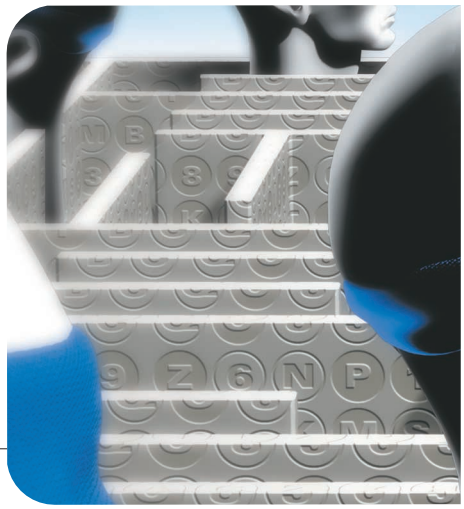
Typing Patterns: A Key to User Identification

As the deficiencies of traditional password-based access systems become increasingly acute, researchers have turned their focus to keystroke biometrics, which seeks to identify individuals by their typing characteristics. However, this field still faces many challenges before it can see full acceptance.



Alen Peacock, Xian Ke, and Matthew Wilkerson
Massachusetts Institute of Technology

Most current access systems prompt users to authenticate themselves with a username and password pair. This method of authentication relies on the password's secrecy and, in some cases, even the username's secrecy. If this secrecy is not breached, the assertion is that these tokens uniquely identify a valid user.

The problems associated with maintaining password secrecy are well understood.[1] Passwords that consist of common words, phrases, or terms associated with a particular user are universally considered weak because of the relative ease with which a third party can guess them or find them via dictionary attacks. Some systems require users to remember obscure token phrases—the more obscure, the better. Of course, obscure also usually implies "hard to remember," which is a usability liability. Consider the users' plight: not only must they choose obscure passwords, but they also must choose them repeatedly. If users access multiple independent systems, they're encouraged to use unique passwords for each one to ensure that the compromise of a single password doesn't compromise them all. In practice, though, many individuals find the burden of remembering many unique, obscure, constantly changing passwords too heavy to carry, so they don't comply fully with policies and recommendations—typically, using the same password for all accounts.[2] However, even if they do follow the best-recommended practices, passwords are still easily transferable from one party to another, whether transferred inadvertently or not: users sometimes write passwords down on paper, store them in accessible text files, accidentally expose them by entering them in the username field, and so on.

Over the past 25 years, researchers have developed authentication systems based on keystroke dynamics with the hope that they would improve traditional password system security while increasing (or at least not decreasing) usability. *Keystroke biometrics* measure typing characteristics that are believed to be unique to an individual's physiology and behavior, and thus difficult to duplicate.

Most academic papers published on keystroke biometric systems since 1980 present independent studies, each with their own samples from unique sets of individuals. The researchers collected these samples through diverse methods, and they vary widely in the mechanics of user input, the granularity of measured data, the amount of input required to train the system and authenticate users, the number of test subjects, and the diversity of these subjects' typing experience. Such nonuniformity alone makes comparison between different studies difficult; add to this difficulty the diversity of keystroke-pattern classification approaches and the application of these technologies to different domains, and the task becomes even more complex.

Although commercial interests have noted the promise of keystroke dynamics and have acquired several patents on related processes, awareness and deployment of the technology has been limited so far. Most published literature is optimistic about the potential of keystroke dynamics to benefit computer system security and usability, but several drawbacks are well known. In this article, we'll address these issues while surveying recent developments, comparing results from the field with both well-known and newly proposed metrics, and examining the potential roadblocks to widespread implementation of keystroke biometrics.

Applications

The first suggested use of keystroke characteristics for identification appeared in 1975,[3] but observations about the uniqueness of an individual's typing characteristics stretch as far back as the end of the 19th century. Telegraph operators at the time could often identify each other by listening to the rhythm of their Morse code keying patterns.[4] Let's look at some of the pertinent and interesting ways in which keystroke dynamics can be applied.

Authentication

The domain of applications that would benefit from more secure authentication without significant burdens on usability is extensive. Applications involving financial transactions are among the most likely to be targeted by attackers. Gartner Group estimates that online retailers in the US lost US$1.64 billion to fraudulent sales in 2002 and rejected another $1.82 billion in legitimate sales that looked suspicious.[5] Consumers share in the desire to keep financial information safe from prying eyes, but their tolerance for inconvenient security solutions is tempered by fraud laws that place the burden of financial loss on retailers. Another application domain is digital rights management. Here, the interests of consumers and content providers are not entirely aligned: content providers want to discourage individuals from sharing accounts, but without alienating legitimate users with additional complexity.

