Fred Cohen & Associates - Analyst Report and Newsletter

Welcome to our Analyst Report and Newsletter

Can we attribute authorship or human characteristics by automated inspection?

There have been many announcements, papers, discussions, and claims regarding attribution solutions, but few literature reviews and little useful technology. Cutting to the chase, content inspection for attribution does not work well, if at all, today. Under naïve deception, these methods perform no better than random guessing. Without deception, for 5000 words of prose per individual and collections of 10 to 20 individuals, current methods correctly identify authorship, at best 80% of the time, and at worst, never. Current approaches are summarized here considering potential use in legal proceedings, terminations, authentication, and attack attribution.[1]

Using authentication for attribution vs. attribution for authentication

Authentication is intended for confirming identification, not for attributing actions to actors. But it is often used as a basis for attribution because it provides some level of demonstration that an individual is using a system at a time. Key questions are; what that level of certainty is, how it is determined, and it is adequate to the requirements at hand.

Authentication methods typically have very low probability of error in statistical terms. To the extent that they involve authentication devices, like hardware tokens, or memorization, like passwords, they have an extremely low probability of false authentications. For example, even a 4 character password has a 1 in 26^4 (just over 2 in a million, or .000002%) - chance of being guessed randomly. For false negatives, authentications with these methods essentially never fail except through typographic errors by users, which happen more often for longer and harder to guess mechanisms.

Attribution methods, including biometrics, have also explored for use in authentication. That is, if we can attribute actions to actors by some method, it would seem reasonable to use that attribution as a basis for authentication of identity. We have to understand the level of certainty associated with attribution methods to evaluate their use for authentication.

Attribution methods sometimes involve collecting data with precision and accuracy at the level of milliseconds using previously implemented and calibrated devices are in place and properly functioning. This is typically unavailable retrospectively. Unless the attribution method was undertaken as part of normal business practices, it is unlikely that the necessary information will be available. But some traces may be found in detailed network logs, from incidental packet sniffing, or from company keystroke loggers used to track user behavior. Some less reliable methods of attribution may also be available.

Types of authentication methods

Authentication has historically been categorized as consisting of something the user has (an authentication device), something the user is (a biometric property), or something the user knows (a password, pass phrase, or query response). Biometrics will be explored separately.

[1] Portions of this article are extracted from F. Cohen, “Digital Forensic Evidence Analysis”, ASP Press, 2009-11

Something the user has

At best, on their own, something a user has can only be used to attribute actions to the presence of the possessed item somewhere, and not to the individual. It has been suggested that the only way to show people are present is with continuous video, and their actual use requires a video of user motions and display outputs. Ignoring disguises, this depends on the standard of proof in the matter at hand and other events and traces available. For something the user has, the examiner should point out that use only indicates that the device was forged or otherwise used, and not that the individual, or even the device, was present. The link to individual presence must be made by some other means.

Something the user knows or can do

Passwords, and variations on them, have been in use since biblical times. While modern variations may include behavioral indicators and query response systems, these are really no different than they have been for millennia. Many flaws and limitations with these systems are widely known and have been widely published over many years. Variations on approaches that use human ability to do things have also been introduced.[2] Some major problems with these sorts of authenticators include;

(1) authentication sequences are often not stored and unavailable to the examiner;

(2) records of a successful authentication may be easily forged;

(3) authenticators are often easily guessed or replayed and regularly used by malicious third parties for unauthorized access;

(4) many automated mechanisms use these sorts of authenticators and can readily perform tasks in place of the individual;

(5) users commonly share identities and authenticators or use group accounts;

(6) passwords are easily observed over the shoulder or in logging devices; and

(7) authenticated individuals may not perform all subsequent acts or remain present.

Attribution methods

Biometrics are generally understood under the authentication approach of what a user is. As such, they fit nicely into the range of attribution methods rather than authentication methods. While they are used for authentication, they provide authentication precisely because of their qualities associated with attribution and are useful only if and to the extent that they provide few false positives and negatives.

In the larger picture of attribution, human behavior forms an increasingly explored area. Observables like keystroke sequences and patterns, word usage patterns, typing errors and quirks, spelling errors, the commands used in the order applied, how editing is done, and other similar things have been increasingly explored as attribution indicators.

These behavioral methods are used, largely experimentally today, for authentication of a known user with previously collected characteristics, under controlled circumstances, and are highly problematic.

[2] F. Cohen, "Algorithmic Authentication of Identification"', Information Age, V7#1 (Jan. 1985), pp 35-41.

