IJDCST @Aug-Sept-2015, Issue- V-3, I-6, SW-20
ISSN-2320-7884 (Online)
ISSN-2321-0257 (Print)

A Novel Social Authenticating Approach for Security using Trustee-Based

Venkata Sai Kumar Bellapu[1], M Samba Siva Rao[2]

[1]M.Tech (CSE), Usha Rama College of Engineering & Technology, A.P., India.

[2]Asst.professor., Dept. of Computer Science & Engineering, Usha Rama College of Engineering & Technology, A.P., India.

Abstract — Among the several backup authentication mechanisms, authenticating users with the help of their friends (i.e., trustee-based social authentication) has been shown to be a promising backup authentication mechanism. A user in this system is associated with a few trustees that were selected from the user's friends. When the user wants to regain access to the account, the service provider sends different verification codes to the user's trustees. The user must obtain at least *k* (i.e., recovery threshold) verification codes from the trustees before being directed to reset his or her password. In this paper, we provide the first systematic study about the security of trusteebased social authentications. In particular, we first introduce a novel framework of attacks, which we call forest fire attacks. In these attacks, an attacker initially obtains a small number of compromised users, and then the attacker iteratively attacks the rest of users by exploiting trustee-based social authentications. Then, we construct a probabilistic model to formalize the threats of forest fire attacks and their costs for attackers. Moreover, we introduce various defense strategies. Finally, we apply our framework to extensively evaluate various concrete attack and defense strategies using three real-world social network datasets. Our results have strong implications for the design of more secure trustee-based social authentications.

Keywords — Social authentication, security model, backup authentication.

I. INTRODUCTION

The most common and traditional technique to authenticate users is asking passwords. Web services (e.g., Gmail, Facebook, and online Bankings) today

most commonly rely on passwords to authenticate users. Unfortunately, two serious issues in this paradigm are: users will inevitably forget their passwords, and their passwords could be compromised and changed by attackers, which result in the failures to access their own accounts.

Therefore, web services often provide users with backup authentication mechanisms to help users regain access to their accounts. Unfortunately, current widely used backup authentication mechanisms such as security questions and alternate email addresses are insecure or unreliable or both. Previous works [5] have shown that security questions are easily guessable and phished, and that users might forget their answers to the security questions. A previously registered alternate email address might expire upon the user's change of school or job. For the above reasons, it is important to design a secure and reliable backup authentication mechanism.

Recently, trustee-based social authentication has attracted increasing attentions and has been shown to be a promising backup authentication mechanism [4], [7], [8]. Brainardet all. first proposed trustee-based social authentication and combined it with other authenticators (e.g., password, security token) as a two-factor authentication mechanism. Later, trustee-based social authentication was adapted to be a backup authenticator [7], [8]. In particular, Schechter et al. designed and built a prototype of trustedbased social authentication system which was integrated into Microsoft's Windows Live ID. Schechter et al. found that trustee-based social authentication is highly reliable. Moreover, Facebook announced its trustee-based social authentication system called Trusted Friends in October, 2011 [8], and it was redesigned and improved to be Trusted Contacts [7] in May, 2013.

However, these previous work either focus on security at individual levels [4] or totally ignore security [7], [8]. In fact, security of users is correlated in trustee-based social authentications, in contrast to traditional authenticators (e.g., passwords, security questions, and fingerprint) where security of users are independent. Specifically, a user's security in trustee-based social authentications relies on the security of his or her trustees; if all trustees of a user are already compromised, then the attacker can also compromise him or her because the attacker can easily obtain the verification codes from the compromised trustees. The impact of this key difference has not been touched. Moreover, none of the existing work has studied the fundamental design problems such as how to select trustees for users so that the system is more secure and how to set the system parameters (e.g., recovery threshold) to balance between security and usability.

II. OUR WORK

In this paper, we aim to provide the first systematic study about the security of trustee-based social authentications. To this end, we first propose a novel framework of attacks that are based on the observation that users' security is correlated in trustee-based social authentications. In these attacks, an attacker initially obtains a small number of compromised users which we call *seed users*. The attacker then iteratively attacks other users according to some *priority ordering* of them.

In an attack trial to a user Alice, if at least k trustees of Alice are already compromised, then the attacker can easily compromise Alice; otherwise the attacker can (optionally) send spoofing messages to Alice's uncompromised trustees to request verification codes, and such spoofing attacks can succeed with some probability. Our attacks are similar to forest fires which start from a few points and spread among the forests. Thus, we call them forest fire attacks.

