The Enhancement of Password Security System Using Keystroke Verification

Taweetham Limpanuparb[1,2]

[1]*Faculty of Science, Mahidol University, Rama VI Rd., Bangkok 10400*
[2]*Mahidol Wittayanusorn School, Salaya, Phuttamonthon, Nakhon Pathom 73170*
Email: taweetham@acm.org, u4705075@student.mahidol.ac.th

ABSTRACT - At present computer security is increasingly important as global access to information and resources becomes an integral part of many aspects of our lives. Reliable methods for user verification are needed to protect both privacy and important data. Password verification has been used for a long time. Due to its simplicity and affordability, this technique had been very successful. However, processing power of computers has increased dramatically which has made the use of password verification insufficient. Enhancement of security in password verification can simply be conducted by increasing password length, changing passwords more often, and/or using meaningless strings as passwords. Nonetheless these methods may not work efficiently because of human memory limitation. Keystroke verification, a biometric method, is based on user typing parameters which can be defined as key hold time and interkey time. These two parameters are collected while users type in their passwords. Novel statistical methods are proposed here to determine whether the keystroke data actually belong to the user.

KEY WORDS – keystroke verification, biometrics

บทคัดย่อ – การรักษาความปลอดภัยในระบบคอมพิวเตอร์มีความสำคัญมากขึ้นในปัจจุบัน เมื่อการเข้าถึงข้อมูลข่าวสารและทรัพยากรบนคอมพิวเตอร์กลายเป็นส่วนหนึ่งของวิถีชีวิตมนุษย์ เราต้องการวิธีการตรวจสอบผู้ใช้ที่ถูกต้องเพื่อปกป้องสิทธิ์ส่วนบุคคลและข้อมูลที่มีค่า การตรวจสอบผู้ใช้ด้วยรหัสผ่านเป็นวิธีการดั้งเดิมซึ่งได้รับการยอมรับมาเป็นเวลานานเนื่องจากเป็นวิธีการที่สะดวกที่สุด อย่างไรก็ตามเมื่อกำลังการประมวลผลของคอมพิวเตอร์ได้เพิ่มขึ้นอย่างรวดเร็วทำให้การตรวจสอบรหัสผ่านเพียงอย่างเดียวนั้นไม่เพียงพอ การเพิ่มความปลอดภัยของระบบรหัสผ่านอาจทำได้โดย การเพิ่มความยาวของอักขระรหัสผ่าน เปลี่ยนรหัสผ่านให้บ่อยขึ้น และ/หรือ ใช้อักขระที่ไม่มีความหมายเป็นรหัสผ่าน อย่างไรก็ตามวิธีการเหล่านี้อาจไม่สามารถใช้ได้อย่างมีประสิทธิภาพ เนื่องจากความทรงจำที่จำกัดของมนุษย์ การตรวจสอบจังหวะการพิมพ์เป็นวิธีการตรวจสอบค่าทางชีววิทยา ซึ่งอยู่บนพื้นฐานของการวัดค่าเวลากดแป้นพิมพ์และค่าเวลาระหว่างแป้นพิมพ์ ค่าเวลาทั้งสองประเภทนี้จะถูกจัดเก็บระหว่างที่ผู้ใช้ป้อนรหัสผ่าน วิธีการทางสถิติแบบใหม่ได้ถูกนำเสนอขึ้นมาเพื่อตัดสินว่าข้อมูลเวลาเหล่านี้เป็นของผู้ใช้จริงหรือไม่

คำสำคัญ - การตรวจสอบจังหวะการพิมพ์, ไบโอเมตริกส์

Honorable Mention Award from Association for Computing Machinery (ACM)
Intel International Science and Engineering Fair 55th
May 9 - May 15, 2004 Portland, Oregon

National First Place Award & Intel Excellence in Computer Science Award
Young Scientist Competition in Computer Science Project (YSC.CS 2004)
February 5 – February 7, 2004 Thailand Science Park

This research was supported by NECTEC's YSC.CS 2004 C008 grant.

