Journal of Mobile Multimedia, Vol. 12, No.1&2 (2016) 159-180
© Rinton Press

PRIVACY-BASED ADAPTIVE CONTEXT-AWARE AUTHENTICATION SYSTEM FOR PERSONAL MOBILE DEVICES

ZHAN LIU RICCARDO BONAZZI

University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis), Sierre, Switzerland
zhan.liu@hevs.ch riccardo.bonazzi@hevs.ch

YVES PIGNEUR

University of Lausanne, Lausanne, Switzerland
yves.pigneur@unil.ch

Over the past decade, mobile devices such as smartphones have become increasingly common as a form of handheld computing platform. The use of mobile applications on these mobile devices is experiencing unprecedented rates of growth. However, when using mobile applications, users are often requested to give context information. Such requests have led to growing privacy concerns. This paper proposes the use of context-awareness to improve single sign-on (SSO) solutions so that mobile users can protect their private information. A privacy-based adaptive SSO (ASSO) may be able to increase users' perceived ease of use of the system and give service providers the necessary authentication security for their applications. The study was based on data gathered from 168 participants as part of the Lausanne Data Collection Campaign. This was led by the Nokia research center in Switzerland and used Nokia N95 phones. The analysis of SVM showed our expectations to be correct. Consequently, a new business model for mobile platforms has been proposed to reinforce our claim that privacy-friendly value propositions are possible and can be used to obtain a competitive advantage.

Key words: Adaptive Single Sign-On, Authentication Security, Ease of Use, Context-Aware, Privacy, Business Model

1 Introduction

In recent years, the penetration of smartphones has reached particularly high levels. As one of the defining technologies of our time, they have had a pervasive influence on the personal lives of individuals in many ways, including giving competence in communications and connectedness, and in terms of privacy issues, confidentiality, and individuality [3]. On the negative side, however, the use of smartphones has created many privacy concerns, especially with regard to the context-based services that often accompany users' Personal Identifiable Information (PII). As a consequence, security problems have arisen, namely in the form of privacy protection.

Information security systems are required to protect PII and ensure users' privacy. Existing research on information security has implied that such security implies a tradeoff between the system developers' efforts to implement privacy-enabling technologies and the cognitive effort required to use such technologies. Let us consider a mobile user trying to access a set of web services, as shown in the top part of Figure 1. The user has to pass a set of access controls for authentication, identification, authorization, and accountability. This security procedure increases users' perceived performance of the protection application, but it negatively affects the ease of use of the system. Inglesant and Sasse [20] have shown how low perception of ease of use can lead to a lack of user compliance with security policies. A SSO solution can increase ease of use, as shown in the middle part of Figure 1. Solutions such as Firefox's built-in password manager increase ease of use. However, they also reduce the amount of effort needed by attackers to access users' accounts, because only the master password has to be broken.

In order to offer stronger authentication, an SSO solution usually requires a shift from an access control list system (e.g., passwords) to a capability-based system (e.g., biometric controls or multi-factor authentication). However, this approach to security lacks flexibility, because a user's biometry cannot be changed over time. To this end, this paper aims to achieve the correct tradeoff between dynamic authentication and ease of use. We are looking for a system that can transparently authenticate a user and dynamically adapt to that user's behavior.

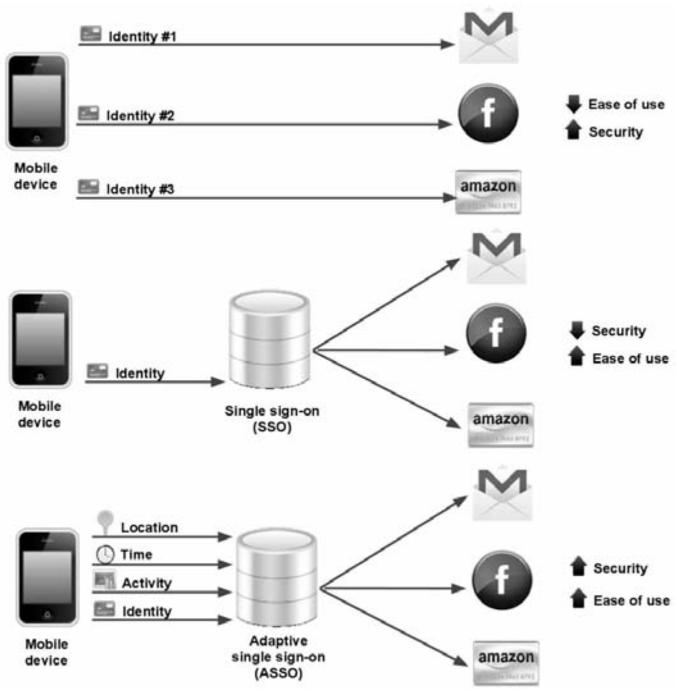


Figure 1. Adaptive Single Sign-On (ASSO) solution for security and ease of use

Since the early 1990s, context-aware mobile services have been of interest to scholars (e.g., [1,16,38,42]). These studies have basically treated context awareness as a trigger for privacy concerns. However, this study takes a different approach. We intend to utilize users' context information to protect their private information, thereby decreasing mobile users' privacy concerns. For the purpose of this paper, we refer to context as any information that can be used to characterize the situation of [...] a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves [1]. Based on this definition, there are four types of primary context: location, identity, activity, and time. These types characterize a situation by answering where, who, what, and when, respectively. Such contextual information can help to identify users' behavior patterns, which in turn would be useful for protecting their private information. Therefore, our research question is: how can context-awareness technologies be used in a privacy management system for mobile phone devices in order to improve authentication security and maintain ease of use?

To answer the research question, we first examined state-of-the-art studies on context awareness and SSO solutions. We then proposed an adaptive SSO (ASSO) that utilizes context awareness to help achieve authentication and ease of use, as shown in the bottom part of Figure 1. A methodology was executed, together with an SVM supervised learning algorithm-based approach. The latter necessitated the collection of huge amounts of location and time data from real mobile users over the course of one year. We found that users' context information is a unique identifier, therefore verifying our proposed solution. This paper further describes the business model of an ASSO and its implications for business practitioners. The audience addressed is mainly composed of stakeholders in mobile services who seek guidelines to develop privacy-friendly business models. In this regard, minimal research has failed to successfully capture the design of business models for mobile services. This study aims to fill this research gap.

The rest of this paper is organized as follows. The section 2 discusses the differences between privacy and security in the extant literature. We then introduce the methods used to carry out this study, including an illustrative scenario, hypothesis development, the scope and magnitude of data collected, and the analysis applied in section 3. The section 4 discusses the findings of our research, and the 5 section illustrates a set of business model considerations that relate to the application of our solution. The section 6 lists the contributions made by this paper and the research avenues it opens up.

2 Literature Review

2.1 Privacy and Security

The concept of "privacy" is usually considered a human right; in other words, it is "the right of the individual to decide what information about himself should be communicated to others and under what condition" [50]. Based on this understanding, Smith et al. [45] identified four factors of online privacy: the unauthorized secondary use of personal information, improper access to personal information, the over-collection of personal information, and errors made when collecting personal information. Mobile devices can be seen as more personal than, for example, a traditional desktop computer. In particular, mobile phone users generate even more personal data, including geographic location data about the physical movement of their mobile devices [21]. However, such location information often reveals the position of a person in real time, rendering the potential intrusion of privacy a critical concern [51]. Mobile devices also store additional personal information such as personal contacts, photos, messages and emails. Combined with such personally identifiable information, location information may have consequences relating to the extent of access, collection, the use or disclosure of personal identifiable information, and privacy concerns in the form of data protection. In the current study, the privacy of personal identifiable information refers to the ability of an individual to control the way that personal identifiable information is gathered or used by an unauthorized party.

