equal Variances Not Assumed	-0.696	965.99	0.486	-0.01579	0.02268

Table 6: The effect of age on overall satisfaction data

Descriptive Statistics

Total						
Age	Frequency	Mean	Standard Deviation	Standard Error	95% confidence interval of the mean	
					Min.	Max.
<20	341	4.1666	0.42105	0.0228	4.1218	4.2115
20~29	559	4.1171	0.40708	0.01722	4.0833	4.1509
30~39	279	4.1049	0.46221	0.02767	4.0504	4.1594
40~49	293	4.1666	0.38337	0.0224	4.1225	4.2107
>50	90	4.1465	0.47025	0.04957	4.048	4.245
Total	1562	4.1367	0.42034	0.01064	4.1159	4.1576

Table 7: The effect of age on overall satisfaction test

ANOVA

Total					
	Sum of Sqaures	Degrees of Freedom	Average Sum of Squares	F	Sig.
Between Groups	1.073	4	0.268	1.52	0.194
Within Groups	274.733	1557	0.176		
Total	275.806	1561			

Table 8: The effect of familiarity with electronic equipment on overall satisfaction data

Group Statistics

	Familiarity with Electronic Equipment	Frequency	Mean	Standard Deviation	Standard Error of the Mean
Total	No	310	4.0842	0.42594	0.02419
	Yes	1251	4.1472	0.41672	0.01178

Table 9: The effect of familiarity with electronic equipment on overall satisfaction test

Independent Sample t Test

		Equal mean of t test				
		t	Degrees of Freedom	Significant (two-tailed)	Mean Deviation	Standard Error Difference
Total	Equal Variances Assumed	-2.372	1559	0.018	-0.06299	0.02656
	Equal Variances Not Assumed	-2.341	466.478	0.02	-0.06299	0.02691

Table 10: Overall satisfaction

		Frequency	*Percentage*	*Effective Percentage*	*Cumulative Percentage*
Efficient	Very Satisfied	383	23.9	24.1	24.1
	Satisfied	1045	65.3	65.7	89.8
	Neutral	163	10.2	10.2	100
	Total	1591	99.4	100	
Missing Values	System Missing Values	9	0.6		
Total		1600	100		

Table 11: Favorite exhibition item

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Efficient	The "Bon Voyage" Projection Wall	225	14.1	14.1	14.1
	Cross-over Dialogue:	228	14.3	14.3	28.4
	Cloud Gallery	65	4.1	4.1	32.5
	Deconstructing the Tongan Ship	140	8.8	8.8	41.2
	Looking Through the Tongan Ship	65	4.1	4.1	45.3
	The AR Clothes Changing System	124	7.8	7.8	53.1
	Breaking Waves	561	35.1	35.2	88.2
	The Nautical Chart Interactive Tabletop	51	3.2	3.2	91.4
	Rebuilding the Tongan Ships: Theater	137	8.6	8.6	100
	Total	1596	99.8	100	
Missing Values	System Missing Values	4	0.3		
Total		1600	100		

Table 12: Satisfaction with hardware equipment

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Efficient	Very Satisfied	433	27.1	27.2	27.2
	Satisfied	964	60.3	60.6	87.9
	Neutral	193	12.1	12.1	100
	Total	1590	99.4	100	
Missing Values	System Missing Values	10	0.6		
Total		1600	100		

Table 13: Satisfaction with interactive installation

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Efficient	Very satisfied	413	25.8	25.9	25.9
	Satisfied	1030	64.4	64.6	90.5
	Neutral	146	9.1	9.2	99.6
	Not Satisfied	6	0.4	0.4	100
	Total	1595	99.7	100	
Missing Values	System Missing Values	5	0.3		
Total		1600	100		

Table 14: The interactive Installation adequately conveys the theme of Tong-an ships

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Efficient	Strongly agree	403	25.2	25.2	25.2
	Agree	1024	64	64.1	89.3
	Neutral	164	10.3	10.3	99.6
	Disagree	7	0.4	0.4	100
	Total	1598	99.9	100	
Missing Values	System Missing Values	2	0.1		
Total		1600	100		

Table 15: The design of the interactive game allows for a understanding of the art work

		Frequency	Percentage	Effective Percentage	Cumulative
Efficient	Strongly Agree	742	46.4	46.4	46.4
	Agree	795	49.7	49.7	96.2
	Neutral	58	3.6	3.6	99.8
	Disagree	3	0.2	0.2	100
	Total	1598	99.9	100	
Missing Values	System Missing Values	2	0.1		
Total		1600	100		

Table 16: The spatial design has the ambiance of a cultural exhibition from a museum

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Strongly Agree	407	25.4	25.5	25.5	
Agree	918	57.4	57.4	82.9	
Neutral	242	15.1	15.1	98.1	
Disagree	31	1.9	1.9	100	
Total	1598	99.9	100		
System Missing Values	2	0.1			
Total		1600	100		

Table 17: Interest towards Tong-an Ships prior to visiting the exhibition

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Efficient	Strongly Agree	248	15.5	15.5	15.5
	Agree	693	43.3	43.4	58.9
	Neutral	522	32.6	32.7	91.6
	Disagree	118	7.4	7.4	98.9
	Strongly Disagree	17	1.1	1.1	100
	Total	1598	99.9	100	
Missing Values	System Missing Values	2	0.1		
Total		1600	100		

Table 18: Interest towards Tong-an Ships after visiting the exhibition

		Frequency	Percentage	Effective Percentage	Cumulative Percentage
Efficient	Strongly Agree	291	18.2	18.2	18.2
	Agree	1013	63.3	63.4	81.6
	Neutral	266	16.6	16.6	98.2
	Disagree	28	1.8	1.8	100
	Total	1598	99.9	100	
Missing Values	System Missing Values	2	0.1		
Total		1600	100		

5. Conclusions

Due to the lack of visual supporting materials, previous exhibitions based on documents had trouble relaying to the audience their historical significance, making the task of educational dissemination difficult. By bringing ancient artifacts to life with new technology and conveying knowledge of their historical background through new media, museums can maximize their educational potential. Documents, memoranda, and other historical artifacts related to the Tong-an ships can all be reproduced in the physical or virtual space of the exhibition area.

"Rebuilding the Tong-an Ships--New Media Art Exhibition," the NPM's first digital interactive art exhibition with the combined theme of maritime history and Taiwanese history, is seminal to the fields of historical education, art and artifact research, and digital display technology. The exhibition focuses on the Tong-an ship theme as the main historical axis, from which the story of a period unfurls in simple terms. As technology advances, cultural art and artifacts, after undergoing digitization, mediatization, and interactivity in accordance with national technological advancement plans for a digital archive, have utterly transformed audiences from passive recipients to active participants. This new aesthetic experience creates deeper feelings and memories, shatters the limitations of traditional display methods, and utilizes the newest technology. In so creating this new aesthetic, we can say that the NPM as the inheritor of Chinese civilization has fully demonstrated its value as the key player of Asian art and culture and fulfilled its function of knowledge transmission.

Having provided an overview and visitor experience analysis of the exhibition, it is evident that this exhibition has much creative potential in the field of new media displays. The exhibition communicated to audiences the possibilities of future exhibitions and established the creation of memorable experiences as the new museum focus.

References

[1] Guang-nan Huang, "Museum Exhibition Conception and Planning", National Taiwan University of Arts Department of Painting and Calligraphy Thesis, 2006.
[2] National Palace Museum, Digital Archive Summary. Retrieved on 12/03/2014 from http://www.npm.gov.tw/da/ch-htm/about.html
[3] Taiwan Contemporary Art Archive, New Media Art: retrieved on 12/02/2014 from http://goo.gl/gcXMZD
[4] National Palace Museum, Rebuilding the Tong-an Ships New Media Art Exhibition: Retrieved on 12/03/2014 from http://theme.npm.edu.tw/exh102/tongan_ships/ch/ch00.html
[5] Yiru Tsai, "Rebuilding the Tong-an ships New Media Art Exhibition," The National Palace Museum Monthly of Chinese Art (366) (2013) p. 4.

Chen-Wo Kuo is currently in charge of information in museum education at the National Palace Museum, Taipei, Taiwan. He is a section chief in the Department of Education, Exhibition and Information Services. He has two Master's degrees. One is MBA (master of business administrative) and the other is CIS (Computer Information System). Besides that, he got his Ph.D. degree in the Dept. of information science at National Chengchi University. His major studies are related to e-government policy-making, digital archives, e-learning, digital divide and culture creative industry.

Chih-ta Lin is currently working as the art director with Geometry Global of Ogilvy Group, where he is in charge of creative marketing. He has a master's degree from the Graduate Institute of Graphic Communication Arts from the National Taiwan University of Arts.

Michelle C. Wang is a research assistant working in the Department of Education, Exhibition, and Information Services at the National Palace Museum. She graduated with a B.A. in English from Wellesley College. Her interests are English and Chinese literature, literary translation, intercultural communication, and museum education.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Real-Time Color Coded Object Detection Using a Modular Computer Vision Library

Alina Trifan[1], António J. R. Neves[2], Bernardo Cunha[3] and José Luís Azevedo[4]

[1] IEETA/ DETI, University of Aveiro
Aveiro, Portugal
alina.trifan@ua.pt

[2] IEETA/ DETI, University of Aveiro
Aveiro, Portugal
an@ua.pt

[3] IEETA/ DETI, University of Aveiro
Aveiro, Portugal
mbc@det.ua.pt

[4] IEETA/ DETI, University of Aveiro
Aveiro, Portugal
jla@ua.pt

Abstract

In this paper we present a novel computer vision library called UAVision that provides support for different digital cameras technologies, from image acquisition to camera calibration, and all the necessary software for implementing an artificial vision system for the detection of color-coded objects. The algorithms behind the object detection focus on maintaining a low processing time, thus the library is suited for real-world real-time applications. The library also contains a TCP Communications Module, with broad interest in robotic applications where the robots are performing remotely from a basestation or from an user and there is the need to access the images acquired by the robot, both for processing or debug purposes. Practical results from the implementation of the same software pipeline using different cameras as part of different types of vision systems are presented. The vision system software pipeline that we present is designed to cope with application dependent time constraints. The experimental results show that using the UAVision library it is possible to use digital cameras at frame rates up to 50 frames per second when working with images of size up to 1 megapixel. Moreover, we present experimental results to show the effect of the frame rate in the delay between the perception of the world and the action of an autonomous robot, as well as the use of raw data from the camera sensor and the implications of this in terms of the referred delay.

Keywords: *Image Processing; object detection; real-time processing; color processing*

1. Introduction

Even though digital cameras are quite inexpensive nowadays, thus making artificial vision an affordable and popular sensor in many applications, the research done in this field has still many challenges to overcome. The main challenge when developing an artificial vision system is processing all the information acquired by the sensors within the limit of the frame rate and deciding in the smallest possible amount of time which of this information is relevant for the completion of a given task.

In many applications in the areas of Robotics and Automation the environment is still controlled up to a certain extent in order to allow progressive advancements in the different development directions that stand behind these applications. There are many industrial applications in which semi or fully autonomous robots perform repetitive tasks in controlled environments, where the meaningful universe for such a robot is reduced, for example, to a limited set of objects and locations known apriori. In applications such as industrial inspection, traffic sign detection or robotic soccer, among others, the environment is either reduced to a set of objects of interest that are color-coded (or color-labeled) or the color segmentation of the objects of interest is the first step of the object detection procedure. This is mainly due to the fact that segmenting a region based on colors is less heavy from the point of view of the computational resources involved than the detection of objects based on generic features. In what concerns color object detection there is little work done in a structural manner. There can be found in literature published work on color segmentation for object detection and common aspects in the research papers that approached this problem can be traced. However, to our knowledge, there is no free, open source available library that allows the complete implementation of a vision system software for color coded object detection.

In this paper we present a library for color-coded object detection, named UAVision. The library aims at being a complete collection of software for real-time color object detection. UAVision can be split into several independent modules that will be presented in the following sections. The architecture of our software is of the type "plug and play", meaning that it offers support for different digital cameras technologies and the software created using the library is easily exportable and can be shared between different types of cameras. We call the library modular as each module can be used independently as a link in an image processing chain or several modules can be used for creating a complete pipeline of an artificial vision system.

Another important aspect of the UAVision library is that it takes into consideration time constraints. All the algorithms behind this library have been implemented focusing on maintaining the processing time as low as possible and allowing the use of digital cameras at the maximum frame rates that their hardware supports. In Autonomous Mobile Robotics, performing in "real-time" is a demand of almost all applications since the main purpose of robots is to imitate humans and we, humans, have the capacity of analyzing the surrounding world in "real-time". Even though there is not a strict definition of real-time, almost always it refers to the amount of time elapsed between the acquisition of two consecutive frames. Real-time processing means processing the information captured by the digital cameras within the limits of the frame rate. There are many applications in which the events occur at a very high pace, such as, for example, industrial production or inspection lines where robots have to repeatedly and continuously perform the same task. If that task involves visual processing, the vision system of such a robot has to keep up with the speed of the movements so that there are the smallest delay as possible between perception and action.

In this paper we present a vision system for color-coded object detection, which has been implemented using the UAVision library and three different types of digital cameras. Three different cameras have been used in order to prove the versatility of the proposed library and to exemplify the range of options that a user has for implementing a vision system using the UAVision library. The vision systems that we present are of two different types, perspective and omnidirectional, and can perform colored object detection in real-time, working at frame rates up to 50fps.

We present the same vision system pipeline implemented with different types of digital cameras, which fundamentally differ one from another. Even though the three vision systems that we present are designed to detect colored objects in the same environment and they share the same software structure, their architectures both in terms of software implementation and hardware are very different.

The first vision system that we present is a catadioptric one and uses an Ethernet camera. Following the same software pipeline, we present also two perspective vision systems implemented using a Firewire camera and a Kinect sensor, respectively. These two perspective vision systems are used with different purposes. Not only the access to these cameras is done differently in each of these cases, but the calibration module of the library provides methods for calibrating them both in terms of camera parameters and in terms of pixel-world coordinates. The calibration for each of these cameras is done by employing different types of algorithms, that will be explained in the next sections.

The UAVision library has been built using some of the features of the OpenCV library, mostly data structures for image manipulation. Although a powerful tool for developers working in image processing, the OpenCV library does not provide all the necessary support for implementing a complete vision system from scratch. To name some of the contributions of our library, it provides support for accessing different types of cameras using the same interface, algorithms for camera calibration (colormetric, catadioptric and perspective calibration) and the possibility of using image masks to process only relevant parts of the image. Moreover, we introduced blobs as data structures, an important aspect in colored object detection.

The authors have not found in literature free, up to date, available computer vision libraries for time constrained applications. We consider this paper an important contribution for the Computer Vision community since it presents in detail a novel computer vision library whose design focuses on maintaining a low processing time, a common constraint of nowadays autonomous systems.

The practical scenario in which the use of the library has been tested were the robotic soccer games, promoted by the RoboCup Federation [1]. The RoboCup Federation is an international organization that encourages research in the area of robotics and artificial intelligence by means of worldwide robotic competitions organized yearly. The RoboCup competitions are organized in four different challenges, one of them being the Soccer League. In this challenge, completely autonomous robots compete in games of soccer, following adjusted FIFA rules. The soccer challenge is divided into four leagues, according to the size and shape of the competing robots: the Middle Size League (MSL), the Standard Platform League (SPL), the Humanoid League (HL) and the Small Size League (SSL) (Fig. 1).

In MSL the games are played between teams of up to five wheeled robots that each participating team is free to build keeping in mind only the maximum dimensions and

weight of the robots. In SPL, all teams compete using the same robotic platform, a humanoid NAO built by Aldebaran [2] and the games take place between teams of up to five NAOs. In HL, the games are played between humanoid robots, that is robots that imitate the human body and the number of players in a game depends on the sub-league in which a team is participating: Kid-Size League, Teen League or Adult League. All the RoboCup games occur in closed, controlled environments. The robots play soccer indoor, on a green with white lines carpet and the color of the ball is known in advance. The color labeling of the environment makes the RoboCup competitions suitable for employing the UAVision library for the vision system of the autonomous robotic soccer players. Based on the library, an omnidirectional vision system has been implemented for the MSL team of robots CAMBADA [3], [4] from University of Aveiro. A perspective vision system that can be used either as a complement of the omnidirectional one or as an external monitoring tool for the soccer games has also been implemented and results obtained with both types of these systems are presented in this paper.

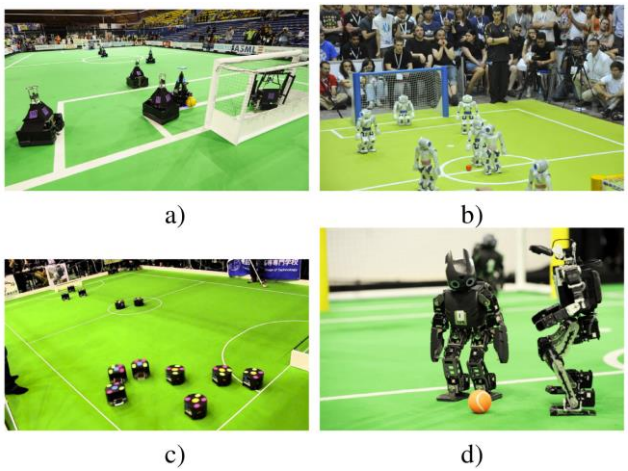


Fig. 1 Game situations in the four soccer leagues of RoboCup: a) MSL, b) SPL, c) SSL and d) HL.

This paper is structured into six sections, first of them being this introduction. Section 2 presents an overview of the most recent work on color coded object detection. Section 3 describes the modules of the vision library, while Section 4 presents the typical pipeline of a vision system developed based on the UAVision modular library. In Section 5 we present the practical scenarios in which the library has been used and the results that have been achieved, as well as the processing time of our algorithms and the proof that the library is indeed suitable for time-constrained applications. Finally, Section 6 concludes the paper.

2. Related Work

The library that we are proposing is an important contribution for the Computer Vision community since so far, there are no free, open source machine vision libraries that take into consideration time constraints. CMVision [5], a machine vision library developed at the Carnegie Mellon University was one of the first approaches to build such a library but it remained quite incomplete and has been discontinued in 2004. Several other machine vision libraries, such as Adaptive Vision [6], CCV [7] or RoboRealm [8], provide machine vision software to be used in industrial and robotic applications but they are not free. UAVision aims at being a free, open-source library that can be used for robotic vision applications that have to deal with time constraints.

In most autonomous robots there is the concern of maintaining an affordable cost of construction and limited power consumption. Therefore, there are many limitations in terms of hardware that have to be overcome by means of software in order to have a functional autonomous robot.

This is the case of most of the robots used by the broad RoboCup community. The most recent innovations done by the RoboCup community are presented next as they are related to the work proposed in this paper. The UAVision library has been used for the implementation of a vision system for a soccer robot and therefore, it is important to highlight the challenges that uprise in this environment and the state of the art available solutions.

The soccer game in the RoboCup Soccer League is a standard real-world test for autonomous multi-robot systems. In the MSL, considered to be the most advanced league in RoboCup at the moment, omnidirectional vision systems have become widely used, allowing a robot to see in all directions at the same time without moving itself or its camera [9]. The environment of this league is not as restricted as in the others and the pace of the game is faster than in any other league (currently with robots moving with a speed of 4 m/s or more and balls being kicked with a velocity of more than 10 m/s), requiring fast reactions from the robots. In terms of color coding, in the fully autonomous MSL the field is still green, the lines of the field and the goals are white and the robots are mainly black. The two teams competing are wearing cyan and magenta markers. For the ball color, the only rule applied is that the surface of the ball should be 80% of a certain color, which is usually decided before a competition. The colors of the objects of interest are important hints for the object detection, relaxing thus the detection algorithms. Many teams are currently taking their first steps in 3D ball information retrieving [10], [11], [12]. There are also some teams moving their vision systems algorithms to VHDL based algorithms taking advantage of the FPGAs versatility [10]. Even so, for now, the great majority of the

teams base their image analysis in color search using radial sensors [13], [14], [15].

While in MSL the main focus of the research efforts relies in building robots with fast reactions, good cooperation skills and complex artificial intelligence, able to play soccer as well as humans do from the point of view of game strategy and cooperation, in SPL the focus is achieving a stable biped walk, similar to the human walk. Most teams participating in this league use color segmentation as basis of the object detection algorithms. The first step in most of the color segmentation algorithms is scanning the image either horizontally, vertically, or in both directions [16], [17] while looking for one of the colors of interest. For this process, the image is almost always sampled down by skipping lots of pixels and thus losing information and considering what might be false positives. Another approach is using the methods provided by the CMVision library [5] which provides a base to segment images into regions of interest, to which probabilities are assigned [18]. The information about a color of interest segmented is validated in most cases if there is a given threshold of green color in the proximity of the color of interest previously segmented [16], [17], [19]. This is done in order to guarantee that the objects of interest are only found within the limits of the field. Therefore, when searching for an orange ball, as an example, the simplest validation method is checking if there is as little as one green pixel before and/or after an orange blob that has been found.

Team description papers like [17], [20], [21] validate the colors of interest segmented if they are found under the horizon line. The notion of horizon is calculated based on the pose of the robot's camera with respect to the contact point of the robot with the ground, which is the footprint. The next step in the detection of the objects of interest is the extraction of certain features or calculation of different measures from the blobs of the segmented colors of interest. It is common and useful to compute the bounding box and the mass center of a blob of a given color [16], [22], [23], as well as to maintain a function of the size of the object relative to the robot [23].

3. UAVision: A Real-Time Computer Vision Library

The library that we are presenting is intended, as stated before, for the development of artificial vision systems to be used for the detection of color-coded objects. The library incorporates software for image acquisition from digital cameras supporting different technologies, for camera calibration, blob formation, which stands at the basis of the object detection, and image transfer using TCP communications. The library can be divided into four main modules, that can be combined for implementing a time-constrained vision system or that can be used individually. These modules will be presented in the following subsections.

3.1 Acquisition Module

UAVision provides the necessary software for accessing and capturing images from three different camera interfaces, so far: USB, Firewire and Ethernet cameras. For this purpose, the Factory Design Pattern [24] has been used in the implementation of the Image Acquisition module of the UAVision library. A factory called “Camera” has been implemented and the user of the library can choose from these three different types of cameras in the moment of the instantiation. This software module uses some of the basic structures from the core functionality of OpenCV library: the *Mat* structure as a container of the frames that are grabbed and the *Point* structure for the manipulation of points in 2D coordinates. Images can be acquired in the YUV, RGB or Bayer color format. The library also provides software for accessing and capturing both color and depth images from a Kinect sensor [25].

Apart from the image acquisition software, this module also provides methods for converting images between the most used color spaces: RGB to HSV, HSV to RGB, RGB to YUV, YUV to RGB, Bayer to RGB and RGB to Bayer.

3.2 Camera Calibration Module

The correct calibration of all the parameters related to a vision system is crucial in any application. The module of camera calibration includes algorithms for calibration of the intrinsic and extrinsic camera parameters, the computation of the inverse distance map, the calibration of the colormetric camera parameters and the detection and definition of regions in the image that do not have to be processed. The process of the vision system calibration handles four main blocks of information: camera settings, mask, map and color ranges.

The camera settings block represents the basic camera information. Among others, this includes: the resolution of the acquired image, the Region of Interest regarding the CCD or CMOS of the camera and the colormetric parameters.

The mask is a binary image representing the areas of the image that do not have to be processed, since it is known apriori that they do not contain relevant information. Using a mask for marking non-interest region is a clever approach especially when dealing with a catadioptric vision system. These systems provide a 360º of the surrounding environment and that includes, most of the times, parts of the robot itself. By using a mask that allows skipping the processing of certain regions, the processing time decreases significantly.

The map, as the name suggests, is a matrix that represents the mapping between pixel coordinates and world coordinates. This mapping between the pixels of the acquired image and real coordinates allows the vision system to characterize the objects of interest, their position or size. The algorithms presented in [9], [27] are being used for calibration of the intrinsic and extrinsic parameters of catadioptric vision systems in order to generate the inverse distance map. For the calibration of the intrinsic and extrinsic parameters of a perspective camera, we have used and implemented the algorithm for the "chessboard" calibration, presented in [28].

For the omnidirectional vision system, by exploring a back-propagation ray-tracing approach and the physical properties of the mirror surface [27], the relation between the distances in the image and the distances in the real world is known. For the perspective vision systems, after having calculated the intrinsic and extrinsic parameters, the map calculation is straightforward, based on the camera pinhole model equations.

The color ranges block contains the color regions for each color of interest (at most 8 different colors as it will be explained later) in a specific color space (ex. RGB, YUV, HSV, etc.). In practical means, it contains the lower and upper bounds of each one of the three color components for a specific color of interest.

The Camera Calibration module of the library allows the storage of all of the information contained in these four blocks in a binary or text configuration file.

UAVision library contains algorithms for the self-calibration of the parameters described above, including some algorithms developed previously within our research group, namely the algorithm described in [26] for the automatic calibration of the colormetric parameters. These algorithms extract some statistical information from the acquired images, such as the intensity histogram, saturation histogram and information from a black and white area of the image, to then estimate the colormetric parameters of the camera.

3.3 Color-coded Object Segmentation Module

The color-coded object segmentation is composed of three sub-modules that are presented next.

Look-up Table

For fast color classification, color classes are defined through the use of a look-up table (LUT). A LUT represents a data structure, in this case an array, used for replacing a runtime computation by a basic array indexing operation.