Portions of this article were presented at 30th Congress on Science and Technology of Thailand, 19-21 October 2004 at Impact Exhibition and Convention Center, Muang Thong Thani, Thailand.

1. Introduction

At present, computer security is increasingly important as electronic crimes become more destructive. Methods for user verification are needed to protect both privacy and important data.

Password verification has been used for a long time and has been a successful method because of its simplicity and affordability. However, processing power of computers has been dramatically increasing so password verification is insufficient. For instance, there are several kinds of password dictionaries; the famous worm in 1989 was widely spread across networks by using random passwords [6].

Enhancement of security in password verification is being conducted by increasing password length, changing passwords more often, and using meaningless strings as passwords. In everyday life, we use passwords for accessing different media and it is difficult to remember all different meaningless passwords for different accounts.

Because of biometrics [1], a group of methods using human characteristics as criteria for user verification, there is no need to remember anything and hardly anything to forge. However, it usually requires additional hardware which can cost a lot and is still under development.

Keystroke verification, a biometric method, does not require any additional hardware, so it costs the least compared with other methods. It can also instantly install to the system so it is one of the most interesting methods for research and development. It may lead to useful application in the near future [2].

Although keystroke verification has been researched for a long time, it still has adaptable shortcomings. It has many approaches but still is not widely used.

There are several techniques [6-8] for keystroke verification e.g. statistical methods, neural network, and fuzzy logic. Each technique has its advantages and disadvantages. In this research, the statistical methods were studied because of the following advantages. Increasing the number of users does not significantly affect speed and complexity of the system. The statistical method is the simplest method yet it yields efficiency.

The performance of neural network technique seems to be superior to the other techniques [4], while neural network have a fundamental limitation in training requirement [9].

Obaidat and Sadoun [5] reported a 0% error rate in user verification using neural network approach, which is the best result in this time, using 7-character-long login name. However, the assumption in this research is impractical [10]. First, imposter patterns were used in training, while in the real situation these patters are not available. Second, a huge training data set of 6,300 per user was used. Third, the training and test pattern were not chronologically separated.

In previous research [8] implied that statistical method yielded superior results than neural work which have been known as the best method. Thus, there are rooms for the improvement of statistical method.

Two types of keystroke data can be described as following.

- **Key hold time**: an exact time that a user hold a key

  = *time stamp when a key is released - time stamp when a key is pressed*

- **Interkey time**: an exact time between pressing and/or releasing of two sucessive keys There are two techniques to obtaine interkey time. Conventional method yields interkey time that can be positive or negative, while adapted method yields only non-negative interkey time.

  = *time stamp when next key is pressed – time stamp when current key is released (conventional)*

  = *time stamp when next key is pressed – time stamp when current key is pressed (adapted)*

  # interkey times = # keys has been pressed -1

  # key hold times = # keys has been pressed.

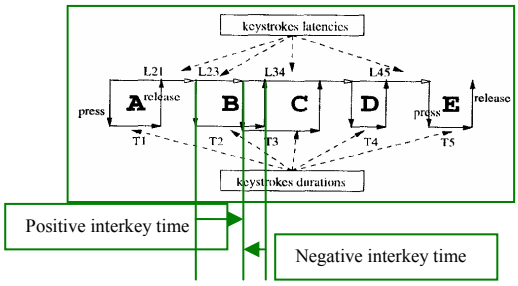

Figure 1. Keystroke measurement [4]

Keystroke measurement was shown in figure 1. This measurement technique [4] was adapted from the former one. It can give positive interkey time although next button was pressed before previous button was released.

Keystroke data obtained from this measurement are key hold time (T_i) and interkey time (L_i) with a total of *2n-1* elements. (Where n is number of pressed keys.) Keystroke data will be served as input for all techniques [4-5].

2. Study I

Using preliminary empirical data, two elementary statistical approaches were proposed and were tested using seventeen subjects. Subjects were asked to use computer program written for collecting their keystroke data. The program consists of four parts - register, training, FRR test, and FAR

test. Verification techniques are newly proposed in this part as following.

