Improving Accuracy of Authentication Process via Short Free Text using Bayesian Network

Charoon Chantan[1], Sukree Sinthupinyo[2] and Tippakorn Rungkasiri[3]
[1] Techopreneurship and Innovation management Program, Graduate school, Chulalongkorn University, Bangkok, Thailand.

[2] Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand.

[3] Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok, Thailand.

Abstract
The internet security problems are a crucial threat to all users in the cyber world. One of the important problems about internet security concerned with user classification and authentication. However, there are multiple components to classify and authenticate users. The first one is using username/password and the second method is OTP or Token. This paper presents a novel method which cans Classify User via Short-text and IP Model (CUSIM) to grant or reject a user in authentication. CUSIM is a Bayesian network model which utilizes the Bayesian Inference to authenticate the user. The objective of this paper is to use the model based on conditional independent with the prior knowledge, i.e. Keystroke dynamics, location used to connect to the internet, and IP address. Finally, a numerical example is provided to illustrate the probability of incorrect authentication and use an algorithm of machine learning to test the efficiency and find out the accuracy, FAR, and FRR. The model results gave better value of accuracy, FAR, and FRR.

Keywords: *Authentication, User Classification, Short Free Text, Keystroke Dynamics, IP Address, Bayesian network.*

1. Introduction

An important Internet security problem involves user authentication and classification, in which an intruder or a scammer can maliciously access to the system to steal or hack important information such as username/password or credit card number. The multi-component authentication is composed of (1) known information, (2) stored information, and (3) information of the person. Mostly the first component is the username and password, the second is likely the OTP (One Time Password) or Grid token. However, a weak point can be found in these components [1]. The above-mentioned problems are critical to the effectiveness of filtering and classifying users in the Internet security system. It is then very important to determine whether it is a real user or an intruder.

Internet security problems have attracted researchers. The multi-component authentication with three components is especially concerned with biometrics and it includes the use of IP address to verify a user [2], when a Bayesian network solution algorithm was proposed to obtain the probability of being a real user. Soon afterwards, Asadamongkol [3] developed the Bayesian network model for user classification, which also involves the multi-component authentication. The model uses the combination of classification between Long Free Text Keystroke dynamics and IP address value. This model does not consider any other components which may render the classification more complete. One of the important modifications that we consider in this paper is introducing location component, which is used to co-verify with the IP and Keystroke Dynamics (KD) method. Concerning KD, we can use it with short-free-text as in [4]. In that research, the combination of two statistical features: the *average/standard deviation* and the *probability score*, is used as input vector of the artificial neural network in classification of users. The proposed method obtains more accurate results. Consequently, our model is built with some information of Keystroke Dynamics from the research result of [4]. We formulate the User Classification Model using the Bayesian network composed of variable functioning as the node in network, i.e. KD and location. This model is formulated as a conditional independence and joint probability distribution (JPD) with prior knowledge in calculation. The objective of this paper is to calculate the optimal Bayesian network equation such as JPD, Conditional Independence (CI), and Conditional Probability Table (CPT) of each variable node with prior knowledge value. The results determine the probability of the user or intruder from the classification by KD combined with IP and location in order to satisfy the user authentication in the Internet security system.

This paper is organized as follows: Section 2 provides a literature review on the KD, IP and Bayesian network. In Section 3, we describe our model. Section 4 presents a numerical example of model. Finally, the results and conclusions are presented in Section 5.

