CASA: Context-Aware Scalable Authentication

Eiji Hayashi[1] **Sauvik Das**[1] **Shahriyar Amini**[1]
Jason Hong[1] **Ian Oakley**[2]

[1]Carnegie Mellon University
5000 Forbes,
Pittsburgh PA, 15213, USA

[2]University of Madeira
Funchal,
9000-390, Portugal

ehayashi@cs.cmu.edu

ABSTRACT
We introduce context-aware scalable authentication (CASA) as a way of balancing security and usability for authentication. Our core idea is to choose an appropriate form of active authentication (e.g., typing a PIN) based on the combination of multiple passive factors (e.g., a user's current location) for authentication. We provide a probabilistic framework for dynamically selecting an active authentication scheme that satisfies a specified security requirement given passive factors. We also present the results of three user studies evaluating the feasibility and users' receptiveness of our concept. Our results suggest that location data has good potential as a passive factor, and that users can reduce up to 68% of active authentications when using an implementation of CASA, compared to always using fixed active authentication. Furthermore, our participants, including those who do not using any security mechanisms on their phones, were very positive about CASA and amenable to using it on their phones.

Categories and Subject Descriptors
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Security; Human Factors.

Keywords
User Authentication; Context-Aware; Mobile;

1. INTRODUCTION
Reliable authentication is an essential requirement for secure systems. Today, passwords are the most common form of authentication. However, passwords are also a major source of vulnerabilities, as they are often easy to guess, re-used, forgotten, shared with others, and susceptible to social engineering [7,19,20,22]. We argue that the commoditization of sensor technologies coupled with advances in modeling people and places offers new opportunities for both simplifying and strengthening authentication. This insight is the basis for what we call *context-aware scalable authentication*, or CASA.

CASA embodies two concepts. First, these cheap digital sensors combined with models of people and places can yield multiple *passive* factors about users' identities. For our specific context, we define a factor as any data that provides information about a user's identity. Passive factors are those that can be acquired without explicit interaction from the end-user (e.g., a user's location or time since last login). In contrast, active factors require explicit interaction (e.g., entering a PIN or scanning fingerprints).

Second, CASA is based on the idea that this passive multi-factor data can be used to *modulate the strength of active authentication needed* to achieve a given level of security. For example, with CASA, we want to enable quick and easy active factors in situations where passive factors indicate a high probability the user is a legitimate user (for instance, being located in home or work where only the user and a small number of trusted people can access). Conversely, we want active factors to be tough and reliable in situations where the passive factors indicate a low probability (such as being located in an unfamiliar place).

In this approach, CASA breaks current underlying assumptions about authentication, by making authentication easier or harder based on passive factors rather than making it uniformly hard for all cases. We argue that today's authentication systems are designed to ensure security in extreme cases; consequently, they overlook common, mundane and ultimately *average* case scenarios that characterize most user authentications. Some people have argued that this conventional approach of always having more security actually leads to less compliance and less security overall (see for example, [7,31]). In particular, Norman argues that "[t]he more secure you make something, the less secure it becomes. Why? Because when security gets in the way, sensible, well-meaning, dedicated people develop hacks and workarounds that defeat the security" [6]. Norman's predictions appear to be well founded in statistics about mobile phone PINs usage. A survey in 2007 found that 61% of people had no PIN on their phones [5].

CASA targets this large population of users who do not secure their devices by attempting to derive solutions that offer them a more appropriate perceived balance between usability and security. By exploring solutions that provide easy access in commonplace everyday situations, such as whilst a user is at home, but require more secure authentication in less common scenarios, this paper points the way towards how to lower the overall burden of having user authentication on mobile devices to increase the compliance rate.

Towards this end, this paper makes several research contributions. First, we present a general Bayesian framework that allows us to choose active factors given passive factors. Second, we examine the feasibility of using location as one possible passive factor, presenting the results of a field study that analyzes users' mobility patterns along with their phone usage patterns. Third, we describe the results of two field studies where we iteratively designed,

Symposium on Usable Privacy and Security (SOUPS) 2013, July 24–26, 2013, Newcastle, UK.

developed and evaluated one implementation of CASA. We believe that our Bayesian framework can be applied to many different authentication systems that use both active and passive factors. We also believe that the data and the process described in this paper will help practitioners and researchers working on user authentication to design their systems.

2. Related work

Existing user authentication systems primarily depend on three types of mechanisms: *what you know*, *what you have* and *what you are*. Passwords are the most commonly used authentication system based on *what you know*. Password authentication has advantages in its simplicity and convenience [19]. However, many studies have found that passwords still place substantial burdens on users, resulting in users adopting insecure practices such as choosing weak passwords or reusing passwords [7,22].

Authentication systems based on *what you have,* include eToken USB devices [1], RSA securID [2], and Google's two-step verification [3]. Such techniques are only partially related to the work in this paper as it involves the user carrying a smartphone, but this is used solely to gather data rather than being used as a token. Finally, authentication systems leveraging *what you are*, or biometrics have been widely commercialized. Common examples uses fingerprint, iris, voice, and face. These kinds of biometrics tend to focus on *physical characteristics* of individuals. In contrast, researchers have also investigated a number of behavioral biometric techniques focusing on individuals' behavioral patterns, e.g., walking gait [29], keyboard typing pattern [30], what applications and features on a mobile phone are being used [23].

This paper is closely related to behavioral biometrics but differs in two key ways. First, we seek to use commodity devices as well as sensors that already exist on many computers today. In particular, this paper examines the potential for using location as a factor modulating authentication on smart phones. Second, we seek to understand how to use passive factors to adjust the level of active authentication to satisfy security requirements rather than solely using the behavioral patterns for authentication.

2.1 Modulating the Level of Authentication

Some online services already modulate authentication level in specific circumstances. For example, many bank web sites ask extra questions when logging in from new network IP addresses. Similarly, Facebook asks additional questions when using an unusual IP address [4]. In this technique, users must identify several of their friends' photos before being allowed to login.

CASA differs from these techniques in that it seeks to expand and extend (and ultimately generalize) the factors used when adjusting authentication level. Our current work also focuses on authenticating primarily with a device rather than an online service, as there are serious privacy issues with storing behavioral data on multiple online services. We do believe that CASA can be used to simplify authentication with remote online services, but we consider this beyond the scope of the current paper.

There are also theoretical investigations of risk-based access control models. These models focusing on handling uncertainties and risks in making access control decisions (e.g., [12,21]). CASA differs from these investigations in that it focuses on passive factors collected by sensors on mobile devices and modulating active actors based on the passive factors.

