When context is better than identity: authentication by context using empirical channels

Chen Bangdao, Long Hoang Nguyen and Andrew William Roscoe

Oxford University Computer Science Department
{Bangdao.Chen, Long.Nguyen, Bill.Roscoe}@cs.ox.ac.uk

Abstract. In mobile computing applications the traditional name-based concept of identity is both difficult to support and frequently inappropriate. The natural alternative is to use the context in which parties sit to identify them. We discuss this issue and the ways in which Human Interactive Security Protocols (HISPs) can play a role in enabling this.

1 Introduction

Context arises in an application when one or more entities are acting in a certain situation. For example, one of the most significant types of context is location, which can influence a wide range of decisions about who to connect to across many applications.

We notice that context serves better than identity in some cases. For example, a customer C wants to pay a shop S. In this scenario, C knows he is in *this* shop and wants to pay it, even though he does not know its identity in a conventional sense.

To understand this better, think of the scenarios in which you would be willing to hand over cash: you might trust a merchant by experience or reputation, you may choose to trust him by context, or you may "trust" him to receive payment because you have already received goods or services from him. Note that there is a weaker need for trust if, as with handing over cash, you know that the damage that can be caused by an abuse of trust is strictly limited (i.e. to losing a defined amount of cash).

Therefore we conclude that when it is difficult for ordinary users to correctly verify the identity of whom they pay, context may be better than identity to be used to help users to authenticate the payment. The difficulties of users may consist of two parts:

A. Users lack the necessary knowledge to correctly verify identity;
B. Users can be lazy, especially when the amount of payment is small.

We investigate how the payer (human) can authenticate that his device (typically mobile phone) is connected to the intended payer. This authentication both provides assurance directly to the human and opportunities for improved transaction security such as better authentication of the human's identity.

We assume that there is a low-bandwidth empirical channel from the payee to the human that is not fakeable. This enables the human to play his part in the protocols in Section 3. This is straightforward when the two are in the same place and various ad hoc solutions also work in remote contexts, such as e-commerce. More discussion about defining proper context is made in Section 2.

We note that the payee and the payer usually both require authentication of the other. The payer needs typically to know he is paying the right entity for the transaction he is trying to complete. The payee, or more accurately, the infrastructure supporting payment (such as the banking system) needs to know that the payer is who he claims to be and that he is entitled to make the payment. As we have already noted, the first of these is frequently best attached to context. The latter on the other hand, is a much more typical process and we note that much technology has been developed for this. It is also worth noticing that in most cases the payee itself does not need to have the payer's identity information, rather the assurance of its bank that it will make (or has already made). Therefore the payee, in authenticating the payer, is acting as a proxy for the bank. Current technology typically gives the payee all information such as the credit card number, password/PIN, or at least makes it easy for this to be obtained. *This is undesirable as it offers opportunity for abuses.*

Similar situations can be found in access control examples: in order to pass a check-point CP of building B, user C must submit his credentials (stored on his own device) to CP. In order to protect C's credentials, C needs to authenticate CP. In other words he needs to know that his device is giving information to precisely this CP. But C does not know the identity of the person who stands in front of CP or there are no personnel at all. Therefore only context that C can draw from the situation can help authenticate CP, for example, the location of CP (given there is no other check-points standing there), the logo of CP, C's recognition of CP based on previous experiences, or somebody else C trusts tells C that this is the correct CP. In this case, there is only context rather than actual identity.

Another example can be found in social networks. Social networks are constantly changing our social styles and habits, and they are often considered to be our virtual presences on the Internet. Although different people may have the same name, but the photos they share, the activities they join, the friends they have, and the profiles they present, provide a sophisticated body of context which can allow people to authenticate "who's who". We will investigate in details of how to properly adapt security to the impact of social networks in our demonstrations.

We can further observe that when context is better than identity, authentication by context brings more security than authentication by identity. To fulfil this requirement, we need to firstly obtain all the required contexts which would introduce the following three challenges:

1. It is difficult to define what proper context is or what is context after all.
2. Quite a few contexts cannot be automatically sensed by machines.
3. Some contexts can be easily forged.