Biometrics

Biometric authentication methods have shown false acceptance rates as follows:[3]

Retina Recognition	1 to 10,000,000
Iris Recognition	1 to 131,000
Fingerprint Recognition	1 to 500
Hand Geometry Recognition	1 to 500
Signature Recognition	1 to 50
Voice Recognition	1 to 50
Face Recognition	No Data Available
Vascular Patterns Recognition	No Data Available

Biometrics cannot be readily or inexpensively changed or revoked, making them suitable for attribution, but introducing problems with the approaches often taken to authentication, where replay or reuse of authenticators is a problem. If and when biometric data becomes available to those wishing to exploit systems, they have long opportunity to create forgeries that may not be detectable by any given method in use. Thus malicious exploitation of biometric authentication and attribution is potentially problematic.

Fist and related approaches

For a long time, it has been conjectured that analysis of text, keystroke timings, writing patterns, and other similar phenomena might be used to attribute actions to actors. Starting at least in 1980,[4] extensions of the work in World War 2 for identification of the "fist" of key code operators was considered for identifying which of a group of known individuals were typing on a computer keyboard. The early work in this arena seemed promising, but it took quite a while to make limited progress, and that progress revealed many limitations of such processes.

Authentication based on behaviors, such as keystroke patterns, are easily forged with methods similar to those used to collect the keystrokes in the first place.[5,6] But even if they weren't, tests that indicated 93% correct recognition on average (i.e., only about 1 in 14 legitimate uses would be declared illegitimate), were only 51% correct for some users (half the time they would be declared illegitimate), and these results were only under very limited conditions (i.e., for distinguishing individuals from groups of about 50 people, with entry of 200 or more characters, use of the same keyboard by each user each time, using entry of free text of their own composition, with accurate timing information to the millisecond, including key press duration, transition times, and special key uses, and collected in structured tasks using identical software and operating environments installed on every computer).[7]

[3] Aykut Guven and Ibrahim Sogukpinar, "Understanding users' keystroke patterns for computer access security", Computers & Security, Volume 22, Issue 8, December 2003, Pages 695-706.

[4] Gaines, R.S., et al. 1980. Authentication by keystroke timing: Some preliminary results. Rand. Report R-256-NSF. Rand Corporation. Available at http://www.rand.org/pubs/reports/R2526/.

[5] F. Cohen, et. al. "Leading Attackers Through Attack Graphs with Deceptions", IFIP-TC11, `Computers and Security', V22#5, July 2003, pp. 402-411(10).

[6] F. Cohen, I. Marin, J. Sappington, C. Stewart, and E. Thomas, "Red Teaming Experiments with Deception Technologies", 2001.

[7] Mary Villani, Charles Tappert, Giang Ngo, Justin Simone, Huguens St. Fort, Sung-Hyuk Cha, "Keystroke Biometric Recognition Studies on Long-Text Input under Ideal and Application-Oriented Conditions", School

This seems unlikely to satisfy the requirements of forensic examination,[8] and would be extremely problematic for larger groups under less controlled circumstances. Suitability for identification is dubious, but for a small population of suspects with known characteristics, if the traces are available with appropriate calibration, results may be leveraged after the fact to confirm that traces are consistent with known subjects, or more consistent with one subject than other subjects. But the quality of such conclusions should be scrutinized and validated, and the potential for deception, if present, is highly problematic.

For authentication, these methods are also problematic, but perhaps less so. For keystrokes, with millisecond timing resolution, timing only of password entry, and using long training times, failure rates of 1 in 20 have been achieved (95% recognition). For an operation where 100 users logged in twice a day, on average, there would be 10 false detections of illicit users per day and some users would produce false detections once a day. It is key to understand that these methods are used for authentication of identified individuals from known populations of limited size under controlled conditions where the mechanism is previously calibrated for the purpose.

Stylometrics, phrasing, and similar document analysis

The basic premise of this approach is that people are formed over time by how they acquire knowledge, skills, and techniques from the people, places, societies, and media they interact with. The fact that they can only use the methods they are aware of, the notion that methods are commonly originated and disseminated through a path, and other similar notions, provide the means to reverse the paths, identify properties, type characteristics, and perhaps even individualize a source. This approach falls under and is compatible with the notion of identifying attacker type[9] and individualizing based on specific characteristics. A reference base is built and pattern recognition is done on traces against baselines. Stylometric,[10] methods used for disambiguation of word sense and plagiarism detection have been admitted in US courts in select cases, but they are readily susceptible to deception.