- **Pressed-Released Key sequence** is the time order sequence of keys that are pressed or released. For example, password "Abc" has various pressed-released key sequence. "[Shift] [A] (A) (Shift) [B] (B) [C] (C)" is a typical pattern, but it is possible to be "[Shift] [A] (Shift) (A) [B] [C] (B) (C)". "[]" shows key which is pressed and "()"shows key which is released. Users have to type exactly the same sequence to be verified by the system. Non-negative interkey time measurement comes after this pressed-released key sequence verification.

Statistical approaches

$$A_j = \begin{cases} 1 \text{ if Prob}(|Z_j|) > 0.975 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{if } \sum_{j=1}^{2n-1} A_j > Round(.05 \times (2n-1))$$

then reject user
else accept user

Figure 2 CAV approach

$$\overline{|Z|} = \frac{\sum_{j=1}^{2n-1} |Z^*_j|}{2n-1}$$

$$\text{if Prob}(\overline{|Z|}) > 0.975$$

then reject user
else accept user

Figure 3 PMAS approach

- **Counting of Abnormal Values approach**, CAV approach is based on assumption that all time values a user typed should be in 95% confident interval of a standard normal curve. However, there can be time values that are not in this interval according to human error when typing a long password. Thus, the number of time values which is not in the confident interval, called abnormal values, is counted. If the number of abnormal values is more that five percent, this method rejects the test data. Otherwise, the test data is accepted as a valid user data. The pseudo code is shown in figure 2.

- **Probability of Mean Absolute of Standard score approach**, PMAS approach is based on an assumption that we should consider overall time values instead of considering some designated abnormal values. Therefore, a mean of absolute values of standard scores is used. If any absolute of standard score is extremely high, it influences the mean value. If mean of absolute standard scores is in 95% confident interval, this method accepts that test data. Otherwise, test data is rejected as shown in figure 3.

Reference patterns for these statistical approaches consisted of 30 attempts data. Both key hold times and interkey times were used. For PMAS method, data in first and fourth quartiles were excluded as outliners to ensure the robustness as a conventional research practice.

- **Adapted method for decreasing type I error (EI) approach** Both PMAS approach and CAV approach may give high type I error, so this approach accepts the user if PMAS approach ***or*** CAV approach accepts the user.

- **Adapted method for decreasing type II error (EII) method** Both PMAS approach and CAV approach may give high type II error, so this approach will accept the user if PMAS approach ***and*** CAV approach accept the user.

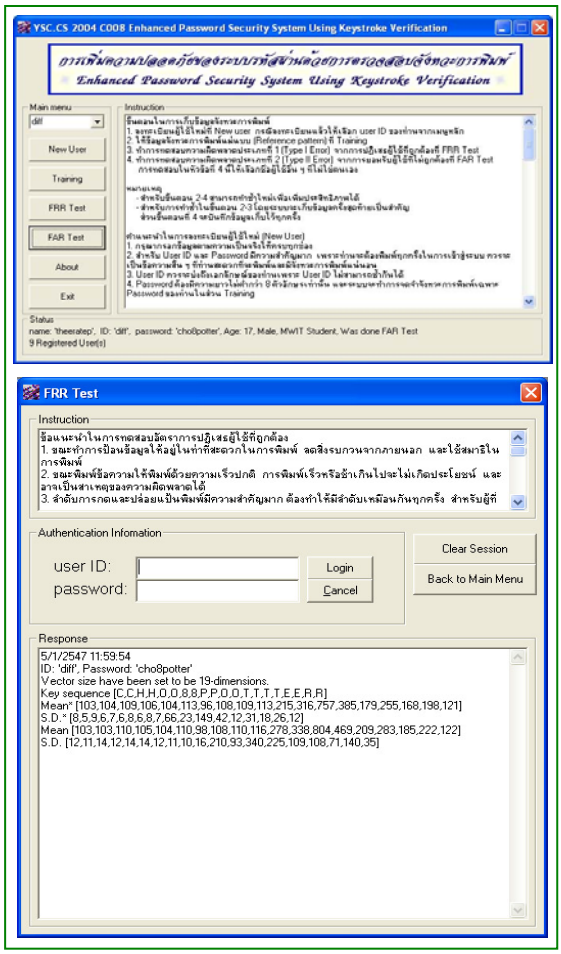

Figure 4 computer program written for study I

Results and Discussion

Table 1 False Rejection Rate (FRR)