This approach has been chosen in order to save significant processing time. The images can be acquired in the RGB, YUV or Bayer format and they are converted to an index image (image of labels) using an appropriate LUT for each one of the three possibilities.

The table consists of 16,777,216 entries (2^{24}, 8 bits for R, 8 bits for G and 8 bits for B) with one byte each. The table size is the same for the other two possibilities (YUV or RAW data), but the meaning and position of the color components in the image changes. For example, considering a RGB image, the index of the table to be chosen for a specific triplet is obtained as

```
idx = R << 16 | G << 8 | B,
```

being the << a bitwise shift to the left operation and | the bitwise OR. The RAW data (Bayer pattern) is also RGB information, however, the position where the R, G and B information are picked changes. These positions are also calculated only once by the LUT module in order to obtain a faster access during color classification.

Each bit in the table entries expresses if one of the colors of interest (in this case: white, green, blue, yellow, orange, red, blue sky, black, gray - no color) is within the corresponding class or not. A given color can be assigned to multiple classes at the same time. For classifying a pixel, first the value of the color of the pixel is read and then used as an index into the table. The 8-bit value then read from the table is called the "color mask" of the pixel. It is possible to perform image subsampling in this stage in systems with limited processing capabilities in order to reduce even more the processing time. If a mask exists, color classification is only applied to the valid pixels defined by the mask.

Scanlines

To extract color information from the image we have implemented three types of search lines, which we also call scanlines: radial, linear (horizontal or vertical) and circular. They are constructed once, when the application starts, and saved in a structure in order to improve the access to these pixels in the color extraction module. This approach is extremely important for the minimization of processing time. In Fig. 2 the three different types of scanlines are illustrated.

A Scanline is a list of pixel positions in the image, whose shape depends on its type. The module Scanlines allows the definition of an object that is a vector of scanlines. The amount and characteristics of these scanlines depend on the parameters (start and end positions, distance - vertical, horizontal or angular between scanlines, etc).

Each Scanline object is used by the RLE module, described next, that produces a list of occurrences of a specific color. Depending on the number of Scanlines objects created in a specific vision system, we will have the same number of RLE lists produced. Then, the Blob module is able to use all these lists to combine RLEs and form blobs, as described next.

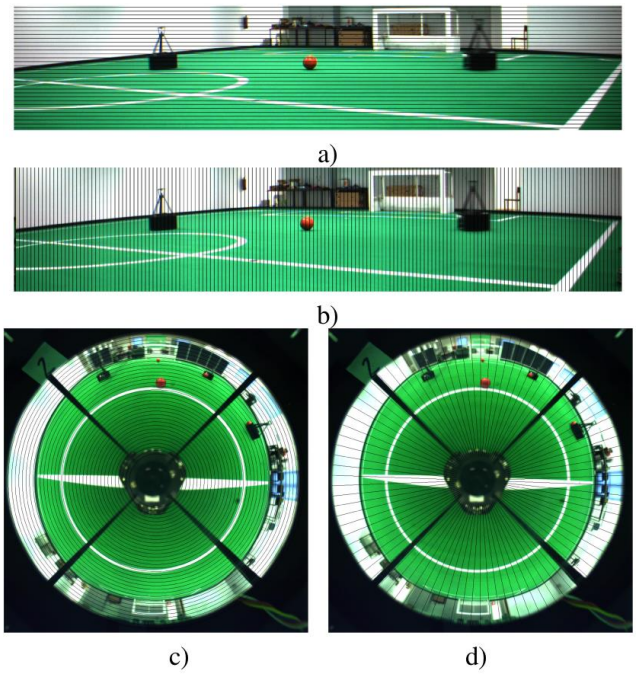


Fig. 2 - Examples of different types of scanlines: a) horizontal scanlines, b) vertical scanlines, c) circular scanlines, d) radial scanlines.

Run Length Encoding (RLE)

For each scanline, an algorithm of Run Length Encoding is applied in order to obtain information about the existence of a specific color of interest in that scanline. To do this, we iterate through its pixels to calculate the number of runs of a specific color and the position where they occur (Fig. 3). Moreover, we extended this idea and it is optional to search, in a window before and after the occurrence of the desired color, for the occurrence of other colors. This allows the user to determine both color transitions and color occurrences using this approach.

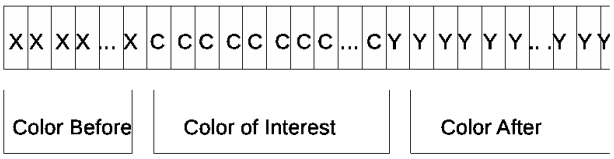


Fig. 3 - RLE illustration.

When searching for run lengths, the user can specify the color of interest, the color before, the color after, the search window for these last two colors and three thresholds that can be used to determine the valid information. These thresholds are application dependent and should be determined experimentally.

As a result of this module, the user will obtain a list of positions in each scanline and, if needed, for all the scanlines, where a specific color occurs, as well as the amount of pixels in each occurrence (Fig. 4).

A A A C C C C C B B B X X C C C B B...
colorList={(3, 5, 5, 5), (13, 3, 0, 2)...}

Fig. 4 - An example of a result of the RLE module. The first quadruple in the list describes the first detection of the color of interest, the second one the second detection of the color of interest and so on. The values presented in each quadruple represent the following information, by order: position of the first found pixel of the color of interest, number of pixels of the color of interest, number of pixels of the color found before the color of interest and number of pixels of the color found after the color of interest. The number of pixels found before or/and after the color of interest are optional and can be specified when searching for transitions between colors.

Blob formation

To detect objects with a specific color in a scene, one has to detect regions in the image with that color, usually designated as blobs. The blobs can be later on validated as objects of interest after extracting different features of the blobs that are relevant for establishing whether they are more than just blobs of colors. Such features can be, among others, the area of the blob, the bounding box, solidity, skeleton.

In order to construct the blobs, we use information about the position where a specific color occurs based on the Run Length module previously described (a practical example can be seen in Fig. 5). We iterate through all of the run lengths of a specific color and we apply an algorithm of clustering based on the Euclidean distance. The parameters of this clustering are application dependent. For example, in a catadioptric vision system, the distance in pixels to form blobs changes radially and non-linearly regarding the center of the image (Fig. 6).

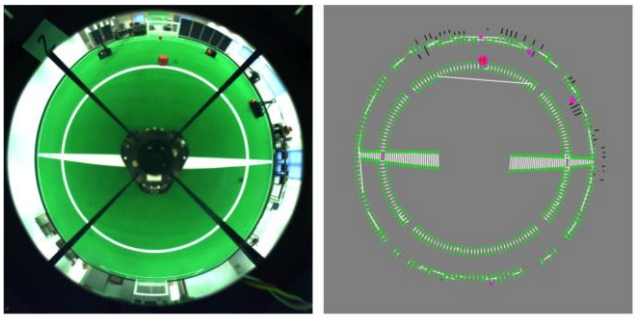


Fig. 5 - On the left, an image captured using the Camera Acquisition module of the UAVision library. On the right, the run length information annotated.

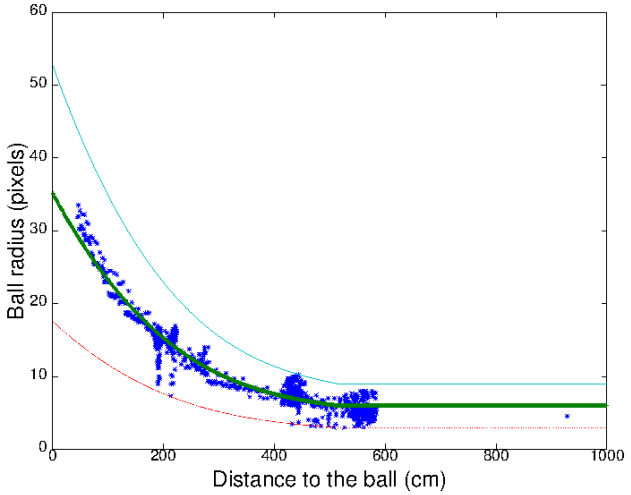


Fig. 6 - An example of a function used for a catadioptric vision system describing the radius (in pixels) of the object of interest (in this case, an orange soccer ball) as a function of the distance (in centimeters) between the robot and the object. The blue marks represent the measures obtained, the green line the fitted function and the cyan and red line the upper and lower bounds considered for validation.

The algorithm for forming blobs that we implemented in the library works as follows: a RLE object contains information about the occurrence of a color of interest on a given scanline. The information contained by the RLE object is: the position of the first found pixel of the color of interest, the number of consecutive pixels of the same color and the position of the last found pixel of the color of interest. The middle point between the first and last found pixels of the color of interest can be calculated for each RLE.

When searching for blobs, the first blob is considered to be the first found RLE and the center of the blob is the center of the RLE. Running through all the found RLE, based on the Euclidean distance between middle points in each run lengths, the color information in adjacent RLE is merged if the calculated Euclidean distance is between a given threshold. This threshold has been experimentally calculated. If the calculated distance is not lower than the threshold, the given RLE is considered to be the starting point of a new blob and the process is repeated.

When joining another RLE to the Blob, the center of the blob, as well as other measurements such as area, width/height relationship, solidity are being calculated. The blobs are then validated as being objects of interest if they respect the size-distance function presented in Fig. 6 and if they are fairly round objects.

TCP Communications Module

Two of the major limitations that have to be considered when working with autonomous robots are the energetic autonomy of the robot and its processing capacities. The broad experience in autonomous robotics within our research group has pointed out the fact that certain calibration tasks, that have to be performed prior to the functioning of a robot, might require too much of the on board resources. Therefore, based on the approach "divide and conquer" we have implemented a TCP Communications module that allows us to remotely communicate with a robot and distribute some of the tasks that are solely related to the calibration processes. Moreover, this module is relevant when developing applications for logging or monitoring, in which the robot performs autonomously and there is the need of having a real-time remote logger/monitor.

The purpose of this module was to simplify and be more intuitive than the POSIX API, while at the same time remain close to the known flow of operations defined by it. The new proposed flow of operations can be seen in Fig. 7. The socket descriptor is no longer the representation of a socket, but rather the object instantiated. Separating the client and server into two distinct sockets, with a single operation to setup both, the "open" operation, the flow of operations is simpler while familiar. On the client side, the socket descriptor request, the socket binding and the connection to the server were all merged into one "open" operation. The rest of the client operations were left unchanged, "send" and "receive" operations to communicate with the server and a "close" operation to close the socket. On the server side, similar to the client, the socket descriptor request, the socket binding and the passive socket marking were merged into one "open" operation, retaining the "accept" operation. The behavior of server and client are now defined in the type of socket instantiated.

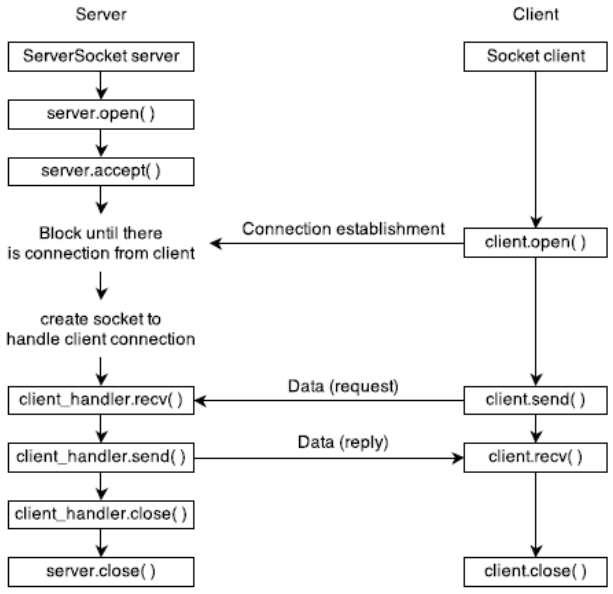


Fig. 7 - Setup of TCP Communications Module sockets for a client-server application.

Performance wise, this library has the same performance as the POSIX API while having the benefit of a simpler API, object oriented and with code legibility.

This module is relevant for the library that we are presenting since it allows the user to have a client and a server that can exchange information via TCP connection. The exchanged information is an image, with or without any annotated information. Based on this module, the vision system that will run on a robot can act as a server and an external application, the client, can request information from the server while it is performing its regular task. By using threads, the communications with the client will not affect the normal flow of operations on the server side. An example of a client application that has been implemented will be presented in Section 5.

4. A Typical Vision System Using the UAVision Library

An example of a typical vision system that can be used when developing a time-constrained robotic vision system is presented in Fig. 8.

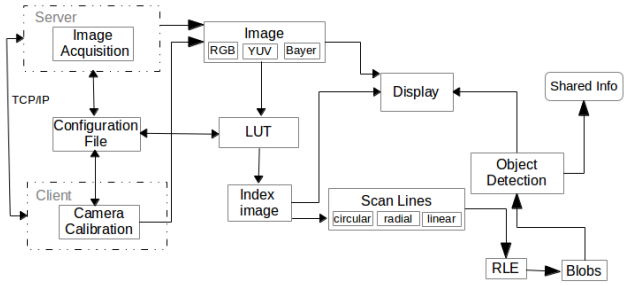


Fig. 8 - Software architecture of the vision system developed based on the UAVision library.

The pipeline of the object detection procedure is the following:

- After having acquired an image, this is transformed into an image of labels using a LUT previously built. This image of color labels, also designated in this paper by index image, will be the basis of all the processing that follows.
- The index image is scanned using one of the three types of scanlines previously described (circular, radial or linear) and the information about colors of interest or transitions between colors of interest is run length encoded.
- Blobs are formed by merging adjacent RLEs of a color of interest.
- The blob is then labeled as object of interest if it passes several validation criteria that are application dependent.
- If a given blob passes the validation criteria, its center coordinates will be passed to higher-level processes and they can be shared.

5. Experimental Results

The UAVision library has been used so far in different applications that are being developed at the University of Aveiro, Portugal. In this section we present the scenarios in which time-constrained robotic vision systems that have been built based on UAVision library have been successfully employed.

The code is written in C++ and the main processing unit available on the robots is an Intel Core i5-3340M CPU @ 2.70GHz 4 processor, running Linux (distibution Ubuntu 12.04. LTS Precise Pangolin). In the implementation of the vision system that we present in the paper, we did not use multi-threading, however both image color classification and the next steps can be parallelized if needed.

The parallelization of our software is not done for a basic reason: the Linux kernel installed on our computers is configured to use a 10ms scheduling interval. At this time interval the kernel determines whether the current process continues using the same core, or performs a context switch to another process. Since none of the modules of our vision system has processing time higher than 10ms, using threads would not improve it. In slower systems or in applications that have to deal with much higher resolution images, parallelization can be an option.

5.1 A time constrained catadioptric vision system for robotic soccer players

CAMBADA team of robots from University of Aveiro, Portugal (Fig. 9) is an established team that has been participating, for a decade, in the RoboCup MSL competitions previously described. These robots are completely autonomous, able to perform holonomic motion and are equipped with a catadioptric vision system that allows them to have omnidirectional vision [29]. In order to play soccer, a robot has to detect, in useful time, the ball, the limits of the field and the field lines, the goal posts and the other robots that are on the field. These robots can move with a speed of 4m/s and the ball can be kicked with a velocity of more than 10m/s, which leads to the need of having fast object detection algorithms. In the RoboCup soccer games, the objects of interest are color coded making thus the soccer games a suitable application for the use of the UAVision library.

Using the modules of the UAVision library, a vision system composed of two different applications has been developed with the purpose of detecting the objects of interest. The two applications are of the type client-server

and have been implemented using the TCP Communication module. The core application, that runs in realtime on the processing units of the robots, was implemented as a server that accepts the connection of a calibration tool client. This client is used for configuring the vision system (color ranges, camera parameters, etc.) and for debug purposes. The main task of the server application is to perform real time color-coded object detection. The camera used by the CAMBADA robots is an IDS UI-5240CP-C-HQ-50i Color CMOS 1/1.8'' Gigabit Ehernet camera [30], providing images with a resolution of 1280 x 1024 pixels. However, we use a region of interest of 1024 x 1024 pixels.

Fig. 9 - An example of a robot setup during a soccer game.

The architecture of the vision system that has been developed for these robots follows the generic one, presented in Fig. 8. For these robots, the objects of interest are: the ball, usually of a known solid color, the green field and the white lines necessary for the localization of the robot and the black obstacle, i.e. other robots that are on the field.

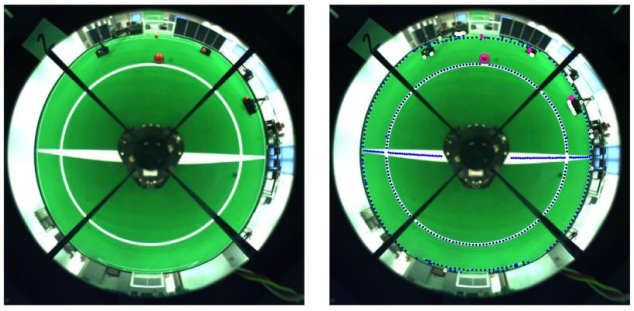
In this vision system, we use the following types and numbers of scanlines:

- 720 radial scanlines for the ball detection.
- 98 circular scanlines for the ball detection.
- 170 radial scanlines for the lines and obstacle detection.
- 66 circular scanlines for the lines detection.

As explained before, the last step of the vision pipeline is the decision regarding whether the colors segmented belong to an object of interest or not. In the vision system developed for the CAMBADA team using the proposed library, the white and black points that have been previously run-length encoded are passed directly to higher level processes, where localization based on the white points and obstacle avoidance based on the black points are performed.

For the ball detection, the blobs that are of the color of the ball have to meet the following validation criteria before being labeled as ball. First, a mapping function that has been experimentally designed is used for verifying a size-distance from the robot ratio of the blob (Fig. 6). This is complemented by a solidity measure and a width-height ratio validation, taking into consideration that the ball has to be a round blob. The validation was made taking into consideration the detection of the ball even when it is partially occluded.

Visual examples of the detected objects in an image acquired by the vision system are presented in Fig. 10 and Fig. 11. As we can see, the objects of interest (balls, lines and obstacles) are correctly detected even when they are far from the robot. The ball can be correctly detected up to 9 meters away from the robot (notice that the robot is in the middle line of the field and the farthest ball is over the goal line) even when they are partially occluded or engaged by another robot. No false positives in the detection are observed.



a) b)

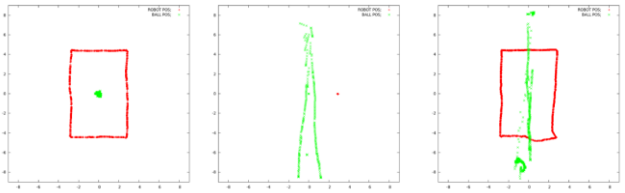
Fig. 10 – On the left, an image acquired by the omnidirectional vision system. On the right, the result of the color-coded object detection. The blue circles mark the white lines, the white circles mark the black obstacles and the magenta circles mark the orange blobs that passed the validation thresholds.

a) b) c)

Fig. 11 – On the left, the index image in which all of the colors of interest are labeled. In the middle, the color classified image (a colored version of a)) and on the right, the surrounding world from the perspective of the robot.

Several game scenarios have been tested using the CAMBADA autonomous mobile robots. In Fig. 12(a) we present a graphic with the result of the ball detection when the ball is stopped in a given position (the central point of the field, in this case) while the robot is moving. The graphic shows consistent ball detection while the robot is moving in a tour around the field. The field lines are also properly detected, as it is proved by the correct localization of the robot in all the experiments. The second scenario that has been tested is illustrated in Fig. 12(b). The robot is stopped on the middle line and the ball is sent

across the field. This graphic shows that the ball detection is accurate even when the ball is found at a distance of 9m away from the robot. Finally, in Fig. 12(c) both the robot and the ball are moving. The robot is making a tour around the soccer field, while the ball is being sent across the field. In all these experiments, no false positives were observed and the ball has been detected in more than 90% of the frames. Most of the times when the ball was not detected was due to the fact that it was hidden by the bars that hold the mirror of the omnidirectional vision system. The video sequences used for generating these results, as well as the configuration file that has been used, are available at [31]. In all the tested scenarios the ball is moving on the ground floor since the single camera system has no capability to track the ball in 3D.



a) b) c)

Fig. 12 – On the left, a graph showing the ball detection when the robot is moving in a tour around the soccer field. In the middle, ball detection results when the robot is stopped on the middle line on the right of the ball and the ball is sent across the field. On the right, ball detection results when both the robot and the ball are moving.

The cameras that have been used can provide 50fps at full resolution (1280 x 1024 pixels) in RGB color space. However, some cameras available on the market can only provide 50 fps accessing directly to the CCD or CMOS data, usually a single channel image using the well known Bayer format. As described before, the LUT in the vision library can work with several color spaces, namely RGB, YUV and Bayer format. We repeated the three scenarios described above acquiring images directly in the Bayer format also at 50fps and the experimental results show that the detection performance is not affected. A detailed study about the detection of colored objects in Bayer data can be seen in [33].

5.2 Delay Measurements between Perception and Action

In addition to the good performance in the detection of objects, both in terms of number of times that an object is visible and detected and in terms of error in its position, the vision system must also perform well in minimizing the delay between the perception of the environment and the reaction of the robot. It is obvious that this delay depends on several factors, namely the type of the sensor used, the processing unit, the communication channels and the actuators, among others. To measure this delay in the CAMBADA robots, we used the setup presented in Fig. 13.

The setup consists of a led that is turned on by the motor controller board and the same board measures the time that the whole system takes to acquire and detect the LED flash, and to send the respective reaction information back to the controller board. The vision system detects the led on and when it happens, the robotic agent sends a specific value of velocities to the hardware. This is the normal working mode of the robots in game play.

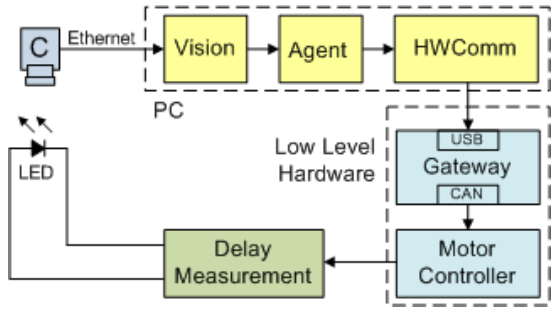
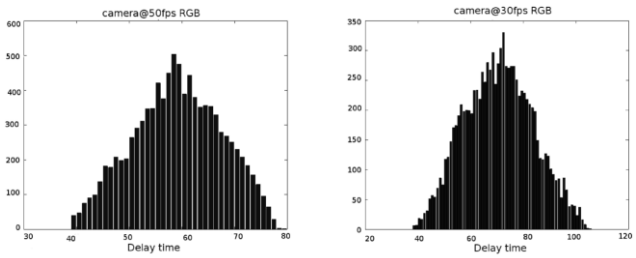


Fig. 13 – The blocks used in our measurement setup. These blocks are used by the robots during game play.

As presented in Fig. 14, the delay time between the perception and reaction of the robot significantly decreases when working at higher frame rates. The average delay at 30fps is 71 ms and at 50fps it is 60ms, which corresponds to an improvement of 16%.



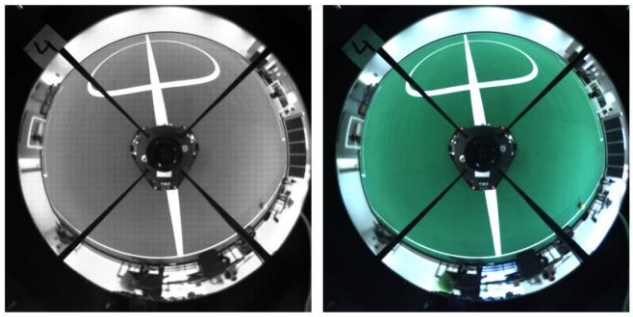
a) b)

Fig. 14 – Histograms showing the delay between perception and action on the CAMBADA robots. On the left, the camera is acquiring RGB frames and is working at 50fps (average = 60ms, max = 80ms, min = 39ms). On the right, the camera working at 30fps (average = 71ms, max = 106ms, min = 38ms).

The jitter verified reflects the normal function of the several modules involved, mainly because there is no synchronism between the camera, the processes running on the computer and the microcontrollers at the hardware level. If our vision system would have been synchronized with all the other processes running in the computational units of the robot (laptop and microcontrollers), for example using an external trigger signal, we would expect a similar delay in all the acquired frames. However, in the application described, we just guarantee that the other processes running in the laptop (decision and hardware communication processes) run after the vision system but

without synchronism with the camera and the software running in the microcontrollers.

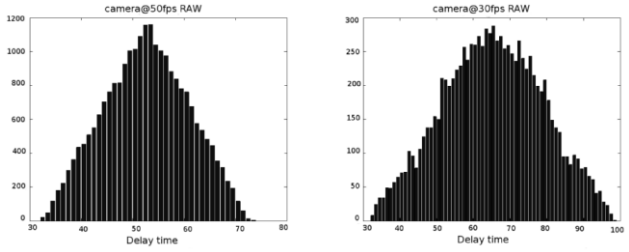
Since some digital cameras do not provide frame rates greater than 30fps in RGB, but allow the access to the raw data of the sensor at higher frame rates, and also because it is expected that the conversion between the raw sensor data and a specific color space in the camera will take some time, we performed some experiments for measuring the delay between perception and reaction of the robot accessing to the raw data. An example of a raw image acquired by the camera and the corresponding interpolated RGB version are presented in Fig. 15.



a) b)

Fig. 15 – In a) an example of a RAW image acquired by the camera (grayscale image containing the RGB information directly from the sensor - Bayer pattern. In b) the image a) after interpolation of the missing information for each pixel in order to obtain a complete RGB image.

