Grip Change as an Information Side Channel for Mobile Touch Interaction

Matei Negulescu, Joanna McGrenere
Department of Computer Science, University of British Columbia
{mnegules, joanna}@cs.ubc.ca

ABSTRACT
In order to reach targets with one hand on common large mobile touch displays, users tilt and shift the device in their hand. In this work, we use this grip change as a continuous information stream for detecting where the user will touch while their finger is still en-route. We refer to this as in the air prediction. We show that grip change detected using standard mobile motion sensors produces similar in the air touch point predictions to techniques that use auxiliary sensor arrays, even in varying physical scenarios such as interacting in a moving vehicle. Finally, our model that combines grip change and the resulting touch point predicted where users *intended* to land, lowering error rates by 41%.

Author Keywords
Adaptive interface; mobile device; predictive model; touch; grip; display size

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

INTRODUCTION

With the growing popularity of mobile devices, researchers have searched for ways of improving input accuracy on touch displays. Of note, researchers have identified a number of factors that affect touch patterns in systematic ways: hand posture used to touch a target (e.g., a one-handed posture that uses the dominant thumb rather than two thumbs to type) [1,4]; the location of the target on the screen [6]; physical activity, such as walking [3] versus sitting at a desk; and even whether or not users are interacting while carrying items in the other hand [8]. These are all significant factors that impact touch patterns and therefore touch accuracy on mobile devices. Adaptive systems take advantage of systematic patterns of user touch interaction under these conditions to adjust where the user touched to where they *intended* to land. For example, Yin et al. [10] and Goel et al. [4,5] showed improved touch accuracy using adaptive soft keyboards that take into consideration systematic offsets caused by hand posture and individual differences.

ACM978-1-4503-3145-6/15/04…$15.00
http://dx.doi.org/10.1145/2702123.2702185

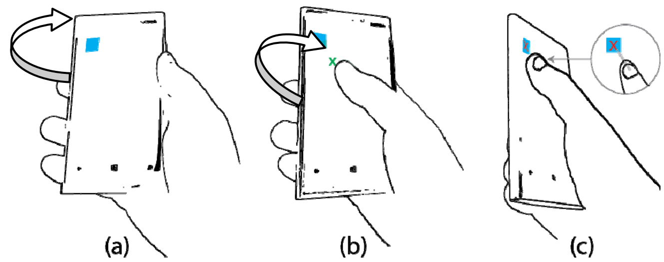

Figure 1. Grip change as the user reaches for the target (blue square): (a) device lays flat in palm, then (b) device is tilted towards the thumb by the other fingers enabling in the air prediction of an area (green circle) around a likely touch point (green X), and finally (c) at thumb touchdown, prediction can adjust actual input to intended target (red X).

The need for these adaptations has, if anything, increased with the variability of form factors spanning compact 3.5 inch displays (e.g. Nokia M8) and phone/tablet hybrids with near-6 inch touchscreens (e.g. Samsung Galaxy Note 3). As mobile displays get larger, the ability to accurately target, especially with one hand, around the screen suffers. Bergstrom-Lehtovirta et al. adaptively modeled users' reach to define a functional area describing the maximum reach of their current hand posture [2].

When users need to reach targets on the edge and outside their functional area with one hand, they change their grip of the phone. As Figure 1 shows, users tend to shift the device in their hand, extending the thumb's reach. As Noor et al. showed, this grip change can be used to predict the touch point [7], or rather a likely area around it (Figure 1b in green), while the finger is still en-route. We refer to this as *in the air prediction.* Beyond in the air prediction, grip change can also be an indicator of where the user *intended* to land at the moment they touchdown. We refer to this as *on touchdown prediction*. The target may be outside the user's functional area even at maximum reach, so their touch point might miss a target but be close to one or more viable targets (Figure 1c). The grip change motion preceding the touch is indicative of intent to reach to further targets. In other words, a model that uses both the grip change and where the user touches down can better predict where the user intended to land than models that only consider one of these factors.