Second, we construct a probabilistic model to formalize the threats of forest fire attacks and their costs for attackers. For each user, our model computes the compromise probability that the user is compromised after a given number of attack iterations. With those compromise probabilities, our model calculates the expected number of compromised users and treats it as the threat. Moreover, our model quantifies the costs of sending

spoofing messages for attackers. Third, we explore various scenarios where seed users have different properties and introduce strategies to construct priority orderings. For instance, one scenario could be that seed users happen to be appointed as trustees of a large number of users. Furthermore, we discuss a few defense strategies. For example, one strategy is to guarantee that no user is appointed as a trustee of a large number of users.

Third, we explore various scenarios where seed users have different properties and introduce strategies to construct priority orderings. For instance, one scenario could be that seed users happen to be appointed as trustees of a large number of users. Furthermore, we discuss a few defense strategies. For example, one strategy is to guarantee that no user is appointed as a trustee of a large number of users.

Results and Impact of Our Work

We apply our framework to extensively evaluate various concrete attack scenarios, defense strategies, and the impact of system parameters using three real-world social networks. First, we find that forest fire attack is a potential big threat. In particular, when all the users with at least 10 friends in these social networks adopt trustee-based social authentications, an attacker can compromise tens of thousands of users in some cases even if the number of seed users is 0; using a small number (e.g., 1,000) of seed users, the attacker can further compromise two to three orders of magnitude more users with low (or even no) costs of sending spoofing messages. Second, our defense strategy, which guarantees that no users are selected as trustees by too many other users, can decrease the expected number of compromised users by one to two orders of magnitude and increase the costs for attackers by a few times in some cases. Third, we find that, in contrast to existing work [1] where the recover threshold is set to be three, it could be set to be four to better balance between security and usability.

In summary, our key contributions are as follows:

- We propose a novel framework of attacks, which we call forest fire attacks.
- We construct a model to formalize the threats of forest fire attacks and their costs for attackers. Moreover, we explore various attack scenarios and defense strategies.
- We apply our framework to extensively evaluate these attack scenarios, defense strategies, and the impact of system parameters using three real-world social networks.

Our results have strong implications for designing more secure trustee-based social authentications.

First, we overview how a trustee-based social authentication system works. Then, we introduce two basic concepts, i.e., social networks and trustee networks.

A. Trustee-Based Social Authentications

A trustee-based social authentication includes two phases:

- Registration Phase. The system prepares trustees for a user Alice in this phase. Specifically, Alice is first authenticated with her main authenticator

(i.e., password), and then a few (e.g., 5) friends, who also have accounts in the system, are selected by either Alice herself or the service provider from Alice's friend list and are appointed as Alice's trustees.

- Recovery Phase. When Alice forgets her password or her password was compromised and changed by an attacker, she recovers her account with the help of her trustees in this phase. Specifically, Alice first sends an account recovery request with her username to the service provider which then shows Alice an URL. Alice is required to share this URL with her trustees. Then, her trustees authenticate themselves into the system and retrieve verification codes using the given URL. Alice then obtains the verification codes from her trustees via emailing them, calling them, or meeting them in person. If Alice obtains a sufficient number (e.g., 3) of verification codes and presents them to the service provider, then Alice is authenticated and is directed to reset her password. We call the number of verification codes required to be authenticated the recovery threshold.

Note that it is important for Alice to know who her trustees are in the Recovery Phase. Schechter et al. showed that users cannot remember their trustees via performing user studies. Thus, a usable trustee-based social authentication system should remind Alice of her trustees.

Next, we provide details about two representative trustees based Social authentication systems which were implemented by Microsoft [24] and Facebook [7], [8], respectively.

1) Microsoft's Trustee-Based Social Authentication: Schechter et al. [24] designed and built a trustee-based social

Authentication system and integrated it into Microsoft's Windows Live ID service. In the Registration Phase, users provide four trustees. The recovery threshold is three. Moreover, users will be reminded of their trustees.