2. Literature Review

The user classification and authentication problems have been one of the most studied problems in the Internet security. Moreover, authentication with multiple components has been studied continuously. Liou and Bhashyam [5] proposed a technique of softoken, in which two components are used in authentication. The first step is logging in with username/password. Then, it must be followed by the softoken technique in which a client software will generate a time synchronized OTP (one time password) as the second component. This technique does not require any hardware to store software applications or files. Sarier [6] proposed a security model for multi-component biometric authentication, which does not incorporate secure sketches. Different template extraction methods are performed in the encrypted domain without requiring a decryption key for the authentication decision. It is shown that classification and authentication with multiple components are more and more in use. Biometrics is therefore another component, used as the third component in the Internet security system. In 2010, Yuan et al. [7] proposed the user authentication based on a biometrics-based scheme for wireless sensor network. Yoon and Yoo [8] proposed a new biometric-based user authentication scheme without password for wireless sensor network, which provides the result with more efficiency and strong security. Apart from that, Keystroke dynamics (KD) is classified as one of the methods of Biometrics, it is concerned with the keystroke which can identify a unique characteristic of each individual. It is then used in user classification and authentication. Revett [9] presented a study indicating that keystroke dynamics is a viable method not only for user verification, but also for identification by means of position, specific scoring matrices, and multiple sequence alignment within the context of a keystroke dynamics-based authentication method. The result yield virtually 100% user authentication and identification. Hu et al. [10] proposed the classification of user's keystroke dynamics profile using *k*-nearest neighbor approach for authentication. The results shows the performance of less than 5% FAR (False Acceptance Rate) and less than 0.005% FRR (False Rejection Rate), and it also shows the improvement of the authentication speed achieved as high as 66.7%. The keystroke dynamics has been studied for user classification with fixed text method, which is mostly applied to username and password [11], [12] and [13]. Similarly, Rybnik et al. [14] proposed an efficient user authentication with KD using short fixed text. Besides, the user authentication with KD is also in application with the free text method. Monrose and Rubin [15] studied authentication via free text, and achieved a classification accuracy of about 23%. In 2005 Gunetti and Picardi [16] presented the use of free text in detecting impostors. The result shows a False Acceptance Rate (FAR) of 5% and a False Rejection Rate (FRR) of 0.005%. Davoudi and Kabir [17] proposed a new distance measure for free text KD between typing sample and set of samples of user, by means of histogram-based density estimation to find out the duration time. The result yields considerable decrease in FAR while FRR remains constant. Roadrungwasinkul and Sinthupinyo [4] also proposed a method of identification by short free text, using the combination of two statistical features, the *average/standard deviation* and the *probability score*, and using the artificial neural network to classify the user; the result is better than the existing method with short free text. In this paper, to our study we apply the result of [4] who uses the keystroke dynamics with short free text as one of the variables of the model.

The Internet security system, besides using the multi components in authentication and verification of users and intruders, also uses IP address for the same purpose. Park et al. [18] proposed an intrusions detection using a PCB and IP address (IDIP), which monitors and checks the intrusion possibilities using IP information of processes. Normally, allocating IP address has two aspects, i.e. static or dynamic. In this respect, static IP address is not changed for a long time and it is updated regularly once a day or a week; while dynamics IP address is changed rapidly because IP addresses will be allocated to multiple users by authentication hosts when the users connect to the internet. Therefore, the location of the IP address is determined after the user connects to the internet connection service providing server, logins internet connection service, and receives an IP address from the authentication server [19]. Aldridge et al. [2] suggested using an IP address along with other component to authenticate a user for better accuracy. In consequence, we can use IP address (Static IP) and Group IP address (Dynamic IP), which are allocated from Dynamic Host Configuration Protocol (DHPC), as another variable of the model. The research of [3] also proposed security verification on websites by using Keystroke Dynamics with IP address via the Bayesian network, which provides the result with accuracy of 97% whilst the authentication with solely KD is only 66.67% accurate.

Apart from those studies, in which IP address is used to identify users, there are other studies that the researchers employ location and device connected to the Internet to

verify and control the use of Internet [20]. For instance, Jaros and Kuchta [21] proposed a new location-based authentication technique, that is STAT I (Space-Time Authentication Technique), which relies on the GPS system in order to determine the location; and STAT II which uses proprietary communication technology IQRF for location determination. Various researches use machine learning techniques to test the efficiency of the model, possibly by different algorithms such as neutral networks or decision tree. Still, there are many researches which study the use of Bayesian network for user classification and authentication [22-25]. To increase accuracy of user classification and authentication, we propose a Bayesian network model which consists of KD, IP, and Location to improve efficiency.

