

Gathering Text Entry Metrics on Android Devices

Steven J. Castellucci, I. Scott MacKenzie

York University, Department of Computer Science and Engineering
4700 Keele St., Toronto, Ontario, Canada M3J 1P3
steven@se.yorku.ca; mack@cse.yorku.ca

Abstract -We developed TEMA, an application to gather Text Entry speed and accuracy Metrics on Android devices. This paper details the features of the application, describes a user study to demonstrate its utility, and establishes entry speed and accuracy measurements for the evaluated text entry techniques. We evaluated and compared four mobile text entry methods: two-thumb QWERTY typing, one-finger QWERTY typing, handwriting recognition, and shape writing recognition. The two QWERTY techniques were the fastest, with no statistically significant difference between them in entry speed or accuracy. Shape writing was slightly slower, but similar in accuracy. Handwriting was the slowest and least accurate technique.

Keywords: Text entry, metrics, entry speed, accuracy, Android

1. Introduction

Mobile devices are often used for SMS text messaging and social networking. An estimated 9.2 trillion text messages will be sent in 2013 (Web-1) and more than 680 million people access Facebook using mobile devices per month (Web-2). Thus, investigating methods for mobile text entry is a significant research topic. To aid evaluation of mobile text entry methods, we created an application to gather user performance metrics on Android devices: Text Entry Metrics on Android (TEMA).

In this paper, we present our motivation for TEMA and its features. After summarizing existing text entry methods, we present a novel study using TEMA to compare four popular mobile text entry techniques.

2. TEMA Development and Features

The prevalence of mobile devices encouraged us to develop text entry techniques for mobile devices. This necessitated a program to gather performance metrics – a mobile equivalent to *TextTest* (Wobbrock and Myers 2006) for the PC. The result, TEMA, is an application to aid mobile text entry researchers using Android devices.

The choice to target Android devices was simple. The Android operating system powers over 900 million (Web-3) mobile devices, including smartphones and tablets. Text entry is accomplished using physical keyboards, soft keyboards, shape writing (Kristensson 2007), handwriting recognition, or voice recognition. This variety exists because anyone can freely develop and distribute an Android text Input MEthod (IME in developer parlance). Android is the only popular mobile platform to allow third-party text entry methods. These IMEs can be used system-wide, without modifying installed applications. Consequently, TEMA can run on a vast number of mobile devices and form factors, each capable of using a variety of IMEs.

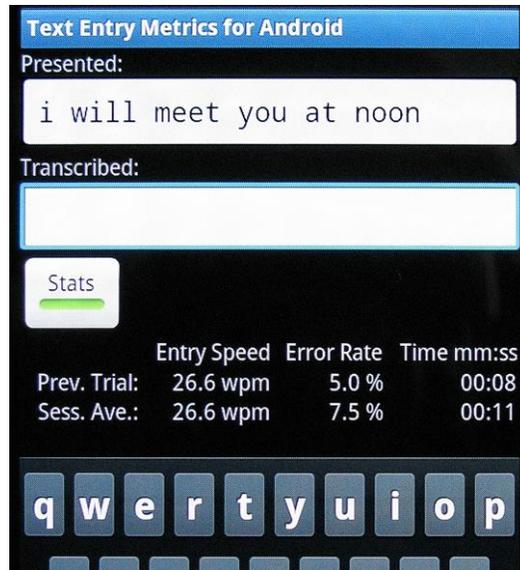


Fig. 1. The TEMA application (above) is available at <http://www.cse.yorku.ca/~stevenc/tema/>.

TEMA (Fig. 1) is a ready-made application to aid researchers gathering text entry metrics on Android devices. It occupies little storage space (only 125 kB) and has the following features:

Entry speed metric: Entry speed is calculated by dividing the length of the transcribed text by the entry time (in seconds), multiplying by sixty (seconds in a minute), and dividing by five (the accepted word length, including spaces (Yamada 1980)). Thus, the result is reported in words-per-minute (wpm).