An information security system can be defined as one that offers protection against threats from potential circumstances, conditions, or events that cause economic hardship; for example, data transaction attacks and the misuse of financial and personal information [7]. According to Pennanen et al. [35], information security consists of three main parts: confidentiality, integrity, and availability. The current study focuses on confidentiality, which refers to limitations in information access and disclosures to authorized parties, as well as the prevention of access by or disclosure to unauthorized parties.

Privacy and security concerns are not new concepts. Indeed, they have been viewed by some researchers as one construct in online purchasing for a number of years (e.g., [26]). While online privacy and security concerns are sometimes inextricably linked, Belanger et al. [7] have argued that these two notions are two distinct constructs and that there is a lack of understanding about their true relationship. Security is sometimes described as a necessary tool for building privacy. In other words, privacy cannot be achieved without a good security foundation.

In context-aware applications, however, a security issue may take place without the occurrence of privacy violations. For example, some cases may lead only to a security issue because only location data is collected by an untrustworthy party. Here, users' identity information is not associated with location information; thus, there is no privacy issue. If location information is linked to users' information, however, then a privacy issue does exist.

It is important to note that privacy and security are not absolute concepts – users hold very different opinions about their level of privacy concerns. Some might not care about their privacy at all. This study assumes that mobile users are pragmatists [2] who often have specific concerns and particular tactics for addressing them. Therefore, protecting their privacy is a valuable benefit for them.

2.2 Context and Context-Aware Applications

In the literature, a lot of researchers have attempted to define context, and the precise nature of context-awareness. The term context-aware first appeared in a study by Schilit and Theimer [43], who described context as locations, the identities of nearby people and objects, and the changes that occur to those objects over time. Later, Schilit et al. [42] offered three categories: who you are, who you are with, and the objects that are around you. Such examples of context were often used in early research into context-aware systems. Hull et al. [19] described context as the environmental aspects of the entire current situation. Dey [17] defined context as the emotional state of users, the focus of their attention, their location and orientation, date and time, and the objects and people in their environment. Chen and Kotz ([12], p.3) referred to context as the set of environmental states and settings that either determine an application's behavior or in which an application event occurs and is interesting to the user. These definitions are often too broad and difficult to apply in specific systems. The definitions given by Ryan et al. [40] and Abowd et al. [1] are similar, and correspond with the own beliefs. Ryan et al. [40] claimed that context involved not only a user's location and identity information, but also his/her environment and time information. However, Abowd et al. [1] went on to argue that the term "environment" should be replaced with "activity". They thought that activity can answer a fundamental question of what is occurring in a particular situation, whereas environment cannot. This study followed Abowd et al. [1]'s definition; thus, context characterizes users' situations by answering where, who, what, and when, respectively.

With the technological advances of today's hand-held devices, collecting context information is no longer an issue [27]. As a result, mobile context-awareness focuses on building applications that can take advantage of contextual information. A context-aware application must have a large and significant ability to perceive the surrounding environment. According to Biegel and Cahill [9], a context-aware application has three main components: a set of sensors for detecting and capturing contextual information, a set of rules that governs behavior according to context, and a set of actuators for generating responses. Accordingly, our ideas on the development of context-aware applications are based on privacy protection in a mobile environment, which not only defines the sensors and actuators, but also provides the corresponding rules that drive behavior. Sensors are the sensor components fitted to mobile devices, such as GPS, Bluetooth, real-time, and WiFi. Actuators, also known as opacity tools, are privacy protection tools. They aim to protect the identity of users, and minimize the effects of revealed personal data by, for example, encrypting private information and blocking potential attacks. Rules are applied to a specific environment. From a regulatory point of view, data privacy laws are present in different business sectors and in different countries, leading to a complex multitude of overlapping and sometimes conflicting regulations that change over time. In this study, we refer to the concept of rules in order to determine which actuator should be used in a given context. As a computing platform, however, a smartphone is both pervasive and personal. The personal nature indicates two important implications. On the one hand, all the elements vary a lot in the mobile environment: users are different and they may use different services or activities at different times and in different places. On the other hand, smartphones store our most personal information, such as photos and passwords, and contain important associated private information (i.e., clues about current location). This increases mobile users' privacy concerns. Therefore, this paper proposes to use the contextual information of mobile phone users to protect their private information.

2.3 Single Sign-On Solution

Traditionally, systems have used databases as their authentication mechanism. In such solutions, the users are given a login name and a password for accessing each system. With heterogeneous systems, users have to manage a set of passwords for each system and log in separately. All the passwords have to be remembered, and even worse, changed frequently, so this is hardly an ideal situation. More recently, single sign-on (SSO) technology has become more widely used for authentication solutions. As the name implies, SSO systems are designed to authenticate once, without further manual interaction; thus, users do not have to repeat the log in process for each system. Several research studies [14,37,47] have also argued that SSO solutions reduce administrative work by resetting forgotten passwords over multiple platforms and applications, and improving on convenience, because users need only to remember a single set of credentials. Recently, new variations of SSO authentication have been developed using mobile devices as access controllers. Users' mobile devices can be used to automatically log them into multiple systems by, for example, building access control systems and computer systems through the use of authentication methods. Such methods include OpenID Connect [41] and SAML [23], which associates the mobile device with an access server. In general, the main advantage of designing an SSO-based system is the ability to retain ease of use whilst also providing improved user-controlled privacy capabilities.

However, SSO is not yet a universal solution. Without careful planning, implementation and verification, SSO products may introduce new security holes. For example, if an SSO account is hacked or an account password is copied, all others under that authentication will be hacked as well [5]. This risk may be reduced when choosing SSO credentials that are not knowledge-based (e.g., when a classic password is used) but are biometric-based (e.g., users' unique behavior movements) or possession-based (e.g., NFC smart cards). Our solution, which is intended to further reduce this risk, focuses on the use of multi-factor adaptive authentication solutions for SSO.

3 Methodology

In this section, we start by presenting an illustrative scenario to address gaps in the literature. We then go on to describe our theoretical model, including the constructs and hypotheses. Finally, we show how the data was collected, as well as offer a description of the method and procedure used for data analysis.

3.1 An Illustrative Scenario

Figure 2 offers a simple scenario to present the ASSO solution.

