Enhancing KLM (Keystroke-Level Model) to Fit Touch Screen Mobile Devices

Karim El Batran
Computer and Information Sciences,
University of Strathclyde
Richmond St, Glasgow, G1 1XH, Scotland, UK.
karim.el-batran@strath.ac.uk

Mark D. Dunlop
Computer and Information Sciences,
University of Strathclyde
Richmond St, Glasgow, G1 1XH, Scotland, UK.
Mark.Dunlop@strath.ac.uk

ABSTRACT
This short paper introduces an enhancement to the Keystroke-Level Model (KLM) by extending it with three new operators to describe interactions on mobile touchscreen devices. Based on Fitts's Law we modelled a performance measure estimate equation for each common touch screen interaction. Three prototypes were developed to serve as a test environment in which to validate Fitts's equations and estimate the parameters for these interactions. A total of 3090 observations were made with a total of 51 users. While the studies confirmed each interaction fitted well to Fitts's Law for most interactions, it was noticed that Fitts's Law does not fit well for interactions with an Index of Difficulty exceeding 4 bits, highlighting a possible maximum comfortable stretch. Based on results, the following approximate movement times for KLM are suggested: 70ms for a short untargeted swipe, 200ms for a half-screen sized zoom, and 80ms for an icon pointing from a home position. These results could be used by developers of mobile phone and tablet applications to describe tasks as a sequence of the operators used and to predict user interaction times prior to creating prototypes.

Author Keywords
Quantitative prediction model; KLM (Keystroke-Level Model); GOMS; Fitts's Law; touch screen interaction.

ACM Classification Keywords
H.5.2 User Interfaces: Input devices and strategies.

General Terms
Design; Measurement.

INTRODUCTION
Quantitative prediction models, such as GOMS (Goals, Operators, Methods, and Selection rules) and KLM (Keystroke-Level Model) have been shown to be useful tools in modeling interaction and in deciding between, and filtering out designs (e.g. [1]). KLM is used to estimate the time taken to complete simple data input tasks by combining small timing constants. On physical-key devices it has been widely used to predict task times (e.g. text entry for Korean Language [10], Text Entry Speed on 12-button Phone Keypads [7] etc.).

Many enhancements to KLM have been proposed in the literature to be able to evaluate different techniques. However, there has been little research that improves user behavior modeling techniques to estimate the time taken to achieve common interactions performed on mobile touchscreen devices. In this paper we report our work on enhancing KLM by extending it with three new operators. In this short paper we report our investigation into modeling three interactions performed on mobile devices and tablets using KLM. Our model is based on suggested times derived from Fitts's Law modelling and analysis of how well these interactions fit Fitts's Law. We believe this gives an enhancement for developers of smart-phone and tablet applications to predict user interaction times without even needing to create prototypes.

KLM-GOMS MODEL
KLM was introduced as part of the wider GOMS-related work of Card, Moran, and Newell on modelling and quantitatively predicting the skilled and error free performance of users interacting with a text editor [1]. KLM is usually applied in situations that require minimal focussed, scripted interaction with a computer interface or software design.

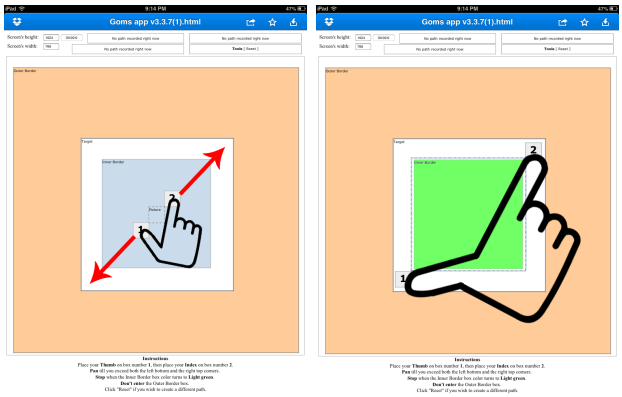

Figure 1 Zoom usability test scenario

MobileHCI '14, September 23 – 26, 2014, Toronto, ON, Canada

http://dx.doi.org/10.1145/2628363.2628385

Assuming error-free expert user interaction, KLM proved to yield precise estimated results, which is a drawback of similar models [10]. Moreover, when the estimated experimental studies results were considerably different from observed values, the estimated difference between two examined designs still proved to be a strong basis for making design choices.

KLM FOR PHONE USERS

Dunlop and Crossan [2] used KLM operator sequences to compare three different text entry methods for traditional physical-key phones (multitap, predictive, and word completion). How and Kan [5] defined 13 operators that directly map onto the phone keyboard interface according to the different input methods. They used videotaped sessions with a small set of subjects and a message typing task in order to gather new times for these operators. Mori et al. [9] studied how the time values of the original KLM operators apply to mobile phone menu navigation and concluded that the operator values fit well and suggested only minor modifications.