Accuracy metrics: Accuracy is evaluated according to the total error rate (TER), corrected error rate (CER), and uncorrected error rate (UER) metrics (Soukoreff and MacKenzie 2004). TER characterizes general input accuracy and is the sum total of CER and UER. CER reflects the errors that the participant corrected during transcription, while UER reflects the errors that the participant did not correct. All three error rates are reported as a percent.

Stats log: The “stats” log summarizes entry speed, the accuracy metrics mentioned above, and intermediate measurements (e.g., presented text, transcribed characters, elapsed time, etc.) for each trial. This information is in comma-separated values (CSV) format and can be opened by most spreadsheet applications.

Event log: The “event” log is also in CSV format and contains time-stamped (in milliseconds) input events for low level, post-study analysis. Both event and stats logs are saved to the Android device’s internal storage. Log files can be transferred to a PC via a USB or wireless connection.

Set of 500 phrases: The text presented for transcription is randomly chosen from a 500-phrase set (MacKenzie and Soukoreff 2003).

Interruption timer: Although interruptions are not recommended during evaluation sessions, TEMA measures the duration of interruptions (e.g., an incoming phone call, etc.) and deducts it from the transcription time.

Trial management: Text entry trials can be refreshed (with a new phrase) or reset (with the same phrase) if a user gets distracted or pauses unnecessarily during a trial; all measurements are reset.

3. Text Entry Methods

Many mobile text entry techniques exist. The ubiquitous QWERTY layout is available as an onscreen keyboard, where users enter text one character at a time by tapping on the desired character’s “key”. A few devices also have a slide-out QWERTY keypad with physical keys for text entry. The physical keypad does not occupy screen space, but provides tactile feedback when typing.

Handwriting recognition requires the user to draw gestures with a finger or a stylus. Each gesture corresponds to a specific character. Some techniques, such as *Unistrokes* (Goldberg and Richardson 1993), *MDTIM* (Isokoski and Raisamo 2000), and *EdgeWrite* (Wobbrock et al. 2003), use gestures that users must learn, but are easily distinguishable by the recognizer. Others, such as *Graffiti* (www.hpwebos.com) and *DioPen* (www.diotek.com), use gestures that resemble handwritten characters. *DioPen* gestures are composed of up to three separate strokes to allow for variations in handwriting input (Web-4).

Shape writing recognition allows entry of entire words with a single, continuous gesture. Users draw a path on the keypad from the first letter of the word, intersecting each subsequent letter. If the shape of the path matches multiple words, the user selects the desired word from a short list. The initial shape writing technique (Zhai and Kristensson 2003) was adapted to the QWERTY layout and released commercially as *ShapeWriter* (www.shapewriter.com). There are currently many other shape writing techniques. *Swype* (www.swypeinc.com) is a popular one.



Fig. 2. The letter “e” being entered using *DioPen* (left).
The word “the” being entered on the *Swype* keypad (right).

4. Method

An earlier pilot study (Castellucci and MacKenzie 2011) was small, with only six participants of varying experience. The study detailed below addresses both concerns and introduces an additional input condition.

4. 1. Participants

Sixteen paid participants (ten male, six female) were recruited from the local university campus. Ages ranged from 18 to 31 years ($\mu = 23$; $\sigma = 3.53$). Two participants were left-handed. Although participants were familiar with the QWERTY layout, none was an expert in onscreen QWERTY keypads, handwriting, or shape writing techniques. Therefore, the results are characteristic of novice, not expert, performance.

4. 2. Apparatus

The TEMA application ran on a Samsung *Galaxy S Vibrant* (GT-I9000M) cell phone running Android OS v2.1. The touch screen measured 4.0 inches diagonally and had a resolution of 480×800 pixels. The phone was held in portrait orientation throughout the study. The phone’s wireless radios were disabled to eliminate disruptions due to incoming calls or text messages.