We present the time results of these experiments in Fig.16. Similar to the results obtained when capturing RGB frames, the time between perception and reaction decreased when the frame rate increased. In this case, we have an improvement of 20% when using the camera at 50fps.



a) b)

Fig. 16 – Histograms showing the delay between perception and action on the CAMBADA robots. The camera is acquiring single channel frames, returning the raw data of the sensor (Bayer pattern). On the left, the camera is working at 50fps (average = 53ms, max = 74ms, min = 32ms). On the right, the camera working at 30fps (average = 66ms, max = 99ms, min = 32ms).

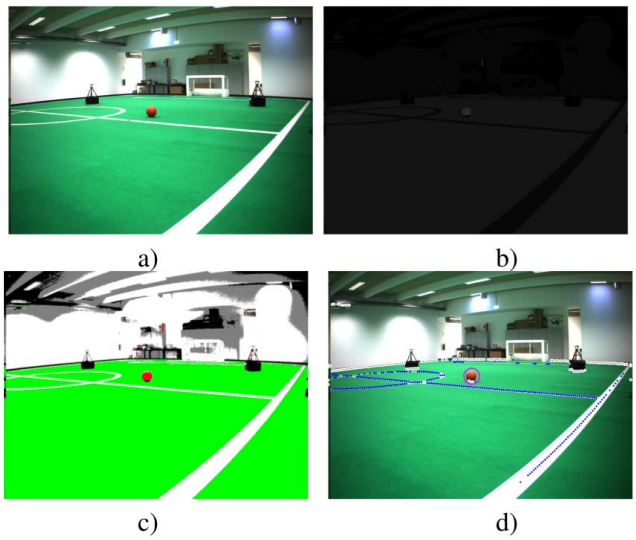
Another important result is the comparison between the delay time when acquiring frames in RGB and when accessing the raw data (Bayer pattern). Based on the graphics that we presented, we measured an improvement of 10% of the reaction time at 50fps and 7% at 30fps when using raw data. If the camera used by the robot could deliver raw data, these results prove that it is more advantageous. Moreover, we repeated the experiments with the robots detecting the ball using similar scenarios as the ones presented in Fig. 12. The experimental results obtained show that there is no loss in the detection performance of the robots and we observed an even better control in terms of motion. We provide in the library website [31] the sequences used in these experiments, as well as the binaries of the vision system used by our robots.

5.3 Perspective vision systems using the UAVision library

These results are complemented with another set of images acquired using a perspective camera that is fixed pointing to the soccer field (Fig. 17). The perspective camera is used for detecting the same objects of interest: the ball, the field lines and the obstacles/other robots and as future work will be used as ground truth for the team. The perspective camera used is a Firewire Point Grey Flea 2, 1 FL2-08S2C with a 1/3" CCD Sony ICX204 [32]. The images that we have used for the following experiments have a resolution of 1280 x 960 pixels. We present results achieved by using the UAVision library for two different cameras, but the architecture of the vision system is the same (the one presented in Fig. 8).

In this vision system, we use the following types and numbers of scanlines:

- 480 horizontal scanlines.
- 640 vertical scanlines.



a) b)

c) d)

Fig. 17 – Results obtained using a perspective camera: (a) - original image, (b) - index image, (c) - color classified image and (d) - image annotated with detected objects of interest.

The UAVision library has also been successfully employed for the detection of aerial balls, based on color information, using a Kinect sensor. This new implemented vision system, following the already presented architecture, was integrated in the software of the robots and a Kinect has been mounted on the goalie of the robotic soccer team. In this application, the depth information from the Kinect sensor has been used for discarding the objects of the color of the ball that are found farther than a certain distance (in this case, 7m were considered) in order to filter possible similar objects outside the field. An example of a ball detection by this system is presented in Fig. 18.

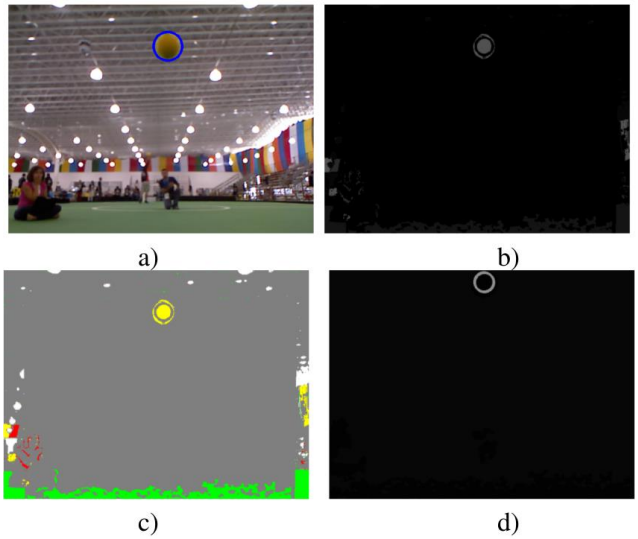


Fig. 18 – Results obtained with a Kinect sensor: a) original image with the ball detected, b) index image, c) color classified image, d)depth image.

5.4 Processing time

The processing time shown in Table. 1 proves that the vision systems built using the UAVision library are extremely fast. Taking as example the omnidirectional vision software of the CAMBADA robots, the full execution of the pipeline only takes on average a total of 15ms, allowing thus a framerate greater than 50fps if the digital camera supports it. Moreover, the maximum processing time that we measured was 18ms, which is a very important detail since it shows that the processing time is almost independent of the scene complexity. The time results have been obtained on a computer with a Intel Core i5-3340M CPU @ 2.70GHz 4 processor, processing images with resolutions from 640 x 480 pixels to 1280 x 960 pixels. In the implementation of these vision systems we did not use multi-threading. However, both image classification and the next steps can be parallelized if needed.

Table 1: Average processing times (in milliseconds - ms) measured for the vision systems developed using the UAVision library and presented in this section. The column "Omni" refers to the omnidirectional vision system used by the CAMBADA robots, using a resolution of 1024 x 1024 pixels. Column "Perpective"' refers to the perspective vision system developed with a Firewire Camera and a resolution of 1280 x 960 pixels. Column "Kinect"' refers to the vision system developed for the goalie using a Kinect camera, working with a resolution of 640 x 480 pixels. The filtering operation only exists for the Kinect vision system, as described in Section 5.3.

Operation	Vision Systems		
	Omni	Perspective	Kinect
Acquisition	1	1	1
Color Classification	5	10	12
Filtering	-	-	16
RLE	4	1	0
Blob Creation	2	0	0
Blob Validation	3	0	0
Total	**15**	**12**	**19**

The LUT is created once, when the vision process runs for the first time and it is saved in the cache file. If the information from the configuration file does not change during the following runs of the vision software, the LUT will be loaded from the cache file, reducing thus the processing time of this operation by approximately 25 times.

The time spent in the color classification step is greater in the perspective vision system, considering a resolution similar to the omnidirectional vision system, due to the fact that in the omnidirectional vision there is a mask defined and not all the pixels are color classified, as described before. This is one important feature of the proposed library.

The time spent by the omnidirectional vision system regarding RLE and Blobs modules are greater than the ones in the perspective vision system due to the existence of more scanlines to process, as presented in the description of the vision systems.

6. Conclusions

In this paper we have presented a novel time-constrained computer vision library that has been successfully employed in the games of robotic soccer. The proposed library, UAVision, encompasses algorithms for camera calibration, image acquisition and color coded object detection and allows frame rates of up to 50fps when processing images with more than 1 mega pixels.

The experimental results in this paper show that the library can be easily used in different vision systems. The

similar vision pipeline was successfully used with three different cameras (Ethernet, Firewire and Kinect) in three different applications (omnidirectional and perspective vision systems for ground object detection and a vision system for aerial balls detection).

The use of this library in other applications is straightforward. The user only has to configure the colors of interest for the application and the validation procedure for the detected blobs.

The algorithms developed for this library take into consideration applications in which the computing resources are limited. As an example, the vision system of the soccer robot presented in this paper can be used in another robot with more limited processing capabilities after a small number of adjustments. These adjustments are related to the sparsity of the scanlines.

In what concerns the future work, the next step will be to use the developed library in other RoboCup Soccer Leagues and the first concern is adding support for the cameras used by the robots in the Standard Platform and Humanoid Leagues and employing the same vision system on them. Moreover, we aim at providing software support for image acquisition from several other types of cameras and complement the library with algorithms for generic object detection, relaxing thus the rules of color coded objects and supporting the evolution of the RoboCup Soccer Leagues.

Acknowledgments

This work was developed in the Institute of Electronic and Telematic Engineering of University of Aveiro and was partially supported by FEDER through the Operational Program Competitiveness Factors - COMPETE and by National Funds through FCT - Foundation for Science and Technology in a context of a PhD Grant (FCT reference SFRH/BD/85855/2012) and the project FCOMP-01-0124-FEDER-022682 (FCT reference PEst-C/EEI/UI0127/2011).

References

[1] RoboCup. http://www.robocup.org. Last visited - September, 2014.

[2] Aldebaran Robotics oficial website. http://www.aldebaran-robotics.com. Last visited - September, 2014.

[3] A. Neves, J. Azevedo, N. Lau B. Cunha, J. Silva, F. Santos, G. Corrente, D. A. Martins, N. Figueiredo, A. Pereira, L. Almeida, L. S. Lopes, and P. Pedreiras. CAMBADA soccer team: from robot architecture to multiagent coordination, chapter 2. I-Tech Education and Publishing, Vienna, Austria, In Vladan Papic (Ed.), Robot Soccer, 2010.

[4] R. Dias, F. Amaral, J. L. Azevedo, R. Castro, B. Cunha, J. Cunha, P. Dias, N. Lau, C. Magalhaes, A. J. R. Neves, A. Nunes, E. Pedrosa, A. Pereira, J. Santos, J. Silva, and A. Trifan. CAMBADA Team Description. RoboCup 2014, Joao Pessoa, Brazil, 2014.

[5] CMVision. http://www.cs.cmu.edu/~jbruce/cmvision/. Last visited - September, 2014.

[6] Adaptive Vision. https://www.adaptive-vision.com/en/home/. Last vis- ited - September, 2014.ADVANCES IN COMPUTER SCIENCE : AN INTERNATIONAL JOURNAL 12

[7] CCV. http://libccv.org/. Last visited - September, 2014.

[8] Roborealm. http://www.roborealm.com/index.php. Last visited - September, 2014.

[9] Antnio J. R. Neves, Armando J. Pinho, Daniel A. Martins, and Bernardo Cunha. An eficient omnidirectional vision system for soccer robots: from calibration to object detection. Mechatronics, 21(2):399–410, mar 2011.

[10] F.M.W Kanters, R. Hoogendijk, R.J.M. Janssen, K.J. Meessen, J.J.T.H. Best, D.J.H Bruijnen, G.J.L. Naus, W.H.T.M Aangenent, R.B.M. van der Berg, H.C.T. van de Loo, G.M. Heldes, R.P.A. Vugts, G.A. Harkema, P.E.J. van Brakel, B.H.M Bukkums, R.P.T. Soetens, R.J.E. Merry, and M.J.G. can de Molengraft. Tech United Eindhoven Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[11] U. P. Kappeler, O. Zweigle, H. Rajaie, K. Hausserman, A. Tamke, A. Koch, B. Eckstein, F. Aichele, D. DiMarco, A. Berthelot, T. Walter, and P. Levi. RFC Stuttgart Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[12] A. Ahmad, J. Xavier, J. Santos-Victor, and P. Lima. 3d to 2d bijection for spherical objects under equidistant fisheye projection. Computer Vision and Image Understanding, 125:172–183, 2014.

[13] M. Huang, X. Ge, S. Hui, X. Wang, S. Chen, X. Xu, W. Zhang, Y. Lu, X. Liu, L. Zhao, M. Wang, Z. Zhu, C. Wang, B. Huang, L. Ma, B. Qin, F. Zhou, and C. Wang. Water Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[14] H. Lu, Z. Zeng, X. Dong, D. Xiong, and S. Tang. Nubot Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[15] A.A.F Nassiraei, S. Ishida, N. Shinpuku, M. Hayashi, N. Hirao, K. Fu- jimoto, K. Fukuda, K. Takanaka, I. Godler, K. Ishii, and H. Miyamoto. Hibikino-Musashi Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[16] A. Trifan, A. J. R. Neves, B. Cunha, and N. Lau. A modular real- time vision system for humanoid robots. In Proceedings of SPIE IS&T Electronic Imaging 2012, January 2012.

[17] T. Rofer, T. Laue, C. Graf, T. Kastner, A. Fabisch, and C. Thedieck. B-Human Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[18] leader@austrian kangaroos.com. Austrian Kangaroos Team Description. RoboCup 2011, Istanbul, Turkey, 2011.

[19] TJArk.ofice@gmail.com. TJArk Team Description. RoboCup 2014, Joao Pessoa, Brazil, 2014.

[20] H. L. Akin, T. Mericli, E. Ozukur, C. Kavaklioglu, and B. Gokce. Cerberus Team Description. RoboCup 2014, Joao Pessoa, Brazil, 2014.

[21] D. Garcia, E. Carbajal, C Quintana, E. Torres, I. Gonzalez, C. Busta- mante, and L. Garrido. Borregos Team Description. RoboCup 2012, Mexico City, Mexico, 2012.

[22] S. Liemhetcharat, B. Coltin, C. Mericli, and M. Veloso. CMurfs Team Description. RoboCup 2014, Joao Pessoa, Brazil, 2014.

[23] E. Hashemi, O. A. Ghiasvand, M. G. Jadidi, A. Karimi, R. Hashemifard, M. Lashgarian, M. Shafiei, S. Mashhadt, K. Zarei, F. Faraji, M. A. Z. Harandi, and E. Mousavi. MRL

Team Description. RoboCup 2014, Joao Pessoa, Brazil, 2014.
[24] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides. De- sign Patterns: Elements of Reusable Object-oriented Software. Addison- Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1995.
[25] Microsoft Kinect oficial website. http://www.microsoft.com/en-us/ kinectforwindows/. Last visited - September, 2014.
[26] Alina Trifan A. J. R. Neves and Bernardo Cunha. Self-calibration of colormetric parameters in vision systems for autonomous soccer robots. In in Proc. of RoboCup 2014 Symposium, 2014.
[27] B. Cunha, J. L. Azevedo, N. Lau, and L. Almeida. Obtaining the inverse distance map from a non-svp hyperbolic catadioptric robotic vision system. In Proc. of the RoboCup 2007, Atlanta, USA, 2007.
[28] Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In in ICCV, pages 666–673, 1999.
[29] A. J. R. Neves, G. Corrente, and A. J. Pinho. An omnidirectional vision system for soccer robots. In Proc. of the EPIA 2007, volume 4874 of Lecture Notes in Artificial Inteligence, pages 499–507. Springer, 2007.
[30] IDS. https://en.ids-imaging.com. Last visited - September, 2014.
[31] UAVision. http://sweet.ua.pt/an/uavision/. Last visited - September, 2014.
[32] Pointgrey. http://www.ptgrey.com/products/flea2/flea2 firewire camera. asp. Last visited - September, 2014.
[33] António J. R. Neves, Alina Trifan, José Luís Azevedo. Time-constrained detection of colored objects on raw Bayer data. Proceedings of the 5th Eccomas Thematic Conference on Computational Vision and Medical Image Processing, VipIMAGE 2015, Tenerife, Spain, p. 301-306, October 2015.

Alina Trifan obtained her bachelor's degree from Universitatea Tehnica Cluj-Napoca, Romania in 2009 and the MSc degree from University of Aveiro, Portugal in 2011. She is currently a Ph.D student working on computer vision algorithms for robotic applications, under the supervision of Prof. Antonio Neves and Prof. Bernardo Cunha. Her research is focused on time constrained algorithms for vision systems of autonomous mobile robots.

António J. R. Neves received the Electronics and Telecommunications Engineering degree from the University of Aveiro, Portugal, in 2002 and the Ph.D. in Electrical Engineering from the University of Aveiro, in 2007. Currently, he is an Assistant Professor at the Department of Electronics, Telecommunications and Informatics of the University of Aveiro, and a researcher at the Intelligent Robotics and Inteligent Systems group of the Institute of Electronics and Telematics Engineering of Aveiro - IEETA. His main research interests are robotics (computer vision and multi-agent robotics), image and video coding (lossless coding, pre-processing techniques, images with special characteristics) and bioinformatics (microarray images and DNA coding). He published more than one hundred papers including several book chapters, journal articles and conference papers in the areas described above.

Bernardo Cunha received the Electronics and Telecommunications Engineering degree from the University of Aveiro, Portugal, in 1982 and the Ph.D. in Electrical Engineering from the University of Aveiro, in 1991. Currently, he is an Assistant Professor at the Department of Electronics, Telecommunications and Informatics of the University of Aveiro, and a researcher at the Intelligent Robotics and Inteligent Systems group of the Institute of Electronics and Telematics Engineering of Aveiro - IEETA. His main research interests are robotics (hardware design, computer vision and multi-agent robotics). He published several papers in the areas described above.

José Luís Azevedo received the Electronics and Telecommunications Engineering degree from the University of Aveiro, Portugal, in 1985 and the Ph.D. in Electrical Engineering from the University of Aveiro, in 1998. Currently, he is an Assistant Professor at the Department of Electronics, Telecommunications and Informatics of the University of Aveiro, and a researcher at the Intelligent Robotics and Inteligent Systems group of the Institute of Electronics and Telematics Engineering of Aveiro - IEETA. His main research interests are robotics (hardware design, electronics and multi-agent robotics). He published several papers in the areas described above.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Integrating Learning Style in the Design of Educational Interfaces

Rana Alhajri[1] and Ahmed AL-Hunaiyyan[2]
[1]Computer Science Dept. Higher Institute of Telecommunication and Navigation, PAAET, Kuwait
ra.alhajri@paaet.edu.kw
[2]Computer and Information Systems Dept. College of Business Studies, PAAET, Kuwait
Hunaiyyan@hotmail.com

Abstract
Understanding learners' characteristics and behaviour using hypermedia systems in education is still a challenge for most developers and educators. This study seeks to understand the influence of learning style on learners' use and preference of a Computer Based Learning (CBL) program which was developed using two navigational structures (linear and non-linear). The study presents the findings of a case study conducted in Kuwaiti Higher Education. Data analysis was used to understand learners' needs and perceptions in using navigational structures. The relationship between learners' needs, learning styles, perceptions, and preferences in using the CBL program in regard to learner's gender is discussed in this paper. We found that both males and females liked to see the two navigational structures in the CBL program. Moreover, we found that although males and females are both prefer using the non-linear structures, the data analysis shows that males actually used linear structures of the program and they are characterized as verbalized, field independent and serialist learners. Females, on the other hand, need to use the non-linear structure, and characterized as visualized, field dependent, and holist learners. It is interesting to find that learners (both males and females) may use specific navigational structures (linear and non-linear) accommodated in a CBL program although it may not be what they prefer.

Keywords: *e-Learning, HCI, learning style, Individual differences, ICT, Hypermedia*

1. Introduction

Opportunities and challenges are emerging for learners, teachers and institutions from the increasing implementation of Information and Communication Technology (ICT) and associated infrastructure. However, designing and development an efficient educational interface within a learning environment is still a challenge for most developers, facilitators, and educators due to the complex understanding of learners' characteristics and behavior that incorporates many pedagogical and technological elements. The computer human interaction (CHI) environment regularly researches factors that affect the success or failure in interaction with computers.
The rapid developments in ICT and the evolving learner characteristics and behaviours require continuous effort to design digital content, both in the physical and virtual 'classroom' spaces [1]. The implementation of ICT has become standard in all aspects of life, including the field of education. Educational institutions, educators, and researchers are calling for providing educational materials that is informative, well designed, with consideration of learners' characteristics and style [2]. There have been numerous research studies on learning style's effect on learners using ICT in teaching and learning. Learners' preferences, perception, and their ability to understand educational programs are determined by their varying skills and abilities.

Learners' performance, perception, and their ability to comprehend course content are determined by their varying skills and abilities. Individual differences such as gender, socio-cultural, and cognitive style may also affect learning motivation and performance. Such differences, known as individual differences of learners, have been found to be important factors to consider in the development of digital learning systems [3].

It is generally accepted that there are individual difference between learners in terms of perceiving, remembering, processing, organizing information and problem solving [4]. It is important that instructors are able to recognize information resources that match learner's needs. In addition, learners should have a flexible interface that accommodates their learning styles, individual preferences, and should also be able to easily identify relevant content and navigation support. In Guo & Zhang, 2009, a framework of individual computer- based learning systems focuses on the learner's cognitive learning process, learning patterns and activities, and the technology support needed. This if properly designed by considering learners' individual differences will provide sufficient learning resources and communication tools to build a collaborative learning environment where both students and instructors gain significant benefits. Gülbahar and Alper (2011) [5] stated that Learning preferences and learning styles are a way to enhance the quality of learning. They stressed that student can adapt learning processes, activities and techniques, if he/she is able to understand his/her own personal characteristics and the consequences of possible different experiences.

The central theme of this research paper is to understand the influence of learning style on learners' use and preference of an interactive Computer Based Learning (CBL) program which was designed and developed in a structure of different navigational structures (linear and non-linear access). The study presented in this paper was conducted on Kuwaiti Higher Education students in Kuwait.

The rest of the paper is organized as follows. In Section 2, we present previous related works which elaborate on individual differences of learners. Section 3 describes the methodology used to conduct the study. Results are discussed in Section 4, and Section 5 concludes the work presented in this study with future work.

2. Individual Differences of Learners

Previous studies demonstrated the importance of individual differences as a factor in the design of computer-based learning. Such individual differences have significant effects on user learning in computer-based learning, which may affect the way in which they learn from and interact with hypermedia systems. Tailoring the process of instruction to match learners' style and to reflect individual differences of learners is a strong challenge under the conditions of the ICT supported education. [2].

Rurato & Borges Gouveia (2014) [6] stressed that when providing educational instruction that takes into account different issues of learners such as: personal and professional life; available technology resources; motivation and learning preferences; will allow to both learners and facilitators the proper way to adopt learning strategies easily, and in turn, enhance the possibility of making the learning experience successful.

Chen (2002) [7] indicated that a non-linear learning approach in hypermedia learning systems may not be suitable to all learners. Learners may have different backgrounds, especially in terms of their knowledge, skills, and needs, so they may show various levels of engagement in course content. Therefore, many studies argue that no one style will result in better performance. However, learners whose browsing behavior was consistent with their own favored styles obtained the best performance results. Individual differences may also affect learning motivation and performance. Such differences of learners have been found to be important factors to consider in the development of digital learning systems. The following sections discuss individual differences of learners such as gender differences; learners' culture; and cognitive styles of learners.

2.1. Learner's Gender

Gender differences are also argued as an important factor that significantly impacts learning in hypermedia learning systems [8]. Studies show that, in general, females have less experience with computers than males [9]. Thus, females tend to experience more disorientation in hypermedia than males [10], and males have been found to outperform females [11]. Several studies also examined gender differences in perceptions of computers and the Web and found significant differences. Male users had more positive attitudes towards computers and the web compared to female users [9]. Conversely, there are studies indicating that there are no gender differences in attitudes towards computers. Young and McSporran (2001) [12] examined gender differences in user learning performance in a hypermedia learning system and found that females favored and performed better with online learning courses. A study conducted by Atan, et al. (2002) [13], indicate that female distance education learners participate equally with their male counterparts in the utilization of computer technology to assist their study requirements as well as in their involvement in information and communication technology (ICT) to support the educational and learning process. Still, other studies have argued that there are no gender differences [14].

2.2. Learner's Culture

Understand ethnicity in different societies, the cultures of different generations, religion, education and literacy and language will undoubtly help to successfully develop any product. Technology developers unfortunately concentrate solely on economic influences, assuming that the world that is becoming more globalized, however the continuing effect of local culture is present.

Designers of multimedia interfaces should be aware of the cultural features of the program in which it is important to have a mechanism to understand the cultural elements of the target user. These mechanisms are needed not only to provide "good" cultural interfaces to learners across multiple cultures, but also to serve as tools for users of a specific culture.

It is important to understand the difference between what is comprehensible to a culture and what is acceptable. Because social norms, values, and traditions vary greatly between cultures, what is acceptable in one culture can be objectionable in another. In addition, it has been argued that what is known in one culture may have little or no meaning in another. This was addressed by Russo and Boor (1993) who gave an example of a trash can from Thailand which looks different from a US trash can. They believed that the Thai user is likely to be confused by the US trash can image [15]. It is believed that culture affects a user's perception and understanding of interface elements.

2.3. Cognitive Style

Cognitive style is an influential factor in users' information seeking, it refers to how the learner process information and represent the individual's mode of thinking, remembering, perceiving, and problem solving [16]. Frias-Martinez, et al. (2008) listed a number of dimensions of cognitive styles; Holism/Serialism; Divergent/Convergent; Field Dependence/Independence; and Imager/Verbalizer, and characterized Field Dependence/Independence and Imager/Verbalizer are especially as they are related to information seeking. He argued that Dependence /Independence concerns with how users process and organize information whereas Imager/Verbalizer emphasizes how users perceive the presentation of information [17].