KLM FOR MOBILE DEVICE INTERACTIONS

Until recently the main use of mobile phones was making phone calls, sending text messages, and sometimes basic calendar tasks. Phones used very small displays and small physical buttons as the primary interface. With modern smartphones these buttons have been replaced with larger touch screens. Moreover, other uses are becoming more popular such as taking pictures, surfing the web, social networking, and playing music, videos and games. This adds several new interaction styles that have not yet been fully described by interaction models.

There has been surprisingly little published research that includes new interaction techniques such as swipe, zoom, or tabbing, for mobiles and tablets. Luo and John followed by Teo and John showed that the method could be soundly applied to handheld devices using stylus-based interfaces [6, 11]. They also presented a tool to automatically generate KLM models from storyboard descriptions and stated that they aimed to apply such research to novel interfaces those using speech and gestures. Holleis et al. adopted and defined a set of operators to give study-based estimates of performance measures. They assumed that developers of mobile applications could then describe tasks as a sequence of the operators they added and predict user interaction times without needing to create prototypes [4]. Closest to our work, as part of their research on user performance of multi-touch gestures on mobile devices, Tran et al. conducted an exploratory study of pinch and spread gestures on smartphones and tablets [12]. Their spread gesture on tablet is equivalent to our Zoom interaction.

CONCEPT AND METHODOLOGY OF PROTOTYPES

The development of three prototypes came in the process of enhancing KLM. These prototypes reflected three common interactions used widely on tablets and mobile devices:

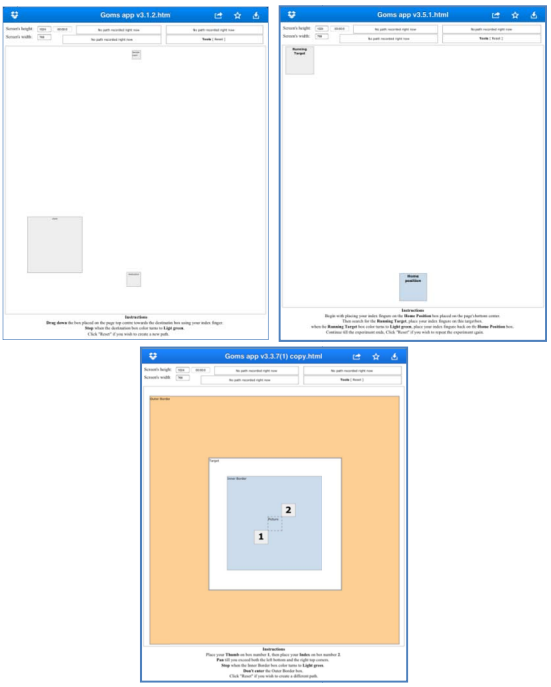

Figure 2. Swipe (left), Tap (right), & Zoom prototypes.

1. Swipe interaction, which enables the user to do different swipe gestures (short or long) to achieve certain tasks (e.g. scrolling through a document etc.).
2. Tap interaction, which estimates the user's tapping response time in comparison to his/her finger travel distance from a home position (e.g. to estimate typing speed or pointing to a target etc.).
3. Zoom interaction, which is used to cause text or other graphics in a window or frame to appear larger on the screen (e.g. zooming, panning a map/photo etc.).

The three prototypes were built to estimate the time taken to conduct these interactions. Fitts's Law [3] as a model is used to analyze the general case of time taken estimation for these interactions. It is primarily used in human–computer interaction and ergonomics to predict the time required to rapidly move to a target area, it is a function of the distance to the target and the size of the target. One of a variety of available formulations, the Shannon formulation is widely used [9] and is defined as:

$$T = a + b * ID, \text{ where } ID = log_2\left(\frac{D}{W} + 1\right).$$

Where

- T is the average time taken to complete the movement. It may also be referred to as MT (mean movement time),
- a represents the incorporating reaction time and/or the time required to execute the operator,
- b represents for the inherent speed of the device,
- D is the distance from start to the center of the target and
- W is the width of the target.