The objective of this paper is to clarify the sort of situation in which context-based authentication is most appropriate, as well as giving examples and one technical solution for achieving it. We will be able to demonstrate various implementations and applications of these ideas.

2 Defining proper context

We borrow the definition of context from [5], as we think this best describes the meaning of context in the above cases:

> Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves.

However, this definition does not outline what is proper context that can be used in a specific application, because the scenarios of applications differ in both the nature of applications and the environments of applications.

While context-aware applications are now widely available, we notice that some contexts, for example, logos, images, biometrics like faces, voices cannot be easily recognized by machines (it may require special hardware support as well as large amount of computation); and some contexts just simply cannot be sensed by machines at all, for example, human trust.

This creates a gap between a system and a human in terms of contribution to the process of bootstrapping security. We are looking for a method of authentication that can capture the properties of the process when we try to establish our trust in daily activities as described in Section 1. Such an authentication may be required in two different sorts of situation: remote authentication and on-site authentication. On-site authentication is the most natural way of allowing humans to make their own judgements: humans can see, hear and touch. While remote authentication, can also provide humans an interface to sense if there is a proper proxy, for example, a phone call with a known individual, a PC and an *https* session, or a known social network web-page protected by *https*[1].

Such a method should be resistent to attacks using fake contexts. For example, locations, which usually appear in the form of addresses or GPS data (longitude, latitude, altitude), can be easily forged by presenting the false data; names, logos, photos can be easily forged as well. In order to achieve more pervasiveness, we do not rely on specific hardware or infrastructure support to solve this problem, instead, we assume that humans, with or without enough knowledge, can make a choice between genuine context and false context on behalf of their own risks.

The authors of [8] give an example of using telephone line to authenticate location. Given that a telephone is located in a place where the user is, a phone

[1] By the time we were writing this paper, Facebook had just released its *https* service on 26th Jan, 2011[1].

call can be made to send a challenge to the user, and the user will send back the challenge together with other information back to the service provider via another electronic connection. Such practices are common in financial services, for example, banks or credit card companies make phone calls to verify suspicious card transactions, and we have to answer a few secret authentication questions to get authenticated. We are interested to notice that even though we do not know the caller's phone number, we are still willing to give our secrets to the caller who claimed to be the bank or the credit card company. The context of when and where we receive this phone call, or the fact that the caller knows our names and the transactions we made or being made, persuades us into believing that this call is authentic, even though it is subject to a man-in-the-middle attack.

The discussion of security of hardware and software is beyond this paper. Our focus is that in either remote or on-site authentication, humans are able to find or create an empirical channel which is resistant to man-in-the-middle attacks. The question of whether we trust a telephone line to provide a good empirical channel is clearly open to different answers depending on the nature of threat and risk.

Before introducing our solution, we firstly assume that in any application that has human presence, humans are capable of evaluating their own risks based on the context and then making their own judgements. Such context can be sensed either by human recognition or sensors.

In this way, the problem of defining what is proper context for each application and the problem of sensing contexts can be generalised as the user's choice of whether or not to trust the data received from the empirical channel. By incorporating humans into the design of protocol, we can avoid the difficulties of defining specific scenarios and provide a homogeneous way of incorporating trust into our protocols. In addition, the process of human interaction can be regulated by the design of implementation in order to eliminate unexpected human errors.

3 Authentication by context

The ease of on-site authentication has already been recognised and supported by various technologies, for example, the development of NFC interface on mobile phones provides users a convenient method to create their "bond of trust". A user can touch his mobile phone to a touch pad to complete an authentication: the proximity is used to create such a bond which is similar to the cash payment scenario described in Section 1.

Remote authentication is made by using a proxy, or a "referee". It is a common practice in our daily life that when P_1 meets an important person P_2, presenting a reference letter from a person P_3 that P_2 knows can effectively earn P_2's trust. Therefore the trust from P_2 between P_3 can be used to bootstrap trust between P_2 and P_1 even if P_2 has never met P_1 before. The creation and delivery of such a reference letter is essentially the communication via an empirical channel. For example, Bob makes a phone call to Alice, by recognizing the

caller's phone number or his voice, Alice knows that she is communicating with Bob. In this case, the telephone is a proxy between Bob and Alice. However, the assumption of knowing the caller's phone number or recognizing his voice does not stand in ad-hoc scenarios.