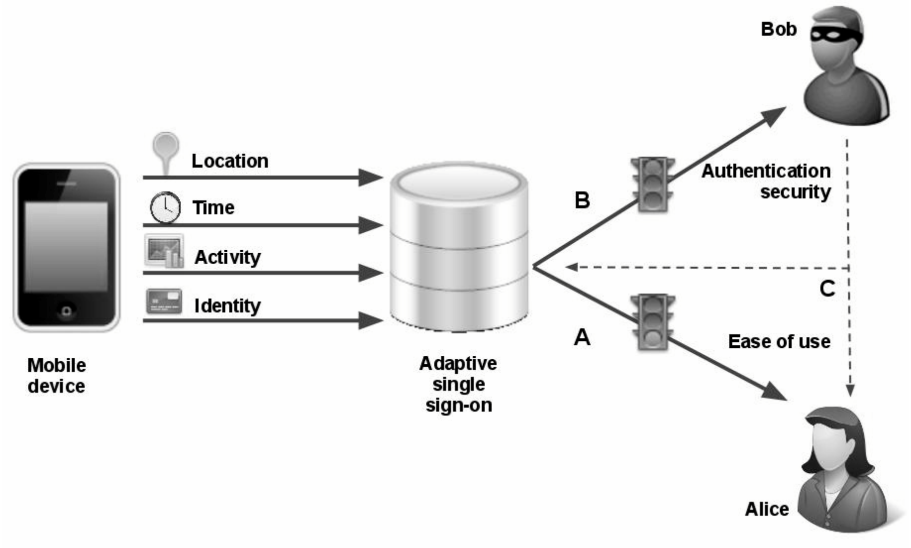


Figure 2. Process involved in an ASSO solution for a context-aware mobile device

A. Alice accesses her Internet accounts

The end user, Alice, has a mobile device with an application called Privacy Manager. This application uses adaptive authentication to combine real-time transaction data with Alice's behavioral profile. The real-time transaction data used to identify Alice includes her current location, speed (activity), and time. Once the data analysis application returns a positive authentication result, Alice can check her e-mail and bank accounts online through the protected channel. Thanks to Privacy Manager she can access her email and bank accounts without having to enter any passwords, as long as the data analysis returns a positive result.

B. Bob cannot access Alice's accounts

Let us assume that a thief (Bob) plans to steal Alice's mobile device to access Alice's bank account. Once Bob steals the device, the real-time transaction data does not match the data stored in Alice's profile. Suppose that Bob knows this authentication method, and tries to follow Alice before stealing the phone. Bob would note her location at any given time and collect personal information about Alice in order to copy her behavior. However, the Privacy Manager applies state-of-the-art obfuscation techniques so that no information about the activity and the identity of Alice is disclosed. This leaves Bob with only half of the information required.

C. Alice goes on holiday

When Alice goes to another city to see a friend, Privacy Manager detects that the behavior disclosed does not match Alice's older profiles. Nevertheless, Alice possesses a trusted means of identification (e.g., a password) to provide user identification. After identification, Privacy Manager creates a new profile and stores data to include Alice's behavior on this day. When Alice comes to this city again to see her friends, her behavior data will be matched with this profile.

3.2 Hypothesis Development

From a cognitive point of view, usability issues arise when users cannot properly manage the information required to sign in to different web services using a large set of different pseudonyms and passwords. The possibility of capturing a change in the identity of a real user (using the features of his or her everyday life behavior) has only been considered as a threat to that user's privacy. We propose to shift from a discretionary access control approach to an attribute-based approach, where the attributes are features of the user's environment and his or her behavior in that context.

This approach assumes that each user has a unique pattern of behavior to provide a high level of control over access to mobile services whilst still maintaining a high level of usability. In previous studies, context-aware technology evoked concerns about privacy. Location-based applications track users automatically on an ongoing basis, generating an enormous amount of potentially sensitive information. From this information, the identity of the owner of the mobile device can be implicitly obtained from the analysis of its location [8,18]. However, we see great potential in such a threat, and believe that context is a unique user identifier.

Context-based authentication is currently used for credit card fraud detection. It relies primarily on artificial intelligence techniques and uses unsupervised learning methods [10]. Machine learning usually refers to evolved behaviors that are based on empirical data, such as that gathered from sensor data or databases associated with artificial intelligence. The information is acquired during authentication through a learning process which authenticates the mobile user. The asserted advantages of machine learning are a level of accuracy that is comparable to that achieved by human experts. Also of benefit are considerable savings in terms of expert labor power, because no intervention from either knowledge engineers or domain experts is needed for the construction of the classifier or for its porting to a different set of categories [44]. User data clustering can be performed on two levels: on the one hand, the best matches and the corresponding data points can be automatically or manually grouped into several clusters so that outliers can be easily detected. The alert will be activated once the number of outliers exceeds the predefined threshold. On the other hand, new trends can be found when regions on the map representing a cluster are identified and used for the classification of new data. To test our hypotheses, we propose a system that collects a set of mobile sensor data and compares them with a known set of users' profiles. Moreover, this study suggests using an escalating procedure to minimize the computational effort of the system for most authentication cases. Thus, in this study, a limited amount of phone sensor data was collected by the context-awareness component. Through machine leaning, the mobile phone can determine whether or not it is dealing with an authorized user. If the result is positive, the user is authorized to access the services (e.g., Amazon, Gmail, Facebook); otherwise, additional contextual information about the user (ranging from time and location through to activity and, eventually, identity) is collected and analyzed before access to any services is granted. Figure 3 presents the structural model. A rectangular element is associated with a variable that can be directly measured, whereas an oval represents a latent concept that has to be measured indirectly by summing the variables with which it is associated.

A high level of authentication security minimizes the number of occurrences in which the user is not allowed to access the system or an unauthorized person is allowed to access the system. We have already stated that location data can be used to infer much about a person, even without the user's name being attached to the data [22]. In this case, let us suppose that the user goes to work every day and comes back following the same routine. In this case, the system would assess the user's location at a certain frequency against the expected pattern (home-work-home). Thus, we derive the first hypothesis: the conjoint effect of time and location increases authentication security (H1).

There may be cases when the home-work-home pattern lacks sufficient variance to discriminate the user from other people. For example, someone physically close to the user could take the phone while it is unattended and access a number of services. Since the phone does not change location, unauthorized access would be possible, even if for a limited amount of time. To address this problem, the system detects when the variance among the collected data is too small, and in this case collects the user's activities (e.g., web pages visited) against known activity patterns. However, due to the complexity of the user's activity information, it is difficult to define such activity patterns. Data on the activity patterns are also difficult to obtain. For this reason, this study has only focused on the combined effect of location and time in this study.

The fourth contextual dimension (i.e., identity) is used when sensor data do not fall into any known pattern. To update the behavioral patterns, access rights are granted to the user following proper identification (e.g., by means of a password, biometric control, or near-field communication card). This kind of identification has already been used by a large set of services; for example, banks make contact with all credit card users following an unexpected buying pattern, and Facebook asks for users to answer a secret question when they try to gain access from a foreign country. If authentication cannot be achieved by the other three dimensions of context, then identity serves as a final step to increase authentication security.

A final consideration is related to how to handle ease of use. It is believed that SVM learning techniques for the classification and eventual use of available solutions for identification would reduce the number of human-computer interactions required for authentication. This would, thus, increase perceived ease of use, and is in line with similar research currently undertaken by banks to develop mobile payment devices that do not use passwords [46]. Thus, we derive the second hypothesis: the conjoint effect of time and location increases ease of use (H2).

The theoretical model is shown in Figure 3.

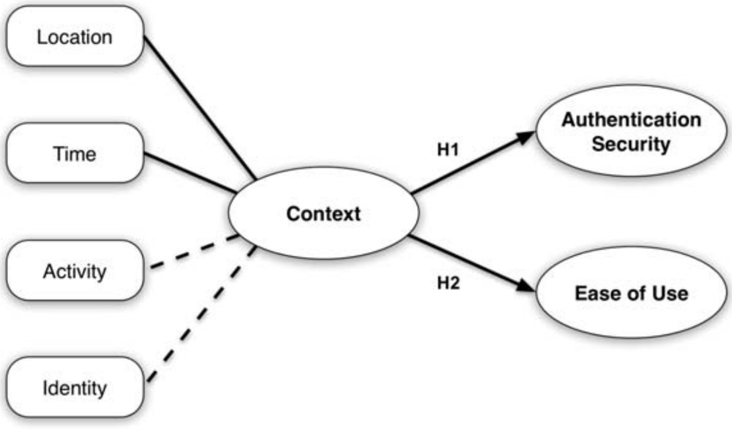


Figure 3. Theoretical Model

3.3 Data Collection and Participants

Users' behavioral data were collected using the Nokia Smartphone N95 and a heterogeneous sample of 168 participants from Lausanne – a second-tier city in Switzerland – over the course of one year. Data collection was carried out as part of the Lausanne Data Collection Campaign (LDCC). Specific data collection mobile software was used. This runs in the background of mobile phones in real time and in a non-intrusive manner, which means that no actions are dependent on the mobile phone itself. This specific software was aimed at making data collection invisible to the participants whilst, at the same time, optimizing the ratio between data collected and power consumed. The mobile application was designed to meet the operating times of one full day, thus enabling participants to use their devices normally during the day and charge the batteries over the night. The objective was to disclose rich and comprehensive data from each individual participant. All collected data was then stored to the mobile device and automatically uploaded to the SQL database server that handles the data when the device detects a known WLAN access point, as well as during the battery charging process.