Keystroke dynamics is highly attractive as an authentication option precisely because of the degree of transparency it offers. The most transparent way to take advantage of it is to collect timing information on data that users already type to log in to the system—that is, username and password. A compelling space for this implementation is the Web, where it is infeasible to outfit client machines with biometric devices. Other multifactor solutions exist, but come with significantly higher infrastructure and usability costs (www.rsasecurity.com/products/securid).

A Web-based authentication solution incorporating keystroke dynamics could replace standard, form-based logins with one that can collect keystrokes. Several research studies have demonstrated the feasibility of implementing a Java applet to perform this function,[6–9] and source code is available.[9] Besides being supported by all recent browsers and operating systems, an applet can keep keystroke timing information private by sending it through an encrypted SSL connection to the server, which performs the processing. Any server responses can then be redirected back to the browser just as if a normal form-based login had occurred.

The first, and so far only, commercial product suite that offers the ability to enhance authentications with keystroke dynamics is BioPassword, distributed by BioNet Systems (www.biopassword.com). The company's flagship product targets the standard Windows login, and several dozen customers have deployed it. The company also offers a related software development kit (SDK) that lets developers integrate the technology into their own Windows applications. By the end of 2004, it plans to release a Web authentication product as well as multiplatform implementations of its SDK.

An interesting issue that has yet to be addressed is the degree to which keystroke dynamics-based authentication solutions scale as the number of users increases. The largest research study conducted in keystroke dynamics collected samples from less than 200 users, and the largest installation of BioPassword has less than 3,000 users. The user base of most consumer Web applications is undoubtedly orders of magnitude larger.

Policy decisions abound in authentication, and they directly impact a system's usability and effectiveness. A primary concern is what to do when the check on the password text succeeds, but the check on the typing pattern fails. Should the user be rejected outright, or should some additional authentication step be performed? A successful supplemental check, such as requiring the user to answer a secret question, could lead to relaxing or adapting the thresholds on matched keystroke patterns, attempting to collect keystroke data once again, or simply allowing access and alerting administrators to closely watch the account. Frequent additional checks carry additional usability costs, so it's important that systems be built with high levels of accuracy to begin with.

Identification and monitoring

Closely related to the problem of authentication is the identification of a user from a set of potential candidates. Imagine a scenario in which physical access to a system could be restricted to a set of users, and the system could decipher which user is at the keyboard.

An identification scheme can also monitor when one user takes over for another on a given machine. Existing research has touched on the idea of detecting changes in identity through continuous monitoring of freely typed text, but only empirically with a very limited sample size.[10] The benefit of monitoring is in its ability to prevent an intruder from taking over a previously authenticated session. A user who forgot to lock down his or her machine before leaving it could, for example, rely on the monitoring system to automatically lock itself down when it detects someone with a significantly different typing pattern.

Keystroke monitoring can also allow a system to detect uncharacteristic typing patterns of valid users caused by drowsiness, distraction, stress, or other factors.[11] In a task for which alertness matters, for example, such an application could automate or augment monitoring of the tasks currently performed by human supervisors.

Several privacy issues correspond to any system designed to constantly monitor users, which we'll discuss

Related work

Various experimental design and techniques have been analyzed in key published research. To see a visual comparison of the works listed below with the work referenced throughout this article, visit http://csdl.computer.org/comp/mags/sp/2004/05/j5toc.htm. We highlight these results because of their contribution to the field, and because each contains enough information to make quantitative comparisons to the other work listed.

D. Umphress and G. Williams, "Identity Verification through Keyboard Characteristics," *Int'l J. Man-Machine Studies*, vol. 23, no. 3, 1985, pp. 263–273.

J. Leggett and G. Williams, "Verifying Identity via Keystroke Characteristics," *Int'l J. Man-Machine Studies*, vol. 28, no. 1, 1988, pp. 67–76.

S. Bleha, C. Slivinsky, and B. Hussein, "Computer-Access Security Systems using Keystroke Dynamics," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 12, 1990, pp. 1217–1222.

S.A. Bleha, J. Knopp, and M.S. Obaidat, "Performance of the Perceptron Algorithm for the Classification of Computer Users," *Proc. 1992 ACM/SIGAPP Symp. Applied Computing*, ACM Press, 1992, pp. 863–866.

M. Brown and S.J. Rogers, "User Identification via Keystroke Characteristics of Typed Names Using Neural Networks," *Int'l J. Man-Machine Studies*, vol. 39, no. 6, 1993, pp. 999–1014.