The difficulty of remote authentication arises when two parties can not easily "find" each other, for example, Bob may not know the phone number of Alice, or Alice can not recognise Bob's phone number. In order to solve this problem, we introduce the idea of using social networks[2] which provides a platform where Bob and Alice both know how to "find" each other.

3.1 The impact of social networks

Social networks are being integrated into every aspect of our daily life, for example, nearly every major news web-site has one or more social network plug-ins to allow readers sharing with their friends, many commercial products or services are using them to promote their business, and many people are using them as their main social communication tools. One significant move for social networks is the integration on mobile phones. For example, today the messages sent via Facebook[3] work exactly the same as SMSs. Such a movement facilitates the online communication and attracts more and more people to join. In the future, we could see that all communications are carried out within social networks, for example, our contact-list may be replaced by the social network friend-list, our SMS service may be replaced by the social network messages, and our phone calls will be or could be linked directly to our social network accounts.

The growing pervasiveness and importance of social networks allows us to make the following assumption: in the near future, most people will have an active social network account, which serves as his or her web-presence in the virtual world. And because of the rich context maintained on social networks, we can search a "friend" by name, email address, phone number, location, school name, company name and other context information on social networks. Therefore we further assume that one can easily find one another via social networks.

The Quantified-I The strength of social networks can be further enhanced by integrating sensor technology. For example, by using a set of on-body sensors, one can display his or her physical conditions on social networks. In this way, a human can become quantifiable online: one social network account may consist of a large amount of sensory data from which one's social patterns or identity can be deducted. This can be exploited to provide more convenient and robust authentication services in security applications.

[2] Unless otherwise stated social networks discussed in this paper all refer to semantic web-based social networks.

[3] Users can activate this service on their Facebook.

3.2 Authenticating online identities

The basic question that needs to be asked when using social network web-pages as empirical channels is "how do I know that what I am seeing on the page comes from the person or other entity that I think".

This divides into two sub-questions: how do I know the (e.g. Facebook) page I am seeing is authentic within the social network, and how do I know it belongs to the person I think it does. The first of these problems can be solved by conventional computer security, for example, the *https* service on social networks. The second of these can be solved by answering the question: is this an established friend for which you are certain of the link between page and person. If so, then a secure access to that page is clearly a good empirical channel. More speculation is the idea that we might use "crowd knowledge" about a web-page that we do not have experience of. We give a solution of how to authenticate the instance found on the social network in the following sections.

3.3 Ratings on social networks

An empirical channel can be established by using a "trusted" proxy, which can refer trust between two parties. But it is difficult in practice that two parties can find such a proxy they both trust. We can, however, to create a general proxy to allow humans to determine the extent of trust between each other. In order to achieve this, we assume a rating system virtually exists, and for each session created for bootstrapping security has a minimum requirement of a trust rating. For example, if we have ratings from 1 to 10 (10 means absolute and complete trust), then a user with rating 3 should not be allowed to join in a session with rating 5.

Rating by trust is a common practice in social network researches. The authors in [7] describe a semantic web-based social network, and they developed algorithms to rate the inferred reputation of a node. In their model, a user can rate each others' trustworthiness in general or with respect to a given topic on a 1-10 scale.

Based on the "6-degree of separation" theory proposed by sociologist Stanley Milgram (similar theories of online society can be found in [2, 3]), given the situation in which a vast majority of people have established their online relationships via social networks, there is always a route from person P_a to person P_b via $P_i ... P_{i+j}$, in which P_i is a friend of P_a and P_{i+j} is a friend of P_b. Therefore two strangers can also view each other's rating. Theoretically a person's rating will be more accurate when the number of his or her friends increases.