Past work has detected physical grip and grip change over time through *adding extra hardware* – additional sensor arrays on the back or side of the device (e.g. [7,9]). Noor et al. used the grip change—as measured while the user reaches for the target—to predict the resulting touchdown point while the finger is still en route (in the air) [7]. However, we wondered whether an auxiliary sensor array is critical to in the air prediction, or whether internal motion sensors (i.e. common accelerometers and gyroscopes) can achieve similar results. Given the potential predictive power of grip change on touch accuracy, it is additionally important to understand how detection of grip change is affected by common physical scenarios such as walking or interacting while in a moving vehicle. This is not addressed in the literature.

We explore using solely the motion sensors internal to today's devices to detect grip change and then use it to make in the air predictions of touch points. We verify prediction results in four different physical scenarios: interacting while sitting at a desk, standing without support, while walking, and while on a moving bus. As proof of concept, we use test our models using a single thumb posture. Our first study shows in the air prediction rates similar to Noor et al. A second study shows that a model that adjusts users' touches using both grip change and their touchdown point lowers touch errors by 41% consistently throughout the four physical scenarios tested.

In short, grip change can be used to continuously provide information about users' intent, in the air and on touchdown, beyond where they actually touch. Predicting where the user will touch prior to touchdown allows for a potential virtual hover space for mobile devices, while touchdown models take into consideration user intent in lowering touch errors.

STUDY 1 – IN THE AIR PREDICTION USING MOBILE MOTION SENSORS

To verify that hand grip change can be reliably detected in the air using solely mobile motion sensors (i.e. accelerometer and gyroscope), we ran a study that largely replicates Noor et al.'s touch target prediction experiment [7]. Participants were asked to touch the single on-screen square target to complete a trial. The target size was chosen to be physically identical to Noor et al.'s setup on the mobile device ($1cm^2$). The target was randomly placed on the surface of the display for each trial with a uniform distribution. As in Noor et al., once the user successfully clicked on the target, the next target was displayed after an enforced 500ms delay.

Participants and Apparatus

Eight participants (mean age 26.3, SD=3.6, 7 males), all right-handed, were asked to perform 1000 trials each (8000 trials in all), using the phone in their right hand and interacting with the thumb (i.e. a one-handed posture). All participants were regular and frequent mobile phone users.

Participants were presented with a custom application on a Nokia Lumia 920 with a 4.5" display at a 1280x720 resolution. Of note is that this device is significantly larger than the one used in Noor et al. (a Nokia N9 with a 3.9" display), which makes a direct comparison of prediction accuracy non-trivial. The application recorded 2D touch points on the display, and 3-axis accelerometer and gyroscope readings at a sampling rate of 50Hz.

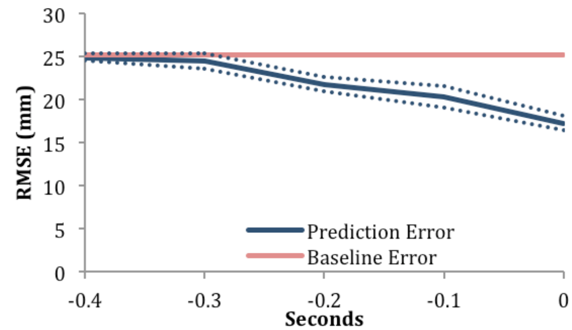

Figure 2. Root mean square error (and ± standard error in dashed lines) of touch point prediction before touch contact using mobile sensor detected hand grip change.

Results

To predict where participants touched based on hand grip change, we used the raw, time ordered accelerometer and gyroscope measures of the 500ms before touch as features. As with Noor et al. [7], we used Gaussian Process (GP) Regression to generate a nonlinear model that maps motion sensor values to 2D screen coordinates. Using Matlab's *gpml* package, the models were built independently for each axis and were defined using a squared exponential covariance function for smoothness and a constant mean function (with a prior of 0). The hyperparameters for these functions were learned from the data. Due to individual differences in how users grip and move the phone in their hand, the model was created on a per-user basis. Participant data was split between training and test sets using 10-fold cross validation.