	False Rejection Rate				Sequence Error
	CAV	PMAS	EI	EII	
Percentage	12.16	31.18	9.02	34.31	8.01
Percentage (combined with Sequence Error)	18.57	35.09	15.84	37.82	N/A

Table 2 False Acceptance Rate (FAR)

	False Acceptance Rate				Sequence Match
	CAV	PMAS	EI	EII	
Percentage	10.65	5.52	10.85	5.33	76.82
Percentage (combined with Sequence Error)	8.18	4.24	8.33	4.09	N/A

Firstly, results suggested that error rates could be lowered because high error rates occurred with some users only. (Detail data was not shown.)

Passwords are most important factor, as FAR of password "gggggggg" is nearly 100%. It suggested that keystroke verification can not well perform with all kinds of password. However, aim of this work is to enhance security of overall password verification system, therefore, some specific cases could be omitted. As implementation problems that error seems too high and valid users may be disturbed by the additional system, therefore the good system should rarely disturb valid users. It means that FRR of such system should be 0% and FAR can be any value which as less as possible, depending on user characteristics and password characteristics.

The results suggested that single constant threshold may not work well. Thus, keystroke data in study I was reanalyzed with revised models that can be described as following.

- **CAV approach** was adapted using two variable thresholds - probability threshold (t_1) and number of abnormal values threshold (t_2) as shown in figure 5. To prove that this model can well perform t_1, t_2 and errors were plotted as shown in figure 6 using data from user ID 13. It was found that FRR was reduced from 33.33% to 0%, while FAR was increased from 0% to 1.67%, where t_1=2 and t_2= 0.990.

- **PMAS approach** was also adapted using two variable thresholds - outliner threshold (t_1) and probability threshold (t_2) as shown in figure 7. To prove that this model also performed better than the model in study I, t_1, t_2 and errors were plotted as shown in figure 8 using data from user ID 13. It was found that FRR was reduced from 73.33% to 0%, while FAR was still 0%, where t_1=3.0 and t_2= 0.835.

According to figure 6 and 8, if t_1, t_2 increase, FRR will be decreased. Unfortunately, as FRR decreases, FAR is increased. However, there should be optimal t_1, t_2 that give the least overall error.

$$A_j = \begin{cases} 1 \text{ if } prob(z_j) > t_1 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{if } \sum_{j=1}^{2n-1} A_j \geq t_2$$