Four categories of data were collected: (1) social interaction data, such as call logs, short message logs and Bluetooth scanning results; (2) location data, including data from GPS, cellular network information, and WLAN access point information; (3) media creation and usage data containing information on the locations where images have been captured, video shot or music played; and (4) behavioral data, such as application usage, activity detection based on acceleration sensor, and regular device usage statistics based on call and short message logs. This study focuses on the location and the corresponding time data in the quantitative research.

Our sample included 168 participants, of which 65% were males and 35% females. The majority of participants were young people aged between 22 to 33 years old. They included a heterogeneous set of real-life social networks with individuals from mixed backgrounds. According to statistical results from LDCC, 63.1% of participants were employed, 8% were not presently employed, 26% were students, and 3% were categorized as 'other'. Moreover, all the participants mentioned that they had prior experience of using a mobile phone and 97% of the sample saw themselves as active Internet users. Table 1 represents the demographic data for the participants.

Table 1. Demographics

		Number	Percentage
Gender	Male	109	64.9%
	Female	59	35.1%
Age	>50	4	2.4%
	44-50	2	1.2%
	39-44	12	7.1%
	34-38	29	17.3%
	28-33	44	26.2%
	22-27	58	34.5%
	16-27	12	7.1%
	<16	2	1.2%
Occupation	Yes	106	63.1%
	No	13	7.7%
	Student	44	26.2%
	Other	5	3.0%
Active Internet user	Yes	163	97.0%
	No	5	3.0%

The participants were asked to use the experimental phone as their primary phone during the course of the study. They were encouraged to stay in the campaign area for twelve months. Considering that a huge amount of sensitive personal data would be collected from each participant, and stored in the secure databases, the raising of privacy concerns became more important. Some privacy policies were applied to the experimental study to protect the participants' personal information. In addition, participants could always access the data records collected and delete part or all of their data. They could also view social patterns inherent to their behaviors using an online visualization tool.

3.4 Classification Algorithm

Our classification algorithm was designed and developed by using SVM with a predict function. SVMs are supervised learning models that have associated learning algorithms to analyze data and recognize patterns. They are widely used for classification and regression analysis [13]. SVM can maximize the margin between two classes, and use non-linear transformations to separate non-linearly separable objects [48]. With the help of an appropriate non-linear kernel function, the algorithm also allows the maximum-margin hyperplane to be fitted into a transformed feature space.

We measured the quality of classifications resulting from an SVM by using the notions of precision and recall. According to the definition of Lingras and Butz [24], in a classification task precision measures the percentage of objects that are correctly classified as “true” as a ratio of all the objects that should really be “true”. Recall describes the proportion of objects that should be classified as “true” and are in fact classified as “true”.

Precision and recall are complementary measures that are commonly used to promote information retrieval system effectiveness. Both precision and recall are based on an understanding and measure of relevance for information retrieval theorists and practitioners. As shown in formula (1), precision is calculated as the number of relevant items retrieved by an information system (i.e., true positives) divided by the total number retrieved (i.e., the sum of true positives and false positives).

$$\textit{Precision} = (\textit{True positives})/(\textit{True positives} + \textit{False positives}) \qquad (1)$$

As shown in formula (2), recall is calculated as the number of relevant items retrieved divided by the total number of relevant items available (i.e., the sum of true positives and false negatives). According to Pevzner and Marti [36], precision can be interpreted as the percentage of boundaries identified by an algorithm that are indeed true boundaries, whilst recall can be interpreted as the percentage of true boundaries that are identified by the algorithm. Thus, we computed precision to test authentication (H1) and recall to test ease of use (H2).

$$\textit{Recall} = (\textit{True positives})/(\textit{True positives} + \textit{False negatives}) \qquad (2)$$

True positive and false negatives in the formula can be explained by the following example. Consider the scenario of Alice and Bob we discussed above, in which it was important to determine whether the mobile phone is being used by the mobile’s owner Alice. Suppose that in Figure 4, in the training mode, all data are marked with different shapes: the shape of each point is its class (the triangle represents Alice and the circle represents Bob) and the location and time are its data. The SVM can help split these points with a non-linear line, of which the left side is predicted as being Alice (i.e., positive) and the right as Bob (i.e., negative) when new data is received. Cases where the user is Alice are labelled as true; however, when the real user is Bob, they are labelled as false. Therefore, we have the four following scenarios as illustrated in Figure 4.

- *True positive*: when the data comes into the left class and the user is Alice;
- *False positive*: when the data comes into the left class and the user is Bob;
- *True negative*: when the data comes into the right class and the user is Bob;
- *False negative*: when the data comes into the right class and the user is Alice.

