Modelling non-Expert Text Entry Speed on Phone Keypads

Andriy Pavlovych *Wolfgang Stuerzlinger*

Department of Computer Science, York University, Toronto, Canada
http://www.cs.yorku.ca/~{andriyp|wolfgang}

ABSTRACT

For mobile phones, previous research has created models that can be used to predict expert performance. However, one important factor that influences the success of new interaction techniques is the users' initial experience with them. In this work, we present a new model to predict text entry speed on 12-button mobile phone keypads for novices. The model is based on Fitts' law, letter digraph probabilities, and a model of the mental processing time before key presses.

KEYWORDS: Text entry, mobile phones, user model.

INTRODUCTION

Widespread use of mobile devices during the last several years has renewed interest in efficient text entry techniques. However, designing new text entry methods for computing systems is usually labour intensive: one typically needs to build a prototype device and to conduct extensive user studies. Thus, a model that predicts the performance of a new method as closely as possible without the need to do either of those time-consuming tasks is valuable.

Existing models for text entry

As of time of writing, two models for text entry on a 12-button keypad are known. Both of them were designed to predict expert (or *peak*) text entry rates for various text entry methods.

Model 1. The first model – a keystroke level model – was presented by Card *et al.* [1]. The model accounts for key press times, mouse movements etc; and it includes a "mental preparation time" before operations. This model does not rely on Fitts' law and assumes that all key presses take an equal amount of time. We adapted this model to phone text entry; and the predictions appear in Table 1.

Model 2. The second model was developed by Silfverberg *et al.* [5]. It contains two parts: a movement model based on Fitts' law and a linguistic model for the distribution of key digraphs in a given corpus. We recomputed the numbers, as we are using a timeout of 1000 ms instead of 1500 ms, and the revised predictions appear in Table 1.

Note that the two models give fairly different predictions, especially for *T9*. Note further, that few people ever reach expert/peak performance, as the average number of text messages sent over phones tends to be rather small.

NEW MODEL FOR TEXT ENTRY ON PHONE KEYPADS

The new model presented here can be viewed as an extension of the model by Silfverberg *et al.* [5]. In particular, the new model is applicable to non-expert users. This is a very useful extension, since, as stated above, few people ever reach the expert level.

Observed time for various key presses

Figure 1 shows times for the keystrokes of different types – single, double, triple, quadruple etc. The data is based on logs from previous experiments [4] and are obtained mainly from novice users.

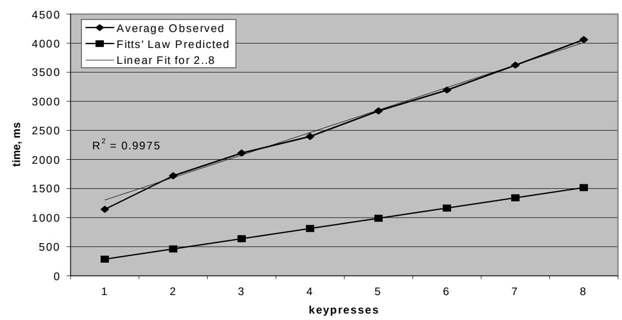

Figure 1 Time needed for multiple key presses (for non-expert users)

While the times for double through octuple presses can be fit very well by a line, the difference between single and double keystrokes is visibly larger. Also, note that the times are much greater than one would expect from Fitts' law.

Key repeat time

It has been doubted for some time and, in fact, experiments have proven that, Fitts' law does not apply when index of difficulty values are small [6]. As an independent verification, we measured the key repeat time on a Nokia 5190 to be 225 ms ($\sigma = 21$) (this value is later referred to as T_{repeat}). The number is an average computed over 10 people. Compare it to 176 ms as the *a* coefficient in the Fitts' law obtained in previous studies [5].

Mental overhead

From the key log data and our observations we tried to identify important events in text entry, other than making key presses:

1. Re-reading the phrase to be entered (most people prefer not to memorize the phrase they have to enter).
2. Figuring out which letter of which word has to be entered next (spelling out the word).
3. Determining which button should be pressed and how many times.

4. Deciding if a second key press is required.
5. For multiple presses, counting the number of presses.
6. Verifying the result.

All these operations could produce the difference between observed values and predictions by models that consider only human motor limitations and letter distributions.

Based on the data collected for *Multitap* and *Less-Tap* [4], each initial keystroke is preceded by 906 ms ($\sigma = 258$) of cognitive delay (D_{init}), computed as the average observed movement time minus the average time predicted by Fitts' law. The first keystroke that falls onto the same key as the previous one is preceded by a 323 ms ($\sigma = 131$) cognitive delay (D_{repeat}), and each key press after that by 125 ms ($\sigma = 143$) (D_{count}).

Time to enter a character (Movement Model)

From the above, we can derive the times required to enter a character using different text entry systems.