2.2 Leveraging Contexts

Several systems have discussed or used contextual information for security [e.g. 15]. For example, proximity has been used both to authenticate users [9,13,21] and to perform pairing operations [24]. Similarly, Seifert et al. proposed TreasurePhones a system that protected information on mobile phones based on a user's location as detected by near field communication technology [33]. Buthpitiya et al. demonstrated that a system could detect anomalous activities (e.g., a phone being stolen) by analyzing a user's location history using an n-gram model [11]. Gupta et al, proposed a context profiler that classifies point of interest into safe, unsafe, and uncategorized places, and evaluated the model with data collected in Lausanne Data Collection Campaign [17]. Riva et al. proposed Progressive Authentication that combined light, temperature/humidity sensor, touch screen, login events, microphone, and Bluetooth receiver using SVM model to authenticate users [28].

The prior work closest to CASA is Jakobsson et al.'s discussion of implicit authentication [23]. Their core idea was to utilize user behavior patterns for authentication. They considered two behavioral features derived from a mobile device: time since the user last checked email and GPS location. The two feature scores were combined through a weighted linear function to calculate an overall "authentication score", which was then compared with a pre-defined threshold to determine whether a user should be authenticated. A fundamental difference between this work and our work is that, while implicit authentication accepts or rejects users only based on behavioral features, our work modulates active authentication based on passive factors, including behavioral features.

CASA has similar goals with these prior works; however, it differs from these past explorations in three important ways. First, we study selecting appropriate active authentication given passive factors rather than authenticating users based on passive factors. Second, our model considers the differences between a user and others while existing works are focused on the user's behavior patterns. For instance, being at a favorite café does not provide strong information about the user's identity because there are many other people who visit the café; however, being at work (a location where we assume a small number of people have access) has increased weight. Finally, in contrast to more theoretical approaches in prior work, this paper presents three field studies with a total of 86 users. This empirical data informs practical system viability and sheds light on users' perspectives.

2.3 Human Mobility Analysis

One of our working assumptions with CASA (which we examine in the first evaluation) is that users' locations can be useful as a passive factor. Prior work has found that there are many predictable patterns in people's mobility patterns [16,26]. For example, Gonzalez et al. [18] analyzed the mobile phone cell tower data of 100,000 people over six months (based on call log and SMS log data). They found that people spent a great deal of time in just a few highly frequented locations. Hayashi et al. [19] presented the results of a diary study investigating where people login to desktop and laptop computers. They found that 84.3% of logins took place at home (59.2%) and work (25.1%).

Combined, this past work suggests that location data may be promising, for two reasons. First, strong, predictable patterns in one's mobility patterns would make location data very useful as a passive factor for authentication. Second, if people primarily use

their mobile devices in just a few places (e.g., home or work), then streamlining the level of authentication for just those places should improve usability. If these places also have reasonably good physical security, then we can improve usability without making significant tradeoffs for security. However, currently, there is little empirical data on where and how often people actually use their smartphones (as opposed to computers). This paper also contributes to this body of knowledge by providing analyses of where people actually use smartphones, using GPS and Wi-Fi data for fine-grained information with ground truth.

3. Framework for active factor selection

In this section, we introduce our probabilistic framework for choosing an active factor given passive factors. Our approach is to use a Naïve Bayes classifier to combine multiple factors, calculating a "risk assessment" value to determine the appropriate level of active authentication required given passive factors.

Most existing user authentication schemes can be considered binary classifiers, classifying a person as a legitimate user ($\hat{u} = 1$) or not ($\hat{u} = -1$). We can also model these schemes probabilistically as shown in Eq. (1) where $\hat{u}$ denotes the prediction (i.e., the result of the user authentication), $P(u = 1|s)$ denotes the probability the requester is the legitimate given the observation *s*, $P(u = -1|s)$ denotes the probability the person is not the legitimate user given the observation *s*, and α denotes the degree to which user authentication is conservative. The α parameter can be set based on one's comfort level with expected costs of false accepts and false rejects.

$$\hat{u} = \begin{cases} 1, & \alpha P(u = 1|s) > P(u = -1|s) \\ -1, & \alpha P(u = 1|s) \leq P(u = -1|s) \end{cases} \quad (1)$$

For instance, for PIN-based authentication, if the system observes that a requester enters the correct PIN, the probability that the requester is legitimate is much higher than the probability he is not. Thus, the system predicts $\hat{u} = 1$ and authenticates the user. Conversely, the system predicts the opposite if the requester enters a wrong PIN.

Many current authentication schemes focus on a single factor that has large differences between the probability distributions of $P(u = -1|s)$ and $P(u = 1|s)$ across the range of values of *s*. In contrast, CASA combines multiple factors that may or may not have as pronounced of a difference between the probability distributions of $P(u = -1|s)$ and $P(u = 1|s)$, but taken together offer benefits over a single factor approach.

In Eq. (2), we show the underlying probabilistic model of multi-factor authenticators such as CASA. Again, *u* denotes whether a user is legitimate (*u*=1) or not (*u*=-1), and s_i denotes the observation value for the *i*-th factor.

$$\hat{u} = \begin{cases} 1, & \alpha P(u = 1|s_1, \dots, s_n) > P(u = -1|s_1, \dots, s_n) \\ -1, & \alpha P(u = 1|s_1, \dots, s_n) \leq P(u = -1|s_1, \dots, s_n) \end{cases} \quad (2)$$

We can reformulate Eq. (2) into Eq. (3) using the sign function, which extracts the sign (positive or negative) of a real number.

$$\hat{u} = \text{sign}\left(\log \frac{\alpha P(u = 1|s_1, \dots, s_n)}{P(u = -1|s_1, \dots, s_n)}\right) \quad (3)$$

Using Bayes' theorem, $P(u|s_1, s_2, \dots, s_n)$ can be reformulated into Eq. (4). Eq. (4) has the term, $P(s_1, s_2, \dots, s_n|u)$, that depends on all the factors simultaneously. In practice, estimating this term is challenging because the number of possible combinations of $(s_1, s_2, \dots, s_n)$ increases exponentially when the number of signals increases. Therefore we simplify Eq. (4) as Eq. (5) by assuming conditional independence between each identifier. This is a standard transformation in building Naïve Bayes classifiers. This simplification allows us to deal with each signal separately. In Eq. (5), $P(u)$ denotes a prior probability of how likely a person is a legitimate user (or not) in general. $P(u)$ will be canceled in the following reformulations.

$$P(u|s_1, s_2, \dots, s_n) = \frac{P(s_1, s_2, \dots, s_n|u)P(u)}{P(s_1, s_2, \dots, s_n)} \quad (4)$$

$$= \frac{\prod_{i=1}^{n} P(s_i|u)\, P(u)}{P(s_1, s_2, \dots, s_n)} \quad (5)$$

Finally, by substituting $P(u|s_1, s_2, \dots, s_n)$ in Eq. (3) with Eq. (5), we obtain a Naïve Bayes classifier (Eq. (6)). Intuitively, the parameter in the sign function increases with the probability that a requester is legitimate and vice versa.

$$\hat{u} = \text{sign}\left[\log\left(\alpha \frac{P(u = 1)}{P(u = -1)}\right) + \sum_{i=1}^{n} \log \frac{P(s_i|u = 1)}{P(s_i|u = -1)}\right] \quad (6)$$

Note that because each factor might not be conditionally independent, Eq. (6) may have approximation errors compared to Eq. (3). However, in practice, we believe the errors will be limited because we can choose largely independent factors (e.g. voice and PIN). Further, in Eq. (6), we can discuss each factor independently by estimating $P(s_i|u = 1)/P(s_i|u = -1)$. Thus, we believe the benefit of the independence assumption outweighs its drawbacks.