then reject user
else accept user

Figure 5 revised CAV approach

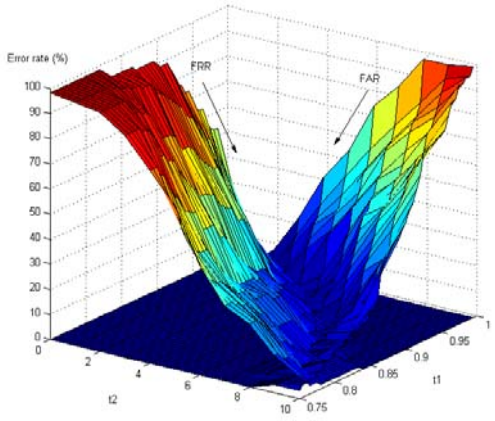

Figure 6 error rates of revised CAV

$$\overline{|Z|} = \frac{\sum_{j=1}^{2n-1} |Z^*_j|}{2n-1}$$

if Prob($\overline{|Z|}$) > t_2

then reject user
else accept user

$\bar{x}^*$ and s* is calculated from data which are in $\bar{X} \pm S\,t_1$

Figure 7 revised PMAS approach

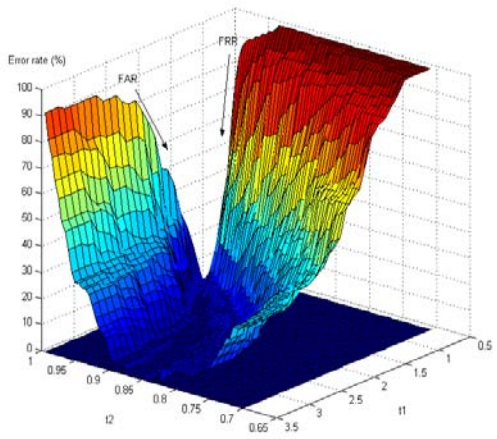

Figure 8 error rates of revised PMAS

For these statistical approaches, EI and EII approaches depend on elementary two approaches, so further study should focus on elementary two approaches and other factors that may affect efficiency of them. Factors that should be considered in further study are number of attempts used to train system, profile updating, and duration of users' participate in the research.

Nonetheless, keystroke measurement in this study performed in Microsoft Windows based environment which may not yield high accuracy as MS-DOS based environment. Most of previous research had used MS-DOS based program to collect their keystroke data. Therefore, in the next study, MS-DOS based program was written for preferred results.

3. Study II

Results in study I suggested that there were rooms for improvement. Therefore, statistical approaches and measurement technique were adapted in this study as following.

- **Pressed Key sequence**
  Pressed-Released Key sequence in study I was reduced to pressed key sequence, due to its high error rate. This technique stored and verified only pressed keys, so it may reduce sequence error.
- **Statistical approaches** It was found that single constant threshold value may not work effectively in keystroke verification. Therefore, the revised models were proposed as shown in figure 5 and figure 7. Revised models which used two threshold values (t_1 and t_2) needed threshold selection technique. For a number of trained data, t_1 and t_2 have to give maximum number of acceptance of trained data. First minimize t_2 then, minimize t_1. Exhaustive search technique was performed to get t_1 and t_2.

i. CAV approach

$t_1 \in \{0.750, 0.755, 0.800, \ldots, 0.995\}$
$t_2 \in \{0, 1, 2, \ldots, \text{number of pressed keys}/2\}$

ii. PMAS approach

$t_1 \in \{0.5, 0.6, 0.7, \ldots, 2.6\}$
$t_2 \in \{0.750, 0.755, 0.800, \ldots, 0.995\}$

The revised approaches were tested using 16 combinations varying in sample size, statistical approaches, and adaptive approaches. "I" denotes that a new attempt was instantly updated to a profile. "A" denotes that only an accepted attempt was updated to a profile as showed in table 3.

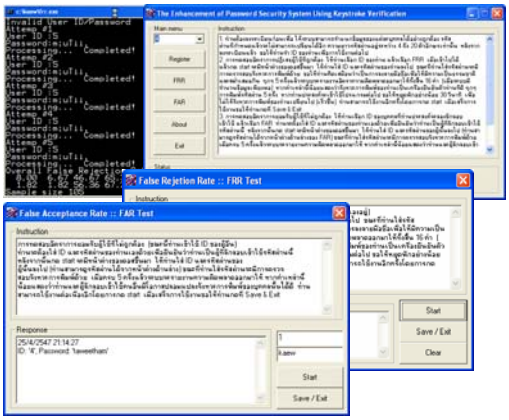

Figure 9 computer program written for study II

Table 3 Experimental Diagram

Data size	30				40				50				60			
Statistical Approaches	CAV		PMAS		CAV		PMAS		CAV		PMAS		CAV		PMAS	
Profile Update methods	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A

Results and Discussion

According to table 4, user ID 1, 6, and 14 used only number as their passwords, thus it is obviously found that their error rate are higher than other users. Accepted attempt - profile updating always gave more error than instant profile updating, due to changing in user typing pattern.

Results were shown here obtained from different statistical approaches, and numbers of trained data. It suggests that trained data size and statistical method that yield the best result vary from user to user.