3. User Classification via Short Free Text and IP Address Model Formulation

We use a Bayesian network to Classify User via Short-text and IP Model (CUSIM), along with prior knowledge:

3.1 Bayesian Network to Classify User via Short-text and IP Model (CUSIM)

Our concept model introduces a user authentication model. The model consists of a six tuples:

CUSIM = (*Real User*, *Location*, *KD*, *IP*, *CU-SIM*, *PDT*), where:

Real User: *The probability of being a user, not an intruder*
Location: *The probability of a same location used to connect to the internet.*
KD: *The probability of getting High, Medium and , Low score)*
IP: *The probability of using IP address*
$\boldsymbol{CU-SIM}$: *The probability of authorized user (yes/no)*
PDT: *A five-tuple* ($P_{User/Intruder}$, $P_{Location}$, P_{KD} ,P_{IP} ,P_{CU-SIM}*), which contains the probability distribution of User/Intruder, Location, KD, IP, CU-SIM.*

There are some relationships between elements in the six tuples, which creates a Bayesian network, as shown in Fig. 1. The relationships are based on probability of being a real user who connects to internet and uses basic information obtained from Keystroke dynamics, Location, IP address; their probability distribution is concluded by prior knowledge from statistical survey and the research results of [4]. The semantics of this graph is described below.

1) IP address is conditionally dependent of location (same location and location change). For example, there are users connecting to the internet from home (same location).
2) CU-SIM (user classification yes or no) is conditionally dependent of KD, IP. For example, if the users connect to internet from home (same location), then the KD can be used to allow the user to access the system based on same IP, the probability of a CU-SIM is indicated to answer "Yes".

The rule of CU-SIM which is used in our calculation is described below.

- If KD score is high and IP is same then CUSIM = "Yes".
- If KD score is high but IP is not same then CUSIM = "Yes".
- If KD score is moderate and IP is same then CUSIM = "Yes".
- If KD score is moderate and IP is not same then CUSIM = "No".
- If KD score is low and IP is same then CUSIM = "No".
- If KD score is low and IP is not same then CUSIM = "No".

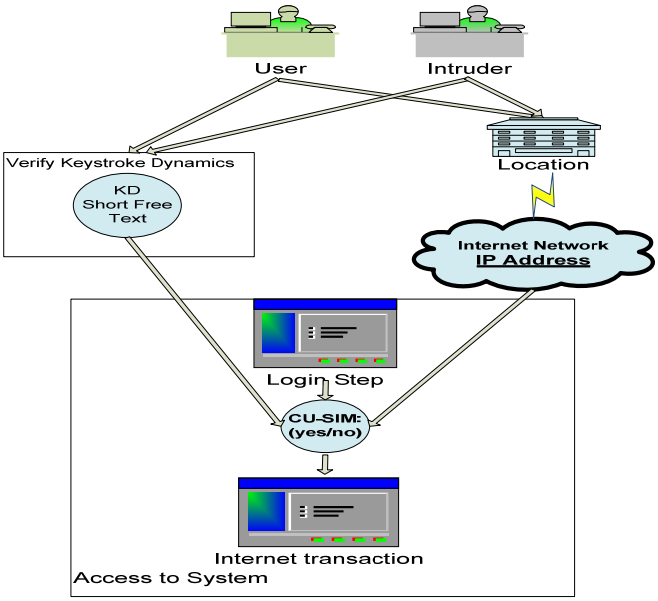

Fig. 1. Bayesian Network of CUSIM.

After we have the structure of the Bayesian network and their probability distributions, we can calculate the probability of being a real user based on concept of Bayesian network inference.