3.1 Selecting an Active Factor

CASA uses this probabilistic model to select an active factor that provides enough evidence to authenticate a user, given a set of passive factors. The model allows us to compare the strength of the evidence using the terms in the sign function in Eq. (6).

We describe one example here to illustrate how we can utilize the framework in choosing active factors. Let's assume we want to choose an active identifier *S* that provides as much evidence when a user is at a café as compared to the user typing her correct PIN at her home. Assuming that location is the only passive factor, the condition that S should satisfy can be written as Eq. (7). The first term in Eq. (6) is canceled. $P_{s,1}(1)$ denotes the probability that the active factor *S* indicates that a person is the legitimate user when a person is actually a legitimate user. $P_{s,-1}(1)$ denotes the same when a person is not the legitimate user. $P_{L,1}(l)$ (or $P_{L,-1}(l)$) denotes the probability the person is at the location *l* when she is the legitimate user (or not). *H* and *C* denote *home* and *café* respectively.

$$\log \frac{P_{S,1}(1)}{P_{S,-1}(1)} + \log \frac{P_{L,1}(C)}{P_{L,-1}(C)} \geq \log \frac{P_{PIN,1}(1)}{P_{PIN,-1}(1)} + \log \frac{P_{L,1}(H)}{P_{L,-1}(H)} \quad (7)$$

Eq. (7) can be rewritten as Eq. (8), which quantifies the security criteria that an active factor *S* should satisfy to have the same level of security as the legitimate user typing her PIN at home, given that the active factor *S* authenticates the person at a café.

$$\log \frac{P_{S,1}(1)}{P_{S,-1}(1)} \geq \log \frac{P_{PIN,1}(1)}{P_{PIN,-1}(1)} + \log \frac{P_{L,1}(H)}{P_{L,-1}(H)} - \log \frac{P_{L,1}(C)}{P_{L,-1}(C)}$$

$$= \log \frac{P_{PIN,1}(1)}{P_{PIN,-1}(1)} + \log \frac{P_{L,1}(H)}{P_{L,1}(C)} \frac{P_{L,-1}(C)}{P_{L,-1}(H)} \quad (8)$$

A legitimate user is more likely to be at her home than to be at café. Thus, $P_{L,1}(H)/P_{L,1}(C) > 1$. In contrast, someone else is

much more likely to be at the café than to be at the user's home, i.e., $P_{L,-1}(C)/P_{L,-1}(H) \gg 1$. Thus, the second term on the right side is positive. Therefore, Eq. (8) indicates that the active factor should provide greater confidence than a standard PIN.

Furthermore, Eq. (8) offers a quantitative guideline for the strength of *S* given the user's location. Our model can also include other passive factors, such as sensor data, time since last login, or number of times logged in at given places. We describe another example of selecting an active factor in our second field study.

4. Empirical Evaluations

To assess the feasibility of CASA, we conducted three different empirical evaluations. In our first evaluation, we investigated the potential of using location as a passive factor. Past work suggests that people spent most of their time in a few locations [8,18,26]. However, there is little empirical data on how frequently people used their smart phones at these locations. We collected this information to evaluate the usefulness of location information for CASA.

In our second study, we conducted a one-week field study of a prototype with 32 participants. This prototype modulated active factors based on their locations. This study helped us understand how well our ideas might work in practice, as well as to obtain feedback from participants.

In our third study, we iterated on both the system design and the study design based on the results of the second study. We conducted a 10-day field study with 18 participants. This prototype took into account location as well as whether the participants used their computers nearby recently.

5. Study #1: Mobility pattern analysis

In this study, we investigated people's mobility patterns along with their phone usage patterns, to evaluate the effectiveness of location information as a passive factor. We recruited multiple Android phone users through Craigslist and e-mails. Participants were asked to install our logging app from the Android Market. Participants were enrolled in a raffle for $50 Amazon gift cards as compensation. Over five months, we collected data from 128 participants. In this analysis, we focused on 36 participants with at least seven days of logs.

5.1 Data Collection

Our app sampled location every three minutes regardless of whether participants were interacting with their phones. Location was obtained through standard Android APIs using Wi-Fi and cell tower information. The standard API also provided the expected error for each location estimate. We discarded location data when the expected errors were greater than 200 meters. Our app also logged the smartphone's running processes every 30 seconds when the smartphone was *not* in sleep mode. The timestamps of these logs let us infer when participants used their phones.

We analyzed location traces from 36 participants. The data collection periods varied from seven days to 140 days. The median length of the data collection was 26.5 days. We divided the latitude and longitude space into discrete 0.002 × 0.002 latitude/longitude grids (each cell was approximately 200 × 200 meters in/near North America) as previously done in [14]. The particular choice of discretization was based on practical considerations balancing the accuracy of Android's positioning system with granularity of the analysis.

Table 1. The distribution of the time spent and the phone activation events at the places where participants spent most of their time. Place 1 to 5 denote the places where participants spent most time (1) to fifth most time (5).

Place	Time		Activations	
	Mean [%]	SD [%]	Mean [%]	SD [%]
1 (Home)	38.9	20.2	31.9	15.6
2 (Workplace)	18.7	12.6	28.9	18.1
3	9.9	8.4	18.5	13.7
4	5.5	4.8	10.8	8.5
5	4.3	4.7	5.2	4.7
Other places	22.6	13.1	4.5	4.6

5.2 Identifying Phone Activation

To track phone use, our app ran a low-level thread that logged active processes every 30 seconds. When the phone was in sleep mode, the thread was automatically paused. Thus, by examining the timestamps of log entries, the phone state could be determined.

Theoretically, intervals between log entries that exceed 30 seconds signified a phone activation event after being in sleep mode once. However, initial trials of this log analysis identified two common sources of error. The first issue was the low priority of the logging thread leading to fluctuations in the sequentially logged times - variations typically in the region of 5 seconds. To deal with this, we considered valid differences between log time stamps to be in the range 30-35 seconds. The second issue was phone activations caused by push notifications (e.g. email arrival). We adopted a conservative approach to mitigate false positives relating to this issue. Essentially, phone activation events were counted only when there were two successive log timestamps after observing at least a 35 second gap. This filtered out short phone activations due to push notifications because the phone would quickly go back to sleep mode after an automatic activation. A consequence of these manipulations was that a certain proportion of valid user activations (e.g. very brief glances and interactions) would not be counted. However, despite this cost, we believe that these manipulations ensured the validity of the study by counting only real user activations of their phones.