3.4 The evaluation of risks and trust

In practice, the concept of trust may vary, for example, in a payment scenario, when Bob needs to pay Alice, Bob will make a payment only if he is sure that the one he is paying is the correct instance of Alice, and the credibility of Alice is ignored in this case. Note this is a distinct difference to the ratings on ebay,

where ratings indicate the credibility of the seller. Therefore the trust here is binary: true or false. For example, this is the answer to the second sub-question in Section 3.1. We call it "binary trust".

The term "binary trust" reflects the process of human evaluation of risks and trust of context, which is the essence of authentication by context. In the following section we will discuss how to formally convert the human evaluation of context into security by using Human Interactive Security Protocols (HISPs).

Note the empirical channels we are discussing in this paper all assume "binary trust". Because the protocol we will introduce does not include the decision making process of why to make the payment or proceed with action. However, in practice, especially in group (size > 2) authentication scenarios, HISPs can be extended by allowing more sophisticated ratings in order to reduce conflicts and speed up the human evaluation process. We will discuss this issue in Section 6.

Is trust transitive? Christianson and Harbison argued in [4] that trust may not be transitive in many aspects, for example, when trust transitivity is happened unintentionally. However, the improved transparency as well as the extended boundary of trust on social networks gives users more convenience and confidence in determining whether trust is transitive, for example, one can explicitly choose to trust information from one another in respect of a certain property.

4 Human Interactive Security Protocols

A new family of authentication protocols that are based on human trust and interaction have been introduced over the past few years. These protocols are often referred to as HISPs. They use a non-fakeable Short Authentication String (SAS) exchanged over a low bandwidth empirical channel (denoted $\longrightarrow_E$) to supplement a normal insecure communications medium, usually a high bandwidth channel (denoted $\longrightarrow_N$) subject to the Dolev-Yao attack model [6].

Those protocols satisfy the requirement that we have laid out in Section 1 and 2, which allow humans to receive, check and compare an SAS over an empirical channel. And the information received from an insecure high bandwidth channel will be authenticated based on the result of human interaction. [9–12] are good examples of HISP.

To demonstrate the use of such protocols in real life, we have selected two distinct scenarios to discuss in details: one is mobile payment; the other is registration of on-body medical sensors, in which a patient wants to register his newly purchased on-body medical sensors to a doctor in a remote hospital.

The following symmetric protocol is modified from a group authentication protocol originally described in [9] in order to adapt the above two scenarios. S and R means the sender (the customer or the patient) and the receiver (the merchant or the doctor). In general, we need to authenticate each party in the first place, which is achieved by using an SAS over $\longrightarrow_E$; and to provide security

to protect private data, which is achieved by establishing a symmetric key using an uncertified public key pk_R transferred over $\longrightarrow_N$.

1. $S \longrightarrow_N R : ID_S, INFO_S, hash(hk_S, ID_S)$
2. $R \longrightarrow_N S : ID_R, INFO_R, pk_R, hash(hk_R, ID_R)$
3. $S \longrightarrow_N R : hk_S, \{k\}_{pk_R}$
4. $R \longrightarrow_N S : hk_R$

5a. $R \longleftrightarrow_E S : digest(hk_S \oplus hk_R, ID_S, ID_R, INFO_S, INFO_R, pk_R, k)$
5b. S and R compares the *digest value* with its own version.

digest value represents the SAS that is manually compared by humans. pk_R is later authenticated in step 5b, and if successful, S and R is both assured of the authenticity as well as the security of the symmetric key k. $INFO_S$ and $INFO_R$ includes details of contexts[4] of S and R respectively, for example, a logo, a picture, a recording of voice or film, a name, the amount of a payment, an account number or a few words of description. $INFO_S$ and $INFO_R$ can be displayed together with an input field of digest value, in a way that S and R can verify these details before entering the digest value. Therefore, a successful comparison of digest values allows the conclusion that the identities and the required contexts of S and R are mutually authenticated. $INFO_S$ and $INFO_R$ can then be used in the rest of the application as authenticated data. For details of the security analysis of the protocol please refer to [9].