3.2 Using Bayesian Network

In CUSIM, there are five modules: User or Intruder verification, Location verification, Keystroke dynamics verification, IP verification module, and Bayesian inference calculator.

The first module gets the probability of being a real user and the second module gets the current location from the user/intruder which connects to the Internet. The third module verifies the score value of KD owner, while the forth module finds the probability of IP changing. Those functions of CU-SIM are to provide the information predefined by a spatial predicate to Bayesian network calculator, which obtains the probability of KD, IP. The basis of the model is the Bayesian inference calculator which will get information from each module. We create the probability distribution tables and calculate the probability using the Bayesian formula. After getting the final probability, we compare it to a predefined threshold value and decide whether it is the real user or intruder. The Bayesian inference calculator uses the following formula to evaluate the probability which we are interested.

P(CU-SIM, Real User, KD, IP, Location) =
P(CU-SIM |K , IP) P(KD| Real User) * P(IP |location) * P(Location | Real User) * P(User or Intruder)*

If we have already known the structure of Bayesian network model and a full sampling of the data is available, we are able to verify the probability distributions by computing statistics from the data samples as figured by the model in Fig. I , for an example, we compute *P*(*IP* | *LOC*). That is:

$$P(CU\text{-}SIM=Yes \mid KD=High, IP=same\ IP) = \frac{P(CU-SIM=Yes,\ KD=High,\ IP=same\ IP}{P(KD=high,\ IP=same\ IP)}$$

$$= \frac{number\ of\ samples\ with\ CU-SIM=Yes,\ KD=high,\ IP=same\ IP}{number\ of\ samples\ with\ KD=high,\ IP=same\ IP} \quad (2)$$

4. A numerical example model analysis

Table 1-5 show a prior knowledge of probability tables used in this paper. Table 1 shows the probability of being a real user which connects to the internet. We use the probability of the intruder attempt to connect to internet instead the user to compare the result. The probability shall be 10% or 30%. Table 2 shows the probability of Short free text Keystroke Dynamics analysis from research result of [4]. Table 3 explain the probability of location provided by the report of the survey result of an Internet user group in Thailand in 2010 by National Electronics and Computer Technology Center (data of August to October 2010). The results show the Internet access behavior of general people in different locations [26]. *IP address* where the probability is obtained from the survey of website accessing [27] with unique IP measured by the number of user's visits on the website since 2007 to 2011, is shown in Table 4.

Table 1: PRIOR KNOWLEDGE OF PROBABILITY OF USER or INTRUDER.

	Probability
User	0.9, 0.7
Intruder	0.1, 0.3

Table 2: PRIOR KNOWLEDGE OF PROBABILITY OF KD [4].

	Probability
High	0.736
Moderate	0206
Low	0.058

Table 3: PRIOR KNOWLEDGE OF PROBABILITY OF LOCATION.

	Probability
Same location	0.9
Change location	0.1

Table 4: PRIOR KNOWLEDGE OF PROBABILITY OF IP.

Year	***Page-views***	***Unique IP***	***P (Unique IP)***
2007	13,760,117	1,288,242	0.094
2008	5,190,856,008	195,827,746	0.038
2009	14,114,880,085	360,668,425	0.026
2010	21,363,242,545	502,433,280	0.024
2011	19,438,576,535	575,665,128	0.030
The average probability of Same IP and Unique IP 5 years (2007-2011)			**0.042**

In each experiment, values of P (Real User), 90% and 70%, have been tested by the conditional independent probability in three scenarios as shown in Table 5-7.

Table 5: THE CONDITION PROBABILTY TABLE OF SCENARIO 1.