5.3 Mobility Pattern Analysis

We identified 55840 phone activation events in our dataset. Participants activated their phones 27.4 times a day on average (SD=19.7). Table 1 shows the distribution of time spent and logins at the places where participants spent most of their time. We first calculated each participant's top five places based on the amount of time spent using location data alone (see the two columns under "Time"). Then, for each participant, we calculated the number of phone activations at each of these places using location data and process data (see the two columns under "Activations").

The results indicated that people spent 57.8% of their time at two locations, which we assume are home and work. This result is aligned with past work investigating people's mobility patterns e.g., [18,19]. However, before conducting this study, it was unclear to us how often people would use their smartphones at home and work, since there would be other devices with network connectivity and larger displays (e.g., desktop and/or laptop computers) at these locations. Nevertheless, our results showed that these top two places accounted for 60.8% of the total phone activation events on average (SD=14.5%). This data indicates that people activate their phones more frequently at their homes and workplaces than at other places.

This result provides supporting evidence that people exhibit strong patterns in where they use their smartphones, suggesting that location could be a very useful passive factor. This result also indicates that we can positively impact both usability and security if we adjust the active factor based on location data coupled with a very trivial model (home, workplace and other places). Again, this approach makes the assumption that a person's home and workplace have reasonably good physical security, and that there are relatively few trusted people that can access those locations.

6. Study #2: Evaluation of CASA prototype

In this field study, we developed and deployed a prototype using CASA framework for Android smartphones. This prototype dynamically selected active factors based on participants' location (i.e., whether they are at home, workplace, or some other places). In this study, we investigated users' reactions to dynamically changing active factors. We also collected empirical data to estimate how much effort our participants could reduce in user authentication when using our prototype. These data help us to understand the design space opened by the concept of CASA, and to improve the prototype for the next design iteration to make it better fit users' needs.

6.1 Participants

We recruited 32 participants using a participant recruitment website at Carnegie Mellon University. Their age ranged from 18 to 40 years old with a mean age of 24. Our participants consisted of 26 students, five full-employed and one non-employed. Twenty-three out of 32 participants were living with others in their homes. We compensated participants $40 for their participation in the study.

Participants were assigned to one of two conditions based on whether they used any security lock on their phones prior to this study. Nineteen participants *not* using a security lock (i.e., PIN or Android Pattern Lock) were assigned to the *PIN* condition. Thirteen participants already using a security lock were assigned to the *password* condition. None of the participants were using passwords to secure their phones. In essence, participants used the same authentication they already used at home and work, and had stronger active authentication at other places.

6.2 Procedure

In the first session, we installed our prototype on participants' Android phones. We asked participants in the PIN condition to choose a PIN. For participants in the password condition, we asked them to choose a password in addition to a PIN.

During the study period, when the participants turned on their phone displays, our prototype selected an active factor based on the participant's location (home, work, and other) and condition (the PIN or password condition) (see Table 2). In the explanation, we explicitly defined "work" as a room or building where the participants spent most of time except home. For instance, for students, "work" means their offices or campus buildings. After participants authenticated, the prototype asked the participants to answer if they were at home, work, or other place (Figure 2 (c)). The answers were used to train the location classifier implemented in the prototype. This classifier is trained on the fly during the study using the ground truth. After one week, we had the second session where we asked participants to complete a post-survey, and conducted a follow up interview that lasted about 15 minutes.

6.3 Prototype with Active Factor Selection

Our prototype used location as a passive factor and selected an active factor from three options: no active factor, a PIN, and a password. First, we describe how we can use CASA in selecting active factors, using the password condition as an example. The participants in the password condition were using PIN or Android Pattern Lock to secure their phones prior to this study. Thus, for the participants in the password condition, we selected active factors so that they would provide the same level of evidence as typing a PIN at workplace, where risks are higher than home, but still lower than other places.

Because location is the only passive factor in our prototype, Eq. (6) can be simplified to Eq. (9) and (10). These equations denote the conditions that active factors should satisfy to provide no less evidence than being at home (Eq. (9)) or at a place other than home and workplace (Eq. (10)), where W, H and O_i denotes *workplace*, *home,* and a place *other* than home and workplace respectively. Note that O_i does not denote the aggregation of places other than home and workplace, but it denotes a single place. $f(l_1, l_2)$ and g(S) are defined as shown in Eq. (11). Intuitively, $\log(f(l_1, l_2))$ means the likelihood that a person is a legitimate user when she is at l_2 compared to when she is at l_1. If it is less likely, $\log(f(l_1, l_2))$becomes positive. Then, the evidence provided by the active factor (the term on the left side) should be greater than that of PIN. If it is more likely, $\log(f(l_1, l_2))$becomes negative. Then, the active factor could be weaker than PIN. As $|\log(f(l_1, l_2))|$ increases, the user's location provides stronger evidence towards authentication. g(S) denotes how strongly an active factor S indicates users' identities.

$$\log g(S) \geq \log g(PIN) + \log(f(W,H)) \quad (9)$$

$$\log g(S) \geq \log g(PIN) + \log(f(W,O_i)) \quad (10)$$

$$f(l_1,l_2) = \frac{P_{L,1}(l_1)}{P_{L,1}(l_2)}\frac{P_{L,-1}(l_2)}{P_{L,-1}(l_1)}, g(S) = \frac{P_{S,1}(1)}{P_{S,-1}(1)} \quad (11)$$

We estimated g(S) based on the entropy of four-digits PINs (~9 bits) and passwords (~18 bits) according to the estimations by NIST [10]. Assuming that the authentication system allows three trials and that a legitimate user always types a PIN and a password correctly, then we have $P_{PIN,1}(1)=1$, $P_{PIN,-1}(1)=3/2^9$, $P_{Pwd,1}(1)=1$, $P_{Pwd,-1}(1)=3/2^{18}$ and $P_{None,1}(1) = P_{None,-1}(1) = 1$. Thus, $g(PIN) = 2^9/3$, $g(Pwd) = 2^{18}/3$ and $g(None) = 1$.