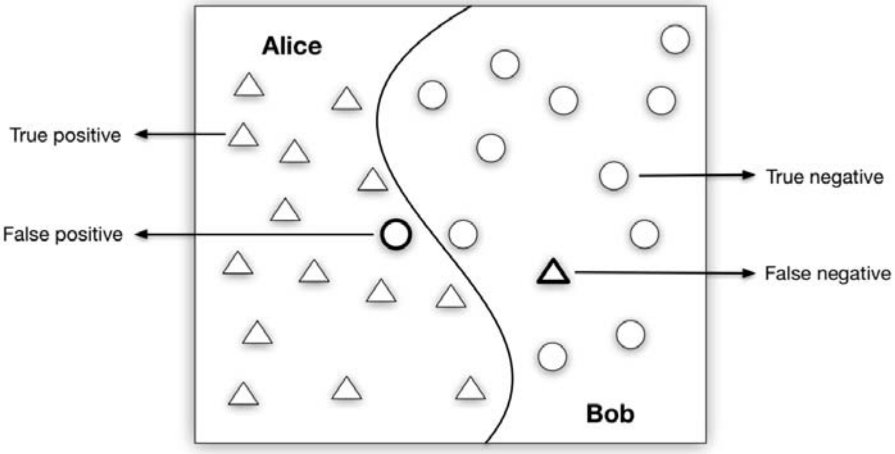


Figure 4. An example of classification with four cases

To acquire a user’s location and time data from the central Nokia databases server, we sent SQL queries over a secure HTTPS prototype. However, this required to store a large amount of data because the results contain all users’ location and time information during one year. In order to reduce both the implementation complexity and meet the storage requirements, we requested the information with individual context and stored each user’s data in a single CSV file. Thus, we collected 168 users’ location and time data into 168 CSV files. We then built the algorithm based on these data files. The script below describes the core of functions and execution process used to obtain the results - Algorithm: Classification based system’s security authentication and ease of use

```
foreach user’s data u1 in u168 do
    Divide the first user dataset into a train and a test set:
    idxtrain_u1<-1:round(nrow(all)/10)
    idxtest_u1<-(round(nrow(all)/10)+1):nrow(all)
    trainSet_u1<-all[idxtrain_u1,]
    testSet_u1<-all[idxtest_u1,]
    Divide the second user (first compare user) dataset into a train and a test set
    if u1 != u2 Then
        idxtrain_u2<-1:round(nrow(all)/10)
        idxtest_u2<-(round(nrow(all)/10)+1):nrow(all)
        trainSet_u2<-all[idxtrain_u1,]
        testSet_u2<-all[idxtest_u1,]
    end if
    else go to next user
    Combine two users’ data frame by rows
    train<-rbind(trainSet_u1, trainSet_u2)
    test<-rbind(testSet_u1, testSet_u2)
    trainDataCall<-subset(train,select=c(-X,-Class))
    trainClassCall<-subset(train,select=Class)
    testDataCall<-subset(test,select=c(-X,-Class))
```

```
        testClassCall<-subset(test,select=Class)
        Running the model again using the train set and predicts classes using the test set in order to verify if the model has good generalization.
        model <- svm(trainDataCall, trainClassCall, type='C',kernel='radial')
        pred <- predict(model, testDataCall)
        A cross-tabulation of the true versus the predicted values yields (the confusion matrix):
        Tab<-table(pred,t(testclasscall))
    end foreach
```

We included four main steps in our algorithm. The first step was to divide the user and compare user data in the training dataset and test dataset. In the experiments of this study, the data collection lasted for one year; thus, we decided to use one month's user data as a training dataset. The second step presented a combination of two users' training and test datasets. In the third step, we used an SVM to create a non-linear classification model. Moreover, the predict function predicts values based on a model trained by the SVM, and returns a vector of predicted labels. In the last step, a cross-tabulation was built to produce the confusion matrix type, which contains the values of true positive, false positive, true negative, and false negative. In order to reduce the algorithm execution time, we used four computers with quad-core processors to calculate in parallel, simultaneously. It took a total of 12 hours to obtain the complete results.

4 Results

The resulting precision and recall are shown in Figure 5 and Figure 6 respectively. These two figures show the frequencies of precision and recall with an interval of 2.5%. We have excluded the ranges that are less than 50% because all results exceed that percentage. The average of precision and recall are indicated as a broken line. As can be seen, the majority of the precision values are between 80% and 100%, with an average of 88.4%. This implies that from among the 100 positive values that were returned from the application, 89 cases were actually the truth (i.e., they correctly identify the real mobile holder). However, on 11 occasions the mobile user was not the real mobile holder. Similarly, most of the recall values were located between 80% and 100%, with an average of 87.7%. Thus, for every 88 true positive values (i.e., correctly identify the real mobile holder), there were 12 false negative values (i.e., fail to identify the real mobile holder). The results suggest that a high level of authentication and ease of use can be achieved. Thus, both H1 and H2 are supported.

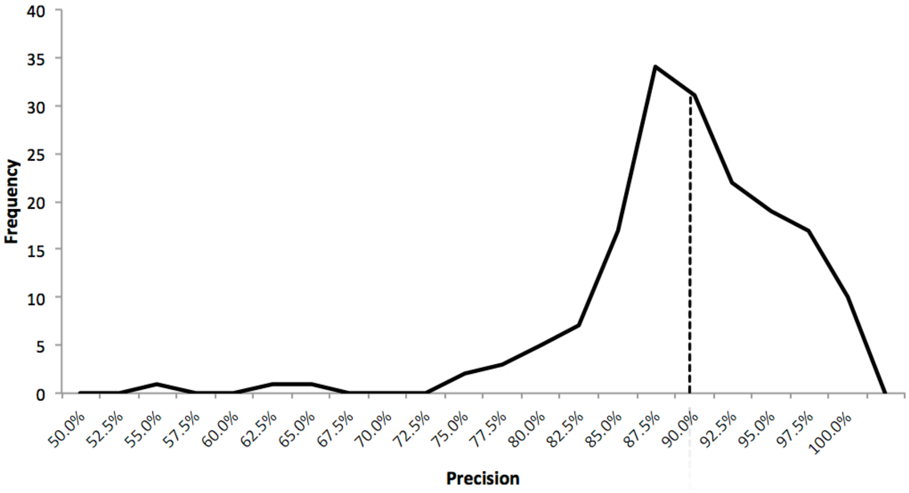


Figure 5. Frequency distribution of relative precision for authentication security

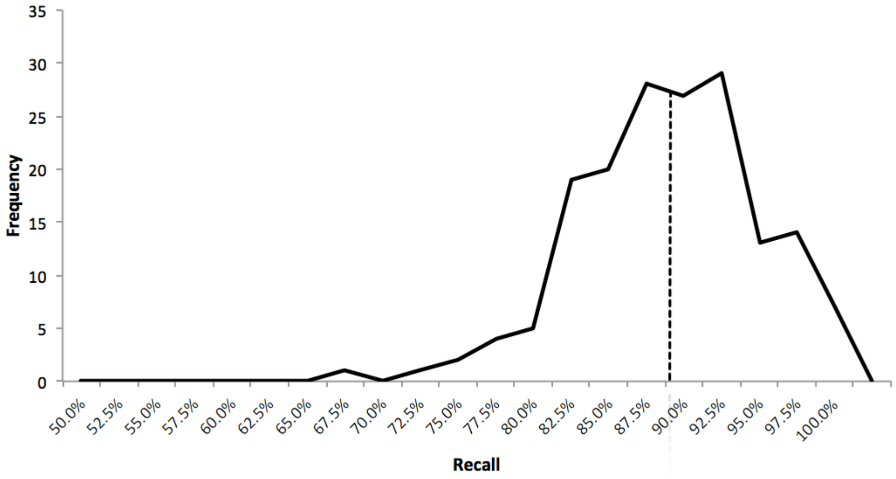


Figure 6. Frequency distribution of relative recall for ease of use

5 Business Model Discussion

The digital economy has provided firms with the potential to experiment with novel value creation mechanisms [52]. Such mechanisms are networked in the sense that they delineate their roles and the economic agents that participate in value creation. In explaining the value chain, the concept of the business model is usually involved. A business model is “a conceptual tool that contains a set of elements and their relationships and allows expressing a company's logic of earning money. It is a description of the value a company offers to one or several segments of customers and the architecture of the firm and its network of partners for creating, marketing and delivering this value and relationship capital, in order to generate profitable and sustainable revenue streams” ([30], p.15).

Therefore, a business model is related to a number of other managerial concepts: it is not just a company's economic model, but also includes operational elements, strategy elements, as well as other key parties in the value chain.