Three of the IMEs included with the phone were evaluated with TEMA: the default QWERTY keypad, *DioPen*, and *Swype*. For each IME, the input language was set to English (US) and options for auto-spacing, auto-capitalization, and word prediction were deactivated. All other options were kept at default values.

4. 3. Procedure

Participants entered ten phrases in each condition. They were instructed to enter text as quickly as possible, to correct errors if noticed immediately, but to ignore errors made two or more characters back.

The QWERTY keypad was used in two conditions. In one, the phone was held with two hands and participants typed with both thumbs (Fig. 3, left). In the *DioPen*, *Swype*, and other QWERTY conditions, participants held the device in their non-dominant hand and use a finger on their dominant hand to perform input (Fig. 3, right).

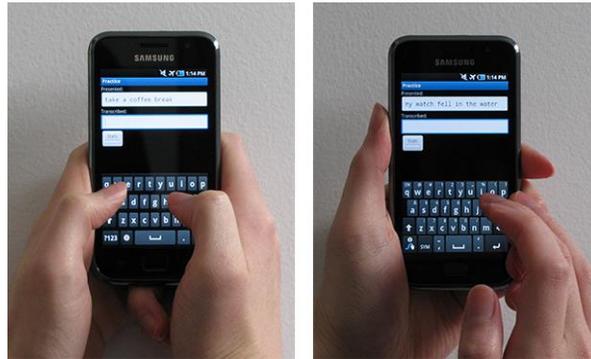


Fig. 3. The above images demonstrate participants' hand positions during the study conditions.

Before each condition, participants were instructed on how to use the corresponding technique. A practice session followed, consisting of three random phrases. Study sessions typically lasted 50 minutes and took place in a quiet office, with participants seated at a desk.

4. 4. Design

The experiment employed a within-subjects factor, technique, with four levels: QWERTY-thumbs, QWERTY-finger, *DioPen*, and *Swype*. The two-thumb QWERTY input condition encapsulates a popular method of mobile text entry. The single-finger QWERTY condition represents an alternative QWERTY input method and allows comparisons with the single-finger handwriting and shape writing input techniques.

The order of testing was counterbalanced using a balanced Latin Square. The dependent variables were entry speed and accuracy. They were measured by TEMA (as detailed previously) and averaged over the ten phrases.

5. Results and Discussion

5. 1. Accuracy

The TER of *Swype* was the lowest, at 7.0%. Interestingly, an evaluation of *ShapeWriter* on a tablet PC revealed a similar TER value of 6.7% (Kristensson 2007, pp. 65-66). The TER of the QWERTY-finger condition was slightly higher at 7.1%. Surprisingly, the QWERTY-thumbs condition was almost double that, at 13.8%. This is considerably greater than the 10.4% TER measured using two thumbs on the *iPhone*'s QWERTY keypad (Arif et al. 2010). Unfortunately, *DioPen* had the worst TER, at 30.4%. In comparison, a *Graffiti* study revealed an error rate of only 19.4% (Költringer and Grechenig 2004).

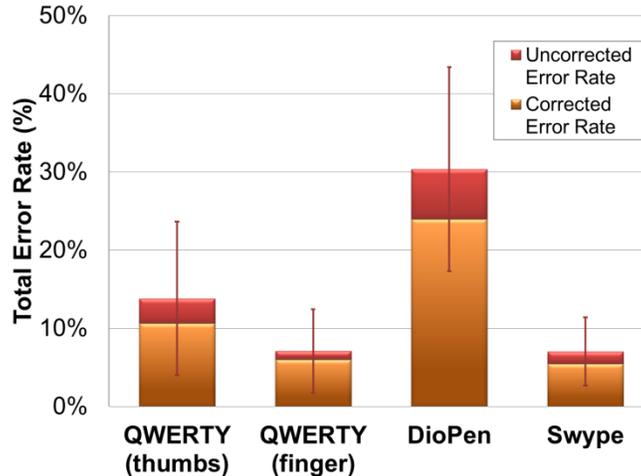


Fig. 4. Accuracy values gathered by TEMA. Error bars represent ± 1 standard deviation of TER.