2) Facebook's Trustee-Based Social Authentication: Facebook's trustee-based social authentication system is called Trusted Friends [8], whose improved version is Trusted Contacts [7]. In the Registration Phase of Trusted Contacts, a

user selects three to five friends from his or her friend list as trustees. The recovery threshold is also set to be three. Facebook does not remind a user of his or her trustees, but it asks the user to type in the names of his or her trustees instead. However, once the user gets one trustee correctly, Facebook will remind him or her of the remaining trustees.

Both trustee-based social authentication systems ask users to select their own trustees without any constraint. In our experiments (i.e., Section VII), we show that the service provider can constrain trustee selections via imposing that no

Users are selected as trustees by too many other users, which can achieve better security guarantees. Moreover, none of these work performed rigorous studies to support the choice of three as the recovery threshold. In fact, our experimental results show that setting the recovery threshold to be four could better balance between security and usability

B. Social Networks and Trustee Networks

We denote a social network as G = (V, E), where each node in V corresponds to a user in the service and an undirected edge (u, v) represents that users u and v are friends. Moreover, in a trustee-based social authentication system, users and their trustees form a directed network. We call this directed network a trustee network and denote it as GT = (VT , ET), where a node in VT is a user in the service and a directed edge (v, u) in ET means v is a trustee of u. One fundamental challenge in trustee-based social authentication is how to construct the trustee network from a social network so that the system is more secure.

III. THREAT MODEL

We first introduce attackers' background knowledge and then a novel family of attacks which we call forest fire attacks.

A. Background Knowledge

We assume that attackers know the trustee network in the target service. The reasonableness of this threat model is supported by two evidences. First, attackers can obtain users' usernames. A username is usually a string of letters, digits, and special characters. Moreover, Bonneau et al. [3] showed that a majority (e.g., 96% in their studies) of websites enable attackers to probe if a string is a legitimate username. Thus, strong attackers, who have enough resources (e.g., a botnet) to perform username probings, can obtain all usernames in the target service. Second, Schechter et al. found, via performing user studies, that users cannot remember their own trustees. Therefore, a usable trustee-based social authentication system must remind users of their trustees. Recall that an account recovery request only requires a username. As a result, an attacker could send account recovery requests with the collected usernames to the service provider which reminds the attacker of the trustees of each user.

Next, we take Facebook as an example to show how an attacker obtains the trustee network. First, Facebook provides an interface1 to test if a user (represented by a username, real name, or email address) is in Facebook. Thus, the attacker can perform username probings to collect Facebook users. Second, the attacker sends account recovery requests to Facebook using the collected names. Recall that Facebook shows all trustees to a user once the user correctly types in one trustee. Moreover, Bilge et al. [4] showed that an attacker can obtain friend lists of around 90% of Facebook users. Thus, the attacker can repeatedly guess the trustees of a user until success. We note that Facebook only allows a user to try around 10 times for typing in the trustees within a short period of time. However, such rate limit cannot prevent a strong attacker from obtaining the trustee network eventually, though it can increase the attacker's cost.

B. Forest Fire Attacks

Our forest fire attacks consist of *Ignition Phase* and *Propagation Phase*.

1) Ignition Phase: In this phase, an attacker obtains a small number of compromised users which we call seed users. They could be obtained from phishing

attacks, statistical guessing's, and password database leaks, or they could be a coalition of users who collude each other. Indeed, a large number of social network accounts were reported to be compromised, showing the feasibility of obtaining compromised seed users.

2) Propagation Phase: Given the seed users, the attacker iteratively attacks other users. In each attack iteration, the attacker performs one attack trial to each of the uncompromised users according to some attack ordering of them. In an attack trial to a user u, the attacker sends an account recovery request with u's username to the service provider, which issues different verification codes to u's trustees. The goal of the attacker is to obtain verification codes from at least k trustees. If at least k trustees of u are already compromised, the attacker can easily compromise u; otherwise, the attacker can impersonate u and send a spoofing message to each uncompromised trustee of u to request the verification code. Schechter et al. found that such spoofing attacks can successfully retrieve a verification code with an average probability around 0.05.

Although the spoofing attacks can help attackers compromise more users, we want to stress that they are optional. We will show in our experiments that an attacker can still compromise a large number of users even if he does not use spoofing attacks to retrieve verification codes in some cases.