M.S. Obaidat, "A Verification Methodology for Computer Systems Users," *Proc. 1995 ACM Symp. Applied Computing*, ACM Press, 1995, pp. 258–262.

D.-T. Lin, "Computer-Access Authentication with Neural Network-Based Keystroke Identity Verification," *Int'l Conf. Neural Networks*, IEEE CS Press, vol. 1, 1997, pp. 174–178.

W. de Ru and J. Eloff, "Enhanced Password Authentication through Fuzzy Logic," *IEEE Expert*, vol. 12, no. 6, 1997, pp. 38–45.

J.A. Robinson et al., "Computer User Verification Using Login String Keystroke Dynamics," *IEEE Trans. Systems, Man and Cybernetics, Part A*, vol. 28, Mar. 1998, pp. 236–241.

S. Haider, A. Abbas, and A.K. Zaidi, "A Multitechnique Approach for User Identification through Keystroke Dynamics," *IEEE Int'l Conf. Systems, Man, and Cybernetics*, IEEE CS Press, vol. 2, 2000, pp. 1336–1341.

Z. Changshui and S. Yanhua, "AR Model for Keystroker Verification," *IEEE Int'l Conf. Systems, Man, and Cybernetics*, IEEE CS Press, vol. 4, 2000, pp. 2887–2890.

M.H. Wong et al., "Enhanced User Authentication through Typing Biometrics with Artificial Neural Networks and k-Nearest Neighbor Algorithm," *Conf. Record 35th Asilomar Conf. Signals, Systems, and Computers*, IEEE CS Press, vol. 2, 2001, pp. 911–915.

F. Bergadano, D. Gunetti, and C. Picardi, "User Authentication through Keystroke Dynamics," *ACM Trans. Information and System Security*, vol. 5, no. 4, 2002, pp. 367–397.

V. Kacholia and S. Pandit, "Biometric Authentication using Random Distributions (BioART)," *Proc. 15th Canadian IT Security Symp.* (CITSS), Government of Canada, 2003; www.cse-cst.gc.ca/en/symposium/symposium.html.

E. Yu and S. Cho, "GA-SVM Wrapper Approach for Feature Subset Selection in Keystroke Dynamics Identity Verification," *Proc. Int'l Joint Conf. Neural Networks*, IEEE CS Press, vol. 3, 2003, pp. 2253–2257.

later. Of course, usability questions abound here as well. Too many alerts due to a single individual's typing inconsistencies would significantly hinder productivity.

Password hardening

A hardened password based on typing patterns can be used to create long-term, cryptographically stronger secrets for login, encryption, and more. Fabian Monrose and colleagues defined a scheme for creating and storing such a hardened password, which is stable over time, leaks no information about the password text, and yet can adjust for changes in the user's typing patterns by expiring older collected samples in favor of newer ones.[12] Their solution thwarts attempts to decipher the password using the server's stored content by a multiplicative factor, but unfortunately (and by design) provides little protection against intruders who already know the password.

Another approach to password hardening takes advantage of the larger input space of keystrokes. Users who knowingly include nonvisible keystrokes in their passwords, such as backspace or shift, force attackers to wade through an expanded search space before they can break in.

Beyond keyboards

The concept behind keystroke dynamics is not limited to the traditional keyboard: any interface in which keys must be pressed can benefit from similar techniques. Such application domains include PIN authentification at automatic teller machines and phone numbers entered through cellular devices. Early studies indicate that there is potential for authenticating users from input on numerical keypads, although the levels of accuracy are expectedly worse than with a keyboard.[13,14]

Evaluating previous research

The first studies on the effectiveness of keystroke characteristics as personal identifiers appeared in 1977 and 1980.[15–17] Over the years, researchers have evaluated many different classifiers in an effort to improve the recognition capabilities of keystroke biometrics, ranging from statistical analysis to neural networks. Delving into the de-

tails of each approach is beyond this article's scope, but in general, each classifier measures the similarity between an input keystroke-timing pattern and a reference model of the legitimate user's keystroke dynamics. The model is built by training each user-provided sample and maintaining varying characteristics depending on the classifier. The time required to generate each model also varies according to the classifier, with neural networks generally taking significantly longer than other approaches.