Other patterns

Movement patterns from mobile devices, user gait patterns, Web click patterns, and writing patterns from pen computing have also been tried,[11] and cognitive methods suggested.[12] But we have yet to see successful attribution with these methods in the peer reviewed literature.

of CSIS, Pace University, Pleasantville, New York, 10570, 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06).

[8] B. Rao, "Continuous Keystroke Biometric System", Media Arts and Technology, September 2005, A University of California, Santa Barbara for partial requirements of Masters of Science in Media Arts and Technology.

[9] F. Cohen, C. Phillips, L. Swiler, T. Gaylor, P. Leary, F. Rupley, R. Isler, and E. Dart, "A Preliminary Classification Scheme for Information System Threats, Attacks, and Defenses; A Cause and Effect Model; and Some Analysis Based on That Model", Sandia National Laboratories, September, 1998, also in Computers and Security and The Encyclopedia of Computer Science and Technology, and at: http://all.net/journal/ntb/cause-and-effect.html

[10] C. Chaski, "Who's At The Keyboard? Authorship Attribution in Digital Evidence Investigations", International Journal of Digital Evidence, V4#1, 2005.

[11] Padmanabhan, B. and Y. Yang. Unpublished manuscript. Clickprints on the Web: Are there signatures in Web browsing data? Available at http://knowledge.wharton.upenn.edu/papers/1323.pdf?CFID=720523&CFTOKEN=57530247.

[12] F. Cohen, et.a,. "A Framework for Deception", http://all.net/journal/deception/Framework/Framework.html

Limitations of human attribution under deception

Under intentional forgery, issues get more complex. Simple forgeries, like deceptive email headers, are common. Moderately skilled attacks may bypass authentication, forge logs, use false user identities, use other user accounts, etc.[13] Sophisticated attackers with financing, access to research and development, and other similar assets, do more complex and realistic forgeries covering redundant indicators. For example, over a network, keystroke activity with timing data and related behaviors have been emulated in deception experiments.[14]

Recent results suggest that naïve deception readily and effectively defeats a wide range of stylometrics approaches.[15] Simple deceptions based on statistical word usage were applied against histogram distance, Manhattan distance, cosine distance, KS distance, cross-entropy, Kullback-Leibler distance, LDA, Gaussian SVM, and Naïve Bayes methods and tested against samples generated from words, 2-3 letter words, 3-4 letter words, word bi-grams, word tri-grams, word stems, parts of speech, word lengths, syllables per word, characters, character bi-grams, character tri-grams, binned frequencies, binned reaction times, and Mosteller-Wallace function words. The sample set had 15 authors with 5,000 words from each, 500 involving imitation and 500 involving obfuscation. Under obfuscation, where individuals sought to change their writing styles, the best performance correctly classified 42% of samples, and the worst was never correct. For imitation, where samples of other writing styles were provided and imitation attempted, the best performance was 23% correct, and the worst was never correct. No method performed significantly better than chance.

Forgery of biometrics has also been widely demonstrated. Eye recognition forgery for various mechanisms is somewhat problematic, but fingerprint mechanisms have been readily overcome with gelatin imprinted with ridges and placed over human digits, hand geometry can be readily forged with a mold, signature forgery with timing and stroke information can be automated using computer output devices, voice recognition is susceptible to various recording and playback mechanisms, and face recognition is problematic when people smile and can be overcome with simple masks.

Most current mechanisms used for storing, analyzing, and making decisions based on these methods are also susceptible to attack and usually readily defeated.

Summary

The problem with most such approaches is that there is little definitive information showing that they are reliable. Significant study is needed to find relevant metrics.[16] Whenever you hear about a breakthrough in attribution, especially based on appearance, behavior, or other similar sorts of information, view it with great skepticism and examine it closely before buying into it. The science to date is almost never able to support these methods as meaningful outside of very limited circumstances, and the hyperbole almost always outstrips the science.

[13] F. Cohen, et. al. "A Preliminary Classification Scheme... ibid.

[14] F. Cohen, I. Marin, J. Sappington, C. Stewart, and E. Thomas, "Red Teaming Experiments with Deception Technologies", 2001, available at: http://all.net/journal/deception/RedTeamingExperiments.pdf

[15] P. Juola and D. Vescovi, “Stylometric Approaches to Author Obfuscation: An Empirical Study”, IFIP TC11.9 Digital Forensics Conference, Orlando, FL 2011-01-31-2011-02-02

[16] C. Chaski, "Who’s At The Keyboard? Authorship Attribution in Digital Evidence Investigations", International Journal of Digital Evidence, V4#1, 2005.