3) Example: Figure 1 shows a forest fire attack to a service with 6 users. Note that a good attack ordering can increase the probability that users are compromised and decrease the number of required spoofing messages. In our example, if the attacker performs attack trials with an attack ordering of u5, u6, u4, the attacker needs to spoof both u4 and u6 to compromise u5, which requires two spoofing messages. However, with the attack ordering of u6, u5, u4, the attacker only needs to spoof u4 to compromise u5, which only requires one spoofing message and could succeed with a higher probability.

4) Compromised Users could be Recovered: Users could recover their compromised accounts to be uncompromised after they or the service provider detect suspicious activities of the accounts. For instance, a trustee of u receiving a spoofing message might report to u, who then changes his or her password; the phenomenon that a trustee requests lots of verification codes for different users within a short period of time is a possible indicator of forest fire attacks, and the service provider could then notify the users, whose trustees have requested verification codes, to change passwords.

Moreover, a recovered account could be compromised again in future attack iterations, e.g., when the trustees of the recovered user are still compromised. The process of being compromised and being recovered could repeat for many attack iterations.

However, these previous work either focus on security at individual levels [4], [24] or totally ignore security [7], [8]. In fact, security of users is correlated in trustee-based social authentications, in contrast to traditional authenticators (e.g., passwords, security questions, and fingerprint) where security of users are independent. Specifically, a user's security in trustee-based social authentications relies on the security of

his or her trustees; if all trustees of a user are already compromised, then the attacker can also compromise him or her because the attacker can easily obtain the verification codes from the compromised trustees. The impact of this key difference has not been touched. Moreover, none of the existing work has studied the fundamental design problems such as how to select trustees for users so that the system is more secure and how to set the system parameters (e.g., recovery threshold) to balance between security and usability.

Algorithm 1 Our Model of Forest Fire Attacks

Input: G_T, k, $p_s^{(t)}(v,u)$, n_s, n, $\mathcal{S}$, $\mathcal{O}$, c_s, c_I, and $p_r^{(t)}(u)$.
Output: $n_c(G_T,k,n_s,n,\mathcal{S},\mathcal{O})$, $c(G_T,k,n_s,n,\mathcal{S},\mathcal{O})$.
begin
  //Selecting seed users in the Ignition Phase.
  $\mathcal{S} \longleftarrow \mathcal{S}(G_T,n_s)$
  //Calculating the compromise probabilities.
  //Ignition Phase.
  for $u \in V_T$ **do**
    if $u \in \mathcal{S}$ **then**
      $p_c^{(0)}(u) \longleftarrow 1$
    else
      $p_c^{(0)}(u) \longleftarrow 0$
    end
    $p_a^{(0)}(u) \longleftarrow p_c^{(0)}(u)$
  end
  //Propagation Phase.
  $t \longleftarrow 1$
  $C \longleftarrow 0$
  while $t \leq n$ **do**
    //Constructing an attack ordering.
    $\mathcal{O}^{(t)} \longleftarrow \mathcal{O}(G_T, p_a^{(t-1)}(V_T))$
    for $i = 0$ *to* $\mathcal{O}^{(t)}.size() - 1$ **do**
      $u \longleftarrow \mathcal{O}^{(t)}[i]$
      Apply Equation 4 to u.
      $p_a^{(t)}(u) \longleftarrow 1 - (1 - p_a^{(t-1)}(u))(1 - p_c^{(t)}(u))$
      $p_a^{(t)}(u) \longleftarrow (1 - p_r^{(t)}(u))p_a^{(t)}(u)$
      $c^{(t)}(u) \longleftarrow$ Apply Equation 10
      $C \longleftarrow C + c^{(t)}(u)$
    end
    $t \longleftarrow t + 1$
  end
  //The expected number of compromised users.
  $n_c(G_T,k,n_s,n,\mathcal{S},\mathcal{O}) \longleftarrow \sum_{u \in V_T} p_a^{(n)}(u)$
  //The expected cost.
  $c(G_T,k,n_s,n,\mathcal{S},\mathcal{O}) \longleftarrow c_I + c_s C$
  return $n_c(G_T,k,n_s,n,\mathcal{S},\mathcal{O})$, $c(G_T,k,n_s,n,\mathcal{S},\mathcal{O})$
end

