# Keylogger Experiments

## Overview

This project concerns keylogger programs and explores the possibility that it might be feasible to make ID/Password pairs more secure by factoring the owner's typing rhythm into the authentication process. To test this idea, I implemented a keylogger using JavaScript and Node.js to gather relevant data over the course of two different experiments. The results of these experiments were quite interesting and serve as the basis for some constructive conclusions.

## Problem Statement

In this report, I seek to contribute to the conversation concerning the biometrics community's hypothesis (provided in the specification) that claims, "a password coupled with the normal rhythmic typing of the owner is stronger than the password on its own," by performing relevant experiments using the keylogger I've written.

## Background Information

Much of the relevant historical information regarding keylogging is covered in Monaco's 2018 paper [1]. He began by detailing the discovery of the first "keylogger" attack, where

researchers at Bell Laboratories discovered they could achieve information gain through monitoring electromagnetic spikes emitted by one of their machine's keyboards. These spikes were determined to be useful to attackers who might try to eavesdrop on the system using a "side-channel attack," or an attack that capitalizes on information the system emits to its environment. Side-channel attacks have greatly evolved since this event, now targeting spatial data pertaining to how far apart keys on a keyboard may be, or temporal data concerning the duration of and time in-between key presses.

Temporal side-channel attacks are speculatively what would be most effective against a keylogger like the one used for this report. Should a company decide to implement such a keylogger in their infrastructure, a potential attacker could listen to how individuals within the ecosystem enter their passwords, or even feel their typing frequencies through moving objects in the work environment, like a desk or table to gain information.

## Relevant System Details

To spare redundancy, I will provide some details about the metrics and measures used by the keylogger. As mentioned in the overview, the system is implemented in JavaScript with a Node.js server that serves HTML pages. Pure JavaScript runs in the background on the frontend to measure the intervals between key presses (milliseconds) and the total time taken for the input entry (milliseconds), which is then sent to the server in JSON format upon submission.

The server compiles this JSON data in a file called <ID>_profile_creation.json during the calibration phase, where the system prompts the user to calibrate their ID/Password combination by entering it seven times. Upon completion, another JSON file is generated,

called <username>_profile.json. This file contains the user's ID, password, the average length of the five shortest intervals (milliseconds) for each keystroke, and average total entry time of the five quickest entries for both the ID and password.

Finally, a successful login attempt must meet **at least one** of the following criteria:

1. $|avgCalibrationTotalTime - loginTotalTime| \leq 50$ (milliseconds)
2. $\text{avgTotalEntryTime} * 0.8 < \text{loginTotalEntryTime} < \text{avgTotalEntryTime} * 1.2$
3. $keystrokeMatches \geq 0.6$ (60%)
    a. This means that the timings for 60% or more of the keystroke intervals must make **one** of the two following statement evaluate to true:

$i.\ avgKeystrokeTime * 0.6 < loginKeystrokeTime < avgKeystrokeTime * 1.4$

$ii.\ |avgKeystrokeTime - loginKeystrokeTime| \leq 50(milliseconds)$

4. $keystrokeMatches = length(ID/Password)$

## Experiment 1

*Methodology:*

The aim of this experiment was to find a length, *L*, where the lengths of the ID and password were greater than or equal to *L*, for which the keylogger layer alone could protect all resources associated with an ID/Password pair and still allow its owner access. The experiment itself was performed in the Jack Cole lab room on a single lab machine and had six total participants: five computer science students and the author of this report. Before starting, each student participant was given a brief explanation of which actions should be avoided, like pressing tab and enter, how some actions are different than normal, like backspace, that they

are all playing the role of an attacker against the rhythmic keylogger system, so they are encouraged to try different key press timings, and six ID/Password pairs for test cases that were pre-loaded into the system by the author. Each participant entered the six pairs into the system five times and the authorization status of each attempt was recorded, for a total of 180 data points. Additionally, this experiment was conducted two hours after the author of this report calibrated each of the six test cases.