To calculate $\log(f(W,H))$ and $\log(f(W,O_i))$ accurately, we need further empirical data collection. However, because our primary purpose in this study was to investigate participants' responses to our concept rather than applying CASA precisely, we approximated these values. We approximate the values in a way so that $|\log(f(l_1, l_2))|$ becomes smaller to avoid overestimating the strength of the evidence provided by location information. We

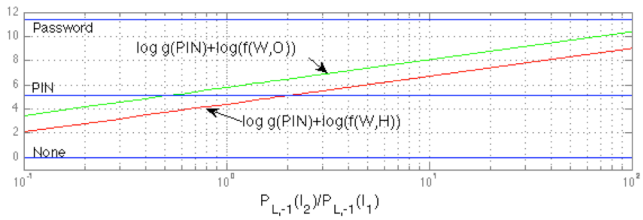

Figure 1. Graphical representations of Eq. (9) and (10). Horizontal line denotes log(g(s)) for each S (None, PIN, or password).

Table 2. Active factors required at different locations in the second study. The prototype required the same active factors as participants were using at their homes and workplaces while required stronger active factors at other places.

Condition	Home	Workplace	Other places
PIN	None	None	PIN
Password	PIN	PIN	Password

(a) PIN **(b) Password** **(c) Questionnaire**

Figure 2. Prototype screenshot. Based on users' locations and the conditions (see Table 2), the prototype either skips authentication, (a) requests PIN, or (b) requests password. After authentication, the prototype showed a questionnaire to obtain ground truth of locations.

discuss the data collection issue more in the discussion section.

For $P_{L,1}(H)$ and $P_{L,1}(W)$, we used 0.389 and 0.187 that were obtained in the first study. For $P_{L,1}(O_i)$, we used 0.099, which was the highest probability among the places other than home and workplace in the first study (Table 1). When $P_{L,1}(O_i)$ becomes higher, the CASA model estimates the evidence provided by being at home and the workplace to be lower. Thus, we used the highest value for all O_i to be conservative. Additionally, we assumed that $P_{L,-1}(l)$ was proportional to the number of people who can physically come into the location. Because we do not have empirical data about $P_{L,-1}(l)$, we make assumptions after showing its effect on the active factor selection.

Figure 1 is a graphical representation of Eq (9) and (10). The diagonal plots show the right sides of the Eq. (9) and (10), and the horizontal lines denotes $\log g(S)$ for each factor (i.e., S=None, PIN or password). Intuitively, the X-axis denotes how many people can access certain locations (home for the red plot and other places for the green plot) compared to the number of people who can access workplaces. The Y-axis denotes confidence about users' identities. When we only consider active factors, the confidences are not relevant to numbers of people who can access certain locations. Thus, the blue plots become horizontal. In contrast, when we consider locations as indicators of users' identities, the confidences become dependent on the likelihood. Thus, the plots become diagonal as shown by the red and green plots.

In Figure 1, satisfying Eq. (9) is equivalent to the condition that the lower diagonal plot is below one of the horizontal lines at given $P_{L,-1}(l_2)/ P_{L,-1}(l_1)$. We assume that the number of people who can access *home* is less than that of *workplace* and more than 1/10 of that of workplace. The lower diagonal plot in the segment $P_{L,-1}(l_2)/ P_{L,-1}(l_1)$ =[0.1, 1] is between the horizontal lines representing *PIN* and *None* under this assumption. Therefore, we select PIN as an active factor that satisfies Eq. (9). Similarly, we assume that the number of people who can access *other places* is more than that of *workplace* and less than 100 times of that of workplace, the upper diagonal plot in the segment $P_{L,-1}(l_2)/ P_{L,-1}(l_1)$ = [1, 100] is between the horizontal lines representing *Password* and *PIN*. Therefore, we select passwords as an active factor that satisfies Eq. (10).

Table 3. The means of the phone activation frequency per day at each location. The numbers in parentheses denote standard deviations. Both the PIN and the password condition activated phones more than 50% of time at homes or workplaces.

Condition	Home	Workplace	Other places
PIN	13.1 (1.4)	2.5 (0.4)	8.1 (1.1)
Password	24.5 (3.2)	7.1 (1.0)	15.7 (2.0)

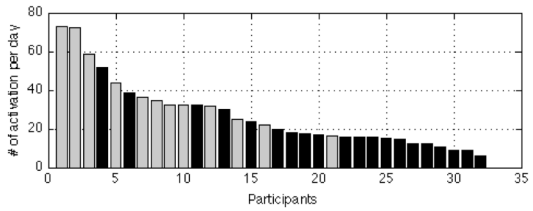

Figure 3. The number of phone activations per day. Gray and black bars denote participants in the PIN-password condition and none-PIN condition respectively.

We made two assumptions above; however, we believe that these assumptions are safe to make considering the ranges. Additionally, our choice of active factors (Table 2) made active authentication more secure than that used by our participants prior to the study. Our prototype required the same active factors as they used prior to this study at their homes and workplaces, and required more secure active factors at other places. Thus, we made the authentication more secure for our participants, compared to pre-study levels.

6.4 User Interfaces

Our prototype estimated the smartphone's location every 150 seconds using standard Android APIs (which uses WiFi access points and cell tower information). The positioning system returns latitude, longitude, and estimated error. We discarded the location if the error was greater than 200 meters.

When a participant turned on her display, our prototype took the latest location information and classified the location as home, workplace, or other, using a 5-nearest neighbors classifier. To minimize misclassifications, especially in areas where ground truth data is sparse, the classifier considered ground truth within a 100 meter radius. Our prototype then requested an active authentication according to participants' locations and the experimental conditions (Table 2). After participants completed the active authentication, the prototype asked participants to confirm their semantic location (home, workplace or others) to use as additional ground truth data for the 5-nearest-neighbor classifier (Figure 2).

6.5 Results

6.5.1 Location Classification

Our prototype asked for the ground truth of locations after each authentication (Figure 2 (c)) and trained the 5-nearest neighbor classifier using all the ground truth collected up to the classification. The classification accuracy was 92%. Most of the misclassifications happened due to our location sampling rate. It would be therefore be possible to improve the classification accuracy by increasing the sampling rate when the accelerometers on the mobile device detect that it is moving.

6.5.2 User Authentication

Our participants activated their phones 33.8 times a day on average. Figure 3 shows the distribution of phone activations per day. The black and gray bars represent participants in the PIN condition and in the password condition respectively. The participants in the PIN and the password condition activated their phone a mean of 23.9 times and 47.3 times a day respectively. The difference between the means was statistically significant with Welch's t-test ($t(14)=2.78$, $p<0.05$). This result might be because those who use their phones more frequently are more likely to have sensitive data on their phone. Table 3 shows that participants in the PIN condition activated their phones 68% of the time at home or work, and participants in the password condition did the same 55% of the time. This indicates that they mostly activated their phones at homes or workplaces.