Table 4 Minimum overall error for each user

ID	Password length	Password	Minimum overall error	
			A	I
1	10	416925342	20.96	7.14
2	7	rujcdh;	5.71	0.00
3	8	deardean	0.00	0.00
5	6	may290	10.91	1.82
6	5	12345	33.33	10.79
7	8	m^;yoo6	18.46	1.82
14	6	281242	17.88	8.78
15	8	dkiditme	40.00	20.00
16	11	magicaddict	0.00	0.00
18	7	2276pnr	22.00	3.33
20	8	kansinee	12.50	4.00
$\bar{X}$	**7.64**		**16.52**	**5.24**

Table 5 Pearson Correlations

		PWDL	DUR	FRR	FAR	OVR
PWDL	ρ	1	(*).151	.029	(**)-.232	-.137
	Sig.	.	.045	.704	.002	.073
	N	176	176	172	176	172
DUR	ρ	(*).151	1	(*).195	.092	(**).205
	Sig.	.045	.	.010	.226	.007
	N	176	176	172	176	172
FRR	ρ	.029	(*).195	1	.049	(**).742
	Sig.	.704	.010	.	.519	.000
	N	172	172	172	172	172
FAR	ρ	(**)-.232	.092	.049	1	(**).706
	Sig.	.002	.226	.519	.	.000
	N	176	176	172	176	172
OVR	ρ	-.137	(**).205	(**).742	(**).706	1
	Sig.	.073	.007	.000	.000	.
	N	172	172	172	172	172

(*) Correlation is significant at the 0.05 level (2-tailed).
(**) Correlation is significant at the 0.01 level (2-tailed).
PWDL – Password length (including shift key)
DUR – Duration of participation in the research (day)
FRR – False Rejection Rate (%)
FAR – False Acceptance Rate (%)
OVR – Overall Error Rate (%)

Duration of participation in the study and password length were considered against errors. Table 5 showed Pearson correlations between errors and corresponding factors.

Duration of participation in the study and FRR has a correlation coefficient of .195 (at α=.05). It also correlated with overall error rate and gave a correlation coefficient of .205 (at α=.01). This suggested that user typing pattern varied from day to day. In other words, the longer duration, the more errors are expected. Results were in agreement with the hypothesis stating that there were correlations between errors and relevant factors.

FAR and password length has a correlation coefficient of -.232 (at α=.01). It showed that as password length increases, FAR tends to reduce. Therefore, the greater password length resulted in more secure system. On the other hand, FRR and password length had no significant correlation. Therefore, system tended to give a constant FRR, as discussed in study I, preferred system should give minimum FRR which does not depend on password. Results were in agreement with hypothesis stating that type I error can be minimized independently on any arbitrary password.

4. Conclusions

Novel keystroke verification approaches were proposed based on statistical methods. Two elementary approaches and two combined approaches were examined using their error rates. Error rates in study I were reduced in study II using two revised elementary approaches.

The prime goal of this research was to develop a verification technique that can be instantly installed in conventional systems, in order to prevent interference with the normal activity of valid users. FRR was minimized in study II and it did not depend on a password length. On the contrary, FAR depended on password length. It suggested that effect on valid user was minimized.

Password is still the main subject that is relevant to error. Some passwords cannot give promising results in keystroke verification. Similar to conventional verification technique, a password such as "ggggggg" or "1234" is not appropriate for providing security. However, keystroke verification can slightly improve security even in this scenario.

The minimum overall error rate or best case was 5.24%, despite there being many passwords such as "1234". This minimum overall error rate showed promising results which indicated that these statistical approaches could be implemented as a complement to the conventional system.

However, further study should be conducted. First of all, selection technique for approach combinations is needed to be developed. Effect of variation from day to day should be decreased to be a constant. Nonetheless, two thresholds values in CAV approach may be further delineated in to a value and a vector.