**Figure 1: Test Cases Used in the Experiment**

| Test Case | ID | Password | Notable Details |
|---|---|---|---|
| 1 | AAAAAA | 123456 | Test Case 1 from the specification |
| 2 | Hello | World | Test to see if typing styles diverge on frequently typed phrases |
| 3 | Az | By | Test for L=2 |
| 4 | AbC | 123 | Test for L=3 |
| 5 | Ball7 | sport | Test for L=5 |
| 6 | CSLabs0 | Program | Test for L=7 |

*Experiment*:

1. Hypothesis:

There is a minimum length for strings, *L*, where an individual can be uniquely identified by their typing style.

2. Constants:
    a. Machine the experiment was performed on
    b. Keyboard used during the experiment
    c. Environment in which the experiment was conducted, including:
        a. Chairs
        b. Noise level
        c. Familiar atmosphere to all participants
    d. Test cases used (Figure 1)
    e. Native English-speaking participants
    f. Time of day (late evening)

3. Variables
    a. Information privy to participants
        a. Each of the five computer science students had identical information
        b. The author of this report had far more knowledge because of the fact he/she wrote the keylogger in use and calibrated the test cases
    b. Typing styles of participants
    c. Skill in adapting to the trial-and-error attack method
    d. The degree of which each participant likes the lab keyboards

4. What is Being Measured:

   The total number of true positives, true negatives, false positives, and false negatives over the 180 total login attempts.

5. Assumption:

   Each participant understands that the keylogger tracks the time between key presses to act as a protection mechanism, given the explanation they received before beginning the login attempts.

*Results*:

The results from the experiment are displayed in the table below (Figure 2). Each cell should be interpreted as the number of successful logins out of five, given the corresponding participant and test case.

The hypothesis of this experiment was disproven because of the false positive seen in cell (Student 1, Test Case 6). This result implies that an ID/Password length *L* does not exist in the scope of this experiment for which the owner of an ID/Password pair can be uniquely identified by their typing style. However, the data from this alpha trial showcase that keyloggers have promise as a potential security factor. The data from all test cases where $L \geq 3$ (all but test case 3) for students one through five produces a false positive rate of 2/125, or $\approx$ 1.6%. One of these two false positives was truly an unexpected outlier (Student 1, Test Case 6), but the other (Student 1, Test Case 2) was an anticipated symptom of a “weak” rhythmic

password in test case 2 (pair Hello/World). Picking phrases that attackers might frequently type increases the probability that the natural typing styles of the owner and attacker may match, so a "secure" rhythmic password is one where the owner can minimize the attacker's familiarity with the passphrase.

**Figure 2: Number of Successful Logins Per Participant Per Test Case**

| | Student 1 | Student 2 | Student 3 | Student 4 | Student 5 | Author |
|---|---|---|---|---|---|---|
| Test Case 1 | 0/5 | 0/5 | 0/5 | 0/5 | 0/5 | 5/5 |
| Test Case 2 | 1/5 | 0/5 | 0/5 | 0/5 | 0/5 | 3/5 |
| Test Case 3 | 4/5 | 0/5 | 0/5 | 0/5 | 1/5 | 4/5 |
| Test Case 4 | 0/5 | 0/5 | 0/5 | 0/5 | 0/5 | 4/5 |
| Test Case 5 | 0/5 | 0/5 | 0/5 | 0/5 | 0/5 | 3/5 |
| Test Case 6 | 1/5 | 0/5 | 0/5 | 0/5 | 0/5 | 4/5 |

Total False Positives: $7/150 \approx 4.67\%$

Total False Negatives: $7/30 \approx 23.3\%$

Additionally, the false negative rate of roughly 23.3% is not concerning. It's a common occurrence for people to forget their usernames or passwords to digital services, but this does not justify abolishing the use of traditional login mechanisms. As it would with a forgotten username or password, forgetting a keylogger rhythm would require the owner to recalibrate their ID/Password pair. This could become be quite annoying should recalibration be required on a

frequent basis, but the author always managed to access the required resource on at least three out of five attempts while maintaining an acceptably low false negative rate.