One possible concern with CASA's approach is that users may be more likely to forget their PINs or passwords because they are used less frequently. However, in this study, we found that participants still typed PINs and/or passwords frequently enough to retain them. As shown in Table 3, participants typed PINs 8.1 times a day on average in the PIN condition. Similarly, in the password condition, participants typed PINs more than 31.6 times a day and typed passwords 15.7 times a day on average. Furthermore, we found that our participants typed correct PINs 96.5% of the time, out of 1034 total authentications using PINs. Additionally, no participant typed the wrong PIN three times successively. For passwords, there were two cases out of 1193 authentications using passwords where participants typed wrong passwords three times successively. However, in both cases they retrieved passwords in the next authentication. These data indicate that, although the frequency of typing PINs and passwords decreased, the memorability of PINs and passwords remained high.

6.5.3 Participants' Receptiveness

In a post-survey, we asked participants about their perceptions of our prototype using a 5-point Likert scale (higher scores being more positive). Participants in both conditions were very receptive to our prototype. Below, the number in the parentheses denotes the median of ratings.

Participants in the PIN condition reported that not requiring a PIN at home and work while requiring a PIN at other places was useful (4) and very easy to understand (5). They also reported that they felt our prototype was secure (4) compared to not having any security lock on their phone. They were neutral (3.5) to using our prototype if it were available on their phones.

Similarly, participants in the password condition reported that requiring a PIN at home or work while requiring a password was neither useful nor useless (3) and easy to understand (4). They also reported that they felt the prototype was more secure (4), as easy to use as requiring a PIN at all the places (4). However, they were neutral (3) to using our prototype.

We further asked the participants in the password condition about the configuration that we used for the PIN condition. (i.e., not requiring PINs at homes or workplaces and requiring PINs at other places). The participants reported that the configuration would be easy to use (4) and as secure as a requiring a PIN at all places (3.5), and they somewhat agreed (4) that they would use the system if it were available on their phones.

Table 4. We categorized attackers into a 2×2 table based on knowledge about target users and technical expertise.

		Knowledge about target users	
		Uninformed	Informed
Technical expertise	Novice	Uninformed Novice	Informed Novice
	Expert	Uninformed Expert	Informed Expert

As these results indicate, participants thought our prototype useful. Although participants were neutral to using our prototype on average, our participants rated the none-PIN configuration as easier to use than the security lock that they used prior to our study, and more or equally secure to the security lock. We further iterated on the system design in our third field study to make it fit better to users' needs.

7. Security Analysis

In this section, we discuss the security implications of CASA with respect to our results from the second study. Through this discussion, we identify potential security risks and possible improvements to the system that we tested in the third study. Table 4 divides possible attackers into four groups based on whether they have information about their target (*informed* or *uninformed* attackers), and whether the attackers have knowledge about CASA as well as information security in general (*novice* or *expert* attackers).

7.1 Uninformed Novices and Experts

An example scenario where an uninformed novice might attack CASA is the case where a legitimate user loses her phone outside of home or work, and some stranger picks it up. In this case, CASA is almost as secure as a system that requires a PIN all the time. The only situation where it would be weaker is if the user loses her phone right next to her home or work, or if an attacker breaks into a person's home or work (again, we assume these places have reasonably good physical security). An expert attacker could try to activate the phone at different places to find the user's home or workplace, in hopes of putting CASA into its simpler mode of authentication. However, CASA can also be configured to always require a PIN after a certain number of trials, making this kind of attack infeasible.

7.2 Informed Novice

Informed novices would be people who know a lot about an individual but not a great deal of technical expertise. Those who living with users, such as family members, could be informed novices. However, our survey results showed that our participants trusted people who they are living with. Furthermore, even if the phone is protected by PIN, existing work has reported that people frequently share their PIN among these people [34]. Thus, even if CASA does not work well when the family members are not trustable, it does not increase the risk significantly. Alternatively, the system can allow people to configure where it does and does not require PIN.

Friends or co-workers could be informed novices as well. They could visit users' homes or workplaces to access the users' phones. The system in study #2 did not have protection for this threat model. In designing the prototype used in the second study, we assumed that homes and workplaces would have reasonable level of physical security and that people would put more weight on ease of access than security in these places. The results of the second user study suggest that these assumptions hold for homes but not for workplaces. Thus, we improved the system to mitigate the risk at workplaces and tested in the field study #3.

7.3 Informed Experts

Informed experts are the most capable attackers against CASA, who can dedicate time and resources to breaking the system. In practice, this group likely represents a small and exceptional case, one that goes outside of the average case that we are focused on, but also a case that CASA should offer some kind of protection against. At the same time, too much emphasis on security may lead people to not have any security, as exemplified by the number of people not using security locks on their phones.

Given the difficulties of defending against a dedicated attacker, the relative rarity of attacks, plus the goal of balancing security and usability, we opted to focus less on prevention mechanisms and more on detection, making it easier for phone owners to see if others were making use of their devices. We implemented and evaluated one possible detection mechanism in the field study #3 in the form of a notification mechanism indicating when and where the smartphone was used.

8. CASA Design Iteration

Based on the previous field study and security analysis, we iterated on the system design of our prototype, to make it more acceptable for users and secure against some of the potential attacks. The results from the previous field study clearly showed that participants with security locks on their phone found requiring passwords on mobile devices too high a burden. Thus, we configured our system to require a PIN at places outside of one's home.

Participants also showed concerns about not requiring user authentication at workplaces. Our survey results showed that, while 68% of participants strongly trusted people who could access their home, only 18% strongly trusted those who could access their workplace. This implies that the approximation that we made in the active factor selection in the system design was not accurate enough because it did not take the level of trust into account. Thus, in this iteration, we assumed that being at workplaces does not provide enough evidence to change the active factor from PIN to none. Instead, we added one more passive factor at workplaces, having smartphones check whether users were using their laptop (or desktop) computers nearby. If the computers were being used, the probability that users were near their smartphones was quite high, assuming that the computers required passwords to be accessed. Thus, the smartphones required no active factor when the computers were used nearby recently.

Our smartphone prototype communicated via Bluetooth with an application installed on the users' computers every 60 seconds to check the last time the keyboard or mouse was used. If the users interacted with their computer within the past 180 seconds, the smartphone prototype did not require a PIN. This modification could address the cases where, for instance, the users leave their smartphones at their workplace unattended.

In addition to improving prevention mechanisms (i.e., user authentication), we also added a notification mechanism. When a user's smartphone was activated, a popup message would appear on their computers. Clicking on the message would show the geo-location of the smartphone on a map. The message disappears after five seconds automatically. Although this approach does not prevent illegitimate accesses, it makes detection easier and could prevent further access to sensitive data.

Finally, we modified CASA to always require a PIN when someone turned on the phones' display more than five times without typing a PIN, to prevent attackers from trying to systematically find a user's home or workplace.