5 Using a HISP

The requirement of using a HISP is simple: we need to establish two communication channels and use a certain amount of cryptography. The mobile phone is a good platform which provides well developed human interfaces and powerful connectivity as well as computing power. The first communication channel is a relatively high speed but insecure electronic connection, for example, Bluetooth, WiFi, GPRS or 3G. The establishment of the second communication channel depends on specific scenarios. We discuss two scenarios in the following sections.

5.1 Mobile payment

The use of a HISP in mobile payment is appropriate because scenarios of mobile payment are typically much more "ad-hoc" than other kinds of e-payment. The mobile phone in this scenario is used as a "trusted device" which can carry payment account details, for example, we can integrate our bank cards onto mobile phones.

In collocated mobile payment, where two parties are close to each other and authentication is on-site, the empirical channel is the direct interactions by or between human(s). We notice those physical contacts or the physical presence

[4] Usually the contexts included in $INFO_S$ and $INFO_R$ can not be easily sensed by machines.

can effectively allow symmetric authentication: if I can see you, then you can see me. And the protocol can run exactly as the one presented in Section 4.

In remote mobile payment the two parties can be connected via Internet, for example, a mobile phone can connect to the Internet via GPRS/3G, WiFi or any possible medium. We suggest using an *https*[5] web-page to construct the empirical channel in remote mobile payment. This is because it is common in practice that the payer knows the payee's web-site or web-page and it is in accordance to the principles we laid out in Section 1. Telephony can be used only if the payer knows the phone number of the payee.

If the payee is a merchant, we can trust the message displayed on its *https* web-site because *https* authenticates the payee (the merchant) to the payer. In case of a peer-to-peer mobile payment, social networks can be used to authenticate the payee to the payer. Note the *https* service only authenticates that the web-page does belong to a certain social network. And the rating of the web-page indicates the level of authenticity of all the information (the context) displayed. The payer can then evaluate the context and make a decision of whether or not to trust the digest displayed on the payee's web-page.

The use of social networks in remote mobile payment has an obvious benefit. It is not economy and practical to have every Internet user to require a public key certificate, but those who runs a small business online can easily display their social network accounts on their personal web-pages, their customers can therefore be able to obtain security via social networks even without using PKI.

Note in mobile payment scenario, $R \longleftrightarrow_E S$ can be changed to $R \longrightarrow_E S$. As we have discussed in Section 1, in a payment scenario, the payer's identity is less important, and current technology provides the payee convenient methods to have the assurance that he or she has been paid by the payer.

An additional benefit is that the payer can download the payment details automatically from the payee once the secure connection has been established.

5.2 On-body sensor registration

The authentication between the patient and the hospital should be symmetric: the patient needs to make sure that his or her medical data are delivered to his or her hospital; the hospital needs to know the data collected are from the correct patient. As in practice the hospital always require the patient to be registered first before receiving any treatment, we can use telephony or postal service to form the empirical channel from the hospital to the patient, the opposite empirical channel can be made by using the hospital's *https* web-site. The bootstrapping of the connection between sensor and mobile phone can be made by using the "resurrecting duckling" method introduced in [10].

Social networks can be used when the sensor is registered to a person rather than a doctor in a hospital.

[5] We are aware of various attacks against *https* or the web-browser, but as long as it is being used as the security solution in current payment systems, using it does not increase the risks in our proposal.

6 Future research: group authentication by context

In this section we discuss a possible approach to authenticate a group by context. The "binary trust" discussed in Section 3.3 is suitable in the implementation of pair-wise authentication by context, for example, the mobile payment scenario and the on-body sensor registration scenario. But in practice the condition may change in a group (size > 2) scenario because members of a group may not readily to "trust" each other. In addition, there would be conflicts, for example, in a group (size is N) with members M_1 to M_N, M_i's "binary trust" value is true to M_j, but this value may be false to M_k. When the value of N increases, the bootstrapping of security for the group may fail frequently.

Note such conflicts can be easily solved when the members of a group are physically close to each other, for example the scenario of on-site authentication. However, the complexity of solving those conflicts will increase in remote authentication scenarios, especially when the group size is big.