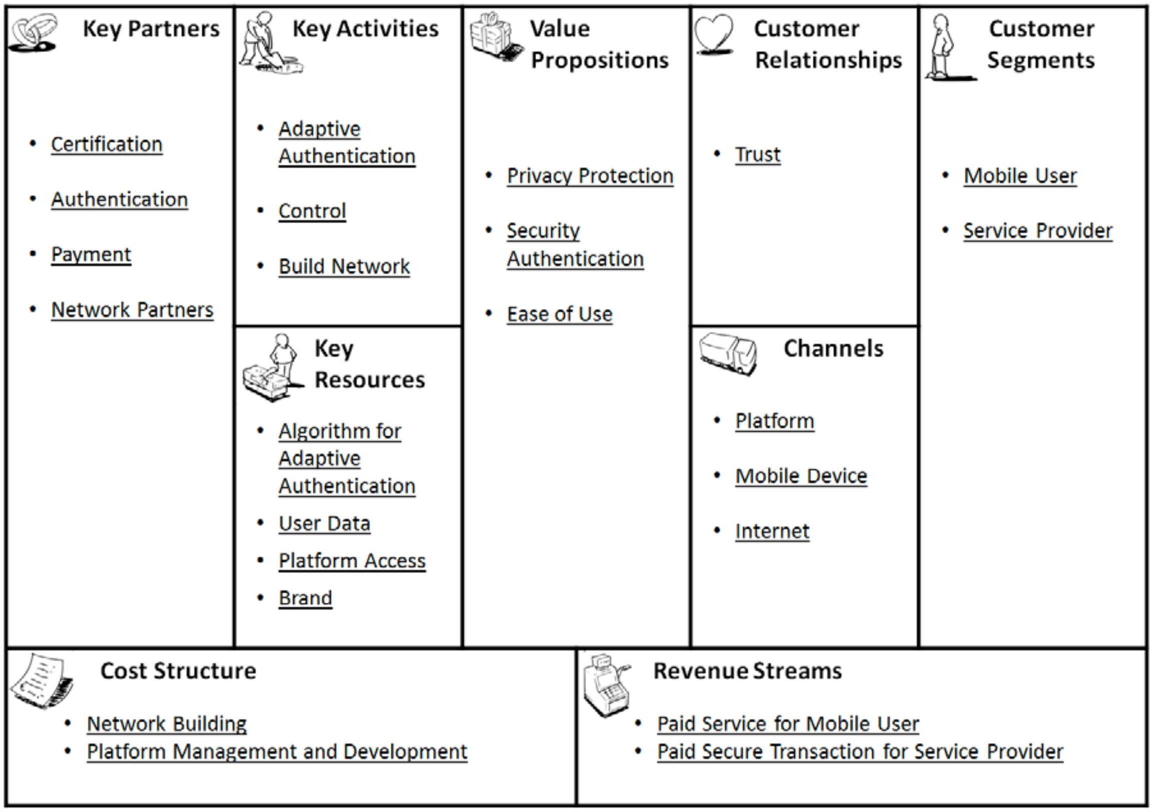
A considerable amount of research about business models has been conducted and reported within the various domains, including business, information systems and strategy. Over the past decade, this is especially true with the growing popularity of digital technologies such as the Internet. More recently, mobile devices have challenged many traditional sectors; consequently, there has been increasing interest in the field of e-business (e.g., [31,32,39]), context-aware mobiles (e.g., [4,15]), mobile service delivery platforms (e.g., [6]) and so on.

Despite such rising interest, academics still fail to agree on exactly what constitutes a business model [29,52]. Many researchers have focused on the business model components that are needed to achieve specific goals. However, there are a few basic components that do emerge in a business model. Based on some key questions that a business model has to address, Osterwalder and Pigneur [31] identified four basic elements: product innovation (what a company has to offer), customer relationships (a company's target customers), infrastructure management (how the proposition can be realized), and financials (the revenue model). Morris et al. [29] identified six components, drawing upon the business model literature: factors related to the offering, market factors, internal capability factors, competitive strategy factors, economic factors and growth/exit factors. In a more recent and relevant study, Reuver and Haaker [15] summarized four key issues in the business model design of context-aware mobile services, namely service domain (i.e., targeting), technology domain (i.e., security), organizational domain (i.e., network openness) and financial domain (i.e., pricing).

To enable business people to easily understand the nature of their business and the essential elements of which it is composed, we used a framework that was proposed by Osterwalder and Pigneur [33]. In this framework, value creation is at the core. The business model canvas can be described by looking at a set of nine building blocks. These blocks were derived from an in-depth literature review of a large number of previous conceptualizations of business models. The framework can serve as a good strategic management and entrepreneurial tool, allowing to test the business model upfront.

This section discusses the business model pattern for a third party in charge of managing the privacy of mobile users and mobile service providers. The following paragraphs use the nine business model elements defined by Business Model Ontology (BMO) to assess the strategic contribution of the ASSO solution for context-aware mobile devices, as shown in Figure 7.

Value proposition: This lies at the center of the business model. It describes which customer problems are solved and why the offer is more valuable than similar products or services from competitors. The context-aware ASSO solution helps to *protect customers' privacy* without sacrificing the level of protection offered. According to Nokia's survey [53], which was drawn from 14 countries and 9,200 mobile phone and Internet users, 82% of respondents viewed privacy as an important topic, whilst more than three quarters of people were concerned about privacy violations. The solution in this study offers value to customers, allowing them to implement privacy in their application, thus reducing their concerns. A second set of values can be drawn from the two performance criteria: *authentication security and ease of use*. On the one hand, a mobile device can authenticate the mobile user through an accurate learning process that has no other costs. Therefore, security authentication is associated with service providers as a main value proposition in the privacy-friendly business model pattern. On the other hand, based on user data analysis associated with the mobile user, ease of use helps manage a privacy profile in one location for multiple services and helps reduce the loss of control that is felt by users when they need a different profile for each service.



Key Partners

- Certification
- Authentication
- Payment
- Network Partners

Key Activities

- Adaptive Authentication
- Control
- Build Network

Key Resources

- Algorithm for Adaptive Authentication
- User Data
- Platform Access
- Brand

Value Propositions

- Privacy Protection
- Security Authentication
- Ease of Use

Customer Relationships

- Trust

Channels

- Platform
- Mobile Device
- Internet

Customer Segments

- Mobile User
- Service Provider

Cost Structure

- Network Building
- Platform Management and Development

Revenue Streams

- Paid Service for Mobile User
- Paid Secure Transaction for Service Provider

Figure 7. Business models for the adaptive single sign on (ASSO) solution

Customer segments: In the BMO, customers are analyzed and separated into groups to help identify their needs, desires, and ambitions (e.g., singles, families). In our pattern, there are two distinct customer segments: the *mobile user who seeks ease of use and the service provider who seeks authentication security*. Since the solution can benefit both mobile users and service providers, our business model pattern is similar to an infomediary between two customer segments.

Customer relationship: This specifies the type of relationship expected by the customer and describes the way it is established and maintained (e.g., promotion, support, individual, or masse). The key to attracting users is to promote the importance of privacy protection and to build a strong *trust relationship* with the customer. In this study, it is defined trust as a willingness to rely on an exchange partner in whom one has confidence [28]. A trust relationship may be built on physical, social, economic, or emotional characteristics. The privacy agent has to demonstrate an awareness of the high value a user has for personal data; it must also show that it cares a great deal for keeping data safe. This relationship is very similar to that of a bank and its customers.

Channel: This illustrates how the customer wants to be reached and by whom the customer is addressed (e.g., the Internet, a store). Service can be personalized either by means of a *platform*, which could be either a *mobile application* or the *Internet*. The ASSO application is based on the mobile device for an end user. Authentication technology would be provided in the form of a service that offers an Application Programming Interface (API) or an application made with a Software Development Kit (SDK).

Key activities: These are used to transform all resources into the final product or service (i.e., through development, production, proprietary process). The key activity of a multi-sided business model is to build and promote a network of users for its platform. To ensure compliance with the users' policies, the privacy risk can be mitigated by implementing and maintaining a set of controls according to security frameworks such as CobiT and ISO 270001, together with privacy guidelines [53]. Moreover, we have introduced the important concept of *adaptive authentication*, which implies unsupervised rule generation. The user can be authenticated by identifying of the user's behavior pattern through machine learning, but if the context-based authentication fails, an adaptive authentication is required.