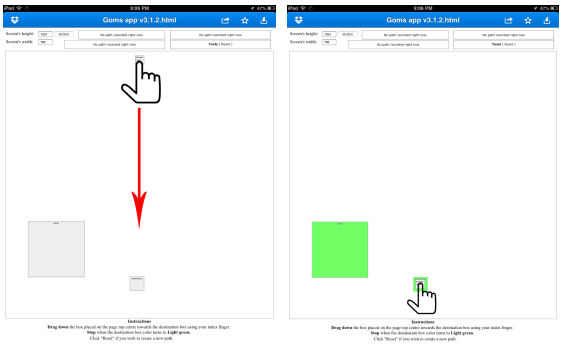

Figure 3 Swipe usability test scenario

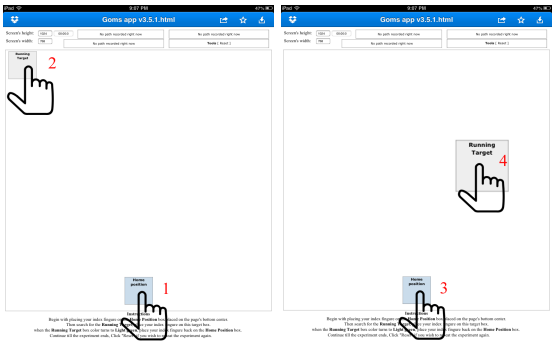

Figure 4 Tap usability test scenario

The constant "1" in the formula made a difference from Fitts's original form, especially for low values of the ratio D/W, given the advantage that the ID is always non-negative. Here we are investigating the ability of modelling the finger touch input for different target sizes and movement distances.

DEVELOPING AND TESTING THE PROTOTYPES

The prototypes were developed in HTML5 / JavaScript and displayed on an iPad Mini (a 7" multi-touch tablet), as shown in figure 2. Three usability studies were conducted to extract completion times (in seconds). For each interaction we tested a range of target size and distances from the starting point to the center of the target (in pixels).

Participants

30 volunteers took part in the Swipe and Zoom usability tests. The participants were aged 22-39, 35% female / 65% male, bachelor graduates, and 70% were familiar with touch screens and 85% were PC users. 21 users conducted the Tap tests, with a similar demographic profile.

Tasks

For the Swipe task, forty eight successive attempts were made. As shown in figure 3, users dragged down a box placed on the top center towards the destination box placed on the bottom center of a web page using the participant's index finger. The participant stopped and removed his/her index finger from the screen when the destination box color turned to light green. This was recorded as a successful attempt.

The Tap task, ten successive attempts were made. As shown in figure 4, users were asked to place their index finger on the "Home Position" box near the bottom of the page. The participant searched for the "Running Target" box and placed his/her index finger on this target box. He/she removed his/her index finger from the "Running Target" box when its color turned to light green. Afterwards, the participant placed his/her index finger back on the "Home Position" box. This was recorded as a successful attempt.

For the Zoom task, twenty six successive attempts were made. As shown in figure 1, each by placing the thumb finger of the participant on the box number 1 and his/her index finger on the box number 2. The participant zoomed till he/she exceeded both the left bottom and the right top corners of the border box, within the target box. He/she stopped and removed his/her thumb and index fingers from the screen when the border box color turned to light green. This was recorded as a successful attempt.

Results

For the Swipe task, as shown in figure 5, target width and target amplitude varied across four levels resulting in IDs of 1 to 4 bits. The target width varied from 50 to 400 pixels, while target amplitude varied from 100 to 800 pixels. Mean movement time ranged from 20ms to 1610ms with each score derived from over 1440 observations. While the Tap task, the target width varied from 50 to 250 pixels, while target amplitude varied from 125 to 900 pixels. Mean movement time ranged from 50ms to 470ms with each score derived from over 210 observations. For the Zoom task, the target width varied from 50 to 150 pixels, while target amplitude varied from 100 to 550 pixels (fixed step of 50 pixels). Mean movement time ranged from 60ms to 1230ms with each score derived from over 1440 observations. Regressing MT on ID yields the following prediction equation for movement time (ms):

- Swipe interaction: $MT = 9.46 + 55.83\ ID$.
- Tap interaction: $MT = 52.12 + 14.62\ ID$.
- Zoom interaction: $MT = 114.86\ ID - 20.45$.

By comparing the results with the previous work of Tran et al. [12], both outcomes are largely in line with slightly different prediction equation and higher correlating MT with ID (R^2 Zoom > 0.884).

POOR FIT FOR ID=4

It was noticed in the Swipe and Zoom usability tests that the 4th ID level does not fit in the straight line, which passes by the other 3 levels of ID. Moreover, the constant "a", representing the incorporating reaction time and/or the time required executing the operator, in the prediction equation for movement time of the Zoom usability test is negative. By excluding ID=4, as shown in figure 5, regressing MT on ID yields the following predictions (ms):

- Swipe interaction: $MT = 44.78 + 34.64\ ID$.

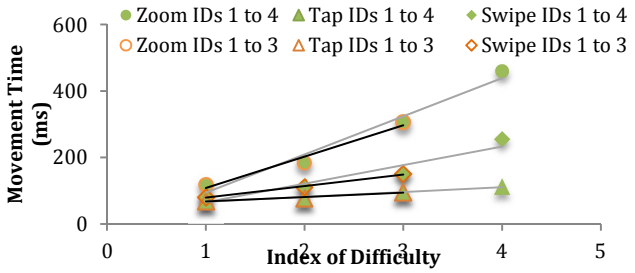

Figure 5 Movement Time vs. Index of difficulty for Swipe, Tap, and Zoom tests with IDs of 1 to 3 and 1 to 4 bits.