*Conclusion:*

Although the hypothesis for experiment one does not hold, the preliminary data collected through the experiment demonstrates that a rhythmic keylogger applied to an ID/Password combination can provide a significant amount of extra security compared to a system without such a keylogger.

## Experiment 2

*Methodology:*

The aim of this experiment was to determine whether the shortness of the ID played a role in authenticating a particular ID/Password combination. The experiment itself was performed in the Jack Cole lab room on a single lab machine and had three total participants, all of whom are computer science students. Before starting, each student was given identical information compared to the students who participated in experiment one. Each participant entered eight ID/password pairs into the system five times and the authorization status of each attempt was recorded, for a total of 120 data points. The key difference between this experiment and the previous is that this experiment will only validate against the ID's rhythm, while the previous validated against both the ID and password's rhythm.

**Figure 3: Test Cases Used for Experiment Two**

| Test Case | ID | Password |
|---|---|---|
| 1 | AB | cd |
| 2 | Jc | Jh |
| 3 | B7 | D5 |
| 4 | AbC | 123 |
| 5 | Oof | Pog |
| 6 | Hello | World |
| 7 | QWERTY | c0mput3r_sc13nc3 |
| 8 | ninjawarrior123 | password1 |

*Experiment:*

1. Hypothesis:

   The shortness of the ID plays a significant role in keylogger-based authentication.

2. Constants:
   a. Machine the experiment was performed on
   b. Keyboard used during the experiment
   c. Environment in which the experiment was conducted, including:
      a. Chairs
      b. Noise level
      c. Familiar atmosphere to all participants
   d. Test cases used (Figure 3)

e. Native English-speaking participants

f. Time of day (mid-afternoon)

g. Each of the three computer science students had identical information going into the experiment

3. Variables

   a. Typing styles of participants

   b. Skill in adapting to the trial-and-error attack method

   c. The degree of which each participant likes the lab keyboards

4. What is Being Measured:

   The total number of true negatives and false positives over the 120 total login attempts.

5. Assumption:

   Each participant understands that the keylogger tracks the time between key presses to act as a protection mechanism, given the explanation they received before beginning the login attempts.

*Results:*

As the results from figure 4 illustrate, the hypothesis of this experiment has been brutally decimated. The best performing test case in this experiment was test case 5, with 0/15

false positives and an ID of length 3. The worst performing test case was case 7, with 11/15 false positives and an ID length of 6.

**Figure 4: Number of Successful Logins Per Participant Per Test Case**

| | Student 1 | Student 2 | Student 3 |
|---|---|---|---|
| Test Case 1 | 2/5 | 0/5 | 0/5 |
| Test Case 2 | 0/5 | 0/5 | 1/5 |
| Test Case 3 | 0/5 | 0/5 | 3/5 |
| Test Case 4 | 4/5 | 1/5 | 0/5 |
| Test Case 5 | 0/5 | 0/5 | 0/5 |
| Test Case 6 | 1/5 | 1/5 | 0/5 |
| Test Case 7 | 4/5 | 4/5 | 3/5 |
| Test Case 8 | 0/5 | 1/5 | 0/5 |

False Positives: $25/120 \approx 20.8\%$

As such, these results suggest that there is no correlation between the false positive rate and the ID length. The ID of test case 7 was registered with a simple, steady rhythm and was easily discovered despite its length. This discovery suggests that complexity of the ID plays a larger role in its rhythmic security than the length does, where complexity is increased by incorporating numbers, symbols and spatially distant characters on the keyboard. I would encourage future research to explore this possibility.

*Conclusion:*

Even though this experiment proved the hypothesis to be false, the results in conjunction with the logic behind the experiment lead to an alternate hypothesis: the complexity of an ID plays a significant role in keylogger-based authentication. Additionally, it's worth pointing out that the second experiment's omission of the keylogger for password authentication likely contributed to the unacceptably high false positive rate.

## Citations for Part 1


[1] Monaco, Vinnie. (2018). SoK: Keylogging Side Channels. 211-228. 10.1109/SP.2018.00026. https://studres.cs.st-andrews.ac.uk/CS4203/Assessment/LoggerPapers/Monaco.pdf