By comparing results from the field with commonly used metrics for measuring accuracy, we can propose new metrics for measuring the usability of keystroke systems. Unfortunately, much of the literature in this area lacks sufficient reported data to measure all features. We therefore restrict the results reported in the following sections to those reports listed in the "Related work" sidebar and those referenced throughout this article that contained enough information to make quantitative comparisons. Notably missing from these comparisons is the only commercial offering, for which published numbers are scant (although public documents at the BioNet Systems Web site do claim that performance is on par with some of the earliest published results).[18]

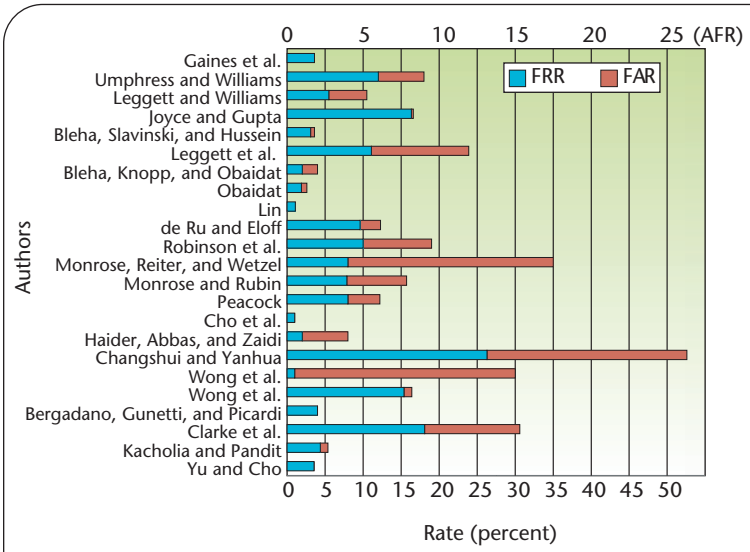


Figure 1. False rejection rate (FRR), false acceptance rate (FAR), and average false rate (AFR) for several approaches. AFR is the average of FRR and FAR, and is shown by the top *y*-axis. Systems with lower FRR, FAR, and AFR are more accurate in discriminating between users, and are thus capable of being more secure.

Classifier accuracy

Three metrics typically describe biometric classifier performance with regard to accuracy:

- *false rejection rate* (FRR), the percentage of valid (genuine) user attempts identified as imposters;
- *false acceptance rate* (FAR), the percentage of imposter access attempts identified as a valid users; and
- *equal error rate* (ERR), the crossover point at which FRR equals FAR.

In much of the literature regarding keystroke dynamics, the imposter pass rate (IPR) and FAR are used interchangeably with FAR and FRR as defined here, respectively. The existence of two terms whose meanings are opposite but both denoted by FAR can cause some confusion. We've adopted the former terminology throughout this article.

Although ERR is a desirable metric in terms of its ability to condense FAR and FRR into one value, few researchers report ERR in their published results. For that reason, we present an alternative approach to combining FAR and FRR: averaging the two values. We call this value the average false rate (AFR). Empirically, AFR closely approximates ERR for those few papers that did report ERR.

In 2000 and 2002, the UK's Biometrics Working Group produced guidelines called *Best Practices in Testing and Reporting Performance of Biometric Devices*.[19] Going forward, we hope that researchers in the field of keystroke-typing patterns will consider these guidelines when reporting results. The main corpus of results we review herein, however, didn't have the advantage of access to these standards, and therefore didn't present data that we could use to produce such useful evaluation criteria as receiving operator characteristic (ROC) or detection error trade-off (DET) curves, which plot pairs of error rates measured at different algorithm parameters. In this light, AFR can be viewed as a useful stopgap for comparing overall classifier accuracy.

Figure 1 graphs the performance of several systems in terms of AFR, FRR, and FAR. Although we make no claim about the validity of a system designed to favor either FAR/FRR over the other,[12] we feel that in the absence of reported ERR, AFR is a good descriptor of a given classifier's overall accuracy in terms of discriminating between users.

Figure 1 shows that the best reported results can achieve an AFR of less than 1 percent, and roughly one-third are capable of AFR near 2 percent—values generally considered to be acceptable for this type of system. The worst performers have average AFR values between 8 percent and 27 percent, and are not likely to provide sufficient accuracy for common usage.