Technology based learning systems provide users with freedom of navigation that allows them to develop learning pathways. Much empirical evidence indicates that not all learners can benefit from these systems. In particular, some learners have problems dealing with non-linear learning. Research into individual differences suggests that a learner's cognitive style has considerable effect on his or her learning. Moreover, in a traditional learning environment, matching a user's cognitive style with content presentation has been shown to enhance performance and improve perception [11]. Simply, cognitive style is known as an important factor that influences learners' preferences. Three divisions of cognitive styles are discussed below.

2.3.1. Field-Dependent versus Field-Independent

Field dependence (FD) and field independence (FI) refers to an analytical or global approach to learning, and is probably the most well-known division of cognitive styles [18]. FI learners generally are analytical in their approach, whereas FD learners are more global in their perceptions. Many experimental studies have argued the impact of FD and FI on the learning process. Jonassen & Grabowski (1993) stated that Field Dependence/Independence is related to the degree to which a user's perception or of information is influenced by the environment [19].

With regard to navigation strategies, some studies suggest that FI users prefer navigational structures such as "index" and "find" to locate specific content [20]. Conversely, FD users tend to see a global picture [18], and prefer to use well-structured tools such as maps or main menus [21]. Additionally, some studies found that FI users relatively enjoy non-linear navigation while FD users seem to prefer a fixed path to navigate computer-based content. Table 1 shows FD and FI categories.

Table 1: FD and FI categories, Source from Chen (2002) [7].

Field dependent learners	Field independent learners
More likely to face difficulties in restructuring new information and forging links with prior knowledge	Able to reorganise information to provide a context for prior knowledge
Their personalities show a greater social orientation	They are influenced less by social reinforcement
Experience surroundings in a relatively global fashion, passively conforming to the influence of the prevailing field or context	Experience surroundings analytically, with objects experienced as being discrete from their backgrounds
Demonstrate fewer proportional reasoning skills	Demonstrate greater proportional reasoning skills
Prefer working in groups	Prefer working alone
Struggle with individual elements	Good with problems that require taking elements out of their whole context
Externally directed	Internally directed
Influenced by salient features	Individualistic
Accept ideas as presented	Accept ideas strengthened through analysis

2.3.2. Visualized versus Verbalized

There are many divisions of cognitive styles, among which Riding (1991) [22] Visualizer / Verbalizer particularly emphasizes the presentation of information. Since multimedia systems incorporate numerous ways to present information, such as text, graphics, sound, animation and video, multimedia content was found to significantly influence users' levels of understanding and enjoyment. The main differences between the two cognitive styles, Visualizers and Verbalizers, are described in [19] are shown in Table 2.

A Visualizer prefers to receive information via graphics, pictures, and images, whereas a Verbalizer prefers to process information in the form of words, either written or spoken. Visualizers prefer to process information by seeing, whereas Verbalizers prefer to process information by listening and talking.

Table 2: The differences between Visualizers and Verbalizers (adapted from Jonassen and Grabowski (1993) [19] and Riding & Rayner (1998) [23]

Visualizers	Verbalizers
Think concretely	Think abstractly
Have high imagery ability	Have low imagery ability
Like graphics	Like reading text or listening
Prefer to be shown how to do something	Prefer to read about how to do something
Are more subjective about what they are learning	Are more objective about what they are learning

2.3.3. Holist versus Serialist Strategy

In the Chen (2002) [7] study, two versions of a hypermedia learning system, the Breadth-first and the Depth-first, were designed with program control paths. In the Depth-first version, each topic was presented in detail before the next topic, which was presented in the same way (i.e., Serialist condition). The material was classified into seven depth levels. In contrast, the Breadth-first version provides a summary of all of the material prior to introducing detail (i.e., Holist condition), and included 12 categories in breadth.

Results showed that users whose cognitive styles were matched to the design of hypermedia learning systems that they preferred achieved higher posttest scores. Field Dependent learners performed better in the Breadth-first version than in the Depth-first version. On the other hand, Field Independent users performed best in the Depth-first version than in the Breadth-first version. The differences that characterize the holist-serialist dimension of style as approaches to learning are listed in Table (3).

Table 3: Characteristics of Holists-Serialists. Source, Jing (2005) [24]

Holist	Serialist
Top-down processor	Bottom-up processor
Global approach to learning	Local approach to learning
Simultaneous processing	Linear processing
Spans various levels at once	Works step by step
Interconnects theoretical and practical aspects	Aspects learned separately
Conceptually orientated	Detail orientated
Comprehension learning bias	Operational learning bias
Relates concepts to prior experience	Relates characteristics within concept
Broad description building	Narrow procedure building
Low discrimination skills	High discrimination skills

3. Methodology Design

In this research, we used quantitative data analysis obtained from a log file generated by using the Computer Based Learning (CBL) program and from a questionnaire. We define our independent values as males and females in addition to the total number of frames pages visited from map frame (non-linear structure) and total number of frames pages visited from index frames (linear structure). We should differentiate between linear and nonlinear approach since it is our main research focus. In a linear structure learners has no facility to jump to out-of-order slides, where with a non-linear structure, learners can access any point of the program, it is a presentation with hyperlinks, learners can navigate to other points in the presentation by simply linking to them.

The central theme of this research paper is to understand the influence of learning style on learners' use and preference of the CBL program which has two structures (linear and non-linear structure). This study tries to confirm the following hypotheses:

Hypothesis 1: Females are more likely use non-linear structure.

Hypothesis 2: Males are more likely use linear structure.

3.1. Participants

We conducted the experiment at the Higher Institute of Telecommunication and Navigation (HITN) in Kuwait. The total number of participants was 86 and their ages ranged between 18 and 26 years. Participants had different computing and internet skills and were classified in terms of gender. Males (N=43) and females (N=43).

3.2. Research Instruments

This research used two instruments, a log file from the CBL program and a questionnaire. The CBL program presents learning materials entitles "an introduction to PowerPoint". The program provides participants with, navigational structures, including a hierarchical map (non-linear structure) and an alphabetical index (linear structure) (Figure 1). In this approach, learners are given control to decide to choose their own learning paths and their favored navigation display.

When a participant clicked on any displayed link in the program, whither from the Map Frame or the Index Frame, a log file records records participant movement and registered visited pages, the clicks then saved in a log file.

A questionnaire was used to capture the users' subjective feelings and perceptions regarding the hypermedia learning environment. The questionnaire responses were made up of 5-point Likert scales which had the following possible responses: "strongly agree", "agree", "neutral", "disagree", and "strongly disagree". There are three questions in the questionnaire: a) "I like the fact that I have the ability to control the pace of instruction using Hierarchal Map", b) "I like the fact that I have the ability to control the pace of instruction using index" and c) "I like the fact that I can see the both frames of navigational structures, map and index frames". Questions a) and b) were used to understand the learners' perception of using map and index in displaying the instruction respectively, while question c) reflects learners' perception whether they like to see the frames of navigational structures, map and index frames, to be displayed.

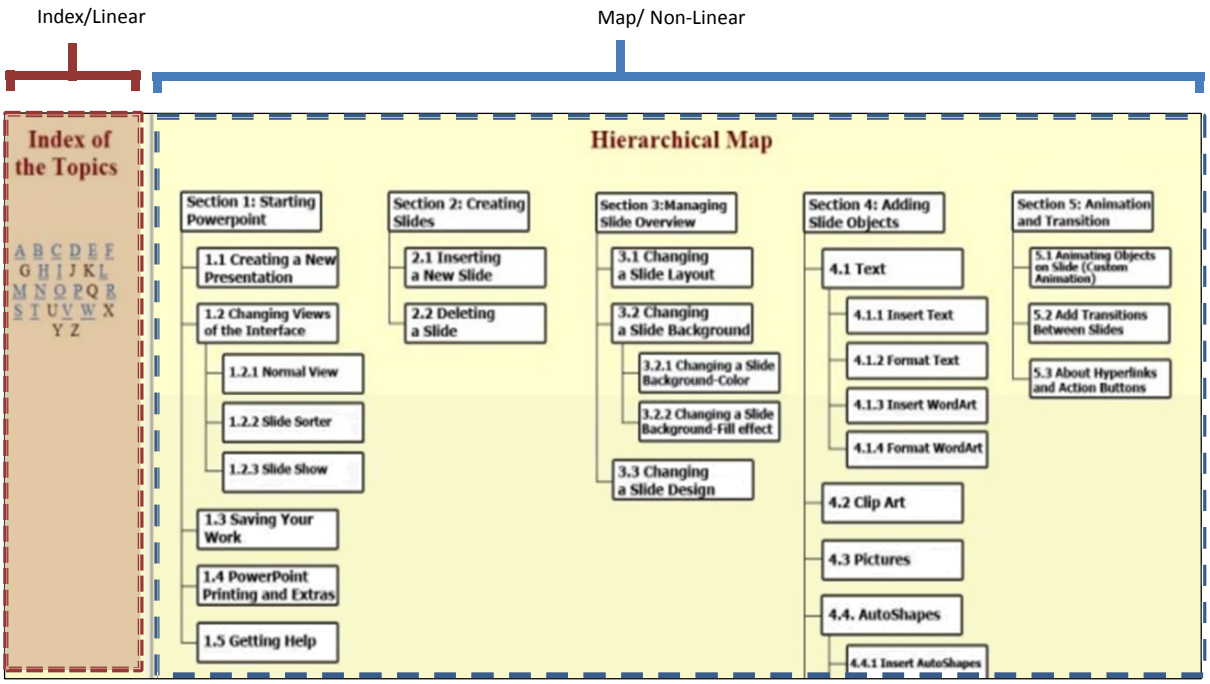


Figure 1: The main page of the CBL program.

3.3. Procedures

The experiment consisted of two phases. All participants were given an introduction to the use of the CBL programs. The students then were given a set of tasks to complete on PowerPoint while utilizing the CBL. In order to capture the behavior and perception of each user, a log file of our CBL program was used to log every hit the participant makes. They were then asked to spend 2 hours interacting with the CBL program using a task. In this way, participants were free to choose their preferred navigational structures, index frames and map frame. Their interactions with the CBL were stored in a log file. The log file recorded participant movement and registered visited pages. Finally, a questionnaire was handled to participants to collect data about learners' perception of using the CBL program.

3.4. Data Analysis

In our study, we used the independent-samples t-test. We defined our independent values as gender and the dependent variables as: a) total number of frames pages visited from map frame (map-pages), b) total number of frames pages visited from index frames (index -pages). T-test was used because it compares the means of two groups, in our case males vs. females.

The novelty of our study is to investigate the learning preferences of different learning styles in using linear/non-linear navigation. A significance level of $p < 0.05$ was adopted for the studies. More specifically, the frequencies of using the non-linear/map and the linear/index between groups were analyzed.

4. Results and Discussion

4.1. Learning Style and Learners' Preferences

In our study, we investigated the learners' preferences in using the linear (index) and non-linear (map) navigational structures. We did this by comparing the number of pages visited by males and females (see Table 4). These pages are those from map and index. We found that there is no significant difference ($p > .05$) in preferences between males and females in using map pages. This means that Hypothesis 1 is rejected. However, females showed a preferences in using the map pages where their mean value = 13.60 which is more than males' mean values = 7.91.

On the other hand, when we tested the gender preferences in using index, we found that there is a significant difference ($p < .05$) between males and females in using index. We found that males preferred using index pages (mean = 12.26) more than females (mean = 3.16). This means that Hypothesis 2 is accepted.

Table 4: Frames pages visited from Map/index for females vs. males

Gender: M=Male, F=Female		Sig.	Mean	Std. Deviation
Total number of frames pages visited from Map	F	.121	13.60	9.284
	M		7.91	6.003
Total number of frames pages visited from Index	F	.000	3.16	3.823
	M		12.26	7.423

Since Field Independence (FI) users prefer navigational structures such as "index" and they prefer to use "find" to locate specific content [20]. Males can be identified as FI learners. Conversely, Field Dependence (FD) users tend to see a global picture [25], and prefer to use well-structured structures such as maps or main menus [10]. Thus, females are those who may identified as FD learners. However, investigation needs to be conducted to prove this.

As the main differences between the two cognitive styles, visualizers and verbalizers are described in Table 2 by Jonassen and Grabowski (1993) [19]. A visualizer prefers to receive information via graphics, pictures, whereas a verbalizer prefers processing information using words. Males can be identified as verbalizers because index frame provide learners the way of navigation when allocating key words for searching. Thus, females are those who may identified as visualizer learners where map frame provide the graphic presentation. However, this may also needs more investigation to be proved.

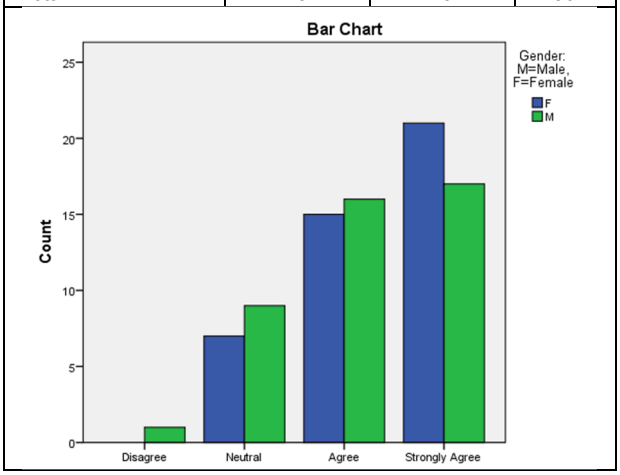
From Table 3, as previously discussed, we focused on the differences between holist and serialist learner. This table shows that the holist learners prefer the global approach to learning, simultaneous processing, spans various levels at once, conceptually orientated, and broad description building. Thus, females tend to have the holist style while using the non-linear structure. On the other hand, males tend to be serialist learners because they use the linear structure which provides local approach to learning, linear processing, works step by step, detail orientated and narrow procedure building which is provided in the index frame.

4.2. Learning Perception of CBL Program

Tables 5, 6, and 7 present results of questionnaire provided from learners. In Table 5, we found that the highest numbers of males (total of N=33) and females (total of N=36) had a positive perception in having the instruction provided by map where they mostly show their perception as "Agree" and "Strongly Agree".

Table 5: I like the fact that I have the ability to control thepace of instruction using Hierarchal Map.

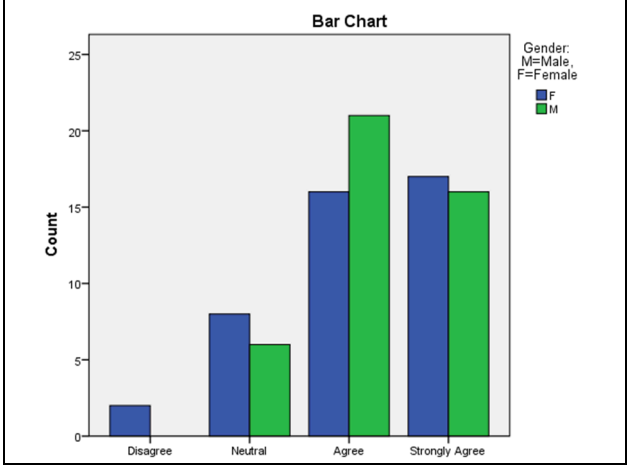
	Gender: M=Male, F=Female		Total
	F	M	
Disagree	0	1	1
Neutral	7	9	16
Agree	15	16	31
Strongly Agree	21	17	38
Total	43	43	86



In Table 6, we also found that the highest numbers of males (total of N=37) and females (total of N=33) had a positive perception in having the instruction provided by index as they mostly show their perception as "Agree" and "Strongly Agree".

Table 6: I like the fact that I have the ability to control the pace of instruction using index

	Gender: M=Male, F=Female		Total
	F	M	
Disagree	2	0	2
Neutral	8	6	14
Agree	16	21	37
Strongly Agree	17	16	33
Total	43	43	86

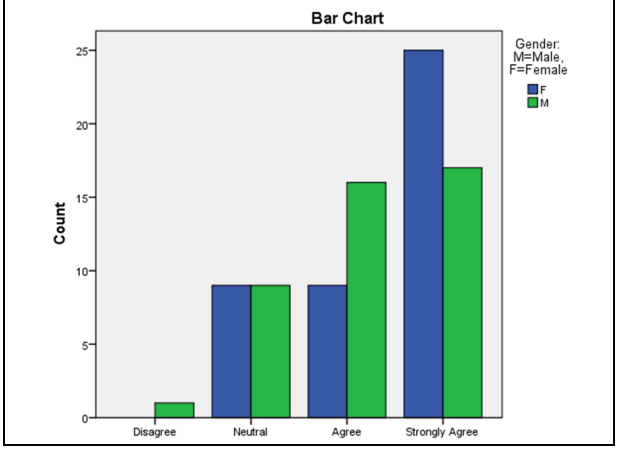


In Table 7, the results of question "I like the fact that I can see the both frames of navigational structures, map and index frames" is provided. We found that females had the

highest number (N=25) of "Strongly Agree" perception. Furthermore, both males and females liked the fact of seeing both frames of navigational structures, map and index frames because most of their responses were shifted to "Agree" and "Strongly Agree".

Table 7: I like the fact that I can see the both frames of navigational structures, map and index frames

	Gender: M=Male, F=Female		
	F	M	Total
Disagree	0	1	1
Neutral	9	9	18
Agree	9	16	25
Strongly Agree	25	17	42
Total	43	43	86



To summarize the previous discussions, it is clear that both males and females liked the fact that they can see both of the navigational structures map and index in the CBL program.

5. Conclusion

This paper highlighted various factors which influence learning when using and designing technology based learning. These factors include individual differences on learners such as gender, culture**,** and cognitive style.

Data analysis was used to understand learners' needs and perceptions in using navigational structures. Such investigation was done to explore the relationships between learners' needs, learning styles, perceptions, and preferences using the CBL program in regard to their gender.

In this study, we conclude that female learners need to use non-linear structure, while male learners counterparts need to use linear structure. The results also shows that females are more visualized, field dependent and holist learners, while males are more verbalized, field independent and serialist learners. The results of the questionnaire helped to understand learners' perception that they prefer having both navigational structures to be presented. This implies that a learner may use specific display accommodated in a CBL program although it may not be what they prefer. These findings emphasis that "*what learners like may not be what they need"* [26].

Understanding individual differences of learners and learner's characteristics will undoubtedly help designers to provide an effective technology based learning materials, in which users can acquire knowledge that will meet their individual needs, resulting in better perception and improved learning patterns. As a future work, further studies can be conducted utilizing data mining to provide a deep understanding of learners' needs and how this may affect their preferences.

The growing number of mobile applications in education adds more complexity to the design and development of educational interfaces with consideration to individual differences of learners. More research and more investigation will be an added value to this field of study for understanding learners and how this technology may affect learner's needs and preferences.

References

[1] R. S. Cobcroft , S. Towers, J. Smith and A. Bruns, "Mobile learning in review: Opportunities and challenges for learners, teachers, and institutions," Queensland University of Technology, Brisbane, 2006.

[2] I. Simonova and P. Poulova, "Plug-in Reflecting Student's Characteristics of Individualized Learning," *Procedia - Social and Behavioral Sciences,* vol. 171, p. 1235 – 1244, 2015.

[3] R. Al-Hajri and A. Al-Hunaiyyan, "Accommodating User Preferences in Designing Computer Based Learning Interfaces," Bangkok, Thailand, 2011.

[4] P. Poulová and I. Šimonová , "Didactic reflection of learning preferences in IT and managerial fields of study," New York, 2013.

[5] Y. Gülbahar and A. Alper, "Learning Preferences and Learning Styles of Online Adult Learners," in *Education in a technological world: communicating current and emerging research and technological efforts*, A. Méndez-Vilas, Ed., Ankara, Turkey, Formatex, 2011, pp. 270-278.

[6] P. Rurato and L. Borges Gouveia, "The importance of the learner's characteristics in distance learning environments: a case study," Barcelona, Spain, 2014.

[7] S. Y. Chen, "A cognitive model for non-linear learning in hypermedia programmes," *British Journal of Educational Technology,* vol. 33, no. 4, pp. 449-460, 2002.

[8] N. Ford, D. Miller and N. Moss, "The role of individual differences in internet searching: An empirical study," *Journal of the American Society for Information Science and Technology,* vol. 52, no. 12, pp. 1049-1066, 2001.

[9] P. Schumacher and J. Morahan-Martin, "Gender, internet and computer attitudes and experiences," *Computers in Human Behavior,* pp. 17(1), 95-110, 2001.

[10] S. Chen and N. Ford, "Modeling user navigation behaviors in a hypermedia-based learning system: An individual differences approach," *International Journal of Knowledge Organization,* vol. 25, no. 3, pp. 67-78, 1998.

[11] N. Ford and S. Chen, "Matching/mismatching revisited: An empirical study of learning and teaching styles," *British Journal of Educational Technology,* vol. 32, no. 1, pp. 5-22, 2001.

[12] S. Young and M. McSporran, "Confident men - successful women: Gender differences in online learning,," Dunedin, New Zealand, 2001.

[13] H. F. Atan, F. Sulaiman, Z. A. Rahman and R. M. Idrus , "Gender Differences in Availability, Internet Access and Rate of Usage of Computers among Distance Education Learners," *Educational Media International,* vol. 39, no. 3-4, 2002.

[14] M. Liu, "Examining the performance and attitudes ofsixth graders during their use of a problem-based hypermedia learning environment," *Computers in Human Behavior,* vol. 20, no. 3, pp. 357-379, 2004.

[15] P. Russo and S. Boor, "How fluent is Your Interface? Designing for International Users," Netherlands, Amsterdam, 1993.

[16] S. Messick, "Individuality in learning," San Francisco, 1976.

[17] E. Frias-Martinez, S. Y. Chen and X. Liu, "Investigation of behavior and perception of digital library users: A cognitive style perspective," *International Journal of Information Management,* vol. 28, p. 355–365, 2008.

[18] H. A. Witkin, C. A. Moore, D. R. Goodenough and P. W. Cox, "Field dependent and field independent cognitive styles and their educational implications," *Review of Educational Research,* vol. 47, p. 1–64, 1977.

[19] D. H. Jonassen and B. L. Grabowski, Handbook of Individual Differences, Learning, and Instruction, New Jersey: Lawrence Erlbaum Associates, 1993.

[20] N. Ford and S. Chen, "Individual differences, hypermedia navigation and learning: an empirical study," *Journal of Educational Multimedia and Hypermedia,* pp. 9, 281–311, 2000.

[21] N. Ford, F. Wood and C. Walsh, "Cognitive styles and online searching," *Online and CD_ROM Review,* vol. 18, no. 2, p. 79–86, 1994.

[22] R. J. Riding, Cognitive Styles Analysis. Learning and Training Technology, Birmingham, 1991.

[23] R. J. Riding and S. G. Rayner, Cognitive styles and learning strategies, London: David Fulton, 1998.

[24] F. P. Jing, *Interface Design for Hypermedia Learning Systems: A Study of Individual Differences and Hypermedia System Features,* London, UK: A thesis submitted for the degree of Doctor of Philosophy, School of Information Systems, Brunel University, 2005.

[25] H. A. Witkin and D. R. Goodenough, "Cognitive Styles: Essence and Origins," New York, 1981.

[26] C. G. Minetou, S. Y. Chen and X. Liu, "Investigation of the Use of Navigation Tools in Web-Based Learning: A Data Mining Approach," *International Journal of Human-Computer Interaction,* pp. 24(1), 48–67, 2008.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

TAPPS Release 1: Plugin-Extensible Platform for Technical Analysis and Applied Statistics

Justin Sam Chew

Colossus Technologies LLP, Republic of Singapore
chewjustinsam@gmail.com

Maurice HT Ling

Colossus Technologies LLP, Republic of Singapore
School of BioSciences, The University of Melbourne
Parkville, Victoria 3010, Australia
mauriceling@colossus-tech.com

Abstract

We present the first release of TAPPS (Technical Analysis and Applied Statistics System); a Python implementation of a thin software platform aimed towards technical analyses and applied statistics. The core of TAPPS is a container for 2-dimensional data frame objects and a TAPPS command language. TAPPS language is not meant to be a programming language for script and plugin development but for the operational purposes. In this aspect, TAPPS language takes on the flavor of SQL rather than R, resulting in a shallower learning curve. All analytical functions are implemented as plugins. This results in a defined plugin system, which enables rapid development and incorporation of analysis functions. TAPPS Release 1 is released under GNU General Public License 3 for academic and non-commercial use. TAPPS code repository can be found at http://github.com/mauriceling/tapps.

Keywords: *Applied Statistics, Technical Analysis, Plugin-Extensible, Platform, Python.*

1. Introduction

There are a number of statistical platforms available in the market today, which can be categorized in a number of different ways. In terms of licenses, some platforms are proprietary licensed (such as SAS [1] and Minitab [2]) while others are open source licensed (such as SOCR [3] and ELKI [4]). Between command-line interface (CLI) and graphical user interface (GUI), some are GUI-based (such as SOFA [5]) or CLI-based (such as ASReml [6]) while others (such as ELKI [4] and JMP [7]) have both GUI and CLI. Among the CLI-based platforms, some platforms (such as SciPy [8]) exist as libraries to known programming languages or develop their own Turing complete programming languages (such as R [9]) while others (such as TSP [10]) implements a domain-specific command language that does not aim to be a Turing complete language. The main advantage of domain-specific command language; of which, a popular example is SQL; is a shallower learning curve compared to a full programming language.