Interaction	Time
Short untargeted swipe, ½ width of 5" screen	70ms
½ screen zoom, 100*100 to 350*350px image	200ms
Icon pointing of size 200*200px at a distance of 700px from home position	80ms

Table 1. Suggested approximate movement time for KLM

Interaction	R^2
Swipe usability test with IDs of 1 to 3	0.995 ($p<0.05$)
Swipe usability test with IDs of 1 to 4	0.911 ($p<0.1$)
Tap usability test with IDs of 1 to 3	0.972 ($p<0.05$)
Tap usability test with IDs of 1 to 4	0.985 ($p<0.05$)
Zoom usability test with IDs of 1 to 3	0.972 ($p<0.05$)
Zoom usability test with IDs of 1 to 4	0.971 ($p<0.05$)

Table 2. R^2 values for interaction execution times.

- Tap interaction: MT = 54.38 + 13.27 ID.
- Zoom interaction: MT = 13.75 + 94.33 ID.

The correlating MT with ID for the Swipe test yields R^2 = 0.995 ($p<0.01$) instead of R^2 = 0.911 ($p<0.1$) (see table 2), which gives a better-fit and more precise prediction.

We believe that the poor fit of ID=4 is indicative the increase in the spread distance required for a Zoom task of this size is too large for comfort, where the target amplitude (D) is between 500px and 550px (about 98 mm) while the target width (W) is 50px. This result parallels what Hoggan et al. found for pinch, where a large pinch distance of over 90mm took significantly longer time to complete tasks and produced an increase in ergonomic failure rate [13].

CONCLUSION

The objective of this paper is to enhance KLM as a quantitative prediction model, by extending it with new operators necessary to describe interactions used on mobile devices and tablets. It was determined, after conducting three usability tests that estimating the movement time for certain interactions is done using Fitts's Law. Furthermore, three equations were formed to estimate the movement time for three interactions; Swipe, Tap, and Zoom. The ID is the variable used to be submitted in any of the three equations in order to calculate the movement time. Based on results, the following approximate movement times for KLM are suggested as shown; a short untargeted swipe will take approximately 70ms, a half-screen sized zoom will take approximately 200ms, and an icon pointing from a home position will take approximately 80ms. We also identified a poor-fit for some high ID operations indicating a possible maximum comfort limit for these tasks.

ACKNOWLEDGMENTS

We are very grateful to our study participants.

REFERENCES

1. Card, S. K., Moran, T. P. and Newell, A. (1983): *The Psychology of Human–Computer Interaction*. Erlbaum.
2. Dunlop, M.D. and Crossan, A. (2000). Predictive Text Entry Methods for Mobile Phones. *Personal Technologies*. 41. 134-143.
3. Fitts's, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Experimental Psychology*. 47(6). 381-391.
4. Holleis, P., Otto, F., Hussmann, H., and Schmidt, A. (2007). Keystroke-level model for advanced mobile phone interaction. *Proc. CHI 2007*. 1505-1514.
5. How, Y. and Kan, M. Y. (2005). Optimizing Predictive Text Entry for Short Message Service on Mobile Phones. *Proc. HCI International 2005*.
6. Luo, L. and John, B. E. (2005). Predicting Task Execution Time on Handheld Devices Using the Keystroke-Level Model. *Proc CHI 2005*. 1605-1608.
7. Pavlovych, A., Stuerzlinger, W. (2004). Model for Non-expert Text Entry Speed on 12-button Phone Keypads. *Proc. CHI 2004*. 351-358.
8. MacKenzie, I. S. (1992). Fitts's' law as a research and design tool in human–computer interaction. *Human–Computer Interaction*. 7. 91–139.
9. Mori, R., Matsunobe, T., and Yamaoka, T. (2003). A Task Operation Prediction Time Computation Based on GOMS-KLM Improved for the Cellular Phone and the Verification of that Validity. *Proc. ADC 2003*.
10. Myung, R. (2004). Keystroke-level analysis of Korean text entry methods on mobile phones. *Int. J. Human-Computer Studies*. 60(5), 545-563.
11. Teo, L. and John, B. E. (2006). Comparisons of Keystroke-Level Model Predictions to Observed Data. *Extended Abstracts of CHI 2006*. 1421-1426.
12. Tran, J. J., Trewin, S., Swart, C., John, B. E. and Thomas, J. C. (2013). Exploring pinch and spread gestures on mobile devices. *MobileHCI 2013*. 151-16.
13. Hoggan, E., Nacenta, M., Kristensson, P. O., Williamson, J., Oulasvirta, A., and Lehtiö, A. (2013). Multi-touch pinch gestures: performance and ergonomics. *Proc. ACM 2013*. 219- 222.