Usability

Two other commonly used metrics in the realm of biometrics are the *failure to enroll rate* (FTR), which describes the percentage of users who lack enough quality in their input samples to enroll in the system, and the *failure to acquire rate* (FTA), which describes the percentage of users for whom

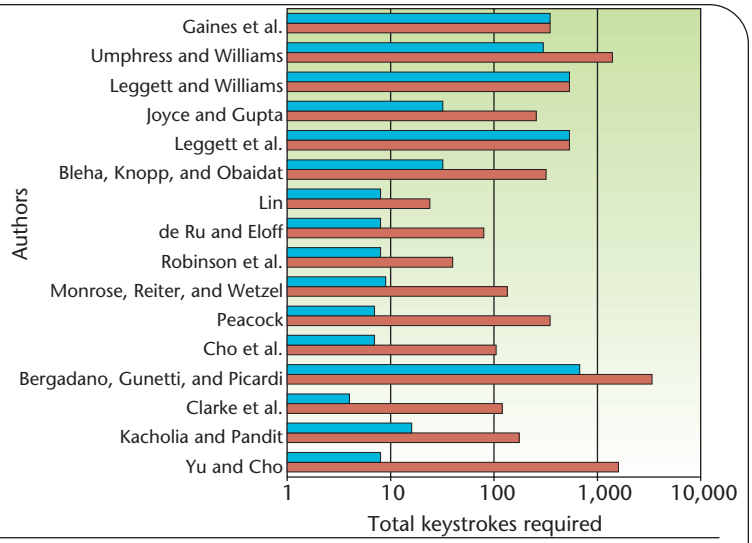


Figure 2. Comparing different keystroke approaches. The cost to a user (in keystrokes) to enroll and to authenticate for a given approach shows that systems that can enroll and authenticate with fewer keystrokes are easier to use. Blue represents the cost to a user to authenticate; red is the cost to a user to enroll.

the system lacks sufficient power to classify, once enrolled. These metrics are proposed primarily as a way to measure classifier accuracy, not as a means of measuring system usability, although they can provide some limited insight into system usability. We found that both FTR and FTA are seldom reported, which might stem from the fact that most studies don't use thresholds for rejecting users during enrollment, which in turn could stem from the relatively small groups of users studied in most reports.

FRR also partially addresses usability, because it's a measure of how often a user might have to reauthenticate after being misidentified by the system. In this section, we compare keystroke biometric systems using two new proposed metrics that further quantify usability:

- *cost to a user to enroll* (CUE) measures the number of keystrokes a user must submit to the system before enrolling as a valid user; and
- *cost to a user to authenticate* (CUA) measures the number of keystrokes a user must submit to the system each time he or she authenticates.

CUE and CUA arise from the need to measure usability in terms of how much work an individual must perform when successfully accessing a system, rewarding classifiers that perform well with less input from the user. Unlike FTR and FTA, these metrics ignore the extra work required of users due to classifier failures, instead focusing on the usability costs associated with successful enrollment and access. If FTR and FTA data were available from these studies, we could combine them with CUE and CUA data to give a more complete picture of overall system usability.

Figure 2 plots CUE and CUA for each of several approaches. (Many of the password-based approaches failed to publish average password length; the data presented in this article assumes a password length of eight characters in such cases.) The graphs show a wide range of requirements both for enrollment (from 24 required keystrokes to nearly 3,500) and for authentication. Authentication costs in Figure 2 fall into three categories: those that require on the order of 10 keystrokes (almost all the systems that monitor only password patterns), those that require tens of keystrokes, and those that require several hundred keystrokes. If we were to eliminate those systems that required more than 1,000 keystrokes to enroll or more than 100 keystrokes to authenticate, we would eliminate a few of the better performers, but would retain many systems that perform accurately and have low usability costs.

Confidence in reported results

There is wide variance in the amount of data researchers have collected to perform their studies and demonstrate their systems' effectiveness. Is a system that can determine the identities of five users with 100 percent accuracy better than a system that can determine the identities of 300 users with 99 percent accuracy? To measure the amount of confidence we can place in reported results, we can compare the various studies according to

- *sample size*, or the number of test subjects taking part in the study;
- *valid access attempts*, or the number of valid authentications attempted; and
- *imposter access attempts*, or the number of imposter authentications attempted.

The results in Figure 3 are not very encouraging: only two of the published studies used more than 50 test subjects, with the majority using less than 25. The lack of extensive test data demonstrates an important deficiency: keystroke biometrics will almost certainly be used for groups with more than 25 members, yet only two of the approaches compared in Figure 3 demonstrated competence on samples large enough to validate their results on large systems. Large sample sets are particularly important for Web-based systems, the largest of which can scale to millions of users. To be fair, harnessing a significant number of human test subjects is a difficult task. Perhaps these numbers indicate the need for a central repository of input data for keystroke biometric analysis. Such a repository would also serve as a source for common benchmarking to compare various approaches. Alternatively, researchers could independently make their own data available upon publication.