Here, we present TAPPS release 1 (abbreviation for "Technical Analysis and Applied Statistics System"; hereafter, known as "TAPPS") as a platform aimed towards technical analyses and applied statistics. TAPPS is licensed under GNU General Public License version 3 for academic and non-profit use. The main features of TAPPS are (1) a thin platform with (2) a CLI-based, domain-specific command language where (3) all analytical functions are implemented as plugins. This results in a defined plugin system, which enables rapid prototyping and testing of analysis functions. TAPPS code repository can be found at http://github.com/mauriceling/tapps.

2. Architecture and Implementation

Python programming language is chosen for TAPPS implementation due to its inherent language features [11], which promotes code maintainability. As of writing, one of the authors is also the project architect and main developer for COPADS (https://github.com/mauriceling/copads), a library of algorithms and data structures, developed entirely in Python programming language and has no third-party dependencies.

This section documents that architecture and implementation of TAPPS in a level of detail sufficient for interested developers to fork the code for further improvements.

TAPPS consists of 4 major sub-systems (Figure 1), supported by a third-party library. Users interact with

TAPPS though its command-line interface. The instructions are parsed into bytecode instructions using Python-Lex-Yacc [12]. The bytecodes are executed by TAPPS virtual machine to move data from the multi data frame object (data holder) to one or more plugins (functionalities), and load the results from plugins back to the multi data frame.

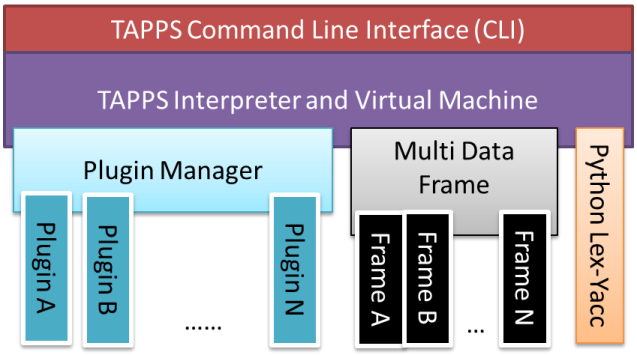


Fig. 1 Architecture of TAPPS.

Multi Data Frame. TAPPS uses the data frame object from COPADS (file: dataframe.py), which is essentially a 2-dimensional data table, as its main data structure (Figure 1). In practice, it is common to have more than one data frame object within a session of use. For example, the user can load one or more CSV files into TAPPS where each CSV file is maintained as a data frame. Furthermore, the user can slice a data frame into multiple data frames based on values or duplicate data frames. Hence, in order to accommodate for multiple data frames, COPADS' multi data frame object is used in TAPPS instead of using COPADS' data frame object directly as multi data frame object is a container for one or more data frame objects.

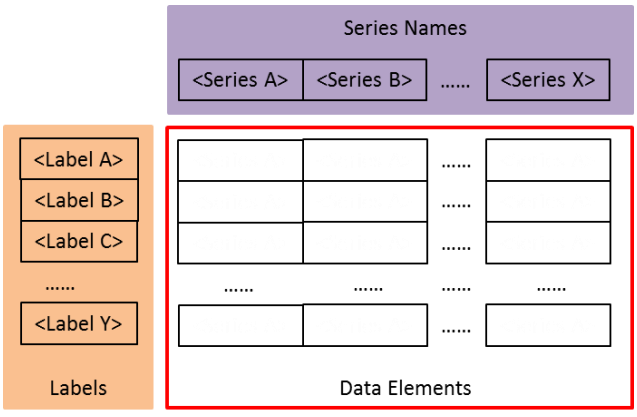


Fig. 2 Schematics of a Data Frame.

The core of a data frame (Figure 2) is a Python dictionary where the each data row is a value list with a unique key. This implies that each data row has a row label. Each data row within the same data frame must contain the same number of data elements. Each data element or value in a row is attributed to a series name, which implies that each series is a column in table form. This ensures that each data frame is either rectangular or square.

Plugins and Plugin Manager. As a thin platform, TAPPS has no analytical functions and does nothing in itself. All analytical functions, including statistical testing routines, are supplied via plugins. Hence, the operational richness of TAPPS is solely based on the repertoire of plugins. Each plugin is implemented as a standalone Python module, in the form of a folder consisting of at least 3 files – (1) the initialization file (file: __init__.py), which allows Python virtual machine to recognize the folder as an importable module; (2) a manifest file (file: manifest.py), which contains attributes to describe the plugin; and (3) an operation file (file: main.py), which must contain a "main" function as the execution entry point to the plugin. In addition, the "main" function takes a single "parameter" dictionary as input parameter, and returns a modified "parameter" dictionary back to TAPPS. All plugins must reside within "plugins" folder in TAPPS (please see Appendix B for a brief description on writing a plugin). At startup, the Plugin Manager will scan all individual sub-folders within the "plugins" folder as potential plugins, and attempts to load each potential plugin.

The "parameter" dictionary forms the communication link between TAPPS and its plugins; hence, an empty "parameter" dictionary for the specific plugin must be included in the main.py file. This is analogous to a customer (TAPPS) sending his/her laptop with instructions (the "parameter" dictionary) to a computer hardware shop (the plugin). The service staff in the hardware shop ("main" function of the plugin) pulls open the laptop, performs requested hardware and software upgrades or services (based on options in the "parameter" dictionary), then closes up the laptop and return the laptop with the service receipt/report as proof of initial service instruction to the customer (TAPPS). Hence, the "parameter" dictionary consists of three major components – the input data frame (the pre-serviced laptop), the results data frame (serviced laptop and service receipt/report), and a set of options that can vary with each plugin (the service requirements).

Command Line Interface. TAPPS command line interface (CLI) facilitates interaction between TAPPS and the user. The CLI can be launched in 2 different modes – interactive mode or script mode. In either mode, execution of the preceding command must complete before the next command can be executed (please see Appendix C for TAPPS language definition). In interactive mode, TAPPS behaves similarly to that of SAS [1] and R [9] CLI where

the user enters each command sequentially, and executes in the order of entry.

In script mode, also commonly known as batch mode, a plain-text script written in TAPPS language to represent a sequential batch of commands is given to TAPPS. This enables plugins implementing computationally intensive statistical algorithms [13], such as Monte Carlo methods [14], to run a job without the need of user intervention. An important feature of a script is that a script can incorporate one or more scripts (Figure 3); thereby, allowing re-use and improving the maintenance of scripts. This is achieved through an "include" pre-processor, "@include", which is syntactically identical to that ("#include") in C language. Operationally, the "include" pre-processor represents the substitution location of another script file. Hence, it is possible for recursive inclusion of script files provided no cyclic inclusion is observed.

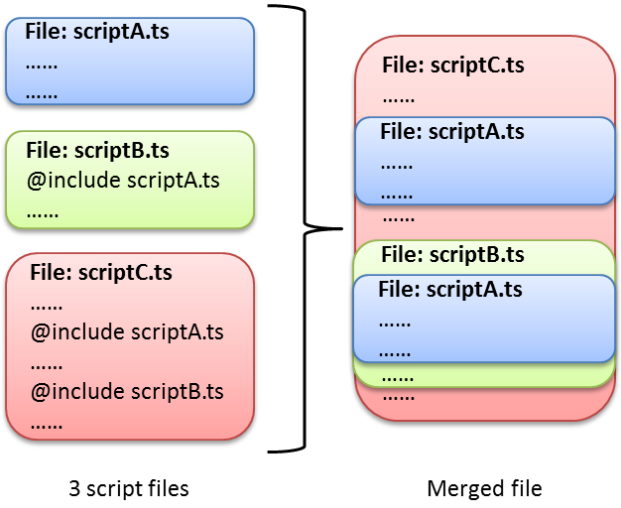


Fig. 3 Import and Merger of Script Files.

Interpreter and Virtual Machine. All TAPPS commands fed through the CLI, either as interactive mode or script mode, are processed by TAPPS interpreter (comprising of TAPPS lexer and parser, which uses Python-Lex-Yacc [12]) into TAPPS bytecodes (please see Appendix D for TAPPS bytecode definition). TAPPS bytecodes follow the *de facto* format [15] by beginning with an instruction (commonly known as opcode) and followed by zero or more operand(s).

The generated bytecodes are input to TAPPS virtual machine (TVM), which consists of 3 main components – (1) a set of environment variables implemented as a Python dictionary, which holds all the environmental settings; (2) a set of session variables implemented as a Python dictionary, which holds all data objects; and (3) TAPPS bytecode executor.

The set of session variables, which is implemented as a Python dictionary, consists of 3 main variables. Firstly, the multi data frame object (`session['MDF']`) to hold all data frames, which can be found within `session['MDF'].frames`. For example, a data frame with the ID of "DataA" will be accessible as `session['MDF'].frames['DataA']`. Secondly, all plugin parameter dictionaries are held as "parameters" (`session['parameters']`), which is implemented as a Python dictionary with the ID of the parameters as dictionary key. For example, a parameter set/dictionary with the ID of "pluginA_test" will be accessible as `session['parameters']['pluginA_test']`. Finally, all potential plugins will be recorded in `session['plugins']`, which is implemented as a Python dictionary. The list of successfully loaded plugins will be listed in `session['plugins']['loaded']` while the list of plugins that fail to be loaded will be listed in `session['plugins']['loadFail']`.

For each successfully loaded plugin, a new dictionary, with "`plugin_<plugin name>`" as key, will be created in session to contain information and the parameter dictionary associated with the plugin. For example, if a "ztest" plugin is successfully loaded, information regarding "ztest" plugin will be accessible as `session['plugin_ztest']`. At the same time, `session['plugin_ztest'] ['main']` will point to ztest's main function, and ztest's parameters dictionary can be found at and be copied from `session['plugin_ztest']['parameters']`.

TAPPS bytecode executor is implemented as large IF statement, which channels the bytecode operands (if any) to their respective supportive executors based on the opcode. Each supportive executor then takes the given operands and operates on the session variables and environment variables.

3. TAPPS Language

TAPPS Language defines 13 types of statements (Appendix C) and can be separated into 2 sections – meta-commands and operatives. Meta-command statements are mainly variables setting and display commands, used to manage the behavior or provide information with regards to the current session. The following explains the 4 meta-command statements:

- Describe: describe a data object or variable, which is more informative than "show" statement

- Set: set options in environment or within plugin parameters dictionary
- Shell: launch a language sub-shell or execute a non-TAPPS statement
- Show: show values of the session or environment variable/item.

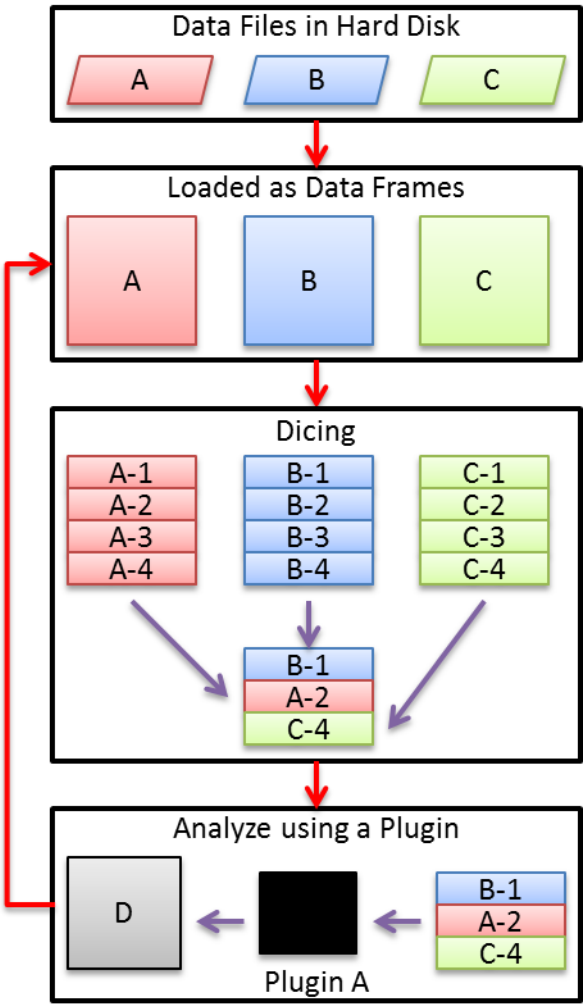


Fig. 4 Load-Dice-Analyze Cycle. Three separate data files (A, B, and C) are loaded into separate data frames (A, B, and C). Each data frame is divided into 4 smaller data frames; for example, Data Frame A is divided into Data Frames A-1, A-2, A-3, and A-4. One data frame from each of the original file (B-1, A-2, and C-4) are merged into a data frame, which is then analyzed by Plugin A. The output of Plugin A is Data Frame D, which can be attached as a data frame for the next cycle.

The operative statements manage the functionalities and operations of TAPPS; more specifically, the Load-Dice-Analyze cycle (Figure 4). Under this cycle, data is first loaded into TAPPS by reading in one or more data files. The next step is Dice, which is the extraction part of the loaded data and/or merge data segments together. For example, several loaded CSV files into data frames can be sliced into multiple smaller data frames by selection criteria before logically merging the sliced data frames into a larger data frame. Finally, the merged data frame can be analyzed by perform technical and/or statistical analyses using a plugin. As the output result of an analysis by plugin is a data frame object, it can be attached into the multi data frame object for a repeat of the Load-Dice-Analyze cycle.

The following explains the 9 operative statements:

- Cast: type cast values into a different data type
- Delete: delete a data frame from multi data frame object, or a plugin parameter set
- Load: load an external file or saved session
- Merge: merge 2 data frames by series or labels
- New: duplicate a plugin parameter set or attach a data frame within a plugin parameter set as a new data frame object
- Rename: rename series or label names within a data frame
- Runplugin: performs an analysis by running a plugin
- Save: save a session
- Select: generate a new data frame from a selection criterion on an existing data frame

The following is an example of a session:

1. Set up the environment (set current working directory to “data” directory relative to the startup directory)

```
TAPPS: 1> set rcwd data
```

2. Load a CSV file (load <TAPPS directory>/data/ STI_2015.csv and attach the data frame as “STI”)

```
TAPPS: 2> load csv STI_2015.csv as STI
```

3. Dice the data (type cast “Open” to from string to float; slice 2 data frames, STI_Low and STI_High, from STI; and merge STI_High and STI_Low into STI_A)

```
TAPPS: 3> cast Open in STI as nonalpha
TAPPS: 4> select from STI as STI_Low where
    Open < 820
TAPPS: 5> select from STI as STI_High where
    Open > 2000
TAPPS: 6> select from STI_Low as STI_A
TAPPS: 7> merge labels from STI_High to
    STI_A
```

4. Set up a plugin parameter dictionary (prepare plugin parameter dictionary from “summarize” plugin, name the parameter dictionary as “testingA”, and set to run “by_series” method on “STI_A” data frame)

```
TAPPS: 8> new summarize parameter as
    testingA
TAPPS: 9> set parameter analysis_name in
    testingA as trialA
TAPPS: 10> set parameter analytical_method
    in testingA as by_series
TAPPS: 11> set parameter dataframe in
    testingA as STI_A
```

5. Run the plugin

```
TAPPS: 12> runplugin testingA
```

6. Attach results data frame

```
TAPPS: 13> new STI_summarize dataframe from
    testingA results
TAPPS: 14> show dataframe

Current Dataframe(s) (n = 5):

  Dataframe Name: STI_High
  Series Names:
Open,High,Low,Close,Volume,Adj Close
  Number of Series: 6
  Number of Labels (data rows): 3919

  Dataframe Name: STI_Low
  Series Names:
Open,High,Low,Close,Volume,Adj Close
  Number of Series: 6
  Number of Labels (data rows): 3

  Dataframe Name: STI_A
  Series Names:
Open,High,Low,Close,Volume,Adj Close
  Number of Series: 6
  Number of Labels (data rows): 3922

  Dataframe Name: STI
  Series Names:
Open,High,Low,Close,Volume,Adj Close
  Number of Series: 6
  Number of Labels (data rows): 6996

  Dataframe Name: STI_summarize
  Series Names:
arithmetic_mean,count,maximum,median,minimu
m,standard_deviation,summation
  Number of Series: 7
  Number of Labels (data rows): 6

TAPPS: 15> describe STI_summarize

Describing Dataframe - STI_summarize

  Series Names:
arithmetic_mean,count,maximum,median,minimu
m,standard_deviation,summation
  Number of Series: 7
  Number of Labels (data rows): 6
  Series Name - arithmetic_mean
  Minimum value in arithmetic_mean:
2673.26085131
  Maximum value in arithmetic_mean:
92459705.7114
  Number of string data type values: 0
  Number of integer data type values: 0
  Number of float data type values: 6
  Number of unknown data type values: 0

  ------- <truncated> ------
```

7. Change current working directory, save a data frame and session before exit (change current working directory to serialize session as <TAPPS directory>/examples, saves STI_A data frame as a CSV file, and serialize current session to <TAPPS directory>/examples /tapps_manuscript.txt)

```
TAPPS: 16> set ocwd
TAPPS: 17> set rcwd examples
TAPPS: 18> save dataframe STI_A as csv
    STI_A.csv
TAPPS: 19> save session as
    tapps_manuscript.txt
TAPPS: 20> exit
```

4. Conclusion and Future Work

In this article, we present TAPPS as a plugin extensible tool for technical analysis and applied statistics, which is driven by a domain-specific command-line language. This article documents the architecture and implementation of TAPPS to document essential concepts, which are needed for further development. An example session of TAPPS use has been given as illustration.

The immediate future work is to develop plugins for TAPPS. For example, hypothesis test routines what are implemented in COPADS [16, 17] and existing Python implementations of econometrical methods [18] can be wrapped into plugins.

Acknowledgement

The authors wish to thank T. Yeo (Union Bank of Switzerland), J. Oon (University of Queensland), and C. Kuo (Yahoo) for their comments on the initial drafts of this manuscript.

References

[1] R. Levesque, “SPSS programming and data management: A guide for SPSS and SAS users (4th edition)”, SPSS Inc., 2007.
[2] R. K. Meyer and D. D. Krueger, “A Minitab guide to statistics (3rd edition)”, Prentice-Hall Publishing, 2004.
[3] I. D. Dinov, “Statistics online computational resources”, Journal of Statistical Software 16, 2006, pp. 1–16.

[4] E. Achtert, H. Kriegel and A. Zimek, “ELKI: A Software system for evaluation of subspace clustering algorithms”, Proceedings of the 20th international conference on Scientific and Statistical Database Management (SSDBM 08), 2008.
[5] J. Knight, "SOFA - Statistics open for all". Linux Journal 201, 2011, pp. 40–41.
[6] A. R. Gilmour, R. Thompson and B. R. Cullis, “Average Information REML, an efficient algorithm for variance parameter estimation in linear mixed models”, Biometrics 51,1995, pp. 1440-1450.
[7] B. Jones and J. Sall, "JMP statistical discovery software", Wiley Interdisciplinary Reviews: Computational Statistics 3, 2011, pp. 188–194.
[8] E. O. Travis, “Python for scientific computing”, Computing in Science & Engineering 9**,** 2007, pp. 10-20.
[9] F. Morandat and B. Hill, "Evaluating the design of the R language: objects and functions for data analysis", Proceedings of the 26th European conference on Object-Oriented Programming, 2012.
[10] B. H. Hall and C. Cummins, “TSP 5.0 reference manual”, TSP international, 2005.
[11] L. P. Coelho, “Mahotas: Open source software for scriptable computer vision”, Journal of Open Research Software 1, 2013, Article e3.
[12] S. Behrens, "Prototyping interpreters using Python Lex-Yacc", Dr Dobb's Journal 2004, 2004, pp. 30-35.
[13] P. Mantovan and P. Secchi, “Complex data modeling and computationally intensive statistical methods”, Springer-Verlag, 2010.
[14] A. Golightly and D. J. Wilkinson, “Bayesian inference for Markov jump processes with informative observations”, Statistical Applications in Genetics and Molecular Biology 14, 2015, pp. 169-188.
[15] K. A. Seitz, "The design and implementation of a bytecode for optimization on heterogeneous systems", Computer Science Honors Theses, 2012, Paper 28.
[16] M. H. T. Ling, “Ten Z-test routines from Gopal Kanji's 100 Statistical Tests”, The Python Papers Source Codes 1, 2009, Article 5.
[17] Z. E. Chay and M. H. T Ling, “COPADS, II: Chi-Square test, F-test and t-test routines from Gopal Kanji's 100 Statistical Tests”, The Python Papers Source Codes 2, 2010, Article 3.
[18] R. Bilina and S. Lawford, “Python for unified research in econometrics and statistics”, Econometric Reviews 31, 2012, pp. 558-591.

Appendix A: List of Statistical Package Available for All 5 Major Operating Systems

A list of 15 statistical packages, out of 57 listed statistical packages (https://en.wikipedia.org/wiki/Comparison_of_statistical_packages; accessed), are available for all 5 major operating systems; namely, Microsoft Windows, Mac OSX, Linux, BSD, and UNIX. Of these 15 packages that are available for all 5 major operating systems, 12 are open source packages. Four of the 12 open source packages (DataMelt, Dataplot, SciPy and Statsmodels) do not have a defined plugin system but by their exposure of internal APIs, presented an implicit plugin system. Of these 12 open source packages, only 2 packages (R and ROOT) have defined plugin system for addition of statistical analysis packages into the system.

Statistical Package	Open Source	License	Interface	Scripting	Plugin System
R	Yes	GNU GPL	CLI/GUI	R, Python, Perl	Yes
ROOT	Yes	GNU GPL	GUI	C++, Python	Yes
DataMelt	Yes	GNU GPL	CLI/GUI	Java, Jython, Groovy, Jruby, BeanShell	Partial
Dataplot	Yes	Public Domain	CLI/GUI	None	Partial
SciPy	Yes	BSD	CLI	Python	Partial
Statsmodels	Yes	BSD	CLI	Python	Partial
ADaMSoft	Yes	GNU GPL	CLI/GUI	None	No
OpenEpi	Yes	GNU GPL	GUI	None	No
PSPP	Yes	GNU GPL	CLI/GUI	Perl	No
Salstat	Yes	GNU GPL	CLI/GUI	Python	No
SOCR	Yes	GNU LGPL	GUI	None	No
SOFA Statistics	Yes	AGPL	GUI	Python	No
Statwing	No	Proprietary	GUI	None	Yes
Brightstat	No	Proprietary	GUI	None	No
TSP	No	Proprietary	CLI	None	No

Appendix B: How to Write a Plugin

A TAPPS plugin is a Python module with the following rules:

1. A plugin has to be a folder named in the following format: `<plugin name>_<release number>`, where `<plugin name>` has to be a single word in small caps and without underscore or hyphenation, and `<release number>` has to be an integer. For example, “generallinearmodel_1” is allowed while “General_Linear_Model_1” or “glm_1.0” or “GeneralLinearModel_1” are not allowed.
2. There must be the following files within the plugin folder – (1) `__init__.py`, (2) `manifest.py`, and (3) `main.py`.
3. Items in `manifest.py` must be filled in, as required.
4. The main file, `main.py`, must contain 3 items – (1) `instruction` string (a multi-line string to explain to users how to fill in the parameters dictionary); (2) a dictionary called `parameters`; and (3) a function called `main`.
5. The `parameters` dictionary must minimally contain the following key-value pairs – (1) `plugin_name`, which must correspond to the name of the plugin; (2) `dataframe`, for attaching a data frame as input to the plugin; and (3) `results`, for the plugin to attach a data frame as analysis output. The following key-value pairs are not mandatory but encouraged – (1) `analysis_name`, for user to give a short description of the analysis; (2) `analytical_method`, for the plugin to know which analysis function to execute is there is more than one analytical option; and (3) `narrative`, for user to give a long description of the analysis.
6. The `main` function (1) must take only `parameters` dictionary as the only parameter, (2) use only information provided in `parameters` dictionary for execution in order to maintain encapsulation, (3) must return the analysis output as a data frame into `results` key of `parameters` dictionary, and (4) must return `parameters` dictionary at the end.
7. The `main` function can call other functions or other modules.

A plugin template (https://github.com/mauriceling/tapps/tree/master/plugins/template_1), containing the essential boilerplate codes of a plugin has been provided. Plugin developers are encouraged to duplicate this folder as basic template for development.