Key resources: Staff, machines, and proprietary knowledge are required to deliver the value proposition. The most important element for the third party is *user data* and control over access to the *data-sharing platform*. An additional resource is represented by the *brand value*, which allows a trusted relationship with the customer segments. Furthermore, the *algorithm for adaptive authentication* is added as a key resource.

Key Partners: For resources or activities, most businesses depend on an *external partner network* (e.g., logistics, financial), which can offer better quality or a lower price on non-essential components. In order to guarantee the trustworthiness and security of the solution and to be able to certify that the application made for this platform is compliant, we had to gain *certification* by an external provider. The third party also had to develop partnerships with a mobile device manufacturer or network operators in order to realize and deploy the product. To offer additional services, our solution also needed to develop a relationship with a mobile user and a service provider.

Revenue streams: These reflect the value that customers are willing to pay and the way in which they will perform the transaction. In our pattern, two revenue streams associated with the two selected customer segments were switched: thus, the end user pays a *fee to use the application*, and the service provider pays a *fee for each secure transaction* (alternatively, it could buy a license to develop a set of ASSO applications).

Cost structure: This comes under the heading of financial information and should be aligned to the core ideas of the business model. *Network building, platform management and development activities* are all costly services.

6 Implications and Contributions

This research offers several theoretical contributions, as well as practical implications. First, we present an improved solution for SSO using the attributes of mobile users; in other words, contextual information. The findings reveal that the conjoint effect of location and time information would increase authentication, whilst retaining a high level of ease of use. In other words, the ASSO solution increases the level of usability in terms of protecting users' privacy without sacrificing the level of protection on offer. To the best of our knowledge, the study is the first to offer empirical evidence that users' context information can be used to protect their private information. Hence, it opens up new understanding about the current literature in this domain.

Second, whilst we are aware that context is about much more than just location, its other elements are still difficult to identify or measure. As a result, most existing studies only focus on location information or location-based features. This study extends the context concept by combining location and time information. We have utilized real mobile users' data to prove that context data is a unique identifier, thus opening up a new avenue for future research on mobile context-awareness.

Third, this paper proposes a new instantiation of the business model pattern that has valuable implications for practitioners. The new model has privacy at the core of its value proposition, whereas previous instantiations considered privacy as a complementary service to be aggregated with other value propositions. Thus, we present new ways to use privacy as a key component of mobile business models.

Finally, the results of this study also have important implications for managerial practices. The findings suggested a radically approach to reignite the growth and innovation capabilities of their enterprises. In most studies, users' privacy issues that rise in context-aware services are usually considered as a burden. However, the ASSO solution made the privacy as a new value proposition in the core of the BMO. On one hand, it will help to build a strong trust relationship with the customer, and this strong trust relationship will increase the revenue. On the other hand, this value proposition increases the key activities construction, but also increases the cost to build these activities. As a result, the revenue increases the profit, and the cost reduces the profit at the same time. Moreover, our research focused on the new instantiation of the business model pattern which would be of great interest and value in understanding the role of intention of management and customer needs in the context of mobile privacy issues. The proposed suggestions highlighted the objectives such as enterprise's vision, core competencies, strategies, infrastructure, organizational structures, trading principles, as well as operational processes and policies.

7 Conclusion

This paper proposed an ASSO for mobile users to protect their private information by using their contextual information. The contextual data for 168 real mobile users was collected by the Nokia Research Center in Switzerland over the course of one year. Our findings revealed that context, which was measured by combining time and location information, could be viewed as a unique identifier. In previous studies, context-aware technology has evoked concerns about privacy. As location-based applications track users automatically on an ongoing basis, it generates an enormous amount of potentially sensitive information about the owner of the mobile device. However, our solution has proved its great potential in terms of this kind of “threat”. Indeed, we have demonstrated that it can increase authentication at a high level of ease of use, therefore suggesting important implications for business practitioners. The new instantiation of the business model pattern that was proposed by Osterwalder and Pigneur [33] is presented in this thesis, together with a discussion of the nine building blocks for the third party in charge of managing the privacy of mobile users and mobile service providers. The new model has the protection of privacy at its core. It also has two performance criteria (i.e., authentication security and ease of use), which make it different from most previous instantiations. In these, privacy was considered a complementary service.

In the current stage of project development, we acknowledge a set of limitations. The first one concerns the choice of the approach to represent the business model. It should be acknowledged that alternative business model frameworks exist, such as that put forward by Bouwman et al. [11] and Wegmann [49]. These frameworks may be more geared to the mobile context since they can use the mobile device or ICT service as a unit of analysis. We also acknowledge that IBM and Vodafone are currently developing a software solution that is similar to the one proposed here. However, since their solutions are proprietary, we could not include information about their performance.

As future extensions of the solution, we intend to use context-aware technology to implement the ASSO in a mobile application to protect users’ personal information on the Android operating system. We bring to mind the mobile device versus central server debate presented in existing research [25]. Thus, it would be interesting to explore two what-if scenarios that arise when one of the two agents takes the lead.

(1) What if the ASSO is owned by services providers in a situation of coopetition? In this case, the application would mostly reside in the central servers of service providers. These providers would use an ASSO API to develop new applications or they would define an ASSO standard. The mobile user would use a client application on the mobile device that would send the context information needed to obtain authentication. This approach would increase the performance of the authentication algorithm, which could take advantage of the central server’s power. It could also decrease learning time by having access to a larger pool of users’ data. Moreover, the updating of a client’s application would be easier to perform and users’ data could reside on the server, so as not to leave any private information on the mobile device in case of malicious attack.

(2) What if a device-centric authentication is preferred to an ASSO platform? In this case, the ASSO would mostly reside on the users’ mobile devices. The authenticating algorithm would be stored within a mobile application that would establish a secure connection with the service provider after user authentication. This approach would address privacy concerns that might arise under the centralized solution, because in this case users’ data are not stored on the server. Additionally, this approach may act as less of a drain on battery power, because the increased computational effort would be compensated for by the reduction of client-server data exchanges required in the first scenario. In a possible extension of this scenario, users could authenticate each other without the need for a third party. This solution would spread users’ data among peers, thus reducing the success of malicious attacks [34].

Acknowledgements

Authors sincerely acknowledge the Lausanne Data Collection Campaign run by Nokia Research Center Lausanne for providing a large amount of data pertaining to the behavior of individuals and social networks that helped us to achieve the research objectives. Special thanks to Dr. Jialu Shan whose suggestion and very constructive criticism have contributed immensely to the evolution of this project.