As Figure 2 shows, the resulting models (in blue) provide a root mean square error (RMSE) smaller than a baseline model that simply picks the center of the available area for its prediction (in red). Predicting 0.1s before the touch provides a RMSE of 20.3mm, and 21.8mm at 0.2s. Though our much larger device makes comparison difficult, our grip change model provides similar results to Noor et al.'s 18mm (at 0.2s before touch) using only internal sensors. However, while motion sensor detected grip change can be satisfactory in some scenarios, more complex interactions (e.g. a virtual hover space) may require combining mobile sensors with high precision auxiliary sensor arrays such as Noor et al.'s.

STUDY 2 – GRIP CHANGE ON THE MOVE

The second study had two goals: (1) show how in the air prediction using motion sensor detected grip change performs in a more representative task and under various physical conditions (e.g. while walking); and (2) show how grip change can provide information on users' intent.

Methodology

The experiment was designed to gather touch and motion data from participants in a common task from every-day device use: selecting a particular target from an array of homescreen icons. In contrast to the first study, the participant selected one target, highlighted red, out of the always-visible array of 5x4 targets that simulate a Google

Android and Apple iOS homescreen. Each target was 99x99 pixels, which corresponds to the Google Android recommended size for icons for the Lumia 920's resolution and physical size.

In order to see how well a grip change prediction model works under common motion scenarios, participants performed the study in four physical conditions:

a) *Sit*: seated at a table, with their phone-handling (right) hand partially resting on the table for stability

b) *Stand*: free standing without any stability support

c) *Walk*: walking around a closed outdoors course at a constant, researcher defined comfortable walking speed

d) *Bus*: seated on a moving bus

Each trial highlighted one out of the 20 targets available at random with a uniform distribution. To simulate realistic, continuous interaction, the application moved to the next trial as soon as the participant touched the screen, without the enforced delay of the first study. The study was designed to be 1 hour in duration, which allowed 200 trials per condition. During the Bus scenario, data was only collected while the bus was moving – participants were asked to pause while the vehicle was stationary.

Participants and Apparatus

Twelve new participants (mean age 26.1, SD=2.2, 7 males), all right handed each completed 200 trials for each of the four conditions (9600 touch trials in all) using a one-handed posture. All participants were regular, frequent mobile users.

The custom experiment application ran on the same Nokia Lumia 920 device as in the first study. The application collected 2D touch coordinates for each trial, and 3-axis accelerometer and gyroscope readings sampled at 50Hz.

Results – In the Air Prediction on the Move

To verify whether target prediction is possible in different physical conditions, we again trained user-specific GP regression models mapping mobile motion sensor sequences to touch coordinates. Figure 3 shows touch prediction results using 10-fold cross validation for each condition, with the baseline model – always guessing in the center of the application area – shown in red. Note that the baseline error is higher (29mm) than the first experiment due to the static set of available targets to select from. As in Study 1, the prediction models created in the stationary conditions (Sit and Stand) have a lower RMSE than the baseline model. At 0.1s before touch, the GP regression model created in the Sit condition had a RMSE of 23.7mm (with Stand at 22.9mm), still lower than the baseline model (29.0mm).

Specifically comparing the results between the two studies shows that Study 1's result was better; it produced an RMSE of 20.3mm at 0.1s before touch, whereas in Study 2 for Sit it was 23.7mm. This can be explained by our more conservative approach in Study 2. We used a more realistic task and did not enforce a delay between trials. At times, more than one target was clicked within the 500ms prior to touchdown, which resulted in a lack of clear segmentation between touches. We also had a more limited data set (200 trials per user, per condition). Altogether, prediction RMSE in the Sit condition is thus higher than the first study. This more realistic scenario shows how in the air prediction using grip change degrades in a freeform homescreen task.