The following code is a section of main file (`main.py`) from ‘`summarize`’ plugin (https://github.com/mauriceling/tapps/tree/master/plugins/summarize_1):

```
instructions = '''
How to fill in parameters dictionary (p):
Standard instructions:
    p['analysis_name'] = <user given name of analysis in string>
    p['narrative'] = <user given description of analysis, if any>
    p['dataframe'] = <input dataframe object from session['MDF'].frames>
Instructions specific to this plugin:
    1. p['analytical_method'] takes in either 'by_series' (summarize by
    series) or 'by_labels' (summarize by labels).'''

parameters = {'plugin_name': 'summarize',
              'analysis_name': None,
              'analytical_method': None,
              'dataframe': None,
              'results': None,
              'narrative': None}

def main(parameters):
    '''Entry function for the 'summarize' plugin.

    @param parameters: set of parameters, including data frame, which are
    needed for the plugin to execute
    @type parameters: dictionary
    @return: parameters
    @rtype: dictionary'''
    # Step 1: Pull out needed items / data from parameters dictionary
    method = parameters['analytical_method']
    dataframe = parameters['dataframe']
    results = parameters['results']
    # END Step 1: Pull out needed items / data from parameters dictionary

    # Step 2: Perform plugin operations
    if method == 'by_series' or method == None:
        (statistics, labels) = summarize_series(dataframe)
        results.addData(statistics, labels)
    if method == 'by_labels':
        (data, series) = summarize_labels(dataframe)
        results.data = data
        results.label = data.keys()
        results.series_names = series
    # END Step 2: Perform plugin operations

    # Step 3: Load dataframe and results back into parameters dictionary
    parameters['dataframe'] = dataframe
    parameters['results'] = results
    # END Step 3: Load dataframe and results back into parameters dictionary
    return parameters

def summarize_series(dataframe):
    '''Function to statistical summaries of each series in the dataframe.'''
    labels = dataframe.data.keys()
    series = dataframe.series_names
    statistics = {'summation': [], 'arithmetic_mean': [],
                  'standard_deviation': [], 'maximum': [],
                  'minimum': [], 'median': [], 'count': []}
    for index in range(len(series)):
        sdata = [float(dataframe.data[label][index]) for label in labels]
        statistics['summation'].append(sum(sdata))            # 1. summation
        arithmetic_mean = float(sum(sdata)) / len(sdata)      # 2. arithmetic mean
        statistics['arithmetic_mean'].append(arithmetic_mean)
        index = int(len(sdata) / 2)                           # 3. median
        statistics['median'].append(sdata[index])
```

```
        sdata.sort()                                            # 4. maximum
        statistics['maximum'].append(sdata[-1])
        statistics['minimum'].append(sdata[0])                  # 5. minimum
        statistics['count'].append(len(sdata))                  # 6. count
        if len(sdata) < 30:                                     # 7. standard deviation
            s = float(sum([(x-arithmetic_mean)**2 for x in sdata])) / (len(sdata) - 1)
        else:
            s = float(sum([(x-arithmetic_mean)**2 for x in sdata])) / len(sdata)
        statistics['standard_deviation'].append((s**0.5))
    return (statistics, series)

def summarize_labels(dataframe):
    '''Function to statistical summaries of each label in the dataframe.'''
    series = ['arithmetic_mean', 'count', 'maximum', 'median', 'minimum',
              'standard_deviation', 'summation']
    data = {}
    for label in dataframe.data.keys():
        sdata = [float(x) for x in dataframe.data[label]]
        temp = ['arithmetic_mean', 'count', 'maximum', 'median', 'minimum',
                'standard_deviation', 'summation']
        temp[0] = float(sum(sdata)) / len(sdata)      # 1. arithmetic mean
        temp[1] = len(sdata)                          # 2. count
        sdata.sort()                                  # 3. maximum
        temp[2] = sdata[-1]
        temp[3] = sdata[int(len(sdata) / 2)]          # 4. median
        temp[4] = sdata[0]                            # 5. minimum
        if (len(sdata) < 30) and (len(sdata) > 1):    # 6. standard deviation
            s = float(sum([(x-temp[0])**2 for x in sdata])) / (len(sdata) - 1)
        elif len(sdata) == 1:
            s = 'NA'
        else:
            s = float(sum([(x-temp[0])**2 for x in sdata])) / len(sdata)
        temp[5] = s ** 0.5
        temp[6] = sum(sdata)                        # 7: summation
        data[label] = [x for x in temp]
    return (data, series)
```

Appendix C: TAPPS Language (Release 1) Definition

The following is TAPPS language (release 1) definition in Backus-Naur Form where reserved words are in bold:

```
statement : cast_statement | delete_statement | describe_statement | load_statement
          | merge_statement | new_statement | rename_statement | runplugin_statement
          | save_statement | select_statement | set_statement | shell_statement
          | show_statement
cast_statement : CAST id_list IN ID AS datatype
delete_statement : DELETE DATAFRAME ID | DELETE PARAMETER ID
describe_statement : DESCRIBE ID
load_statement : LOAD CSV FILENAME AS ID | LOAD NOHEADER CSV FILENAME AS ID
               | LOAD SESSION FROM FILENAME
merge_statement : MERGE SERIES id_list FROM ID TO ID
                | MERGE LABELS FROM ID TO ID | MERGE REPLACE LABELS FROM ID TO ID
new_statement : NEW ID PARAMETER AS ID | NEW ID DATAFRAME FROM ID plocation
rename_statement : RENAME SERIES IN ID FROM id_value TO id_value
                 | RENAME LABELS IN ID FROM id_value TO id_value
runplugin_statement : RUNPLUGIN ID
save_statement : SAVE SESSION AS FILENAME | SAVE DATAFRAME ID AS CSV FILENAME
select_statement : SELECT FROM ID AS ID | SELECT FROM ID AS ID WHERE binop value
                 | SELECT FROM ID AS ID WHERE ID binop value
set_statement : SET DISPLAYAST ID | SET CWD FOLDER | SET SEPARATOR separators
              | SET FILLIN fillin_options | SET PARAMETER ID IN ID AS ID
```

```
              | SET PARAMETER DATAFRAME IN ID AS ID | SET RCWD ID | SET OCWD
shell_statement : PYTHONSHELL
show_statement : SHOW ASTHISTORY | SHOW ENVIRONMENT | SHOW HISTORY | SHOW PLUGIN LIST
               | SHOW PLUGIN ID | SHOW SESSION | SHOW DATAFRAME | SHOW PARAMETER
id_list : ALL | ID | id_list DELIMITER ID
value : number_value | id_value
binop : DELIMITER | GE | LE | EQ | NE
datatype : ALPHA | NONALPHA | FLOAT | REAL | INTEGER
fillin_options : NUMBER | ID
id_value : ID | STRING
number_value : NUMBER
plocation : RESULTS | DATAFRAME
separators : DELIMITER | COMMA | COLON | SEMICOLON | RIGHTSLASH | BAR | DOT | PLUS | MINUS
           | TIMES | DIVIDE | GT | LT
```

Appendix D: TAPPS Bytecode (Release 1) Description

The following table describes TAPPS bytecodes (release 1):

Opcode	Operand(s)	Description
cast	<data type>, <series names>, <data frame name>	Type cast all data elements for a specific set of series names within a data frame.
deldataframe	<data frame name>	Deletes data frame from multi data frame object.
delparam	<parameter name>	Deletes plugin parameter(s) dictionary.
describe	<data frame name>	Describe a data frame – more informative than "show".
duplicateframe	<original data frame name>, <new data frame name>	Duplicate/deep-copy an existing data frame and attach the copy into multi data frame object with a new data frame name.
greedysearch	<data frame name>, <new data frame name>, <binary operator>, <value>	Select data from a data frame into another data frame, where a data element in any series, identified by the same data label, holds true for the selection criterion.
idsearch	<data frame name>, <new data frame name>, <series name>, <binary operator>, <value>	Select data from a data frame into another data frame, where a data element in a specific series, holds true for the selection criterion.
loadcsv1	<file name>, <data frame name>	Load CSV file, with headers as series name, and attach the data frame.
loadcsv2	<file name>, <data frame name>	Load CSV file, without headers, and attach the data frame.
loadsession	<file name>	Loads a previously saved session from file.
mergelabels1	<source data frame name>, <destination data frame name>	Merge data, without replacing any data in destination data frame, from source data frame to destination data frame.
mergelabels2	<source data frame name>, <destination data frame name>	Merge data from source data frame to destination data frame. Existing data within the destination data frame will be replaced with data from source data frame, if the label(s) in the destination data frame is/are found in the source data frame.
mergeseries	<series names>, <source data frame name>, <destination data frame name>	Merge data from one or more series from source data frame to destination data frame.
newdataframe	<new data frame name>, <parameter name>, {dataframe | results}	Duplicate/deep-copy the input data frame or output results from a parameters dictionary, and attach as a new data frame.
newparam	<plugin name>, <new parameter name>	Duplicate/deep-copy parameters dictionary, for use, from a loaded plugin.
pythonshell	No operand	Launch a Python interactive shell.
renameseries	<data frame name>, <original series names>, <new series names>	Rename series within a data frame.
renamelabels	<data frame name>, <original label names>, <new label names>	Rename label(s) within a data frame.
runplugin	<parameter name>	Run plugin using parameters dictionary; name of the plugin is coded within the parameters dictionary.
savecsv	<data frame name>, <file name>	Saves a data frame into a CSV file.
savesession	<file name>	Saves the current session into a file.
set	<variable>, <value>	Set options in environment or within plugin parameters dictionary.
show	<item>	Show values of the session or environment variable/item.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Pictographic steganography based on social networking websites

Feno Heriniaina R.[1], Xiaofeng Liao[2]

[1]College of Computer Science, Chongqing University
Chongqing, 400044, China
fenoheriniaina@cqu.edu.cn

[2] College of Computer Science, Chongqing University
Chongqing, 400044, China
xfliao@cqu.edu.cn

Abstract

Steganography is the art of communication that does not let a third party know that the communication channel exists. It has always been influenced by the way people communicate and with the explosion of social networking websites, it is likely that these will be used as channels to cover the very existence of communication between different entities. In this paper, we present a new effective pictographic steganographic channel. We make use of the huge amount of photos available online as communication channels. We are exploiting the ubiquitousness of those social networking platforms to propel a powerful and pragmatic protocol. Our novel scheme exploiting social networking websites is robust against active and malicious.

Keywords: *Information hiding, steganography, online social networking, pictogram*

1. Introduction

The art of covered writing or steganography has a long history and has always been influenced by the way people communicate. It is often understood as the prisoners` problem where two inmates Alice and Bob, imprisoned in two different cells are trying to formulate a plan for escape. The only way to communicate with each other is through a channel that is monitored by the Warden. Alice and Bob should then find a way for communicating under monitoring without raising the Warden`s suspicion. The expansion of the virtual world has created a wide range of possibilities in the world of secret communication. Although, cryptography is enough to keep the content of a given information unreadable from praying eyes. If the Warden (also referred as person in the middle) who monitors the communication is not favorable to encrypted data, and whenever he knows that something encrypted and suspicious is within the channel, he might decide to stop and interrupt the communication (Fig. 1). Given such situation, the communicating parties should rather use steganography to make the very existence of the message transmission unnoticed.

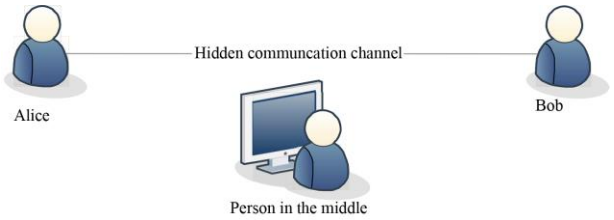


Fig. 1. Hidden communication channel between two entities (Alice and Bob) with the presence of a warden (the person in the middle).

Steganography has been recorded in the literature since the 1953 [1-2] and has evolved a lot to adapt into the 21st century. There are different types of technical steganography: text, audio, images and video [3] and among the most commonly used and studied to date are all based on images.

This is no big surprise because of the huge amount of photos already available online and the new uploads generated by social networks and photo services users. Publishing and sharing images online have now become very easy. Yet the trend in enhancing the image quality and in making the process of high quality content creation much simpler is not stopping, as almost all mobile phone manufacturers are now trying to provide a good embedded camera as a main component of their products. Above all that, higher and better computation resources are available on most mobile handheld devices nowadays, and steganography based on images would be more promising than ever.

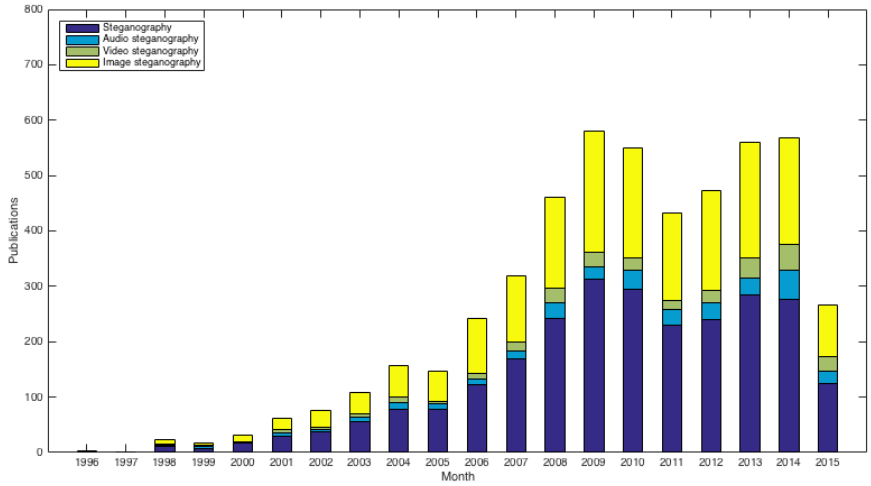


Fig. 2. Growth of the field of steganography witnessed by the number of articles annually published by IEEE that contains the keywords "steganography", "audio steganography", "video steganography", and "image steganography".

Many techniques and research papers about image-based steganography have been made available to the public to date. A simple search query with the key words "image steganography" at ieeexplore.ieee.org, combining journals and conference research papers from 1996 to 2015 gave a total number of 1855 in September 2015. A comparison using other search keywords related to steganography is represented in the figure 2. In term of digital image steganography, many are based on the exploitation of the least significant bits embedding as this is one of the easiest to implement [4] and the identification and replacements of the redundant bits [5]. Almost all these methods recorded in the literature focused on how to hide enough information within an image because an image consists of a large number of individual pixels that can be imperceptibily modified to encode a secret message. But the transmission of a single stego-image would require a great covering technique to go unnoticed from a warden that could apply steganalysis [6], it is important that the produced stego image has no drastic change in the image spectrum in order to preserve the quality as similar to the cover image as possible [7]. Nowadays, sharing photos and other digital media is part of our daily life. The intense use of social networking websites and the millions of photos uploaded everyday have opened new opportunities to an old communication technique to meet the digital world. So far, no one to the best of our knowledge has presented a digital image steganographic method that is leaving the cover image untouched to vehiculate a secret message and respect the most important property of steganography, i.e: undetectability.

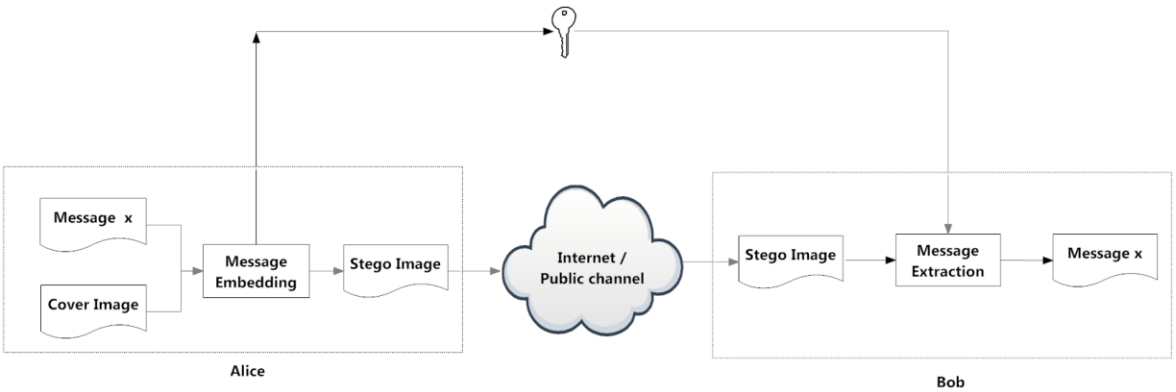


Fig. 3. General summary of image steganography

In the remaining of this paper, we will introduce a pictographic steganography scheme using online social networking with Facebook as our case study. In section 2 we provide an overview on the social networking mechanism. In section 3 we review the current work related to image steganography. Section 4 introduces our image steganography using pictograms. In section 5 we provide a security analysis of the proposed protocol. Section 6 is a study on the performance of the presented pictographic protocol. In section 7 we summarized the main advantages of the proposed steganographic system. The last section 9 is devoted for the discussion and conclusion.

2. Social networking system: creation, utilization and notification.

Online social networking has drastically changed the way Internet users interact with one another. One main characteristics of social networking sites is the free user registration, the user generated content, sharing and collaboration, and the convergence of intelligence.
Facebook (https://facebook.com) is one of the internationally used online social networking websites. Despite its restrictions in some areas, Facebook has users in almost all parts of the globe ranging from the economically poorest to the richest countries. Creating a personal account on Facebook is totally free and a registered user can upload images, post contents and connect to any other Facebook users without physical border limitation. Facebook users can set the security settings of their accounts based on their preferred privacy requirements [8]. This will define the users` account visibility to other Facebook users and the general Internet users. Given a working Facebook personal account, one can create a Facebook page and post in some news, stories and other updates targeting some specific people. Those could click on the “Like” button and become followers of the page which is also referred as a Facebook page in this paper. Following a page will allow them to receive instant notifications each time there is an update.

Using Facebook as a case study for our communication channel between Alice and Bob located in different places, we present the following stego-system to allow them communicate without anyone knowing the existence of the used communication channel. For the scheme to be effective, both Alice and Bob should know the communication protocol prior the real communication using the presented stego-system. They should also know how to read the images, which we will introduce later, as only in knowing how to read the image will one be able to retrieve the secret message.

3. A review of image steganography

Embedding and extracting algorithm are the techniques through which the secret messages are embedded and extracted. Being an important part of a steganography system, there are three fundamental ways that determine its internal mechanism, i.e: cover synthesis, cover selection and cover modification.

- Steganography by cover synthesis: a cover is generated for the sole purpose of concealing a secret message.
- Steganography by cover selection: from a set of data, a cover that conceals best the conveyance of the secret information is chosen.
- Steganography by cover modification: regular data is changed or substituted with the secret information. Depending on the type of cover and the amount of hidden data, the substitution method can degrade the quality of the original cover drastically.

The latter is the most extensively studied steganography model to date [9-12]. Many of the released steganography techniques are based on the least significant bits replacement and / or matching [13-14]. Some are a combination of cover selection and modification [15-17]. Such steganography techniques involve the modification of the cover image to produce the stego-image [14]. In general, the steganography techniques based on images released to date are embedding the message into the cover in order to get the stego-file. The extraction of the hidden message is later feasible for those who know the extraction algorithm and have access to the key (Fig 3).

In this era of the digital world, the Internet is generally the public channel used for transmitting the message in stego format from Alice to Bob. However, there could exist a third party observer in this public channel. A pure steganography scheme will make an assumption that the third party observer also known as the warden does not have any idea of the embedding and decoding algorithm. Robustness then is referred as the difficulty of removing the hidden information from the stego object. While removal of secret data might not be a much serious problem as its detection, robustness is a desirable propriety when the communication channel is distorted by random errors or by systematic interference with the aim to prevent the use of steganography [18]. The warden can also interfere between the communicating parties. Based on its involvement, we assume three different roles for the warden: a passive warden, an active warden and a malicious warden [10, 18].
Usually, a passive warden simply examines the message and tries to determine if it potentially contains a hidden message. The warden has read only access to the message being sent. If he suspects the existence of a secret

message, then the document is stopped. Otherwise it will go through without any further inspection [18].
But an active warden and a malicious warden are often times merged into one. An active warden refers to a third party that attempts to disrupt the steganographic channel by modifying the message. In case of an image file to be transferred, the warden might try to compress or resize the file. Under such condition, any steganographic scheme would be broken unless it is robust to such processing. A malicious warden does not necessarily intend to entirely disrupt the stego channel, but rather uses it to determine whether or not steganography is taking place [10, 19].

Finally, embedding information on photos uploaded on most social networking websites is tough as different processing such as compression is made before having the photos available worldwide, in this paper we introduce this steganographic scheme using pictograms which is resistant to both compression and cropping as long as the main content of the image remains visible.

4. Image steganography using pictograms:

The proposed scheme has three main steps: converting the message into pictogram, uploading the images and from the receiver side only reading the pictogram is required.

4.1 Converting a text into pictogram:

Both Alice and Bob should know how to read the pictogram images and only their creativity is the limit on how to represent things using images. The example below presents a sample of a word conversion into an image.

P denotes a finite set of possible plain text,
I denotes a finite set of images, y is each photo $\in I$,
Enc is the encoding rule,
Dec is the decoding rule,
Each $\mathrm{Enc}: \mathrm{P} \rightarrow \mathrm{I}$ and $\mathrm{Dec}: \mathrm{I} \rightarrow \mathrm{P}$ are functions such that $\mathrm{Dec}(\mathrm{Enc}(\mathrm{x})) = x$ for every word element $x \in P$.

Example:

$$Enc(car) = y_{photo\ of\ a\ car} \quad (2)$$

$$Enc(love) = y_{photo\ of\ a\ heart} \quad (3)$$

$$Enc(money) = y_{photo\ of\ a\ bank\ note} \quad (4)$$

In the above example, we chose a straightforward encoding from a word to an image representative of its meaning. Instead of directly $Enc: P \rightarrow I$, an additional substitution could precede this step giving a rather uncorrelated meaning to the encoding output. A shifting method based on the ordering or location of the words in a chosen dictionary is proposed as presented in figure 3. The key for the encryption and decryption of the image (message) is the combination of the shift cipher and the selected dictionary.
The structure of the dictionary can be considered like one of those empirical hardcopy books mainly used in the year nineties; such an example is the Oxford dictionary 2nd edition. It will be used to locate the exact position of the word derived from the image pictogram. The shift cipher is used for the mapping of the plain word to its cover (encrypted version).
For a correct shift mapping, a word should first be located in the chosen dictionary. For a shifting based on the location of the words in the dictionary, the exact position of the word on the page will be localized. The shift number will be used to map this exact position on the corresponding page-shift; which is the page number of the word location added with the shift number and then modulo n. n is the total page number of the dictionary and $\mathbb{Z}_n$ forms the arithmetic modulo n of set $\{0, \ldots, n\text{-}1\}$.

S denotes a finite set of possible shift cipher. For each $s \in S$ there is an encryption rule $e_s \in E$ and a corresponding decryption rule $d_s \in D$,

$$e_s: P \rightarrow I \text{ and } d_s: I \rightarrow P,$$

are functions such that $d_s(e_s(x)) = x$. (1)

Additionally, because each event is dated on most social networking websites, including our case study Facebook, the date of upload could be used as mask to be added to the shift cipher each time an image needs to be uploaded.

Combining the encoding-encrypting parts above, we have $\mathrm{Enc}\,(e_s(x))$ and the decoding-decrypting is $d_s(Dec(y))$.

4.2 Uploading the images in a manner that facilitates the reading

When uploading a batch of images on Facebook, by default the ordering is done according to the image file name in the local machine with the assumption that the arrangement is also set based on that. Each update given to a Facebook page has to have a time stamp when published and the availability to the public is based on the first uploaded first visible. Let us assume Alice would like to send the message "money loves car" to Bob. M be the message and x_i each word that makes it. The number of images to upload to transcribe the message M is i and it is the same as the number of $x_i$needed to be uploaded. After a minute selection of the images representatives of the message to be sent, if the image-upload is done in a batch, Alice should rename the images based on how the ordering of the reading should be done. In the previous example, were $M = x_1 + x_2 + x_3$, the name of the images to upload together in one batch should be organized so that

the image to be viewed first has the smallest name number. If Alice chooses to use a name with numbers, it should look something like: IMG_1, IMG_2 and IMG_3.
Doing so will permit the correct reading of the message transmitted in a general manner such that:

$$\text{Dec}(y_1)+ \text{Dec}(y_2)+\cdots+ \text{Dec}(y_n)= x_1 + x_2 + \cdots + x_n \quad (5)$$

4.3 Reading the pictogram

In the next section, we introduce two protocols that can be used for this proposed steganography using pictograms on Facebook. The way for retrieving and reading the message will vary based on the chosen protocol. A summary of the two protocols are presented in the table below:

Table 1: the protocols for knowing whether a new message has been made available

Protocol One	Protocol Two
Alice has a Facebook account.	Alice has a Facebook account.
Bob has a Facebook account.	-
Alice creates a Facebook page.	Alice creates a Facebook page.
Bob follows (likes) Alice`s Facebook page.	Bob knows the links to Alice`s Facebook page.
Each time Alice posts something on her page, Bob should receive a notification.	Bob has to check periodically Alice`s Facebook to find out whether new posts have been published.

If Alice and Bob chose to use the proposed protocol one, both Alice and Bob should have a Facebook account. Alice will need to create a Facebook page and Bob should follow her page. Each time Alice makes an update (writing a post or uploading a photo) on her Facebook page, Bob should instantly get a notification information. This notification is the trigger for Bob to view what was the update about and to reconstruct the message if some pictographic images were uploaded.

In the case the protocol two is selected, only Alice needs to have a Facebook account and to create a Facebook page. Bob should visit Alice`s page periodically and see whether or not some updates that cover-up a message were uploaded. Although this method seems lacking the notification system to inform Bob, it has the advantage of strengthening the proposed steganographic system due to the absence of direct connection between Alice and Bob.

5. The proposed protocols

We introduce and analyze two schemes from which we compare based on how they preserve discretion. Prior to the establishment of communication, both Alice and Bob should already have access to the shift cipher and the dictionary. The difference between the schemes proposed is in the way Bob gets to know whether a new message has been released.