9. Study #3: Iterative Evaluation

In addition to using the modified prototype described in the previous section, we also modified our study protocol to improve ecological validity. In the second study, participants were always asked their locations after they activated their phones to train the location classifier. However, for a real version of CASA, this data collection should happen only for a short period when users start using the system. Furthermore, in field study #2, some participants commented that answering where they were located every time they typed in a PIN was tedious. To address this problem, we divided the study into a training period and an evaluation period.

In this study, we clearly separated a training collection period and an evaluation period. On the first day, we had the first session where we explained our system and installed our prototype on their Android phones and laptops. The first five days was a training period where the system asked for participants' semantic locations (home, workplace or other places) as well as requiring PINs if necessary, as we did in the previous study. After the training period, we had five to nine days of the evaluation period, where the system stopped asking the questions. After the evaluation period, we had the second session, where we asked participants to fill our survey and conducted interviews for 20 minutes.

We recruited 18 participants using the university's participant recruitment website. We recruited participants who had Android phones and laptops with Bluetooth, which they used at their workplaces. None had participated in previous studies. Their age ranged from 21 to 40 years old with a mean age of 26.3. Our participants consisted of 12 students and six fully-employed. Sixteen of 18 participants were living with others in their homes. Seven were using security locks on their Android phones prior to our study. All the participants used passwords to log into their computers and to unlock screensaver on their computers. We compensated participants $60 for their participation.

9.1 Results

For the basic analyses, we found results similar to those from field study #2. Thus, we will describe the analysis of the modified parts.

9.1.1 Logins at Workplaces

Figure 4 shows the distribution of the average numbers of the phone activations at workplaces per day for each user. The black parts denote the number of cases where their computers were active and the phones did not require a PIN to be activated, and the gray parts denote the number of cases where the phones required PIN. On average, the participants activated their phone 5.5 times a day at their workplaces. Out of the 5.5 times, the phones did not require PINs 2.9 times, while they did 2.6 times.

9.1.2 User Perceptions

In our post-survey, we asked our participants to rate the three features in our system both in Likert scales and freeform responses. In the followings, the numbers in parentheses denote medians of Likert scale responses where 1 denotes very negative and 5 denotes very positive.

The participants were generally positive about our system. They reported that changing the authentication method based on

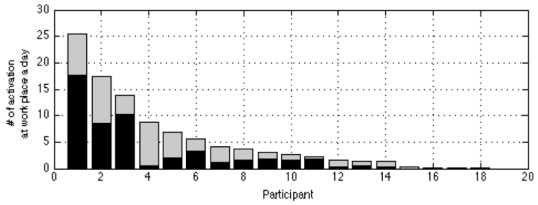

Figure 4. The average numbers of activations at workplace per day. The black portion denotes the cases where the system did not require a PIN because user's computers was being used. The gray portion denotes the cases where the system required a PIN.

location was useful (4.5), and that changing it based on computer usage was useful (4) as well. They also answered that changing authentication method based on location and computer usage was easy to understand (5 and 4.5 respectively). They also reported that, compared to always requiring a PIN to unlock the phone, it was not less secure not to require a PIN at home (4) or based on computer usage (3.5). Furthermore, they reported that they wanted to use the features if the features were available on their phones (4 for the location-based modification and 3.5 for the computer-usage-based modification). These ratings were better than the ratings in the field study #2, which implies that the modifications in this iteration made the system a better fit to users' needs.

One interesting finding was that participants who were not using a security lock prior to this study also reported that they wanted to use CASA (4). P17 commented, "It is annoying to use security locks all the time, but whereas if I had such a system which requires pin only at unsecure places its usefulness adds more value when compared to the annoyance caused by it. So, I will definitely use it." The ratings and this comment indicate the that current all-or-nothing approach where users have to either enable security lock all the time or disable the lock completely, does not meet users' needs well. Furthermore, the users indicated that they would adopt the system even if it would undermines usability so long as they see an appropriate balance between usability and security.

Participants were also very positive about the notification feature. They rated the feature useful (4), and stated that they wanted to use the notification feature if it was available on their phones and computers (4). P13 commented, "I think it's a great way to help with privacy. I use both my computer and phone a lot and it would be very useful to have security." On the other hand, P2 was concerned about distractions, saying, "The notification system is very useful […]. But at the same time, if you just unlock your phone and quickly get back to work, the notifications on the screen can be annoying at times." Interestingly, despite the P2's comment, users are unlikely to see the notifications when they activate their phones by themselves. Users would be looking at their phones when they activate their phones; thus, they are less likely to see the notification on their computers' displays because it disappears after five seconds. Therefore, although there are some cases where the notifications distract users as pointed by P2, we believe that the distraction would be minimal.

10. Discussion

In this paper, we proposed a generic framework for active factor selection based on passive factors. Then, we investigated one possible implementation in this design space by building, evaluating, and iterating on a prototype that made use of location data as the passive factor. We believe that our prototype demonstrated the feasibility and usefulness of our framework. Investigating other points in this large design space will be beneficial in developing authentication systems that provide good security while putting minimum burden on users.

Nevertheless, our work has several limitations. For example, for each factor, CASA needs estimates of $P(s|u = -1)$, the probability that a person trying to be authenticated is not legitimate. Most active factors, such as passwords, have both theoretical and empirical estimates of this probability. In contrast, passive factors, which have not been investigated in the context of user authentication, have limited data. More investigation of passive factors is necessary to rigorously understand this space.

As exemplified in our iterative design process, we can start with a reasonably good system design based on the CASA framework using approximations. Then, we can iterate on the system design, improving the approximation based on data obtained in user studies.

Another limitation of our studies is the treatment of workplace. In our prototypes, we assumed that there would be reasonable physical security at workplaces. This assumption is appropriate for many office workers, but may not be for those who do not have offices or other dedicated space at their workplaces. A possible solution for this issue is to ask users to configure places where they think they have reasonable physical security. It may also be possible to estimate the level of security of a place based on analysis from publicly available sources, such as foursquare.

One line of future work is to evaluate other passive factors and user models. Prior work has investigated the security of some passive factors, such as behavioral biometrics. However, the security of other passive factors is not clear, especially when malicious attackers try to impersonate legitimate users. Furthermore, in this paper, we used a very simple model (two passive factors modeling three locations and computer usage). Our model had the benefit of being simple to implement and simple to understand. It is clearly possible to build more sophisticated models, combining more passive factors (e.g. last login time, number of times logged in at a given location). However, this approach raises new questions about how well users can understand what the system is doing, and could lead to frustration in case of poor prediction.

Lastly, we believe it is worth investigating new "good enough" forms of active authentication. For example, most active authentication schemes today are designed for high accuracy in differentiating between legitimate and illegitimate users. By leveraging multiple passive factors, it is possible to relax this constraint, requiring only "good enough" accuracy.

