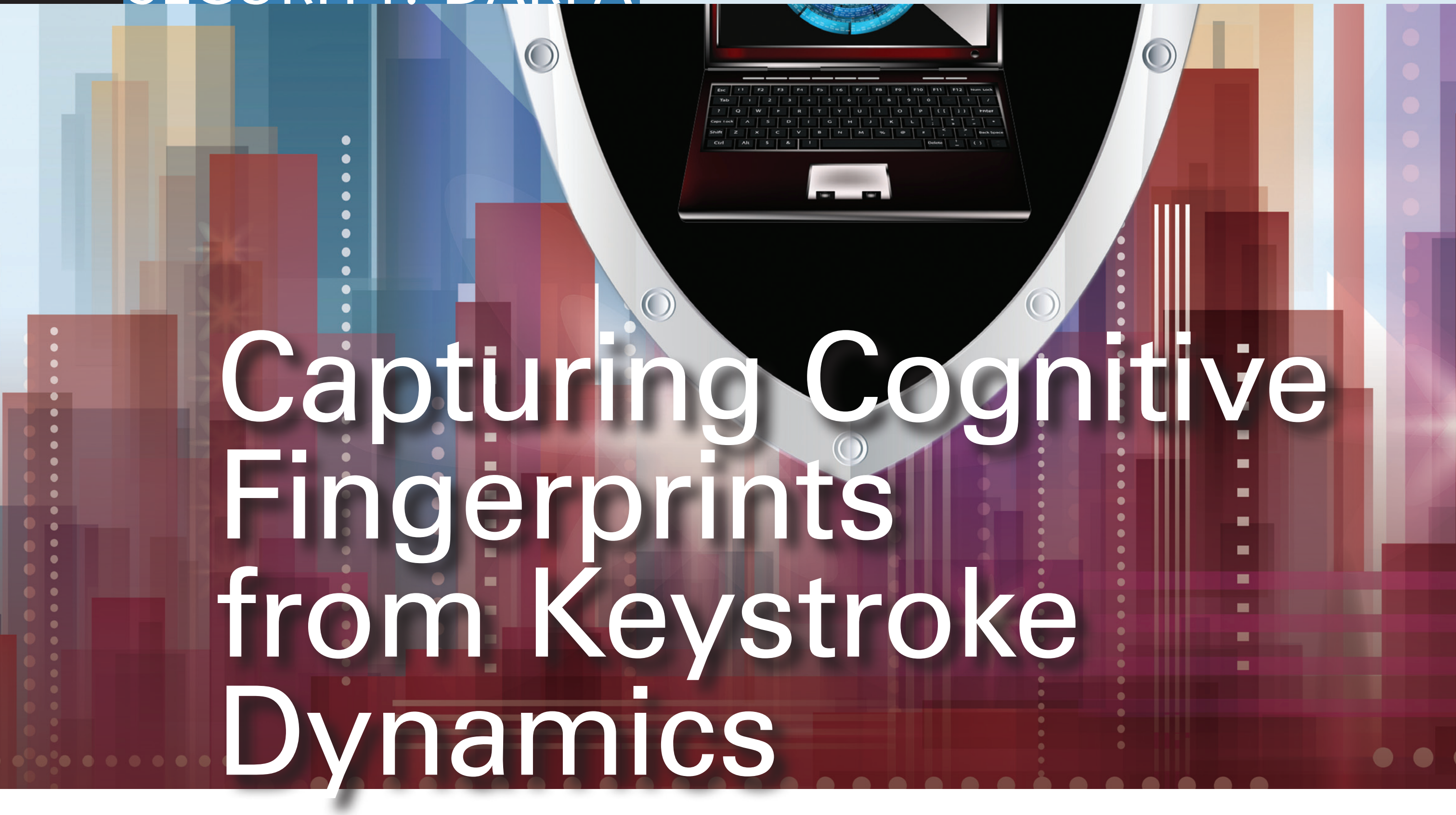

Capturing Cognitive Fingerprints from Keystroke Dynamics

J. Morris Chang, Chi-Chen Fang, Kuan-Hsing Ho, and Norene Kelly, *Iowa State University*
Pei-Yuan Wu, *Princeton University*
Yixiao Ding, Chris Chu, Stephen Gilbert, and Amed E. Kamal, *Iowa State University*
Sun-Yuan Kung, *Princeton University*

The authors present an authentication system that applies machine learning techniques to observe a user's cognitive typing rhythm. Results from a large-scale experiment at Iowa State University show the system's effectiveness.

