Testing the temporal accuracy of keystroke logging using the sound card

Frid, Johan; Wengelin, Åsa; Johansson, Victoria; Johansson, Roger; Johansson, Mikael

Published in:
[Publication information missing]

Published: 2012-01-01

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Testing the temporal accuracy of keystroke logging using the sound card

BACKGROUND

• Writing research has seen an increased use of keystroke logging
• Keystroke logging programs log the writing process in a continuous and non-obtrusive way
• They enable researchers to collect fine-grained data because they log every keystroke in relation to a timestamp (in milliseconds), which indicates the time that a specific key was used.
• For the researcher interested in for example word-internal processing it’s important to know the degree of precision and accuracy that can be achieved by the program.

METHOD

• We propose a method of measuring the accuracy of keystroke timestamps using a recording of the sounds made by key presses.
• Sound cards fit the purpose well since they typically have much better temporal resolution than computer keyboards and they are readily available in most computers
• Key presses produce noise patterns that are easily temporally located in an acoustic waveform.
• The timestamps of the noise patterns can then be compared with the corresponding timestamps reported by the keystroke logging program.
• Specifically, the differences between the two timestamps of each keystroke, provides an estimate of the accuracy of the program.

RESULTS

- We tested the accuracy of different keystroke loggers, including the latest version of the keystroke logging program ScriptLog as well as two prototypes of a new ScriptLog version implemented in C++ and Java respectively.
- Due to the increased use of web-based written communication another keystroke logger was implemented in Javascript, and ran in a recent version of Firefox.
- Each test case consisted of 50 key presses of the ‘space’ bar, and was run on identical hardware and operating system.

### TECHNICAL SETUP

- Windows XP, Dell Latitude E6500 (2008), Dell USB keyboard (2010)
- Noise location points located by visual inspection of spectrogram calculated with a 0.001 length window. Visibility was limited to 0.1s and 16000 Hz on a 1470ps wide view as pixel density was 14 pixels.

### IDENTIFICATION OF A KEY PRESS NOISE BURST

### RESULTS

<table>
<thead>
<tr>
<th>point-by-point</th>
<th>interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ScriptLog</td>
<td>0.005</td>
</tr>
<tr>
<td>JavaScript prototype (Firefox)</td>
<td>0.003</td>
</tr>
<tr>
<td>Java prototype</td>
<td>0.003</td>
</tr>
<tr>
<td>C++ prototype</td>
<td>0.003</td>
</tr>
<tr>
<td>SoundCard</td>
<td>0.29E-05</td>
</tr>
</tbody>
</table>

- **point-by-point** compares the individual timestamps of one keystroke.
- **interval** compares the length of the interval between two consecutive keystrokes.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.
- **sd** is the standard deviation of the distribution of all differences between a program measured and sound measured timestamp. **range** and **maxdiff** are other properties of this distribution.

- We find significant differences between the variances of the prototypes and ScriptLog (example: for Java: F=0.287, p<0.001)
- This implies that a reimplemented version will provide improved timing accuracy
- This method can be implemented as part of any keystroke logging program in order for the user to test the accuracy in his/her own computer environment.

Johan Fridl\(^1\), Åsa Wengelin\(^2\), Victoria Johansson\(^1\), Roger Johansson\(^1\), & Mikael Johansson\(^1\)

\(^1\)Lund University, Sweden
\(^2\)University of Gothenburg, Sweden