Figure 3 also shows the number of valid and imposter ac-

cesses attempted on each system. Valid attempts fall roughly into three categories: those with less than 100 attempts, those with close to 200 attempts, and one with close to 500 attempts.[12] Imposter attempts have four divisions: a few with no imposter attempts, many with 100 or 1,000 imposter attempts, and one with over 70,000 imposter attempts. Larger numbers provide more convincing proof of workable, secure systems. It's also worth noting that a small handful of the best-combined performers maintain reasonable (if not stellar) performance in relation to the current body of work.

Intellectual property and the current market landscape

The biometric security market is expected to grow from US $1 million in 2004 to $4.6 billion in 2008. So why has keystroke dynamics, a cheap biometric alternative, not become a booming part of this trend? Either the marketplace for such software is simply not demanding this level of security or the current implementations don't meet performance and usability standards.

For a technology that has made little progress outside academic circles, the field of keystroke dynamics has a substantial body of patents covering every possible interpretation and incarnation as a product. At worst, these patents negatively impact academic research to improve the biometric's viability. In 2001, one of the authors of this article was asked to cease and desist from making keystroke collection and classification tools available on the Internet by one of the holders of these patents. Undoubtedly, other researchers have encountered similar obstacles.

The first patent that actually defined keystroke dynamics as a method was John D. Garcia's 1986 patent on a personal identification apparatus.[20] It describes a general method for verifying whether someone is part of a predetermined group by using a similarity measure. This similarity measure is based on a vector of time delays between successive key press events. Garcia covers additional extensions of the core claim, including

- inhibiting access if there is no similarity;
- generating timing vectors through multiple entries;
- removing entries that are strongly statistically dissimilar from other entries;
- generating a timing vector by computing a covariance matrix; and
- generating a timing vector as a function of an authorized individual's consistency.

The Garcia patent also describes a personal identification apparatus with the same claims as the method as well as some improvements. One claim covers possible input devices such as a typewriter keyboard, numerical keypad, or piano keyboard. The Garcia patent demonstrates the forethought necessary on the inventor's part. He describes both a method and apparatus so that the invention's interpretation is broad.

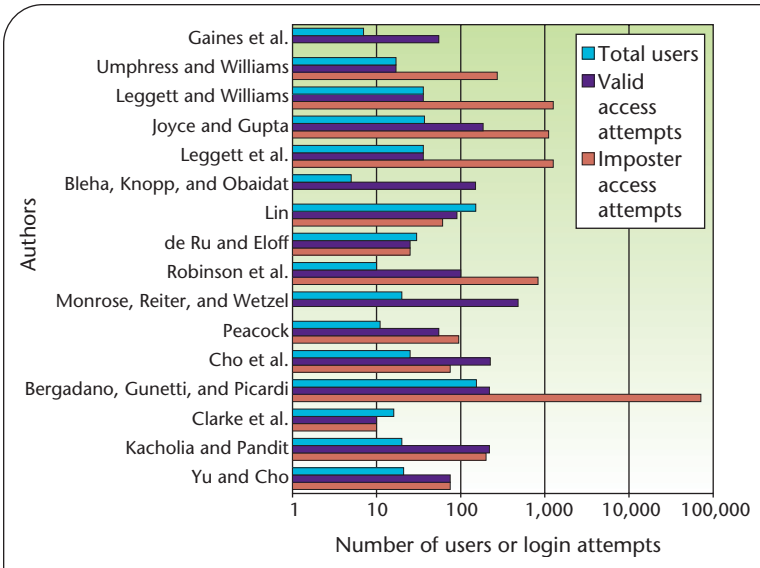


Figure 3. Confidence in test results. The involvement of more users and more valid/imposter logins lends credence to reported results, but even the largest studies in the keystroke dynamics field to date fall short of proving competence on large systems.

Future patents in keystroke dynamics will build off the Garcia and successive patents by claiming an improvement to the invention, but such improvements will depend on the claims' wording and structure. If a patent is based on a previous chain of patents, a product developer licensing the patent might need to license from the foundation patents as well.

The 1989 patent by James R. Young and Robert W. Hammon patent describes a method and apparatus for verifying an individual's identity,[21] and is the main patent currently used in BioNet Systems' BioPassword.[22] Young and Hammon deviate from Garcia by describing a method specifically in a digital computer system using a keyboard as well as a device for verifying identity based on keystroke dynamics. Instead of timing vectors, they describe a template for a user claiming to be a particular individual. In addition to the time periods between keystrokes, this template uses keystroke pressure, an expensive addition in practice. One interesting improvement is that each feature can be an updated average of features from a predefined number of keystrokes.