Conventional authentication systems verify a user only during initial login. Active authentication performs verification continuously as long as the session remains active. This work focuses on using behavioral biometrics, extracted from keystroke dynamics, as "something a user is" for active authentication. This scheme performs continual verification in the background, requires no additional hardware devices, and is invisible to users.

Keystroke Dynamics

Keystroke dynamics—the detailed timing information of keystrokes when using a keyboard—has been studied for the past three decades. The typical keystroke interval time, referred to as a *digraph*, is expressed as the time between typing two characters. A user's keystroke rhythms are distinct enough from person to person for use as biometrics to identify people. However, keystroke rhythm has generally been considered less reliable than physical biometrics, such as fingerprints. The main challenge is the presence of within-user variability.

Owing to this within-user variability of interval times among identical keystrokes, most research efforts have focused on verification techniques that can manage such variability. For example, researchers proposed a method called *degree of disorder* to cope with time variation issues,[1,2] arguing that although the keystroke typing durations usually vary between each digraph, the

1520-9202/13/$31.00 © 2013 IEEE

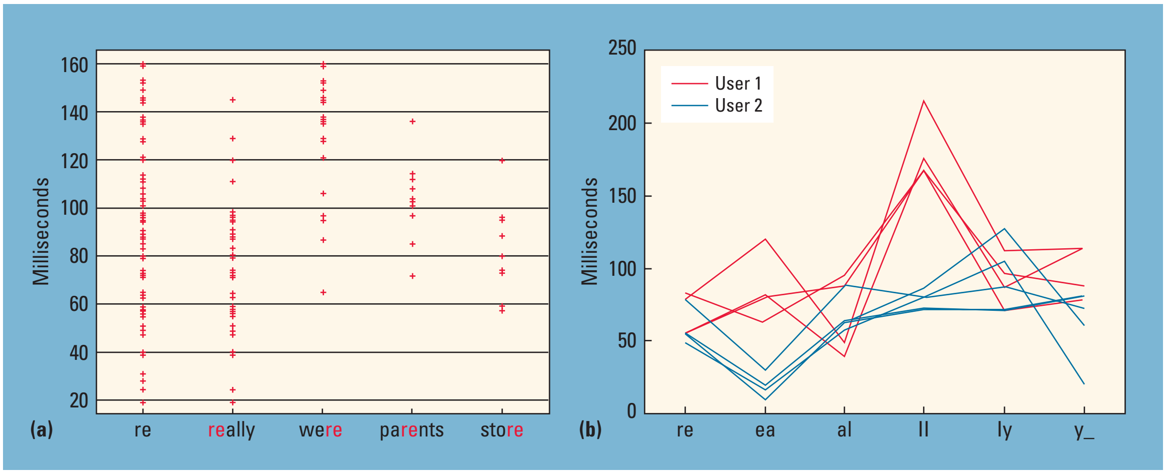

Figure 1. A cognitive factor can affect the typing rhythm of a specific word: (a) the digraph "re" from the same user and (b) two users typing the same word, "really."

order of the timing tends to be consistent. This suggested that the distance of the order between two keystroke patterns can be used to measure the similarity.

A recent survey on biometric authentication using keystroke dynamics classified research papers on the basis of their feature-extraction methods, feature-subset-selection methods, and classification methods.[3] Most of the systems described in the survey were based on typing rhythms for short sample texts, which are dominated by users' physical characteristics (such as how fast your fingers can move) and are too brief to capture a "cognitive fingerprint." In the current keystroke-authentication commercial market, some products combine the timing information of the password with password-based access control to generate a hardened password.[4]

Here, we present a biometric-based active authentication system that continuously monitors and analyzes various keyboard behaviors performed by the user. We extract the features from keystroke dynamics that contain cognitive factors, resulting in cognitive fingerprints. Each feature is a sequence of digraphs from a specific word. This method is driven by our hypothesis that a cognitive factor can affect the typing rhythm of a specific word. Cognitive factors have been largely ignored in previous keystroke dynamics studies.