Scenario 1.		
P(Location | Real User)		
Real User	**user**	**intruder**
Same-Location	0.7	0.7
Change-Location	0.3	0.3
P(KD | Real User)		
Real User	**user**	**intruder**
High	0.7	0.7
Medium	0.2	0.2
Low	0.1	0.1
P(IP | Location)		
Location	**Same-Location**	**Change-Location**
Same-IP	0.7	0.7
Change-IP	0.3	0.3

Table 6: THE CONDITION PROBABILTY TABLE OF SCENARIO 2.

Scenario 2.		
P(Location | Real User)		
Real User	**user**	**intruder**
Same-Location	0.9	0.1
Change-Location	0.1	0.9
P(KD | Real User)		
Real User	**user**	**intruder**
High	0.7	0.3
Medium	0.2	0.5
Low	0.1	0.2
P(IP | Location)		
Location	**Same-Location**	**Change-Location**
Same-IP	0.9	0.1
Change-IP	0.1	0.9

Table 7: THE CONDITION PROBABILTY TABLE OF SCENARIO 3.

Scenario 3.		
P(Location | Real User)		
Real User	**user**	**intruder**
Same-Location	0.9	0.1
Change-Location	0.1	0.9
P(KD | Real User)		
Real User	**user**	**intruder**
High	0.8	0.05
Medium	0.15	0.15
Low	0.05	0.8
P(IP | Location)		
Location	**Same-Location**	**Change-Location**
Same-IP	0.9	0.1
Change-IP	0.1	0.9

We applied the Bayesian Network equation using joint probability distribution and conditional independent equation. We made up various contextual information such as user's KD, type of location and IP preference in accordance with the survey results of [26] and [27]. After obtaining the information, the Bayesian inference calculator is calculated.

The data in training set and testing data were generated by bootstrap method. In this paper, we use WEKA for testing the performance of CUSIM and find out the ratio of FAR and FRR value. We use technique resample to balance data for training, and use classification method by BayeNet with genetic search algorithm for each experiment. The results are shown in Table 8.

Table 8: THE EXPRIMENT RESULTS TABLE (ACCURACY, FAR, FRR RATE).

P (User/Intruder) = 70:30		
Scenario 1		
Accuracy	**FRR:P(User| CUSIM=No)**	**FAR:P(Intruder| CUSIM=Yes)**
1	**0.000**	**0.000**
Scenario 2		
Accuracy	**FRR:P(User| CUSIM=No)**	**FAR:P(Intruder| CUSIM=Yes)**
0.99	**0.000**	**0.029**
Scenario 3		
Accuracy	**FRR:P(User| CUSIM=No)**	**FAR:P(Intruder| CUSIM=Yes)**
0.96	**0.000**	**0.103**
P (User/Intruder) = 90:10		
Scenario 1		
Accuracy	**FRR:P(User| CUSIM=No)**	**FAR:P(Intruder| CUSIM=Yes)**
1	**0.000**	**0.000**
Scenario 2		
Accuracy	**FRR:P(User| CUSIM=No)**	**FAR:P(Intruder| CUSIM=Yes)**
0.98	**0.000**	**0.080**
Scenario 3		
Accuracy	**FRR:P(User| CUSIM=No)**	**FAR:P(Intruder| CUSIM=Yes)**
0.93	**0.057**	**0.100**

We focus on the conditional independent probability of Keystroke dynamics. In a scenario 1, the conditional probability of P (KD|Real User = High), P (KD|Real User = Medium), and P (KD|Real User = Low) of the user and the intruder is the same value. The value of probability to be sorted in a descending order from P(High) to P(low) value. The results of CUSIM in Table 8 show the best value of the accuracy, FAR, and FRR of each situation. In scenario 2, the conditional probability of P (KD|Real user= High), P (KD|Real user = Medium), and P (KD|Real user = Low) of the user and the intruder are different but the probability values of P(KD|Real User) of the user are still same as in scenario 1. The results of CUSIM in Table 8 show that the accuracy and FAR values decrease. In scenario 3, the conditional probability of P (KD|Real user = High), P (KD|Real user = Medium), and P (KD|Real user = Low) of the user will be contrast with value of the intruder in an opposite direction. The results of CUSIM in Table 8 show that the accuracy and FAR values will decrease more than those in scenario 1 and 2. The implication of an experimental shows that the values of Accuracy, FAR, and FRR depend on the conditional probability of Keystroke Dynamics when we know whether it is the real user or the intruder. The values of FAR and FRR obtained from CUSIM is better than the method of [4] with the values FAR and FRR are 0.046 and 0.045, respectively.