An Analysis of Variance (ANOVA) revealed a significant effect of technique on total error rate (TER) ($F_{3,36} = 41.66, p < .0001$). However, Scheffé post hoc analysis indicated a significant difference only between *DioPen* and all other conditions. Condition order had no significant effect on TER ($F_{3,12} = 0.83, ns$).

DioPen's UER of 6.4% indicates participants missed (or ignored) many errors. The corresponding event logs revealed multiple attempts to enter characters (i.e., participants entered an incorrect character, backspaced, entered the same incorrect character, backspaced, etc.). This suggests participants could not reliably draw the required gestures. Considering the *DioPen* gesture alphabet (Web-4), the errors generally fall under three categories: incomplete loops (e.g., "c" inputted instead of "o"), incorrect proportions (e.g., "h" or "r" inputted instead of "n"), and poor timing (e.g., "l" inputted instead of "i"). The frequency of these errors would likely decrease with practice, as users perfect their gestures.

5. 2. Entry Speed

The QWERTY-finger entry rate of 20.9 wpm is the fastest in our study. The QWERTY-thumbs entry rate was just slightly lower at 20.8 wpm. Both values exceed the 15.9 wpm reported for two-thumb text entry on the *iPhone*'s QWERTY keypad (Arif et al. 2010). *DioPen* was the slowest technique at 7.0 wpm. This is probably related to the high rate of gesture misrecognition. A *Graffiti* study yielded a rate of 9.2 wpm (Költringer and Grechenig 2004). Our *Swype* entry speed of 16.7 wpm is consistent with a *ShapeWriter* study that reported 15 wpm (Kristensson 2007, pp. 65-66).

There was a significant effect of technique on entry speed ($F_{3,36} = 71.17, p < .0001$). However, there was no significant difference between the two QWERTY conditions. This is surprising, as many believe two-thumb input to be a faster method of text entry. This study focused on novice performance. Perhaps expert users learn to better coordinate input with two thumbs, resulting in faster input. Every other pairwise comparison of techniques satisfied the 5% threshold for significance. Again, counterbalancing proved effective ($F_{3,12} = 2.34, p > .05$).

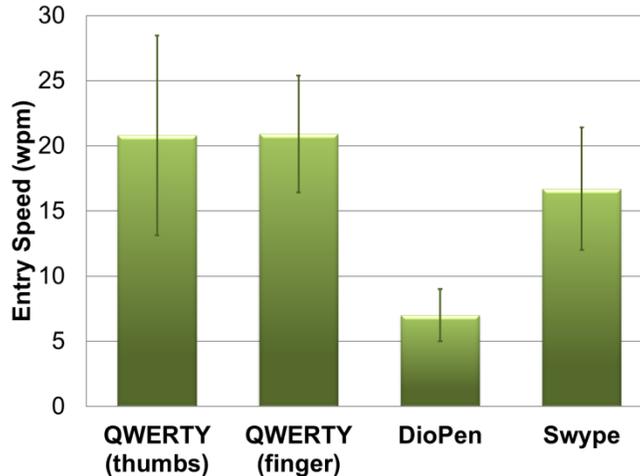


Fig.5. Entry speed values gathered by TEMA.
Error bars represent ± 1 standard deviation.

5. 3. Participant Feedback

We also recorded participants' qualitative feedback about the techniques. The QWERTY-finger condition was rated most favorable overall. Interestingly, *Swype* rated ahead of QWERTY-thumbs. Some participants found using two thumbs awkward. One participant complained the side of his thumbs inadvertently hit adjacent keys. Another participant did not like using the touchscreen at all, stating, "It's difficult with long finger nails."

With *Swype*, some participants stated that having to keep contact with the touchscreen occluded the keypad and made locating keys difficult. With experience, users might forego visual scans of the keypad, draw word paths faster, and increase performance.