Searching for Cognitive Fingerprints

Physical biometrics rely on physical characteristics, such as fingerprints or retinal patterns. The behavioral biometric of keystroke dynamics must incorporate cognitive fingerprints to advance the field, but the cognitive fingerprint doesn't have a specific definition. We hypothesize that natural pauses (delays between typing characters in words) are caused by cognitive factors (for example, spelling an unfamiliar word or pausing after certain syllables),[5–9] which are unique among individuals. Thus, a cognitive factor can affect the typing rhythm of a specific word.

In this research, each feature is represented by a unique cognitive typing rhythm (CTR), which contains the sequence of digraphs from a specific word. Such features include natural pauses among the CTR's timing information (digraphs, for example) and could be used as a cognitive fingerprint. Conventional keystroke dynamics don't distinguish timing information for different words and only consider a collection of digraphs (such as trigraphs or *n*-graphs). Cognitive factors have been ignored.

Figure 1a shows a collection of digraphs observed for one user. It might seem as if the collection of digraphs represents a part of a keystroke rhythm, but in reality, the digraphs are clustered around different words. For example, we can separate the collection of digraphs "re" according to four different words (*really*, *were*, *parents*, and *store*). This shows that examining digraphs in isolation might result in missing some important information related to specific words. Figure 1b shows two users who both typed the word "really" several times, illustrating the typing rhythm for each.

This observation confirms our hypothesis: a cognitive factor can affect the typing rhythm of a specific word. Thus, we extract CTRs from keystroke dynamics and use them as features (cognitive fingerprints) for active authentication. Each

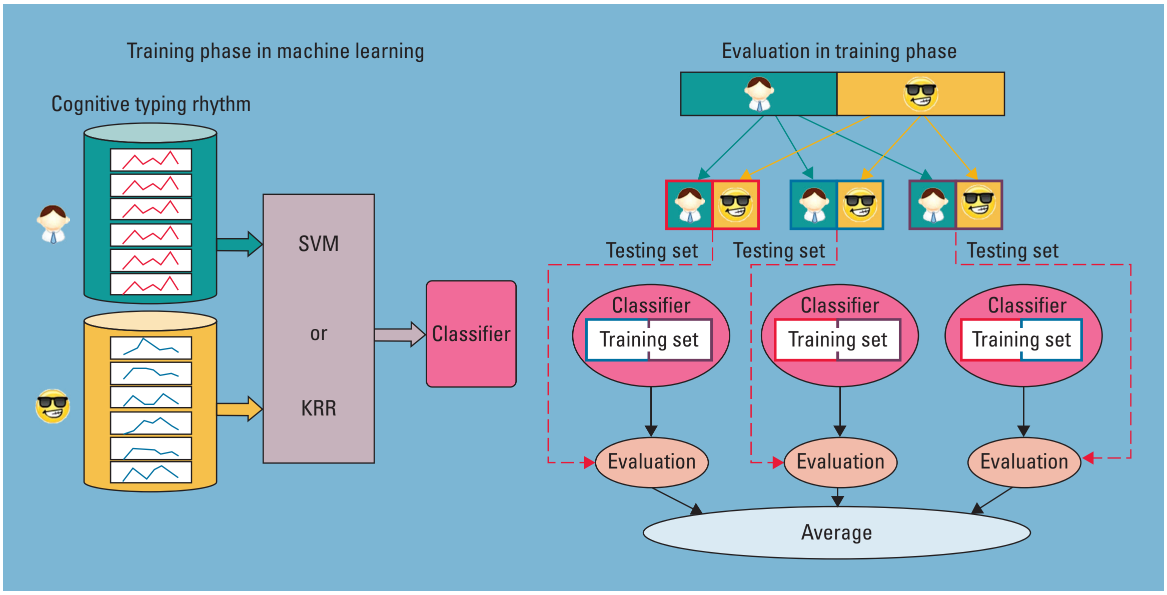

Figure 2. Training and cross-validation in machine learning: (a) training phase for building a classifier and (b) evaluation to obtain the confidence of each classifier.

feature is a sequence of digraphs of a specific word (instead of a collection of digraphs). For each legitimate user, we collect samples of each feature and build a classifier for that feature during the machine-learning training phase.