Table 8 shows the situation that the probability value of user is greater than an intruder. For example, scenario 1 shows when the system indicates that the current user is the owner of keystroke dynamics (KD=High) and the current user uses the laptop to connect to the internet from home (same location) and the network information shows the same value of IP address. The CUSIM will compare the data to the threshold. If the answer from CUSIM is "Yes" the system will grants authentication to the user. On

the contrary, if the answer from CUSIM is "No", an intruder disguises to be the user and invades into the system. The system will not allow doing any transactions. In this case, the calculation result depends upon the prior information. The statistics shows that the accuracy is increased, and it will decrease when IP address is changed.

5. Conclusions

In this paper, we propose CUSIM which is a method of user classification and authentication by resorting to the third component where the verification is operated with keystroke dynamics; Location and IP address used to connect to the internet. The Bayesian network and machine learning are the tools we used to evaluate the efficiency of the model. This is an effective way to classify and authenticate a user in the world of Internet system. With other methods used nowadays, if the information is hacked or stolen, it means the higher risk in terms of security, where a phony user can operate any transaction in lieutenant of the real user. Thus, CUSIM takes the context of user's information to connect to the internet into consideration, and uses the Bayesian network inference to calculate the probability of real user or intruder and uses an algorithm of the machine learning to classify of the user. Experimental results clearly demonstrate the efficiency of our approach.

References

[1] M. Pearce, S. Zeadally, and R. Hunt, "Assessing and improving authentication confidence management", Information Management & Computer Security, Vol.8, No. 2, 2010, pp. 124-139.

[2] A. Aldridge, M. White, and K. Forcht, "Security considerations of doing business via the Internet: cautions to be considered", Internet Research: Electronic Networking Applications and Policy, Vol. 7, No. 1, 1997, pp. 9-15.

[3] K. Asadamongkol, "Innovation of Identification Plug-Ins on Website by Keystroke Dynamics", M.S. thesis, Technopreneurship and Innovation Management program, Graduate School, Chulalongkorn University, Bangkok, Thailand, 2010.

[4] W. Roadrungwasinkul, and S. Sinthupinyo, "A Combination of Statistical Features and Neural Networks to Classify Users Based on Free Text", IRAST International Congress on Computer Applications and Computational Science (CACS 2010)., 2010.

[5] J. Liou and S. Bhashyam, "A feasible and cost effective two-factor authentication for online transactions"., Software Engineering and Data Mining (SEDM) 2nd International Conference, 2010 , pp. 47 – 51.

[6] N.D. Sarier, "Practical Multi-factor Biometric Remote Authentication", In Biometrics: Theory Applications and Systems (BTAS), Fourth IEEE International Conference, 2010, pp.1-6.

[7] J. Yuan, C. Jiang, and Z. Jiang, "A biometric-based user authentication for wireless sensor networks", Wuhan University Journal of Natural Sciences, Vol. 15, no. 3, 2010, pp. 272-276.

[8] E. Yoon, and K. Yoo, "A New Biometric-based User Authentication Scheme without Using Password for Wireless Sensor Networks", Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE) 20th IEEE International Workshops, 2011,pp. 279-284.

[9] K. Revett, "A Bioinformatics Based Approach to Behavioral Biometrics", Frontiers in the Convergence of Bioscience and Information Technologies, 2007, pp.665-670.