References

1. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., and Steggles, P. Towards a better understanding of context and context-awareness. Handheld and ubiquitous computing, Springer (1999), 304–307.
2. Ackerman, M., Darrell, T., and Weitzner, D.J. Privacy in context. Human–Computer Interaction 16, 2-4 (2001), 167–176.
3. Addo, A. The adoption of mobile phone: How has it changes us socially. Issues Bus. Manag. Econ 1, 3 (2013), 47–60.
4. Al-Qirim, N. Context-Aware Mobile Business Model Discovery. Procedia Computer Science 10, (2012), 1180–1187.
5. Anchan, D. and Pegah, M. Regaining single sign-on taming the beast. Proceedings of the 31st annual ACM SIGUCCS fall conference on User services, ACM (2003), 166–171.
6. Becker, A., Mladenowa, A., Kryvinska, N., and Strauss, C. Evolving taxonomy of business models for mobile service delivery platform. Procedia Computer Science 10, (2012), 650–657.
7. Belanger, F., Hiller, J.S., and Smith, W.J. Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. The Journal of Strategic Information Systems 11, 3 (2002), 245–270.
8. Beresford, A.R. and Stajano, F. Location privacy in pervasive computing. IEEE Pervasive computing 2, 1 (2003), 46–55.
9. Biegel, G. and Cahill, V. A framework for developing mobile, context-aware applications. Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications, Percom 2004, IEEE (2004), 361–365.
10. Bolton, R.J. and Hand, D.J. Statistical fraud detection: A review. Statistical science 17, 3 (2002), 235–249.
11. Bouwman, H., Zhengjia, M., Van der Duin, P., and Limonard, S. A business model for IPTV service: a dynamic framework. info 10, 3 (2008), 22–38.
12. Chen, G., Kotz, D., and others. A survey of context-aware mobile computing research. Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College, Hanover, NH, USA, 2000.
13. Cortes, C. and Vapnik, V. Support-vector networks. Machine learning 20, 3 (1995), 273–297.
14. De Clercq, J. Single sign-on architectures. In Infrastructure Security. Springer, Bristol, UK, 2002, 40–58.
15. De Reuver, M. and Haaker, T. Designing viable business models for context-aware mobile services. Telematics and Informatics 26, 3 (2009), 240–248.
16. De Vos, H., Haaker, T., and Teerling, M. Consumer value of context aware and location based mobile services. BLED 2008 Proceedings, (2008), 50–62.
17. Dey, A.K. Context-aware computing: The CyberDesk project. Proceedings of the AAAI 1998 Spring Symposium on Intelligent Environments, (1998), 51–54.
18. Freudiger, J., Manshaei, M.H., Hubaux, J.-P., and Parkes, D.C. On non-cooperative location privacy: a game-theoretic analysis. Proceedings of the 16th ACM conference on Computer and communications security, ACM (2009), 324–337.

19. Hull, R., Neaves, P., and Bedford-Roberts, J. Towards situated computing. First International Symposium on Wearable Computers, IEEE (1997), 146–153.
20. Inglesant, P.G. and Sasse, M.A. The true cost of unusable password policies: password use in the wild. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2010), 383–392.
21. King, N.J. and Jessen, P.W. Profiling the mobile customer–Privacy concerns when behavioural advertisers target mobile phones–Part I. Computer Law & Security Review 26, 5 (2010), 455–478.
22. Krumm, J. A survey of computational location privacy. Personal and Ubiquitous Computing 13, 6 (2009), 391–399.
23. Lewis, K.D. and Lewis, J.E. Web single sign-on authentication using SAML. International Journal of Computer Science Issues 2, (2009), 41–48.
24. Lingras, P. and Butz, C.J. Precision and recall in rough support vector machines. IEEE International Conference on Granular Computing, IEEE (2007), 654–654.
25. Liu, Z., Bonazzi, R., Fritscher, B., and Pigneur, Y. Privacy-friendly business models for location-based mobile services. Journal of theoretical and applied electronic commerce research 6, 2 (2011), 90–107.
26. McCole, P., Ramsey, E., and Williams, J. Trust considerations on attitudes towards online purchasing: The moderating effect of privacy and security concerns. Journal of Business Research 63, 9 (2010), 1018–1024.
27. Mizouni, R., Matar, M.A., Al Mahmoud, Z., Alzahmi, S., and Salah, A. A framework for context-aware self-adaptive mobile applications SPL. Expert Systems with applications 41, 16 (2014), 7549–7564.
28. Moorman, C., Zaltman, G., and Deshpande, R. Relationships between providers and users of market research: The dynamics of trust. Journal of marketing research 29, 3 (1992), 314–328.
29. Morris, M., Schindehutte, M., and Allen, J. The entrepreneur’s business model: toward a unified perspective. Journal of business research 58, 6 (2005), 726–735.
30. Osterwalder, A. The business model ontology: A proposition in a design science approach. 2004.
31. Osterwalder, A. and Pigneur, Y. An e-Business Model Ontology for Mobile E-Business. 15th Bled Electronic Commerce Conference e-Reality: Constructing the e-Economy, (2002), 429–436.
32. Osterwalder, A. and Pigneur, Y. Modeling value propositions in e-Business. Proceedings of the 5th international conference on Electronic commerce, ACM (2003), 429–436.
33. Osterwalder, A. and Pigneur, Y. Business model generation: a handbook for visionaries, game changers, and challengers. John Wiley & Sons, New York, USA, 2010.
34. Pathak, V. and Iftode, L. Byzantine fault tolerant public key authentication in peer-to-peer systems. Computer Networks 50, 4 (2006), 579–596.
35. Pennanen, K., Kaapu, T., and Paakki, M.-K. Trust, Risk, Privacy, and Security in E-Commerce - Google Search. Proceedings of the ICEB + eBRF Conference, (2006).
36. Pevzner, L. and Hearst, M.A. A critique and improvement of an evaluation metric for text segmentation. Computational Linguistics 28, 1 (2002), 19–36.
37. Radha, V. and Reddy, D.H. A Survey on single sign-on techniques. Procedia Technology 4, (2012), 134–139.
38. Rao, B. and Minakakis, L. Evolution of mobile location-based services. Communications of the ACM 46, 12 (2003), 61–65.
39. Raphael, A. and Zott, C. Value creation in e-business. Strategic management journal 22, 6/7 (2001), 493–520.
40. Ryan, N., Pascoe, J., and Morse, D. Enhanced reality fieldwork: the context aware archaeological assistant. Bar International Series 750, (1999), 269–274.

41. Sakimura, N., Jones, M., de Medeiros, B., and Mortimore, C. Openid connect basic client profile 1.0-draft 28. OpenID Connect Specs, 2013. http://openid.net/specs/openid-connect-basic-1_0-28.html.
42. Schilit, B., Adams, N., and Want, R. Context-aware computing applications. in First workshop of Mobile Computing Systems and Applications, IEEE (1994), 85–90.
43. Schilit, B.N. and Theimer, M.M. Disseminating active map information to mobile hosts. Network, IEEE 8, 5 (1994), 22–32.
44. Sebastiani, F. Machine learning in automated text categorization. ACM computing surveys (CSUR) 34, 1 (2002), 1–47.
45. Smith, H.J., Milberg, S.J., and Burke, S.J. Information privacy: measuring individuals' concerns about organizational practices. MIS quarterly, (1996), 167–196.
46. Sterngold, J. Say Goodbye to All Those Passwords-Businessweek. BusinessWeek: Online Magazine, 2011.
47. Suriadi, S., Foo, E., and Jøsang, A. A user-centric federated single sign-on system. Journal of Network and Computer Applications 32, 2 (2009), 388–401.
48. Vapnik, V. The nature of statistical learning theory. Springer Science & Business Media, 2013.
49. Wegmann, A. On the Systemic Enterprise Architecture Methodology (SEAM). Proceedings of the 5th International Conference on Enterprise Information Systems, (2003), 483–490.
50. Westin, A.F. Privacy and freedom. Washington and Lee Law Review 25, 1 (1968), 166–170.
51. Xu, H., Teo, H.-H., Tan, B.C., and Agarwal, R. The role of push-pull technology in privacy calculus: the case of location-based services. Journal of Management Information Systems 26, 3 (2009), 135–174.
52. Zott, C., Amit, R., and Massa, L. The business model: recent developments and future research. Journal of management 37, 4 (2011), 1019–1042.
53. Nokia Siemens Networks: Privacy Survey 2009. 2009.