Building an Authentication System

We developed two authentication systems based on two different machine-learning techniques. The first one uses an off-the-shelf support vector machine (SVM) library,[10] and the second one employs a library developed in-house, based on kernel ridge regression (KRR).[11] We used these libraries to build each classifier during the training phase.

Although we can't know the patterns of all imposters, we use patterns from the legitimate user and some known imposters to build each classifier so it can detect a potential imposter. In machine learning, this is known as a two-class (legitimate user vs. imposters) classification approach. We built a trained profile with multiple classifiers for each legitimate user. Then, during the testing phase (authentication), we gave a set of testing data to the trained profile for verification. Each classifier under testing yielded a matching score between the testing dataset and trained file. The final decision (accept or reject) was based on the sum of scores from all classifiers.

The two systems had different basic machine-learning libraries (SVM and KRR) but shared the same feature selection and fusion method. Using the fusion method, we evaluated each classifier to determine the confidence level of its decision. We conducted this evaluation during the training phase using datasets from each legitimate user and from imposters (see Figure 2). We separated the dataset into k equal-sized subsets. Each time, we used $k - 1$ subsets as training data, and we used the remaining subset for testing. We repeated the testing k times, until each subset had been used to test the model. This technique is called *k-fold cross-validation* (or *rotation estimation*).

The test results let us estimate the probabilities of the classifier's true acceptance (P_{ta}) and false acceptance (P_{fa}) rates. For example, after testing with the dataset from a legitimate user, there were N acceptances out of M samples, so P_{ta} is N/M. The confidence of the acceptance decision (W_a) is expressed as the ratio of P_{ta} to P_{fa}. The confidence of the rejection decision (W_r) is expressed as the ratio of the probability of true rejection $(1 - P_{fa})$ to the probability of false rejection $(1 - P_{ta})$.

After the training, in the trained profile, we have W_a and W_r for each classifier. During the testing phase, each classifier generates a decision (acceptance or rejection). Either W_a or W_r will be applied to this decision. The final decision is based on the sum of the scores from all involved classifiers.

A Large-Scale Experiment

We developed a Web-based software system to collect the keystroke dynamics of individuals in a large-scale testing project conducted at Iowa State University (ISU). This system provided three simulated user environments: typing short sentences, writing short essays, and browsing webpages. We stored the users' cognitive fingerprints in a database for further analyses and applied machine-learning techniques to authenticate users by performing pattern recognition.

During November and December of 2012, we sent email invitations to 36,000 members of the ISU community. There were 1,977 participants who completed two segments, each lasting approximately 30 minutes, resulting in approximately 900 words for each participant for each segment. In addition, 983 participants (out of the 1,977) completed another segment of approximately 30-minutes in length, in which we collected approximately 1,200 words for each participant. We then developed 983 individual profiles (trained files). Each profile was trained under two-class classification, in which one legitimate user had 2,100 collected words, and the imposter training set was based on collected words from the other 982 known participants. Each profile was tested with the data of the 1,977 participants (with a testing dataset of 900 words per participant).

Figure 3 shows the results. Figure 3a summarizes the performance comparison of the two verification systems, and Figure 3b shows the detection error trade-off chart from the KRR-based system. In this experiment, each legitimate profile had been tested using the dataset collected from the same user; seven (out of 983) users were recognized as imposters using the SVM library, so it correctly identified the other 976 users, and 17 (out of 983) users with the KRR library, so it correctly identified 966 users. Also, we tested each profile with the other 1,976 participants, and the false-accept rate was 0.055 percent for both SVM and KRR.