Method and Apparatus for Verification of a Computer User's Identification, Based on Keystroke Characteristics is the 1996 Marcus E. Brown and Samuel J. Rogers patent that also describes both a method and apparatus, with many claimed improvements over Young and Hammon.[23] The first improvement is a description of a training signal purified by discarding the portions that cause the signal to variate outside a given threshold; the signal is then com-

pared with an input signal for similarity. Brown and Rogers also describe several additional improvements based on specific classifiers such as neural networks and Euclidean distance measures.

The most recent patent to be accepted in the area of keystroke dynamics security is the 2002 Zilberman patent, which claims a *Security Method and Apparatus Employing Authentication by Keystroke Dynamics*.[24] Zilberman's improvements include a keyboard device that dispenses physical tokens as a unique key to access the system, a data matrix for storing timing characteristics, and an embedded microcontroller for authentication.

Whether seen as a device, apparatus, process, or method, keystroke dynamics has clearly been well thought out by inventors positioning themselves for the crossing of this technology into the steadily growing information security market. The success of any commercial offering will either herald the acceptance of keystroke dynamics into the marketplace or signal the need for more research and better implementations.

Privacy and security issues

Should keystroke dynamics gain acceptance in the marketplace, issues of privacy and security must be carefully evaluated. Of the most concern are databases that maintain users' keystroke-timing patterns. With this information, attackers can subvert authentication systems that rely on keystroke biometrics.

Attackers might also be able to guess a sequence of typed characters from its corresponding timing pattern. To illustrate, one experiment deciphered encrypted passwords sent through version 2 of the SSH protocol an average of 50 times faster than brute-force methods by using weaknesses in the protocol and a database of user keystroke profiles.[25] Even when user-specific keystroke profiles are not available, generic keystroke profiles created from any representative population subset have been found to weaken security.[25]

Systems that monitor typing patterns must also guard against privacy breaches. If the monitoring process produces records, these records must be protected by a policy regarding their use and a mechanism to prevent unauthorized access both to the records and to live monitoring. But such safeguards do not protect against covert monitoring and tracking of individuals.

Although academics and inventors have pursued keystroke biometric systems for more than a quarter of a century, the field is still maturing. Lack of a shared set of standards for data collection, benchmarking, and measurement have prevented, to some degree, any growth from collaboration and independent confirmation of techniques. Moreover, patents encumber many of the most basic strategies. Finally, until privacy concerns regarding the building of keystroke biometric databases are resolved, wide adoption of this technology could meet opposition from civil libertarians and privacy advocates.

Still, keystroke biometrics hold great promise for creating systems that are both more secure and more usable than their predecessors. Because keystroke biometrics can be collected without the need for special hardware, and because software to perform identification and authentication has shown great potential, keystroke biometrics could be poised to become a standard method of proving identity, online and off. Keystroke biometrics has an advantage over most other biometric authentication schemes: user acceptance.[16] Because users are already accustomed to authenticating themselves through usernames and passwords, most proposed keystroke biometric methods are completely transparent.

Continuing research and commercial activities in the field and the keyboard's popularity as the primary input device for applications ensure that the technology will not fade into history. As the keystroke biometrics field matures, observers should watch for several trends that will indicate when the technology is ready for more widespread adoption:

- greater depth in performance measurement, with the average study involving at least thousands of users;
- the creation of data sets that can be shared between studies, enabling researchers to focus on perfecting their classification methods instead of burdening them with the task of building usable sample data;
- introduction of schemes that ensure the privacy of collected biometric data; and
- expiration or relaxation of existing intellectual property claims, with the resultant competition that this will foster.

Whether keystroke dynamics ultimately becomes a ubiquitous part of the security landscape will be determined not only by how much we trust these systems to uniquely identify individuals and provide a comfortable authentication process, but by how much we trust systems that collect the immutable piece of ourselves known as a biometric. □

References

1. M. Kotadia, "Gates Predicts Death of the Password," *CNET News.com*, Feb. 2004; http://msn-cnet.com/2100-1029_3-5164733.html.
2. A. Adams and M.A. Sasse, "Users Are Not the Enemy," *Comm. ACM*, vol. 42, Dec. 1999, pp. 40–46.
3. R. Spillane, "Keyboard Apparatus for Personal Identification," *IBM Technical Disclosure Bulletin*, vol. 17, no. 3346, 1975.
4. J. Leggett et al., "Dynamic Identity Verification via Keystroke Characteristics," *Int'l J. Man-Machine Studies*, vol. 35, no. 6, 1991, pp. 859–870.
5. R. Richmond, "Fed up with Fraud," *The Wall Street J.*