Acknowledgements

I would like to express my sincere appreciation and gratitude to teachers, institutions, and individuals who gave me a lot of support. I would like to thank Associate Professor Dr.Montip Tiensuwan, statistics professor, who gave me valuable advice, Ms.Laokhwan Ngamprasit, designate computer teacher, who oversaw this project, and Ms.Puneee Lumwanwong, the mathematics teacher who gave helpful suggestions. I would like to thank Assistant Professor Dr.Bundit Thipakorn and Dr.Krissanapong Kirtikara for their invaluable support. I wish to thank all teachers in Mahidol Wittayanusorn School, especially in Computer Department and all staff in the Computer Olympics Project who gave me useful knowledge which is the base in this project. I also would like to thank the Young Scientist Competition in Computer Science and Engineering Project (a collaboration between National Electronics and Computer Technology Center (NECTEC), Intel Microelectronics (Thailand) Ltd, and many universities in Thailand) for their crucial support.

I wish to express a deep sense of appreciation and immense thanks to my family for their unending encouragement. Last but not least, I would like to thank all human subjects who gave their valuable time to participate in this research without any hesitation.

References

[1] Miller B., "Vital signs of identity [biometrics]", *IEEE Spectrum*, Vol. 31, Issue: 2, February 1994, pp. 22 -30.

[2] Shepherd S.J., "Continuous authentication by analysis of keyboard typing characteristics", *European Convention on Security and Detection*, 16-18 May 1995, pp. 111 -114.

[3] Renee Napier, William Laverty, Doug Mahar, Ron Henderson, Michael Hiron, and Michel Wangner, "Keyboard user verification: toward an accurate, efficient, and ecologically valid algorithm", *Int. J. Human-Computer Studies* (1995) 43, pp. 213-222.

[4] Daw-Tung Lin, "Computer-access authentication with neural network based keystroke identity verification", *International Conference on Neural Networks, 1997*, Vol. 1, June 1997, pp. 174 -178,

[5] Obaidat M.S. and Sadoun B., "Verification of computer users using keystroke dynamics", *IEEE Transactions on Systems, Man and Cybernetics, Part B*, Vol. 27 Issue: 2, April 1997, pp. 261 -269

[6] De Ru, W.G. and Eloff J.H.P., "Enhanced password authentication through fuzzy logic", *IEEE Expert*, Vol. 12, No. 6, November-December 1997, pp. 38 -45.

[7] Fabian Monrose and Aviel D. Rubin, "Keystroke dynamics as a biometric for authentication", *Future Generation Computer System*, 16 (2000), pp. 351-359.

[8] Haider, S., Abbas, A., and Zaidi A.K., "A multi-technique approach for user identification through keystroke dynamics, Systems", *IEEE International Conference on Man, and Cybernetics*, 2000, Vol. 2 , 8-11 October 2000, pp. 1336 -1341.

[9] Fabian Monrose; Aviel D. Rubin, Keystroke dynamics as a biometric for authentication, Future Generation Computer System 16 (2000) 351-359.

[10] Enzhe Yu; Sungzoon Cho, *GA-SVM wrapper approach for feature subset selection in keystroke dynamics identity verifeation***,** Neural Networks, 2003. Proceedings of the International Joint Conference on , Volume: 3 , July 20 - 24, 2003, 2253 -2257.

[11] สมจิต วัฒนาชยากูล. *สถิติพื้นฐานสำหรับนักวิทยาศาสตร์*. กรุงเทพ ฯ: สำนักพิมพ์ประกายพรึก, 2546.

[12] สานนท์ เจริญฉาย. *โปรแกรม C คำนวณสถิติ*. กรุงเทพ ฯ: โรงพิมพ์มหาจุฬาลงกรณ์ราชวิทยาลัย, 2545.

Taweetham Limpanuparb was born in Bangkok, Thailand in 1986. He has been a scholar in Junior Science Talent Project (JSTP) since 2001. He graduated from Mahidol-Wittayanusorn School (the first science high school of Thailand) in 2004. He is currently a scholar in the Development and Promotion of Science and Technology Talents Project (DPST) at Faculty of Science, Mahidol University. Both JSTP and DPST encourage him to continue with his study and research till he finishes post doctorate. His areas of interest are widely in science and technology on a multidisciplinary basis.