Protocol One: Both Alice and Bob should each own a Facebook personal account. Alice will create a public Facebook page where she can upload photos. Being a Facebook user, Bob would follow Alice by liking her Facebook page. When Bob starts following Alice`s page, he will be notified each time the page gets updated and he should just visit and check what was updated to reconstruct the message if something was sent. For this given protocol channel, what matters most is the photos that are uploaded on Alice`s Facebook page as they drive the information. It is important that Bob is able to view the photos in a timely and arranged manner.

Protocol two: Alice should own a Facebook personal account and create a public Facebook page where she could upload photos. The photos uploaded being the main message in the form of pictogram, Bob only needs to know the url address to Alice`s page, and needs to visit the given page in a timely manner. All posts in Facebook and including Facebook pages are time stamped, and knowing the ordering of the image uploaded will allow Bob to reconstruct the complete correct message by viewing the images in order.

6. Security of the proposed steganographic system

Our main focus while referring to steganography is to provide a secure channel where a successful attack would consist of a warden being able to identify that a given image uploaded on the network drives a hidden message to the viewer. Law enforcements and intelligence agencies have always been having difficulties deciding which electronic channel to scan and intercept because of the huge volume of traffic [20].
The use of social networking website will just increase the level of this difficulty as the amount of photos being uploaded on those online photo sharing, and social networking websites every day is exceeding 2 terabytes with a peak of 300 000 images served per second in 2008 [21]. Also, although we are using images to drive the message from Alice to Bob, we don`t introduce any additional artifact to those images. The well-known and standard steganalysis algorithms: regular and singular group analysis [22], sample pair analysis [23] and difference in the images histograms [24], are not applicable to our scheme as we are not introducing any artifacts in the core of the images. Relying on the cover channel undetectability, using a popular Facebook page as a network cover channel should not create any suspicions and strengthen the hiddenness of the proposed steganography system.

To improve the complexity of the decoding of the image (reading of the pictogram), further scrambling the ordering of the display of the images in one Facebook page album could be achieved by first uploading the photos in a batch in the correct reading order and then later changing the upload time and dates by editing each image description. Supposing a warden has access to all the pictures that make the complete message and has the correct dictionary allowing $\mathrm{Dec}(\mathrm{Enc}\,(\mathrm{x})) = x$. It is still required that the correct ordering of the image be found too, to properly make sense of all the images. In this case, only one arrangement out of n- permutations of the images which make the complete message will correspond to the correct message. So if we have the image set $E(y_1, \dots, y_n)$ that makes the message $x_1, \dots, x_n$, then we have :

$$p(n,k) = \frac{n!}{(n-k)!}, \quad (6)$$

$$with\ k = n, \quad (7)$$

$$p(n,k) = \frac{n!}{0!} = n! \quad (8)$$

With n images to make the complete message and $n!$ possible permutations, the longer the message is the better its security on an exhaustive search becomes.

If we considered that a robust steganographic approach should define its security based on the chosen secret key as presented in “La cryptographie militaire” In other words, if we assume that the protocol in use is known to a warden, and so the security must lie on the choice of the secret key (shift cipher for the encryption/decryption and the dictionary for encoding/decoding word to image) that Alice and Bob have managed to share based on Kerckhoffs principle [25]. An additional step for implementing a one-time dictionary would be a good the remedy.

For this scheme to be most effective, we consider the worst case of a warden that could intervene as soon as a known communication channel is suspected to exist between Alice and Bob. Considering our case study, the platform and service provider is at the highest level in knowing whether a direct relationship exists between Alice and Bob. With the stegosystem depending on Facebook, it could be very easy for them to analyze users`log data containing information about browsing habits [26].

When creating a Page, we are giving Facebook information about our interests. For the proposed method one, where Bob should like Alice`s page in order to instantly get notification each time Alice has some updates, a direct connection has to be established. Unlike some users who tend to feed their online profile with a tone of information that aim at providing a complete and accurate representation of themselves; in the proposed stegosystem, avoiding direct connection between Alice and Bod is a must in order to increase the degree of separation.

The proposed protocol one does require a connection between Alice and Bob with this later needing to like the Page. But, the method two does not necessitate Bob to have this direct connection with Alice. Bob could simply access Alice`s Facebook page without being logged in to his own Facebook personal account. Although possible individualization could be achieved based on the connecting IP address, used operating system, fingerprinting and browser information [27], no direct identification of Bob within the social nodes could be easily realized.

Another potential imminent third party threat on Bob`s side is the internet service provider, which has access to all the url requests he makes. But if being very active on Facebook, visiting different pages repeatedly is just a normal activity that users do especially when it comes to a Page that is frequently updated and getting high traffic.

To reduce the risk of the channel interruption and clear any doubts or suspicions of any communication between Alice and Bob, we found out that the more active and engaging the content of a Facebook page is, the more disguised and innocent looking the communication channel remains. This is very important because Alice should constantly maintain her page but not only bring some updates when an important message needs to be transmitted.

It is also really important to choose a topic that interests a great number of audience and that deserves a great amount and varied image galleries. In our experiments, we have categorized Facebook pages in two categories; one for posts that gives joy and another that shows violence and sadness. It has been noticed that most pages that bring good feelings are receiving more likes (followers) and are highly viral on the network. Creating pages that will have engaging followers will make the proposed scheme hard to detect and despite its simplicity it makes it more efficient.

7. Performance of the proposed steganography scheme

The performance of the covered channel proposed in our scheme could be analyzed based on the three (3) main indexes, characteristics of a steganographic covert channel: capacity, robustness and undetectability. Capacity represents the data sent per time unit by using a given method. Robustness refers to the amount of manipulation the stego-file can withstand while still being able to preserve the information to be transmitted. Undetectability represents the main security of any steganographic

channels; in other word, it is the ability to keep the communication channel hidden.

Capacity: with pictograms being ideograms, they are very flexible in terms of complexity. In our proposed scheme, we derived words from what the images represent. In this sense, each image can correspond to one word. And as images can be sent in a batch to make a gallery in a given social networking website, a presumably long message could be transmitted at each creation of a gallery. Moreover, unlike other image based steganography where each stego-image (which is the cover image + the message) is fed at the limit till avoidance of statistical steganalysis, this proposed scheme does not require any artifacts to be applied on the cover image making it free of any risks from the most renowned steganalysis in the literature to date.

Robustness: unlike the other image based steganography methods that embed information within the cover, the proposed scheme enables a message transmission while leaving the cover image untouched. Compression, moderate contrast and brightness manipulations, which are really common practices won`t have any effect on the proposed scheme. Even cropping is tolerated up to the point that the main content of the image expressive of the ideogram is still visible.

Undetectability: using images that are presumably free of any artifacts, no steganalysis related work has been made available to fight against pictogram images online to date to the best of our knowledge. This Pictographic Steganography Based on Social Networking Websites requires no extra binary to be added to the cover image. Only identification of the ideogram with no further processing of the image is enough whether on the sender or the receiver side. Social networking websites like Facebook are extensively used for posting photos. Releasing pictograms on such platform would go unnoticed especially if the Page is well maintained and has a lot of visitors. The implementation of the proposed protocol does not have to be limited on a single Facebook Page. Multiple Pages could be used to enable a diversification of the image used.

8. The main advantages of the proposed steganographic system

Communication from one to many: Social networking websites main purpose is to connect people without border limitation. This could be used in the proposed steganographic system where Alice would publish the images that convey the message and anyone knowing the exact protocol and have the key shall be able to retrieve and make a proper reconstruction of its content (Fig. 4).

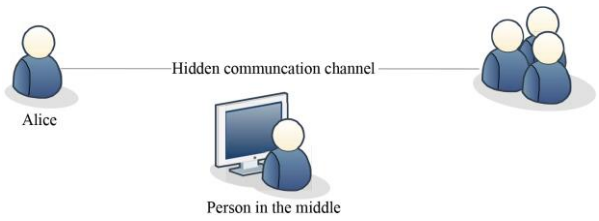


Fig. 4. Hidden communication channel from one to many.

Pervasive platform: Many social networking platforms have emerged in the recent years and photos of different kinds are submerging the Internet. Photo sharing is now a part of the social netizens daily activities. Because the proposed steganographic scheme does not introduce any artifacts in the images to drive the information, an attempt in investigating all nodes connected on a social networking website and trying to make sense of all the submitted images would be practically infeasible. Also it is quasi-impossible to correctly retrieve the images that form the complete message without the key and the dictionary.

Enough images to exploit: One main resource we need in this scheme is a bank of image. There are many free images available online, but as we have selected Facebook for this case study, we could also get the images from Facebook. For a long text to be converted into pictograph, we need a great amount of images and that is of no issue. Facebook allows its users to download and save the images that are publicly available on a Facebook Page. Using an original content although preferable (as it is more likely to be beneficial for some other Facebook users making the page popular, and a popular Facebook page will reduce suspicions) is a bit time consuming. So, in the presented method, Alice (as she is the one who should upload the images) could rip some images from other Facebook pages of her choice by simply retrieving the link to a Facebook page, opening the selected page in a browser, going to the album gallery, selecting the album to rip, clicking on the first photo, and refreshing the page. From there, to automate the download of all the images in the selected gallery starting from what was selected, Alice could simply use a solution for web automation browser add-on like iMacro (http://wiki.imacros.net/) to run the script below straight from a browser such as Firefox, Google Chrome or Internet Explorer.

```
VERSION BUILD=8601111 RECORDER=FX
TAB T=1
TAG POS=1 TYPE=A ATTR=TXT:Next
TAG POS=1 TYPE=A ATTR=TXT:Download
```

To discharge any doubt that placing images that are not related to the event of the author in a Facebook Page is somewhat odd and confessing that these might be pictogram communications. Alice and Bob are and should

not be limited to only just use one Facebook Page for their communication channel. Multiple pages relating to different interests could be used and is a good practice to void the raise of suspicion.

9. Conclusion

In this paper we have presented an image steganographic scheme for social networking websites using pictograms mixed with other innocuous images. The secret message is concealed in a form of pictograms leaving a third party observer unaware of the very existence of the communication channel. The global reach of online social networking websites forms the main platform for this proposed method. And so, by using pictograms to vehiculate secret messages on social networking pages, we could have a mass delivery. The security of the proposed steganography relies on the abundance of images on those social networking websites, the omnipresence of the pictogram within the selected online social network, the complexity of any attempt to brute force the meaning of the pictograms without the key and the dictionary. A higher level of security could be achieved by adding an extra step for using a one-time use dictionary in the scheme. In our case study, the page popularity would reduce the creation of suspicion about the existence of the steganography channel.
It would also enable a communication from one to many and make it look as if is they are ordinary connections. Apart from being simple and very pragmatic, the case study presented in this paper of a steganographic system based on social networking website using pictogram could be implemented in most similar websites with just a minor adjustment.

References

[1] Myres, John Linton. Herodotus, father of history. Clarendon Press, 1953.
[2] Waterfield, Robin, and C. Dewald. "Herodotus: The Histories." Herodotus:(translation (1998).
[3] Sumathi, C. P., T. Santanam, and G. Umamaheswari. "A Study of Various Steganographic Techniques Used for Information Hiding." arXiv preprint arXiv:1401.5561 (2014).
[4] Kawaguchi, Eiji, and Richard O. Eason. "Principles and applications of BPCS steganography." Photonics East (ISAM, VVDC, IEMB). International Society for Optics and Photonics, 1999.
[5] Anderson, Ross J., and Fabien AP Petitcolas. "On the limits of steganography." Selected Areas in Communications, IEEE Journal on 16.4 (1998): 474-481.
[6] Westfeld, Andreas, and Andreas Pfitzmann. "Attacks on steganographic systems." Information Hiding. Springer Berlin Heidelberg, 2000. APA
[7] Provos, Niels. "Defending Against Statistical Steganalysis." Usenix security symposium. Vol. 10. 2001.
[8] Facebook Help Center. 2015. Basic Privacy Settings & Tools. Retrieved April 12, 2015 from https://www.facebook.com/help/325807937506242/
[9] Özera Hamza, Avcıba Ismail, Sankura Bülent and Memonc Nasir. 2010. Steganalysis of Audio based on Audio Quality Metrics.
[10] Fridrich J. 2009. Steganography in digital media principles algorithms and applications. 1st Ed. Campbridge University Press.
[11] Johnson N.F and Jajodia S. 1998. Exploring steganography: seeing the unseen. IEEE Computer, (pp. 26-34).
[12] Srivastava, Lee A. B., Simoncelli E. P. and Zhu S-C. 2003. On Advances in Statistical Modeling of Natural Images. Journal of Mathematical Imaging and Vision.
[13] Fridrich Jessica, Goljan Miroslav and Du Rui. 2001. Reliable Detection of LSB Steganography in Color and Grayscale Images.
[14] Fridrich and Goljan M. and Du R. 2001. Detecting LSB steganography in color and grayscale images. IEEE Multimedia 8, pp. 22–28.
[15] Nabaee H. and Faez K. 2010. An efficient steganography method based on reducing changes. 25th International Conference of Image and Vision Computing New Zealand (IVCNZ), (pp.1 – 7).
[16] Toony Z., Sajedi H. and Jamzad M. 2009. A high capacity image hiding method based on fuzzy image coding/decoding. Computer Conference, CSICC 2009. 14th International CSI, (pp. 518 – 523).
[17] Sajedi H. and Harif. 2008. Cover Selection Steganography Method Based on Similarity of Image Blocks.
[18] Bohme R. 2010. Advanced Statistical Steganalysis (2010 edition). Springer.
[19] Marincola J. M. 1996. Herodotus: The Histories.
[20] Landau S., Kent S., Brooks C., Charney S., Denning D., Diffie W., Lauck A., Miller D., Neumann P. and Sobel D. 1994. Codes, Keys and Conflicts: Issues in U.S. Crypto Policy. Rep. of a Special Panel of the ACM U.S. Public Policy Committee.
[21] Beaver D. 2008. 10 billion photos. (October 2008). Retrieved April 12, 2015 from https://www.facebook.com/notes/facebook-engineering/10-billion-photos/30695603919
[22] Fridrich J., Goljan M. and Du R. 2008. Detecting LSB steganography in color, and grayscale images. IEEE Multimedia, vol. 8 no.4, pp. 22 – 28.
[23] Dumitrescu S., Wu X. and Wang Z. 2003. Detection of LSB steganography via sample pair analysis. IEEE Transaction on Signal processing, vol. 51, no. 7, pp. 1995-2007, 220.
[24] Zhang T. and Ping X. 2003. A new approach to reliable detection of LSB steganography in natural images. In Signal processing (Vol. 83, pp. 2085-2093, 2003).
[25] Kerckhoffs A. 1883. La cryptographie militaire (Vol. ser. 9). J. des Sciences Militaires.
[26] Nikiforakis N. and Acar G. 2014. Browse at your own risk. Spectrum, IEEE, Volume: 51, Issue: 8 , pp. 30 - 35.
[27] Eckersley P. 2010. How Unique is Your Web Browser? Proceedings of the Privacy Enhancing Technologies Symposium (PETS 2010). pp. 1-18

Feno Heriniaina R. received his Master's degree from the College of Software Engineering in 2011and is presently a Ph. D candidate at the State Key Lab. of Power Transmission Equipment & System Security and New Technology, College of Computer Science, Chongqing University.

Xiaofeng Liao received the B.S. and M.S. degrees in mathematics from Sichuan University, Chengdu, China, in 1986 and 1992, respectively, and the Ph.D. degree in circuits and systems from the University of Electronic Science and Technology of China in 1997. His current research interests include neural networks, nonlinear dynamical systems, bifurcation and chaos, and cryptography. He is a senior member of the IEEE.

ACSIJ Advances in Computer Science: an International Journal, Vol. 5, Issue 1, No.19 , January 2016
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Evaluating the Impact of Image Delays on the Rise of MMI-Driven Telemanipulation Applications: Hand-Eye Coordination Interference from Visual Delays during Minute Pointing Operations

Iwane Maida[1], Hisashi Sato[2] and Tetsuya Toma[1,3]

[1] Graduate School of System Design and Management, Keio University
4-1-1 Hiyoshi, Kohoku-ku, Yokohama 223-8526, Japan
i-maida@z6.keio.jp, t.toma@sdm.keio.ac.jp

[2] Kawasaki Interface Techno Co., Ltd.
Kawasaki 216-0001, Japan
hisashi.s@ki-techno.com

[3] Keio Photonics Research Institute, Keio University
7-1 Shin-Kawasaki, Saiwai-ku, Kawasaki 212-0032, Japan
t.toma@kpri.keio.ac.jp

Abstract

Robots with high freedom of movement can be used for minute manipulation, enabling a wide range of real-world applications. However, the use of telemanipulation systems driven by man-machine interaction (MMI) is currently limited to experimental trials, partly because the image delay during transmission interferes with the operator's hand-eye coordination. This study examined how delays affect operation efficiency by pointing target size, and the degree of difficulty when performing telemanipulation operations. We conducted tests in which subjects performed pointing operations while visual delays interfered with their hand-eye coordination. Our findings show statistically that regardless of target size, delay variance of 200 ms or less hardly affects operation efficiency, and that difficulty increases for a target diameter of between 4 and 2 mm.

Keywords: *MMI-Driven Telemanipulation; Image Delay; Hand-Eye Coordination; Minute Pointing; Operation Efficiency*

1. Introduction

Recent advances in robotics technology and ICT (information and communication technology) have enabled telemanipulation driven by man-machine interaction (MMI) using robots with high freedom of movement. Robot-driven telemanipulation is gaining attention for the potential it has to enable work in places unsuited to operators such as disaster-stricken areas or extreme environments. It is an area with a wide range of potential applications that could solve real-world problems such as ensuring operators' safety and creating more balanced distributions of human resources.

For example, medical services provided to users around the world can vary widely in quality and quantity since the concentration of medical resources available varies according to the size of cities in which hospitals are located. With their greater population density, larger cities tend to have greater concentrations of various resources. Since this trend is natural, it calls for effective use of the medical resources of large cities rather than attempts to forcibly correct differences through political measures. In this regard, advances in research on telemanipulation could provide a way to supplement human medical resources in places where they are lacking, such as in Japan's *genkai shūraku* (remote villages on the brink of extinction due to their aging populations).

Combining the functions of remote transmission (communications) and manipulation (robotics) enables composite functions for which several real-world applications have been announced. For example, these functions could be applied to conventional surgery to match telesurgery needs to doctors with the required skills. Remote transmission could also be applied to education, enabling doctors in cities to remotely provide skill education to doctors in remote areas [1]-[4]. These possibilities are therefore creating growing expectations that telesurgery and other telemanipulation applications will become a practical reality [5].

A past study has examined telesurgery using an experimental MMI-driven surgery robot [6]. As the world's first demonstration of telesurgery, it represents a major breakthrough and has had a major impact on medical professionals. However, telesurgery has not achieved widespread use. One reason is that surgery

imposes severe restrictions on operation precision and operation time, so it is not enough to merely show that remote robotic surgery that was once impossible is now possible.

The major problem with remote robotic manipulation for applications such as telesurgery is that a visual delay arises between the work sites. Telemanipulation is an activity requiring the operator to manipulate objects using visual information presented on a monitor. Therefore, the visual delay arising from the transmission delay has a major effect on operation efficiency and operation precision.

Before MMI-driven telemanipulation can achieve widespread use, researchers will need to quantify the maximum amount of delay that can be tolerated without affecting operation efficiency (which also affects safety), and how minute the operation can be before the delay makes efficient operation difficult.

Accordingly, this study examined visual delay (which affects arm positioning), and pointing target size (which corresponds to arm positioning precision). For each of several target sizes, we attempted to find the range of visual delay times (the operation efficiency range) over which operation efficiency can be considered the same, and the target size threshold at which operator positioning efficiency declines sharply.

2. Effects of Visual Delays on Arm Positioning

To provide background information for this study, this section describes the problems related to movement disturbances caused by a delay in the image. It discusses prior studies that have tried to reduce the transmission delay, and prior studies that have examined the movement problems that the delay creates. We describe the validity of the test parameters, the psychological approach used for testing, and the study results and observations, quantitatively demonstrating the effect of a visual delay on telemanipulation.

2.1 Hand-eye coordination

Humans perceive and evaluate stimuli from the outside environment, applying the results to body movements. This process of applying perceptually generated information to ongoing movement control is called hand-eye coordination [7]. Surgery was previously mentioned as a typical example of a telemanipulation application. It too is a process involving movement of the hands and arms relying on hand-eye coordination.

The most important aspect of coordinated movement is the temporal integration of the perception of the sensory organs. One fundamental task for arm movement is pointing—the task of moving the arm from starting coordinates and positioning it at target coordinates. Pointing includes the process of converting external coordinates obtained by vision into internal body coordinates.

Hand-eye coordination functions by integrating systems such as vision and the somatosensory system. The hand arrival process driven by hand-eye coordination is composed of two types of movements: feedforward movements (which are ballistic movements) and feedback movements (which are corrective movements).

For the task of pointing (one type of arm positioning movement), the precision of feedback movements is an important element in enabling the operator to perform positioning accurately. Tasks such as surgery that require minute positioning are particularly prone to effects from visual delays. Therefore, disturbances to hand-eye coordination reduce operation efficiency during pointing operations in environments designed for remote robotic work controlled by an operator, such as master-slave systems.

A reduction in operation efficiency caused by a visual delay means a delay in the progress made during the scheduled operation time. Since the operation being performed will often have time constraints, the operator will often need to boost operation efficiency in subsequent processes, increasing the operator's psychological burden and risk factors such as process errors. This tendency is particularly marked in tasks such as telesurgery that involve minute operations and pressing time constraints. The next section presents prior studies on amounts of visual delay during telemanipulation, and discusses how delays affect ease of operation.

2.2 Prior studies

Since telemanipulation is performed by an operator relying on visual information presented on a monitor, a delay in the image arising from a transmission delay is a major impediment to working. Transmission delays are usually caused by three factors: (1) the communication line's transmission delay, (2) the processing delay caused by information compression and extraction, and (3) the response time of the actuators that move the robot. The transmission line delay is the longest delay of the three.

Prior studies on telesurgery have reported various total delay values. The total delay reported for the world's first demonstration of telesurgery previously cited was 155.0 ms[6]. Studies of telesurgery in Asia have reported total delays of 278.3 ms [8] and 582.4 ms [9]. The amount of delay during telemanipulation depends on several different elements, such as the communication path, transmission format, video device compression/extraction process, and robot reaction time. Previous tests in low latency mode have set visual delays of up to 640 ms since delays of 700 ms or longer make it difficult to temporally integrate the eyes and hands. Smooth movements become impossible at delays of this length, and the adoption of a 'wait-and-

move' strategy [10] is needed to enable accurate movements.

2.3 Effects of visual delays on hand-eye coordination

We have reviewed how hand-eye coordination is an important element for enabling smooth performance of an operation, and how temporal sensory integration of visual information and somatic sensation is important for the smooth working of hand-eye coordination. Then, what type of effect does a visual delay generated by remote transmission have on hand-eye coordination during arm movement? This section discusses this issue.

A prior study on the relationship between visual delays and somatic sensation has shown a sharp deterioration in operation efficiency resulting from environments introducing delays of up to 500 ms [11]. Visual delays of 700 ms or longer have been reported to result in test subjects adopting a 'wait-and-move' strategy to perform movements and smooth movements becoming impossible [10].

However, studies on operations performed with visual delays have not adequately described the effect of target minuteness on operation efficiency, or the minuteness threshold at which the operation efficiency of an operator declines. This study tested the efficiency of pointing tasks performed with visual delays of up to 600 ms (the maximum delay for which smooth movements can be expected). The study method below was used to describe this area.

3. Study Method

For each of several pointing target sizes, the aim of this study was to find the range of visual delay variance over which difference in operation efficiency can be considered significant, and the target size threshold at which operator positioning efficiency declines sharply. To achieve this aim, we needed to quantitatively reproduce the visual delays experienced during telemanipulation on an MMI-driven master-slave system. This section describes the placement of the test apparatus used to reproduce the delays experienced during telemanipulation, describes the operating principle of the image delay generator used, and presents the test plan used to measure performance in the visually delayed environment.

3.1 Test apparatus placement

Figure 1 shows the placement of the test apparatus used for testing. The image of the work area was photographed by a video camera and transmitted to a monitor via a delay generator. Using this apparatus configuration, the visual delay generated between the master and slave in a telemanipulation environment was reproduced on the monitor and used for pointing tasks in the work area. The camera and monitor used were commercial models with short processing delays, minimizing the amount of uncontrolled visual delay generated. The amount of delay presented when the delay generator was set to its minimum delay amount was 98 ms.

To determine the causal relationship between image delays and ease of operation, we controlled distances to keep the visual distance (the distance from the eyes to the monitor display surface) the same as the distance from the work area to the monitor. We also adjusted the size of the pointing targets displayed on the monitor to make them actual-size. The camera was placed in a position away from the test subject's eye line to prevent it from becoming a distraction. Efforts were also made to minimize external disturbances caused by differences from normal sensory experiences.

3.2 Test apparatus and measuring instruments

Table 1 lists the test apparatus and measuring instruments this study used to describe the effect of visual delays on pointing.

Since the aim of this study was to describe the effect of visual delays on ease of operation, the test administrators took steps to ensure a user-friendly environment for test subjects by quantitatively controlling factors such as the distance from the test subject to the monitor, and the size of the targets displayed on the monitor.