Classroom Ed., Apr. 2003; www.wsjclassroomedition.com/archive/03apr/BIGB_retailer.htm/.
6. S. Cho et al., "Web-Based Keystroke Dynamics Identity Verification Using Neural Network," *J. Organizational Computing and Electronic Commerce*, vol. 10, Dec. 2000, pp. 295–307.
7. M. Tapiador and J.A. Sigüenza, *Fuzzy Keystroke Biometrics on Web Security*, tech. report, Escuela Técnica Superior de Informática, Univ. Autonoma de Madrid, Cantoblanco, Mar. 2000.
8. A. Peacock, "Learning User Keystroke Latency Patterns," Apr. 2000; http://pel.cs.byu.edu/~alen/personal/CourseWork/cs572/KeystrokePaper/.
9. X. Ke et al., "Keystroke Dynamics: A Software-Based Biometric," 2004; http://web.mit.edu/xke/Public/kd/.
10. D. Song, P. Venable, and A. Perrig, "User Recognition by Keystroke Latency Pattern Analysis," Apr. 1997; http://citeseer.nj.nec.com/song97user.html.
11. F. Monrose and A.D. Rubin, "Keystroke Dynamics as a Biometric for Authentication," *Future Generations Computing Systems*, vol. 16, no. 4, 2000, pp. 351–359.
12. F. Monrose, M.K. Reiter, and S. Wetzel, "Password Hardening Based on Keystroke Dynamics," *Proc. 6th ACM Conf. Computer and Comm. Security*, ACM Press, 1999, pp. 73–82.
13. T. Ord and S.M. Furnell, "User Authentication for Keypad-Based Devices Using Keystroke Analysis," *Proc. 2nd Int'l Network Conf.* (INC 2000), Inst. of Electrical Eng., 2000, pp. 263–272.
14. N.L. Clarke et al., "Advanced Subscriber Authentication Approaches for Third Generation Mobile Systems," *3rd Int'l Conf. 3G Mobile Comm. Technologies*, vol. 489, May 2002, pp. 319–323.
15. G. Forsen, M. Nelson, and R. Staron, *Personal Attributes Authentication Techniques*, tech. report RADC-TR-77-1033, Griffis Air Force Base, 1977.
16. R. Gaines et al., *Authentication by Keystroke Timing: Some Preliminary Results*, tech. report R-256-NSF, RAND, 1980.
17. R. Joyce and G. Gupta, "Identity Authentication Based on Keystroke Latencies," *Comm. ACM*, vol. 33, no. 2, 1990, pp. 168–176.
18. *Biopassword Keystroke Dynamics*, tech. rep., Net Nanny Software Int'l, 2001.
19. A.J. Mansfield and J.L. Wayman, *Best Practices in Testing and Reporting Performance of Biometric Devices*, tech. report version 2.0, UK Govt. Biometrics Working Group, Aug. 2002.
20. J. Garcia, *Personal Identification Apparatus*, US patent number 4,621,334, Washington, DC, 1986.
21. J. Young and R. Hammon, *Method and Apparatus for Verifying an Individual's Identity*, US patent number 4,805,222, Washington, DC, 1989.
22. *Biopassword: Questions and Answers*, tech. report, BioNet Systems, 2002; www.biopassword.com/home/FAQs/BP_General_FAQs_112502.pdf.
23. M. Brown and S. Rogers, *Method and Apparatus for Verification of a Computer User's Identification, Based on Keystroke Characteristics*, US patent number 5,557,686, Washington, DC, 1996.
24. A.G. Zilberman, *Security Method and Apparatus Employing Authentication by Keystroke Dynamics*, US patent number 6,442,692, Washington, DC, 2002.
25. D.X. Song, D. Wagner, and X. Tian, "Timing Analysis of Keystrokes and Timing Attacks on SSH," *10th Usenix Security Symp.*, Usenix Assoc., 2001; www.usenix.org/events/sec01/song.html/.

Alen Peacock *works in the Massachusetts Institute of Technology's Lincoln Laboratory, where he researches topics in airborne mobile networks composed of heterogeneous data links. His research interests include security in highly mobile networking environments and the application of learning and cognitive processes to ad hoc network topology formation, management, and control. He has an MS in computer science from Brigham Young University. Contact him at peacock@ll.mit.edu.*

Xian Ke *is a program manager with Microsoft's Security Business Unit. Her research interests include practical computer security applications and information systems. She has an MEng in computer science from the Massachusetts Institute of Technology. Contact her at xke@mit.edu.*

Matt Wilkerson *is completing a degree in management science at the Massachusetts Institute of Technology, after which he will attend Columbia University's masters program in computer science. His research interests include information security systems and streaming video applications. Contact him at mwilker@mit.edu.*