11. Conclusion

In this paper, we introduced Context-Aware Scalable Authentication (CASA), which envisions using multiple passive factors to modulate active factors to authenticate users. We proposed a generic probabilistic framework that enables the selection of appropriate active authentication factors given a set of passive authentication factors. We also developed prototypes exploring one point in this design space, investigating the feasibility and effectiveness of our proposed framework. The results of three field studies demonstrated that the prototypes could select active authentication factors based on passive factors while balancing security and usability of user authentication.

In the first user study, we observed that the participants logged into their phone 60% of the time at their homes or workplace. This

data indicated that there was substantial potential to improve both the usability and the security of a user authentication by choosing active factors based on users' locations. In the second study, we developed a prototype that changes active factors (no authentication, PIN, and password) based on users' locations. Through a field study, we observed that our prototype improved the security of user authentication at less frequently visited places (which consisted of 32% to 45% of all user authentications) without affecting usability of the rest of the user authentication at home or workplaces. Finally, in the third study, we showed an iterative design process to improve both usability and security of the prototype based on the results in the second study. We added computer usage as a passive factor at workplaces and implemented detection feature to mitigate the risk of very strong attacks that cost too much to be prevented. Our participants were very positive about our system. Participants including those who did not use any security lock prior to study, showed strong interest to use our system on their phones.

Although there is ample opportunity for further investigation, we believe that this paper proposes a novel authentication framework and demonstrates its feasibility and the usefulness. We hope this work stimulates future research towards our vision of developing user authentication systems that require minimum but sufficient active factors.

REFERENCES

[1] eToken. http://www.aladdin.com/etoken/.

[2] RSA securID http://www.rsa.com/node.aspx?id=1156.

[3] Advanced sign-in security for your Google account. http://googleblog.blogspot.com/2011/02/advanced-sign-in-security-for-your.html

[4] Facebook Social Authentication. http://facebook.com/blog/blog.php?post=486790652130

[5] Lax Passwords Expose Quater of PC Users to Theft. http://www.washingtonpost.com/wp-dyn/content/article/2007/10/09/AR2007100901896.html

[6] When Security Gets in the Way. http://jnd.org/dn.mss/when_security_gets_in_the_way.html

[7] Adams A. and Sasse A. M. 1999. Users are not the enemy. *Commun. ACM* 42, 12 (December 1999), 40-46.

[8] Amini S., Lindqvist J., Hong I. J., Lin J., Sadeh N., and Toch E. 2011. Caché: Caching Location-Enhanced Content to Improve User Privacy. In *Proc. of MobiSys.*

[9] Bardram J. E., Kjær R. E., Pedersen MØ. 2003. Context-Aware User Authentication Supporting Proximity-Based Login in Pervasive Computing. In *Proc. of UbiComp*.

[10] Burr W. E., Dodson D. F., and Polk. W. T. 2006 Electronic authentication guideline. Tech report, NIST

[11] Buthpitiya S., Zhang Y., Dey A. and Griss M, n-gram Geo-Trace Modeling, In *Proc. of Pervasive Computing*.

[12] Cheng P., Rohatgi P., Keser C., Karger P., Wagner G., and Reninger A. 2007. Fuzzy Multi–Level Security: An Experiment on Quantified Risk–Adaptive Access Control. In *Proc. of IEEE Symposium on Security and Privacy*

[13] Corner M. D. and Noble B. D. 2003. Protecting applications with transient authentication. In *Proc. of MobiSys*.

[14] Cranshaw J, Toch E., Hong J. I., Kittur A., and Sadeh N. 2010. Bridging the gap between physical location and online social networks. In *Proc. of UbiComp.*

[15] Fischer I., Kuo C., Huang L., and Frank M. 2012. Short Paper: Smartphones: Not Smart Enough? In *Proc. of SPSM.*

[16] Froehlich J. and Krumm J. 2008. Route Prediction from Trip Observations. Society of Automotive Engineers.

[17] Gupta A., Miettinen M., Asokan N., amd Nagy M. 2012. Intuitive security policy configuration in mobile devices using context profiling. In *Proc. of PASSAT.*González M. C., Hidalgo C. A., Barabási L. A. 2008. Understanding individual human mobility patterns. Nature 453, 779-782.

[18] Hayashi E. and Hong J. I. 2011. A diary study of password usage in daily life. In *Proc. of SIGCHI*.

[19] Herley C. 2009. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proc. of NSPW*.

[20] Hulsebosch J. R., Salden H. A., Bargh S. M., Ebben P. W. G, and Reitsma J. 2005. Context sensitive access control. In *Proc. of SACMAT*.

[21] Inglesant P. G. and Sasse A. M. 2010. The true cost of unusable password policies: password use in the wild. In *Proc. of SIGCHI.*

[22] Jakobsson M., Shi E., Golle P., and Chow R. 2009. Implicit authentication for mobile devices. In *Proc. of USENIX.*

[23] Kalamandeen A., Scannell A., Lara E., Sheth A. and LaMarca A. 2010. Ensemble: cooperative proximity-based authentication. In *Proc. of Mobisys*.

[24] Komanduri S., Shay R., Kelley P. G., Mazurek M. L., Bauer L., Christin N., Cranor L. F., and Egelman S. 2011. Of passwords and people: measuring the effect of password-composition policies. In *Proc. of SIGCHI.*

[25] Krumm J. 2008. A Markov Model for Driver Turn Prediction. Society of Automotive Engineers.

[26] Ni Q., Bertino E., and Lobo J. 2010. Risk-based Access Control System Built on Fuzzy Inferences. In *Proc. of ASIACCS*

[27] Riva, O., Qin, C., Strauss, K., Lymberopoulos, D. 2012. Progressive authentication: deciding when to authenticate on mobile phones. *In Proc. of USENIX*.

[28] Orr, R.J. and Abowd, G.D. 2000. The Smart Floor: A Mechanism for Natural User Identification and Tracking. ACM Press, New York, New York, USA.

[29] Peacock, A., Xian K., Wilkerson, M. 2004. Typing patterns: a key to user identification, Security & Privacy, IEEE, vol.2, no.5, pp.40-47, Sept.-Oct. 2004

[30] Philip G. Inglesant and M. Angela Sasse. 2010. The true cost of unusable password policies: password use in the wild. In *Proc. of SIGCHI.*

[31] Shay R., Komanduri S., Kelley P. G., Leon P. G., Mazurek M. L., Bauer L., Christin N., and Cranor L. F. 2010. Encountering stronger password requirements: user attitudes and behaviors. In *Proc*. of SOUPS.

[32] Seifert J., De Luca A., Conradi B. and Hussmann H. 2010. TreasurePhone: Context-Sensitive User Data Protection on Mobile Phones. In *Proc. of Pervasive*.

[33] Supriya Singh, Anuja Cabraal, Catherine Demosthenous, Gunela Astbrink, and Michele Furlong. 2007. Password sharing: implications for security design based on social practice. In *Proc. of the SIGCHI*