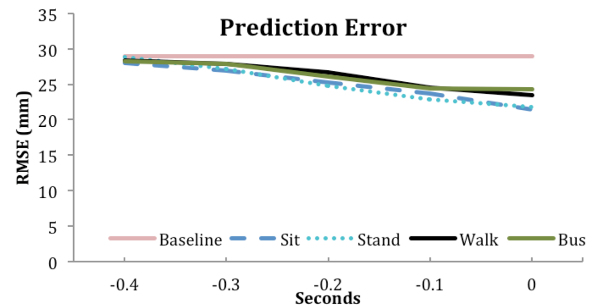

Figure 3. Root mean square error for the predicted touch points in Study 2's four different physical conditions.

Lastly, unpredictable motion such as interacting in a bus, leads to higher prediction RMSE (24.4mm at 0.1s before touch) than sitting at a table (23.7mm at 0.1s), though critically, still lower than baseline models (29.0mm). This shows that, even with a relatively small amount of prior data, touch prediction using hand grip change is viable in high motion, unpredictable physical conditions.

Results – On Touchdown Prediction of Touch Intent

The secondary goal of the study was to verify whether grip change can improve touch performance by providing insight into targeting intent at the point of touchdown.

In order to separate user intent from their inherent touch patterns, we first built a baseline regression model mapping the 2D touch coordinates to the center of the highlighted target (referred to as the *XY Model*). We use the center of the highlighted target as an approximation of where the user intended to touch. This model takes into consideration individual touch behavior but uses no other information other than touchdown to infer user intent.

Next, we built a GP regression model that uses where the user touched, and the grip change as measured by the motion sensor readings of the 500ms prior to touchdown (*XY+Grip Model*). Again, the model attempts to map these features to the center of the highlighted target. We test the two models against the raw touch points (*Raw*). The models' goal is to improve on Raw accuracy by inferring where the user intended to touch rather than where they actually touched. To build and test the models, we use 10-fold cross validation.

An analysis of variance with *Model* (XY, XY+Grip, Raw) and *Physical Condition* (Sit, Stand, Walk, and Bus) as within-subject factors found a significant main effect of *Model* on touch accuracy ($F_{2,22} = 25.162$, $p < 0.001$). Bonferroni correction showed significant differences between all pairs of *Models*. As Figure 4 shows, participants had a Raw accuracy of 83% over all physical conditions. The

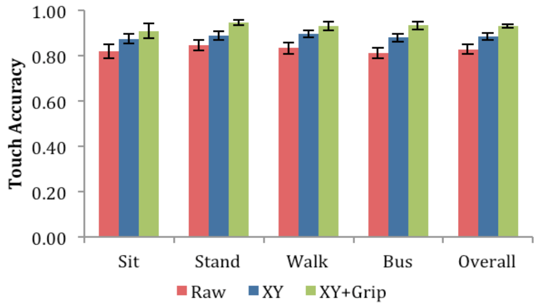

Figure 4. Touch accuracy comparing the three models per physical condition; Overall aggregates across all the data.

XY Model which took into consideration user-specific touch behaviour improved accuracy to 88%. Our XY+Grip Model further boosted accuracy to 93% consistently over all physical conditions—a 41% reduction in error rates over the XY Model alone.

CONCLUSIONS AND FUTURE WORK

In contrast to more static models that take into consideration individual touch behaviour (e.g. hand posture), we consider grip change as a continuous information stream, useful at multiple points in the interaction sequence. Prior to touchdown, grip change detected using standard mobile motion sensors can help estimate a touchdown point, useful for a virtual hover space and providing interactive support (e.g. continuous feedback of a user's touchdown point). On touchdown, grip change that preceded a touch point can be a useful predictor of where users intended to land, rather than where they touched. Although performance degrades in high motion scenarios, our results show that grip change can make reasonable touch predictions and reliably improves touch accuracy, reducing error rates by 41%.