It is reasonable to use more sophisticated rating systems in scenarios of group authentication by context, for example, a rating system with values from 1 to 10. And then the group session will be given a value α. We call it threshold value. Therefore the condition will be changed to $\min(\text{rating}) \geqslant \alpha$. This method has two benefits:

1. It can increase the efficiency of bootstrapping a group by reducing the amount of conflicts as well as mistakes, for example, a scale of 1 to 10 can be rounded up to 0 or 1, if the threshold value of a session is 8, which is rounded up to 1 if we use "binary trust", then we may mistakenly allow members with value of 6 or 7 to join in this session.
2. It provides more context for the user to complete the evaluation process. Because in practice each value in the rating system may have a specific corresponding meaning or description. This is useful because in a remote group session, each member will receive a list of members of his or her group from the initiator, in order to symmetrically authenticate each other, one member M will be required to evaluate all the other members' context information, the ratings and their corresponding descriptions can be served as short but concise context information for M to conduct more accurate and faster evaluation.

However, it also brings a new challenge to future research on security: quantifying the level of risks of a specific task according to the ratings of trust on social networks. Researches on solving this challenge would have a significant impact on the future implementations of HISPs.

7 Conclusion

When the authentication by identity is not available or inconvenient, authentication by context can be exploited by using a HISP, which formalises the human evaluation of context. The advantage as well as the limitation of a HISP is the

use of the empirical channel. Current researches on HISPs focuses on on-site authentication, where empirical channels are the easiest to find or create. Remote authentication frequently suffers from the lack of mutual trust as well as authentic context information. However, with the growing pervasiveness of social networks in our daily life, a new form of trust-based communication system is emerging. By allowing ratings of trust over social networks, we can effectively transform social networks or social network related communication into empirical channels. This allows the wider use of HISPs.

References

1. A. Rice. A Continued Commitment to Security. `http://blog.facebook.com/blog.php?post=486790652130`.
2. L. Adamic. The small world web. In *Research and Advanced Technology for Digital Libraries*, volume 1696 of *Lecture Notes in Computer Science*, pages 852–852. Springer, 1999.
3. R. Albert, H. Jeong, and A. Barabasi. Diameter of the world-wide web. *Nature*, 401:130–131, 1999.
4. B. Christianson and W. Harbison. Why isn't trust transitive? In *Security Protocols*, volume 1189 of *Lecture Notes in Computer Science*, pages 171–176. Springer, 1997.
5. A. K. Dey. Understanding and using context. *Personal Ubiquitous Comput.*, 5:4–7, January 2001.
6. D. Dolev and A. Yao. On the security of public key protocols. *Information Theory, IEEE Transactions on*, 29(2):198 – 208, mar 1983.
7. J. Golbeck and J. Hendler. Accuracy of metrics for inferring trust and reputation in semantic web-based social networks. In *Engineering Knowledge in the Age of the Semantic Web*, volume 3257 of *Lecture Notes in Computer Science*, pages 116–131. Springer, 2004.
8. T. Kindberg, K. Zhang, and N. Shankar. Context authentication using constrained channels. In *Mobile Computing Systems and Applications, 2002. Proceedings Fourth IEEE Workshop on*, pages 14 – 21, 2002.
9. L. H. Nguyen and A. W. Roscoe. Efficient group authentication protocol based on human interaction. In *In Proceedings of the Workshop on Foundation of Computer Security and Automated Reasoning Protocol Security Analysis (FCS-ARSPA*, pages 9–33, 2006.
10. F. Stajano and R. Anderson. The resurrecting duckling: Security issues for ad-hoc wireless networks. In *Security Protocols*, volume 1796 of *Lecture Notes in Computer Science*, pages 172–182. Springer, 2000.
11. S. Vaudenay. Secure communications over insecure channels based on short authenticated strings. In *Advances in Cryptology C CRYPTO 2005*, volume 3621 of *Lecture Notes in Computer Science*, pages 309–326. Springer, 2005.
12. F.-L. Wong and F. Stajano. Multi-channel protocols. In *Security Protocols*, volume 4631 of *Lecture Notes in Computer Science*, pages 112–127. Springer, 2007.