Multi-press Input Methods. Methods based on repeated key pressed include *Multitap* and *Less-Tap* (see e.g. [4]). For such methods, the time to enter a character is modeled as:

$$T_{\text{char}} = D_{\text{init}} + T_{\text{Fitts}} + N_1 \cdot (D_{\text{repeat}} + T_{\text{repeat}}) + N_2 \cdot (D_{\text{count}} + T_{\text{repeat}}) + [T_{\text{timeout}}]$$

T_{Fitts} is the time needed to move the finger from the preceding key to the current key, as predicted by Fitts' law. N_1 is the number of "second" presses (present – one or absent – zero). N_2 is the number of key presses after the second. T_{timeout} is time that the user would have to wait for if current character is located on the same button as the previous one.

Obviously, the coefficients decrease in magnitude with practice. We used the results for *Multitap* of MacKenzie *et al*. where *Multitap* and *LetterWise* were analyzed in a longitudinal study [3]. Figure 2 demonstrates our estimate of the value for D_{init} and a power-law extrapolation for more sessions. Note that the delay is is still 200 ms at the 30$^{\text{th}}$ session! By varying the coefficient values the model can be adapted to model performance of users of various levels of expertise.

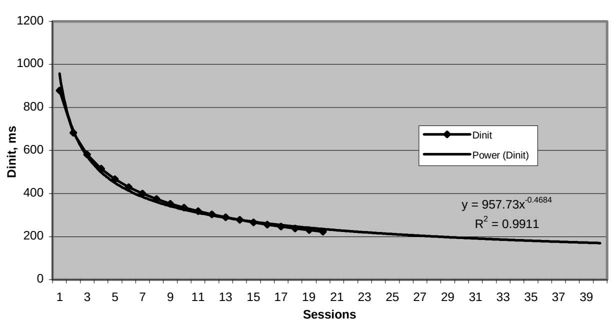

Figure 2 Coefficient D_{init} as a function of time

Predictive Input Methods. Predictive Input Methods include *T9*, *iTap* and several others. For simplicity, we assume that disambiguation is perfect (as presses of the 'next' key amount to less than 1% of the total) and that the verification time is included in the time to prepare for the next key press (D_{init}). Both simplifications cause the model to overestimate, but we believe that the factor is small.

For predictive methods , the model is:

$$T_{\text{char}} = D_{\text{init}} + T_{\text{Fitts}}$$

We did not find as much experimental data on predictive methods, but we find it reasonable to expect that the keystrokes in those systems be preceded by roughly the same amount of cognitive delay as in multi-press methods.

Linguistic Model

A linguistic model contains information about the frequency of different letter-pairs (digraphs). Here, it is based on the letter-pair data from the British National corpus. The model is represented by a matrix (26 English letters plus SPACE). Each cell p_{ij} in the 27×27 the matrix is the probability of the corresponding letter pair in the corpus. See e.g. [5].

Combining the Models

Combining all the above, we obtain a new model to predict the average time to enter a character in the corpus for a text entry system and a language:

$$T_{\text{char\_in\_corpus}} = \Sigma\Sigma\, (p_{\text{ij}} \cdot T_{\text{char ij}})$$

This number can easily be converted to the common measure of words per minute.

VERIFYING THE MODEL

Model's prediction of entry speed in words per minute for various text entry methods is shown in Table 1. That table contains also a comparison with previous models. Fitts' law coefficients from [5] were used, since [4] uses the same telephone handset.

Technique	*NEW Model*	*Model [1].*	*Model [5]*	*[2] Novices*	*[4]*
Multitap	6.97	18.35	22.3	7.98	7.15
Less-Tap	8.01	23.47	26.8		7.82
T9	10.07	24.97	40.6	9.09	

Table 1 Model predictions and experimental results

Note that the new model, in its prediction, comes much closer to values observed in experiments with novices.

DISCUSSION

We presented a model to text entry speed on 12-button telephone keypads, which through inclusion of factors for mental overhead quite accurately predicts the performance for non-expert users. The values computed by the model are reasonably consistent with those experimentally observed.

REFERENCES

1. Card, S.K., Moran, T.P., and Newell, A. *The psychology of human-computer interaction*, Hillsdale, NJ: Lawrence Erlbaum, 1983.
2. James, C. L., and Reischel, K. M. Text input for mobile devices: Comparing model predictions to actual performance, *ACM CHI 2001*, pp. 365-371.
3. MacKenzie et al. LetterWise: Prefix-based disambiguation for mobile text input. *ACM UIST* 2001, pp. 111-120.
4. Pavlovych, A., Stuerzlinger, W. Less-Tap: A Fast and Easy-to-learn Text Input Technique for Phones. *Graphics Interface 2003*, pp. 97-104.
5. Silfverberg et al. Predicting text entry speeds on mobile phones. *ACM CHI 2000*, pp. 9-16.
6. Soukoreff et al., Using Fitts' law to Model Key Repeat Time in Text Entry Models. Poster at *GI 2002*.