Taking surgery as a representative example of the minute pointing operations examined by this study, we devised a simple manual dexterity model of surgery for our examination of minute tasks.

Our simple manual dexterity model reproduced the positioning movements and target minuteness of actual surgery. To set minuteness, we set the size of the dot targets used by considering the average size of dot-target blood vessels and mucosae. To set positioning movement directions, we placed the dot targets so that directions of positioning movements would be as random as possible. We set the maximum distance between targets at 150 mm after interviewing hospital staff to determine the range of movements performed during surgery, and considering factors such as range of wrist movement.

Our simple manual dexterity model consisted of circular dot targets used to recreate pointing minuteness.

The targets were placed on the vertices of a regular pentagon with sides of 100 mm in length. Three model difficulty levels were used, each having a different target diameter (2, 4, or 6 mm). Figure 1 shows the sequence of dots. Operation time was measured in hundredths of a second, and measured time data was rounded to the nearest tenth of a second to create the analysis data.

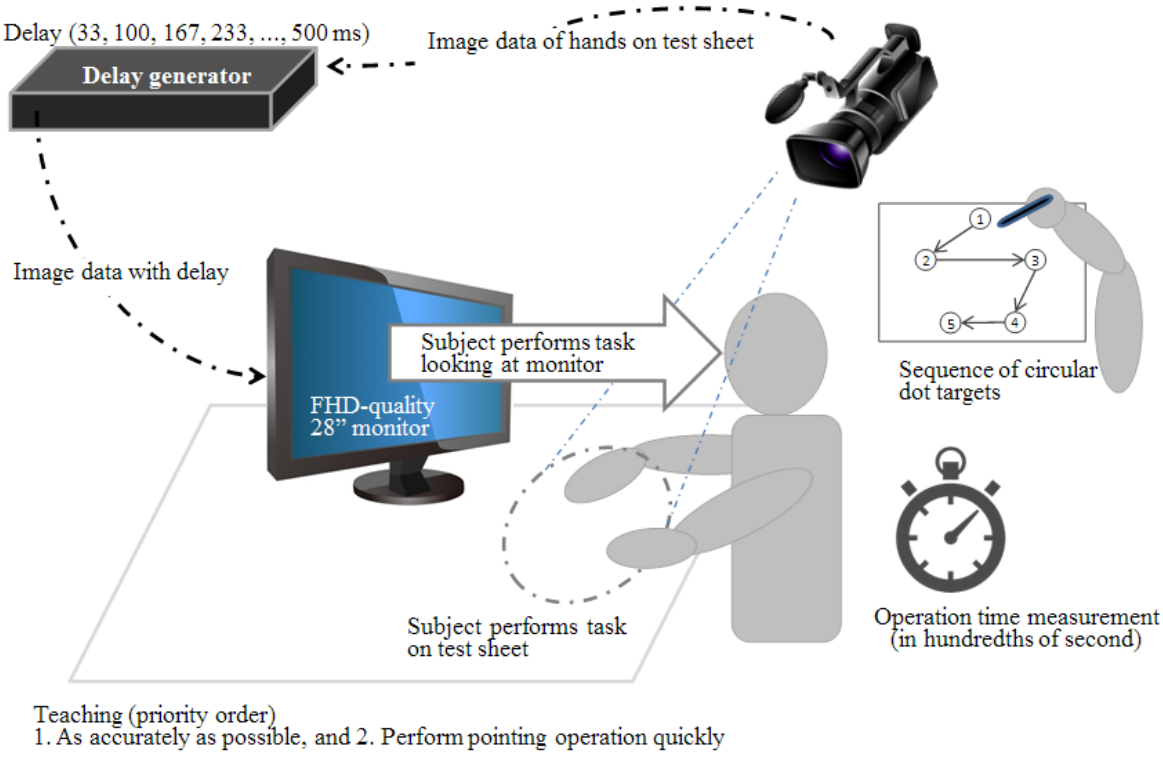


Fig. 1 Test apparatus configuration

Table 1: Test and measurement equipment

	Item	***Manufacturer and model***	***Comments***
Test equipment	Camera	Astro AH-4410-A	Full High Definition (FHD) (60i) output
	Monitor	IBM T221	28-inch FHD monitor
	Delay generator	[Created by authors]	For image delay (1 to 15 frames)
Measurement equipment	Simple manual dexterity model	[Created by authors]	• Circles of 2, 4 or 6 mm in diameter • Circular dot targets placed at vertices of regular pentagon with sides of 100 mm
	Stopwatch	Casio HS-80TW	Can measure thousands of second

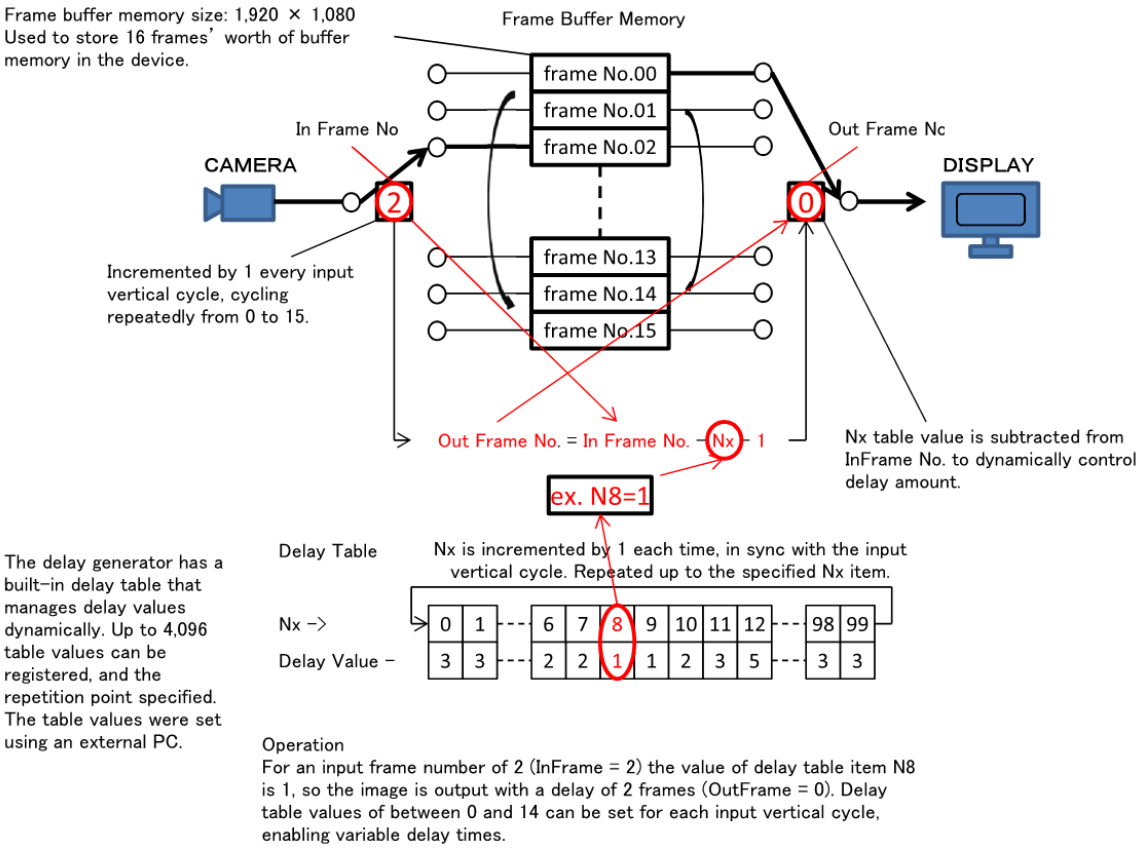


Fig. 2 Delay generator operating principle

3.3 Operating principle of image delay generator

Since the aim of this study was to precisely evaluate the effect of an image delay, a function to quantitatively control the image for presentation on the monitor was an indispensable requirement. To meet this requirement, we created our own image delay generator for this study. To enable test conditions to be set precisely and minimize the processing delay, we used highly responsive RAM (random access memory) as the memory medium for temporary storage of the image. Using this basic design, we created an image delay generator that controlled the image frame by adding a buffer to the internal memory storing the image data. Figure 2 illustrates the principle used to generate delays.

Our delay generator was able to create delays ranging from 1 frame (33 ms) to 15 frames (500 ms), in 1-frame increments. During the design phase, we set the maximum amount of delay not requiring delay compensation for telemanipulation at 500 ms (15 frames).

The standard FHD frame size of 1,920 × 1,080 was input, with the newest frame replacing the last frame. The image reproduced the visual delay experienced during telemanipulation by being output with a delay from the current frame equal to the number of frames set by the test administrator.

3.4 Test plan

The aim of this study was to precisely evaluate the effect of image delays and assess the outlook for the rise of MMI-driven telemanipulation applications. Describing operation efficiency in terms of visual delay levels and dot size differences was indispensable for achieving this aim. We therefore set three groups of dot sizes as variables (2, 4, and 6 mm in diameter), and tried to eliminate individual differences in movements between test subjects by using a random dot placement for every 10 subjects. The image presentation stimulus had 8 delay levels (33, 100, 167, 233, 300, 367, 433, and 500 ms) for hand-eye coordination interference. To reduce the effect of learning on the presented stimulus, we conducted a total of 16 trials in which subjects were presented with a combination of two delay stimulus patterns having conditions of gradually increasing and decreasing delays. To minimize the effect

of fatigue, breaks of at least 1 minute were provided between each trial.

To determine operation efficiency and difficulty levels during pointing operations with visual delays, we performed the following two analyses: (1) We compared groups of the same dot size to determine the delay level at which operation time became sufficiently different. (2) We compared operation times for different dot sizes using the same delay level. The first analysis was performed by comparing each test subject to themselves. The second analysis was performed by comparing different test subjects.

To compare differences in delay level among and between test subject groups, we used multiplex analysis. We used Bartlett's test to check for equal variances. When equal variances could not be found, we selected the Friedman test or Kruskal-Wallis test (nonparametric test methods) to test the differences among different delay levels. Only when this analysis did not detect a significant difference for all processes, we used Scheffe's paired comparison (a type of multiplex analysis).

3.5 Test subjects

To enable precise evaluation of the effect of image delays on ease of operation, we needed to reduce external disturbances during testing. Right-handed males in their 20s with no visual abnormalities were used as the test subjects for this study.

3.6 Test procedure

Before starting the test, test subjects were instructed to complete the operations in as little time as they could manage while still pointing precisely and without errors, taking the operation efficiency of the remote environment into consideration. For ethical considerations, subjects were told they could stop the test if they felt unwell, and subject consent was received before testing started.

Each subject started the pointing operation at a start signal from the test administrator taking measurement. The operation consisted of pointing at (making contact with) the specified five points in the correct sequence. If a subject was unable to point within the circular border of a given target dot, he was not permitted to skip it and proceed to the next dot. He was required to successfully point at each dot before proceeding. To determine the effect of visual delays on positioning, the test administrator visually assessed whether contact had been made with each target dot and instructed the subject whether he was permitted to proceed to the next dot. Ultimately, the time it took the subject to point at dots 1 to 5 was measured and counted as one trial.

4. Analysis Results

We performed two analyses for this study to quantitatively determine the effect of visual delays on arm positioning. One analysis was performed to determine the identification space of operation efficiency at various delay levels. The other analysis was done to identify the pointing difficulty threshold for minute pointing tasks. The analyses done for this study are described below.

4.1 Analysis 1: Differences in delay variance range for each dot size group

We used the test data obtained from the test plan and test subjects to determine statistically that there was little generalization in response to the presented stimuli. Next, we used Bartlett's test to check for equal variances, to determine the directionality of the statistical process. We found no equal variances, so used the Friedman test (a nonparametric test method) on different delay levels.

Testing at the 5% significance level for the degree of operation efficiency resulting from the amount of delay, we obtained a 1% significant difference. From this result, we inferred that we had ensured validity for performing multiplex analysis, and continued multiplex analysis. Since we compared a large number of levels for our tests (eight levels), we used Scheffe's paired comparison, which has a demanding detection rate for nonparametric multiplex comparisons. Our statistical analysis detected a 1% or 5% significant difference for delay level differences of between 8 and 10 or more frames for the 2, 4, and 6 mm diameter groups (Figures 3, 4, 5).

Figure 3 shows that for the 2 mm dot size group, operation time increased sharply as the amount of delay increased, with a sharper upward slope than for the 4 and 6 mm dot size groups. The delay variance range in operation efficiency for the 2 mm group was 200 ms. This range was 67 ms narrower than the range for the other groups (the 4 and 6 mm groups).

Figures 4 and 5 show that for the 4 and 6 mm dot size groups, operation times became longer as delays became longer. However, in terms of difficulty, the slopes of the graphs for the 4 and 6 mm dot size groups were nearly the same. The efficiency delay variance range in operation for these groups was 266 ms.

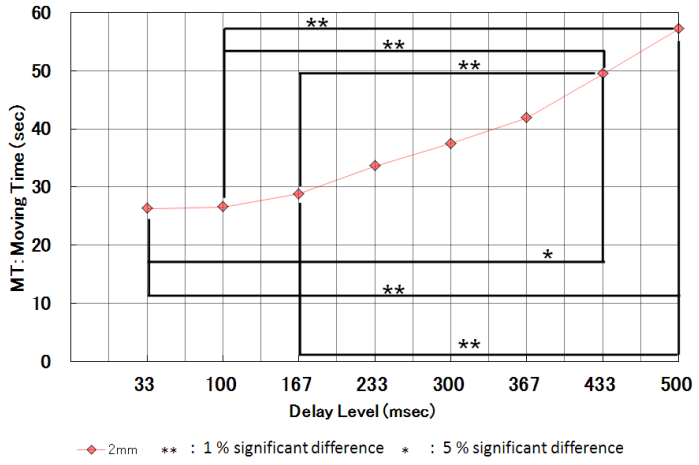


Fig. 3 Delay variance range in operation efficiency for 2 mm diameter group

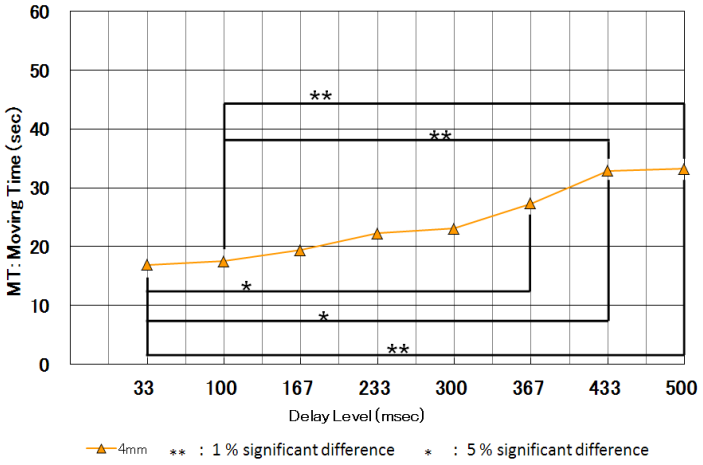


Fig. 4 Delay variance range in operation efficiency for 4 mm diameter group

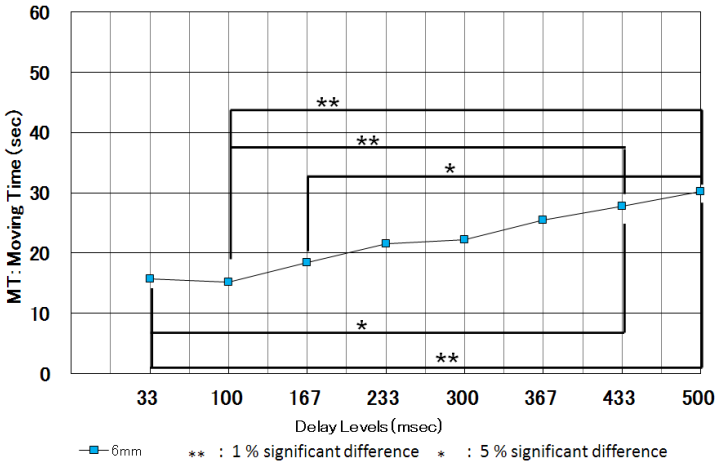


Fig. 5 Delay variance range in operation efficiency for 6 mm diameter group

4.2 Analysis 2: Dot size and target size threshold at which operation efficiency declines (group-to-group comparisons)

Figures 3, 4, and 5 show that there were large differences in operation time at the same delay level when comparing the 2 and 4 mm diameter groups to the 6 mm diameter group. The 4 and 6 mm diameter groups appear to have nearly no difference in operation time. We performed a statistical process to check whether this inference was correct. We found no equal variances using Bartlett's test, so for this analysis we used the Kruskal-Wallis test (a nonparametric test method) to test the differences among different delay levels.

For the Kruskal-Wallis test done at the 5% significance level for the effect that differences in delay level had on operation efficiency, we obtained a 1% significant difference for the entire process. From this result, we inferred that we had ensured validity for performing multiplex analysis, and continued multiplex analysis. We used Scheffe's paired comparison for our tests.

Figure 6 shows the statistical analysis results. When the 2 and 6 mm dot size groups were compared using this analysis, 1% or 5% significant differences were found between all the delay levels. When the 2 and 4 mm dot size groups were compared, 1% or 5% significant differences were also found for half of the compared delay levels. We found no significant difference in operation time between the 4 and 6 mm dot size groups. These findings show statistically that for pointing tasks with delays, operation efficiency is affected by the set minuteness, and for target diameters of 4 to 2 mm, there is a threshold at which operation efficiency declines significantly.

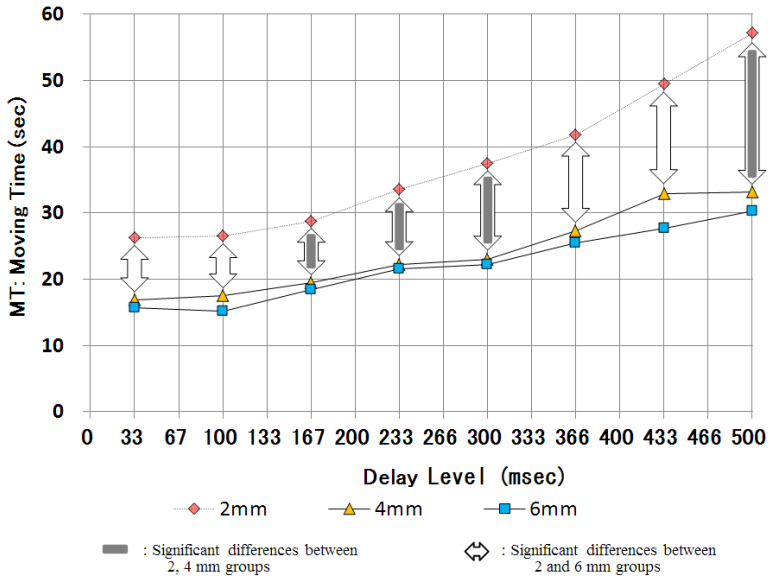


Fig. 6 Difference of difficulty between various target sizes at same delay levels

5. Test Discussion

This section discusses the aims of this study—the delay variance range for each dot size, and the target size threshold at which efficiency sharply declines.

5.1 Delay variance range for each dot size group

The key finding about the delay variance range for each dot size group is that regardless of dot size, significant differences in operation efficiency were detected only when there were delay level differences of 6 to 8 or more frames. These frame delays correspond to a delay variance range of 266 to 333 ms. According to Fitts's law, the index of difficulty (*ID*) for a pointing task is given by $ID = \log_2(A/W + 1)$ [12]-[13], where A is the distance to the pointing target and W is the width of the target. As indicated by the formula, the index of difficulty increases as the target size becomes smaller. Fitts's law is also given as $MT = b(ID) + a$, indicating that movement time MT increases as the dot size becomes smaller. When a pointing task is performed using targets of the same size and within delay variance range described by this study, the findings of this study quantitatively suggest that the task will be

within the same operation efficiency at the 5% significance level.

5.2 Relationship between dot size and delayed pointing task difficulty

The key finding about the relationship between dot size and pointing task difficulty is that when each group was compared at the same delay levels, the 2 mm dot size group had significant differences in operation times relative to the 4 and 6 mm dot size groups. When the 4 and 6 mm dot size groups were compared at the same delay levels, no significant differences were detected. These findings quantitatively suggest that the threshold at which pointing tasks become difficult for humans owing to the effect of delays is between a dot size of 4 and 2 mm.

6. Conclusion and Future Outlook

This study examined the effect of delays (a major problem for MMI-driven telemanipulation) on the operation efficiency and difficulty of minute pointing tasks. We statistically showed that for each dot size group, regardless of target size, delay variance of 200 ms or less hardly affects operation efficiency.

When comparing the operation times for each dot size for the same delay levels, we found that statistically significant differences arose between the 2 mm dot size group and the other groups, thereby demonstrating that there is a wall of difficulty between a dot size of 4 and 2 mm.

These findings have enabled us to make a quantitative assessment that pointing tasks with delays have dot size-dependent difficulty differences, and to anticipate the outlook for the rise in MMI-driven telemanipulation when this information is applied to real-world applications.

Acknowledgments

Through the Japan Society for the Promotion of Science (JSPS), this study was supported by the Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program) conceived by the Council for Science, Technology, and Innovation. The findings of this study were also partly supported by a Keio University doctorate support program.

We would like to take this opportunity to express our gratitude for the wealth of advice received when preparing this paper, from Takashi Maeno (Keio University Graduate School of System Design and Management Professor) and Taketoshi Hibiya (SDM Research Institute Executive Advisor).

References

[1] S. Suzuki, N. Suzuki, A. Hattori, K. Konishi, S. Ieiri, M. Hashizume, “Tele-surgery Simulation for Surgical Planning”, Training and Education, Medical Imaging Technology, Vol. 25, No. 1, 2007, pp. 25-30.

[2] S. Suzuki, N. Suzukia, M. Hashizume, Y. Kakeji, K. Konishib, A. Hattoria, M. Hayashibe, “Tele-training simulation for the surgical robot system “da Vinci, CARS 2004”, Computer Assisted Radiology and Surgery, Proceedings of the 18th International Congress and Exhibition, Vol. 1268, 2004, pp. 86-91. DOI: 10. 1016/j. ics. 2004. 03. 160.

[3] S. Suzuki1, N. Suzuki, A. Hattori1, M. Hayashibe, K. Konishi, Y. Kakeji, M. Hashizume, “Tele-surgery simulation with a patient organ model for robotic surgery training”, The International Journal of Medical Robotics and Computer Assisted Surgery, Vol. 1, No. 4, 2005, pp. 80−88.

[4] S. Suzuki, “Tele-surgery simulation experiment of da Vinci surgery training between Japan and Thailand”, International Journal of Computer Assisted Radiology and Surgery, Vol. 1, No. 1, 2006, pp. 514.

[5] The University of Texas MD Anderson Cancer Center: AT&T's $1 Million Contribution to Seed Telesurgery Program, MD Anderson News Release, 2012, http://www. mdanderson.org/newsroom/news-releases/2012/at-t-s-1-million-contribution-to-seed-telesurgery-program.html.

[6] J. Marescaux, J. Gagner, M., et al., “Transatlantic robot-assisted telesurgery”, Nature, Vol. 413, 2001, pp. 379−380.

[7] G. R. Vandenbos, The APA Dictionary of Psychology, American Psychological Association, 1 edition, July 15, 2006, pp. 586.

[8] M. Mitsuishi, Medical Robotics, J Jpn Coll of Angiol, Vol. 46, 2006, pp. 759-767.

[9] S. Ieiri, M. Hashizume, “Development of Surgical Assisted System for Remote Medicine”, Vol. 49, 2011, pp. 673-677, doi: http: //doi. org/10. 11239/jsmbe. 49. 673.

[10] Hoffmann, R. Errol, “Fitts Law with Transmission Delay”, In Ergonomics, Vol. 35, 1992, pp. 37-48.

[11] W. Fujisaki, Effects of Delayed Visual Feedback on Grooved Pegboard Test Performance, Front Psychol, Vol. 3, No. 61, 2012, pp. 1-14. DOI: 10. 3389/fpsyg. 2012. 00061.

[12] I. S.MacKenzie, “A note on the information-theoretic basis for Fitts' law”, Journal of Motor Behavior, Vol. 21, 1989, pp.323-330.

[13] I. S.MacKenzie, “Fitts' law as a research and design tool in human-computer interaction”, Human-Computer Interaction , Vol. 7, 1992, pp.91-139.

Iwane Maida earned his master degree of System Design and Management from Keio University in 2011. He is a research associate at Keio University since 2011. His research field includes perceptual psychology, human interface, systems engineering, and educational technology. He obtains "Keio University Doctorate Student Grant-in-Aid Program" from 2014. Contributed as a team leader of this research and wrote this paper.

Hisashi Sato graduated from Japan Electronics College in 1985, and joined Keisoku Giken Co., Ltd. where he developed video storage devices and signal converters as the general manager of Visualware Division. He recently founded Kawasaki Interface

Techno Co., Ltd. which mainly develops light transmitting devices by using the video signal processing technologies.

Tetsuya Toma is an associate professor of Graduate School of System Design and Management in Keio University, Japan since 2008. He received Ph.D. in System Design and Management in 2014. Dr. Toma also has a certificate of Project Management Professional (PMP) and has been a director of Japan Chapter of Project Management Institute (PMI) since 2010. His research field is communication design.