Most participants were frustrated by *DioPen's* unreliable input. One participant mentioned that *DioPen* was difficult to use because its gesture alphabet did not resemble his own handwriting. Another participant stated, "It's just easier to type [rather than write]."

6. Conclusion

TEMA is a ready-made application to aid researchers gathering text entry metrics on Android devices. It includes hundreds of phrases for text entry, measures timings, calculates performance metrics, and generates easily viewable log files for post-study analysis.

The conducted study demonstrated TEMA's utility. Despite the perceived advantage of two-thumb input, there was no statistically significant difference between the two QWERTY conditions with respect to either novice entry speed or accuracy. Shape writing was slightly slower, but not significantly less accurate. Handwriting was both slow and error-prone.

Participant feedback was generally positive for both QWERTY conditions and the shape writing condition. However, shape writing requires constant contact with the touchscreen while entering a word. Some participants found this necessitated excessive concentration during input. Handwriting was largely disliked by participants. Its frequent gesture misrecognition was quite frustrating.

We hope other researchers will find TEMA and the metrics derived from our user study beneficial to their mobile text entry research. TEMA may be downloaded from the following URL: <http://www.cse.yorku.ca/~stevenc/tema/>.

References

- Arif, A.S., Lopez, M.H. and Stuerzlinger, W. (2010). Two new mobile touchscreen text entry techniques "poster at GI 2010," Ottawa, Canada, May 31-June 2, pp. 22-23.
- Castellucci, S.J. and MacKenzie, I.S. (2011). Gathering text entry metrics on android devices "Ext. Abs. CHI 2011," pp. 1507-1512.
- Goldberg, D. and Richardson, C. (1993). Touch-typing with a stylus "Proc. CHI 1993," Amsterdam, The Netherlands, pp. 80-87.
- Isokoski, P. and Raisamo, R. (2000). Device independent text input: A rationale and an example "Proc. AVI 2000," Palermo, Italy, pp. 76-83.
- Költringer, T. and Grechenig, T. (2004). Comparing the immediate usability of Graffiti 2 and virtual keyboard "Ext. Abs. CHI 2004," Vienna, Austria, April 24-29, pp. 1175-1178.
- Kristensson, P.O. (2007). Discrete and continuous shape writing for text entry and control "Department of Computer and Information Science," Linköping, Sweden.
- MacKenzie, I.S. and Soukoreff, R.W. (2003). Phrase sets for evaluating text entry techniques "Ext. Abs. CHI 2003," Ft. Lauderdale, FL, United States, pp. 754-755.
- Soukoreff, R.W. and MacKenzie, I.S. (2004). Recent developments in text-entry error rate measurement "Ext. Abs. CHI 2004," Vienna, Austria, pp. 1425-1428.
- Wobbrock, J. and Myers, B. (2006). Analyzing the input stream for character-level errors in unconstrained text entry evaluations. ACM TOCHI, 13, pp. 458-489.
- Wobbrock, J.O., Myers, B.A. and Kembel, J.A. (2003). EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion "Proc. UIST 2003," Vancouver, Canada, pp. 61-70.
- Yamada, H. (1980). A historical study of typewriters and typing methods: From the position of planning Japanese parallels. Journal of Information Processing, 2, pp. 175-202.
- Zhai, S. and Kristensson, P.-O. (2003). Shorthand writing on stylus keyboard "Proc. CHI 2003," Ft. Lauderdale, FL, United States, pp. 97-104.

Web sites:

- Web-1: <http://www.portioresearch.com/blog/2012/12/happy-birthday-sms!.aspx>, consulted 20 Feb. 2013.
- Web-2: <http://newsroom.fb.com/content/default.aspx?NewsAreaId=22>, consulted 20 Feb. 2013.
- Web-3: <http://www.engadget.com/2013/05/15/900-million-android-activations/>, consulted 15 May. 2013.
- Web-4: <http://help.diotek.com/data/diopen/android/10/page42.html>, consulted 20 Feb. 2013.