Our study trained a model to infer intent by having full knowledge of the intended target. A question remains of how well such models can be trained and perform during everyday use where the intended target may not be known; some touches will hit an actual target, but one that was not intended. While our second study did not filter out any movement+touch sequence, more work is needed to verify the impact of such accidental touches in real world use cases. Additionally, our models only considered movement data that ends with a touch. More work is needed to filter out false positive movements and to design interactions that will fail gracefully (i.e., never permanently modify the interface state based on an expected touchdown).

Furthermore, our models were evaluated offline, and it remains to be seen how users adapt their touches when encountering a system that is itself adapting based on inferred intent (i.e., online). Continuous visual feedback of the model's predicted touch point may be critical to help users correct their targeting motion.

Finally, grip change was detected through data gathered on a relatively large smartphone and with a single thumb hand posture. While the current trend is one of increasing display sizes, more work is needed to verify how grip change varies for different phone-to-hand ratios, and whether smaller devices and other common hand postures have detectable and useful grip variations Past work on automatically detected hand posture using internal sensors (e.g. [4,5]) may be critical pre-processing for adaptive grip change models.

In conclusion, our studies show that grip change is a promising information side channel that is detectable with internal motion sensors, even when there is movement due to physical activity. Models using mobile sensor-detected grip change make useful in the air predictions and may complement techniques that use auxiliary sensor arrays to create more complex interactions. Additionally, the grip change measured right before touchdown provides insight into where the user intended to land. Leveraging this information significantly improves touch accuracy beyond models that adjust based solely on the touchdown point.

ACKNOWLEDGEMENTS

This work was funded by NSERC and the GRAND NCE.

REFERENCES

1. Azenkot, S. and Zhai, S. Touch behavior with different postures on soft smartphone keyboards. *Proc. of MobileHCI'12*, 2012, 251–260.
2. Bergstrom-Lehtovirta, J. and Oulasvirta, A. Modeling the Functional Area of the Thumb on Mobile Touchscreen Surfaces. *Proc. of CHI'14* ACM (2014), 1991–2000.
3. Goel, M., Findlater, L., and Wobbrock, J. WalkType: using accelerometer data to accomodate situational impairments in mobile touch screen text entry. *Proc. of CHI'12*, ACM (2012), 2687–2696.
4. Goel, M., Jansen, A., Mandel, T., Patel, S.N., and Wobbrock, J.O. ContextType: Using Hand Posture Information to Improve Mobile Touch Screen Text Entry. *Proc. of CHI'13*, ACM (2013), 2795–2798.
5. Goel, M., Wobbrock, J., and Patel, S. GripSense: Using Built-in Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones. *Proc. of UIST'12*, 2012, 545–554.
6. Henze, N., Rukzio, E., and Boll, S. 100,000,000 taps: analysis and improvement of touch performance in the large. *Proc. of MobileHCI'11*, ACM (2011), 133–142.
7. Mohd Noor, M.F., Ramsay, A., Hughes, S., Rogers, S., Williamson, J., and Murray-Smith, R. 28 Frames Later: Predicting Screen Touches from Back-of-device Grip Changes. *Proc. of CHI'14*, ACM (2014), 2005–2008.
8. Ng, A., Brewster, S.A., and Williamson, J.H. Investigating the Effects of Encumbrance on One- and Two- Handed Interactions with Mobile Devices. *Proc. of CHI'14*, ACM (2014), 1981–1990.
9. Wimmer, R. and Boring, S. HandSense: Discriminating Different Ways of Grasping and Holding a Tangible User Interface. *Proc. of TEI'09*, ACM (2009), 359–362.
10. Yin, Y., Ouyang, T.Y., Partridge, K., and Zhai, S. Making touchscreen keyboards adaptive to keys, hand postures, and individuals: a hierarchical spatial backoff model approach. *Proc. of CHI'13*, ACM (2013), 2775–2784.