[10] J. Hu, D. Gingrich, and A. Sentosa, "A k-Nearest Neighbor Approach for User Authentication through Biometric Keystroke Dynamics", Communications 2008 ICC '08 IEEE International Conference, 2008, pp. 1556-1560.

[11] R.S. Gaines, W. Lisowski, S.J. Press, and N. Shapiro, "Authentication by Keystroke Timing: some Preliminary Results", Technical Report R-2526-NSF: Rand Corporation, 1980.

[12] R. Joyce, and G. Gupta, "Identity authentication based on keystroke latencies", Communication of the ACM 33, Vol. 2, 1990, pp. 168-176.

[13] Z. Syed, S. Banerjee, Q. Cheng, and B. Cukic, "Effects of user habituation in keystroke dynamics on password security policy", IEEE 13th International Symposium on High-Assurance Systems Engineering, 2011, pp.352-359.

[14] M. Rybnik, P. Panasiuk, and K. Saeed, "User Authentication with Keystroke Dynamics Using Fixed Text", Biometrics and Kansei Engineering (CBAKE) 2009 International Conference, 2009, pp. 70-75.

[15] F. Monrose, and A. Rubin, "Authentication via keystroke dynamics", Proceedings of the 4th ACM conference on Computer and communications security, 1997, pp. 48–56.

[16] D. Gunetti, and C. Picardi, "Keystroke analysis of free text", ACM Trans. Inf. Syst. Secure, 2005, Vol. 8, No. 3, pp.312–347.

[17] H. Davoudi, and E. Kabir, "A new distance measure for free text keystroke authentication", Computer Conference (CSICC) 14th International CSI, 2009, pp. 570-575.

[18] J. Park, B. Ahnl, and H. Cho, "Intrusion Detection Using a PCB and IP address", Communications, Computer and Signal Processing, PacRim 2007 IEEE Pacific Rim Conference, 2007, pp. 223-226.

[19] M. Cho, and H. Rim, "A Location Information Retrieval System Using IP Address", Advanced Language Processing and Web Information Technology ALPIT 2007, 2007, pp. 480-485.

[20] J. Winterbottom, and C. Bryce, "The Internet Location Service", Intelligence in Next Generation Networks (ICIN), 14th International Conference, 2010, pp. 1-7.

[21] D. Jaros, and R. Kuchta, "New Location Based Authentication Techniques in the Access Management", Wireless and Mobile Communications (ICWMC) 6th International Conference, 2010, pp. 426 – 430.

[22] A. Cufoglu, M. Lohi, and K. Madani, "A Comparative Study of Selected Classifiers with Classification Accuracy in User Profiling", Computer Science and Information Engineering 2009 WRI World Congress, 2009, Vol. 3, pp.708 – 712.

[23] L. Huijuan, C. Kejie, and L. Bai, "Bayesian Network Based Behavior Prediction Model for Intelligent Location Based

Services", Industrial Electronics and Applications ICIEA 2007, 2nd IEEE Conference, 2007, pp. 703 – 708.

[24] A. Cufoglu, M. Lohi, and K. Madani, "Classification Accuracy performance of Naïve Bayesian(NB) Bayesian Networks(BN) Lazy Learning of Bayesian Rules (LBR) and Instance-Based Learner (IB1) comparative study", Computer Engineering & Systems ICCES 2008, International Conference, 2008, pp.210 – 215.

[25] M. Bartlett, I. Bate, and J. Cussens, "Learning Bayesian Networks for Improved Instruction Cache Analysis", Machine Learning and Applications (ICMLA) Ninth International Conference, 2010, pp.417 – 423.

[26] Thailand: National Electronics and Computer Technology Center, National Science and Technology Development Agency, Ministry of Science and Technology of Thailand, "Internet User Profile of Thailand 2010", at: http://nstda.or.th/prs/index.php/53-4, (accessed December 2011).

[27]Website Ranking", at: http://www.stats.in.th/?cmd=stats_global&list=y, (accessed December 2011).