	SVM	KRR
FAR	0.055	0.055
FRR	0.007	0.0177
Training time	15 m/user	29 s/user
Testing time	0.6 s/user	3.5 ms/user
Size of training file	20 MB/user	1 MB/user

(a)

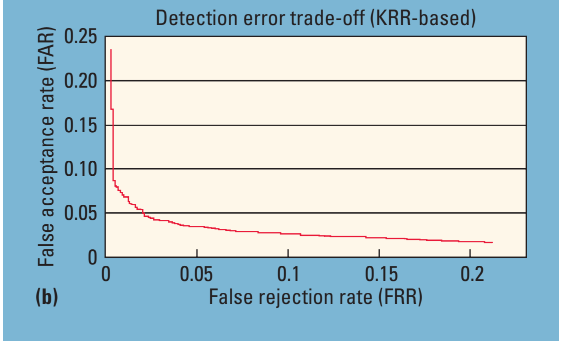

Figure 3. Experiment results: (a) performance comparison of the two verification systems and (b) the detection error trade-off chart from the kernel-ridge-regression-based system.

In summary, the proposed scheme is effective for authentication on desktop devices. Moreover, because of the increasing popularity of mobile devices, it's interesting to find the cognitive fingerprint and apply our authentication system on mobile devices. In the future, we'll study keystroke dynamics on different platforms.

References

1. F. Bergadano et al., "User Authentication through Keystroke Dynamics," *ACM Trans. Information and System Security*, Nov. 2002, pp. 367–397.
2. D. Gunetti and C. Picardi, "Keystroke Analysis of Free Text," *ACM Trans. Information and System Security*, vol. 8, no. 3, 2005, pp. 312–347.
3. M. Karnan et al., "Biometric Personal Authentication Using Keystroke Dynamics: A Review," *Applied Soft Computing*, vol. 11, no. 2, 2011, pp. 1565–1573.
4. F. Monrose et al., "Password Hardening Based on Keystroke Dynamics," *Proc. 6th ACM Conf. Computer and Communications Security*, ACM, 1999, pp. 73–82.
5. C.M. Levy and S. Ransdell, "Writing Signatures," *The Science of Writing: Theories, Methods, Individual Differences, and Applications*, Lawrence Erlbaum, 1996, pp. 149–162.
6. D. McCutchen, "A Capacity Theory of Writing: Working Memory in Composition," *Educational Psychology Rev.*, vol. 8, no. 3, 1996, pp. 299–325.
7. D. McCutchen, "Knowledge, Processing, and Working Memory: Implications for a Theory of Writing," *Educational Psychologist*, vol. 35, no. 1, 2000, pp. 13–23.
8. T. Olive, "Working Memory in Writing: Empirical Evidence from the Dual-Task Technique," *European Psychologist*, vol. 9, no. 1, 2001, pp. 32–42.
9. T. Olive et al., "Verbal, Visual, and Spatial Working Memory Demands During Text Composition," *Applied Psycholinguistics*, vol. 29, no. 4, 2008, pp. 669–687.

12. C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intelligent Systems and Technology,* vol. 2, no. 3, 2011, article no. 27.
13. S.Y. Kung, *Kernel Methods and Machine Learning*, Cambridge Univ. Press, 2013

J. Morris Chang *is an associate professor of electrical and computer engineering at Iowa State University. Contact him at morris@iastate.edu.*

Chi-Chen Fang *is a PhD student in Department of Electrical and Computer Engineering at Iowa State University. Contact him at cfang@iastate.edu.*

Kuan-Hsing Ho *is currently a master student of electrical and computer engineering in Iowa State University. Contact him at pm426015@iastate.edu*

Norene Kelly *is a doctoral student in the Human Computer Interaction program at Iowa State University. Contact her at nbkelly@iastate.edu.*

Peiyuan Wu *is currently a PhD student in the Department of Electrical Engineering at Princeton University. Contact him at peiwu@princeton.edu.*

Yixiao Ding *is currently a PhD student of electrical and computer engineering at Iowa State University. Contact him at yxding@iastate.edu.*

Chris Chu *is a professor in the Electrical and Computer Engineering Department at Iowa State University. Contact him at cnchu@iastate.edu.*

Stephen Gilbert *is the associate director of the Virtual Reality Applications Center (VRAC) at Iowa State University and is an assistant professor of industrial and manufacturing systems engineering in the human factors division. Contact him at gilbert@iastate.edu.*

Ahmed E. Kamal *is a professor of electrical and computer engineering at Iowa State University. Contact him at kamal@iastate.edu.*

S.Y. Kung *is a professor in the Department of Electrical Engineering at Princeton University. Contact him at kung@princeton.edu.*