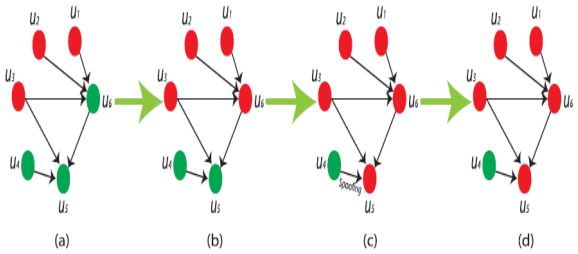

Fig. 1. Illustration of a forest fire attack to a service with 6 users. The shown graph is the trustee network. Recovery threshold is three. Users u5 and u6 have adopted the trustee-based social authentication. The attack ordering is u6, u5, u4. (a) u1, u2, and u3 are compromised seed users. (b) u6 is compromised because three of his or her trustees are already compromised. (c) u5 is compromised because the attacker already compromises his or her trustees u3 and u6 and obtains a verification code from u4 via spoofing attacks. (d) u4 is not compromised because he or she hasn't adopted the service.

IV. Related Work

A. *Social Authentications*

Depending on how friends are involved in the authentication process, social authentications can be classified into trusteebased and knowledge-based social authentications. In trusteebased social authentications [4], the selected friends aid the user in the authentication process. Knowledge-based social authentication, however, asks the user questions about his or her selected friends, and thus friends are not directly involved.

Trustee-Based Social Authentication Systems: Authentication is traditionally based on three factors: something you know (e.g., a password), something you have (e.g., a RSA SecurID), and something you are (e.g., fingerprint). Brainard et al. [4] proposed to use the fourth factor, i.e., somebody you know, to authenticate users. We call the fourth factor trustee-based social authentication. Originally, Brainard et al. combined trustee-based social authentication with

some other factor as a two-factor authentication mechanism. It was later adapted to be a backup authenticator [7], [8]. For instance, Schechter et al. [2] designed and built a prototype of trustee-based social authentication system which was integrated into Microsoft's Windows Live ID system. Moreover, Facebook designed

(b)-(d) Propagation Phase

Trusted Friends in October, 2011 [8], and it was improved to be Trusted Contacts [7] in May, 2013.

Knowledge-Based Social Authentication Systems*:* Such social authentications are still based on something you know. Yardi et al. [9] proposed a knowledge-based authentication system based on photos to test if a user belongs to the group (e.g., interest groups in Facebook) that he or she tries to access. Facebook recently launched a similar photo-based social authentication system [10], in which Facebook shows a few photos of a friend of a user and asks the user to name the friend. Such system essentially relies on the knowledge that the user knows the person in the shown photos. However, recent work has shown, via theoretical modeling [12] and empirical evaluations [11], that photo-based social authentications are not resilient to various attacks such as automatic face recognition techniques, questioning their use as a backup authentication mechanism.

V. CONCLUSION

In this paper, we provide the first systematic study about the security of trustee-based social authentications. First, we introduce forest fire attacks. In these attacks, an attacker first obtains a small number of compromised seed users and then iteratively attacks the rest of users according to a priority ordering of them. Second, we construct a probabilistic model to formalize the threats of forest fire attacks and their costs for attackers. Third, we introduce a few strategies to select seed users and construct priority orderings, and we discuss various defense strategies. Fourth, via extensive evaluations using three real-world social network datasets, we find that forest fire attack is a potential big threat. For instance, with a small number (e.g., 1,000) of seed users, an attacker can further compromise two to three orders of magnitude more users in some scenarios with low (or even no) costs of sending spoofing messages. However, our defense strategy, which guarantees that no users are trustees of too many other users, can decrease the number of compromised users by one to two orders of magnitude and increase the costs for attackers by a few times in some cases. Moreover, the recovery threshold should be set to be 4 to better balance between security and usability. A few future directions include evaluating forest firest attacks on real social authentication systems such as Facebook's Trusted Contacts, designing new attack and defense strategies, and optimizing forest fire attacks given a time constraint.

REFERENCES

[1] L. A. Adamic and E. Adar, "Friends and neighbors on the web," *Social Netw.*, vol. 25, no. 3, pp. 211–230, 2003.

[2] (2013, May). *BadRank* [Online]. Available: http://pr.efactory.de/epr0. Shtml

[3] J. Bonneau and S. Preibusch, "The password thicket: Technical and market failures in human

authentication on the web," in *Proc. 9th Workshop Econ. Inform. Security (WEIS)*, 2010.

[4] J. Brainard, A. Juels, R. Rivest, M. Szydlo, and M. Yung, "Fourth-factor authentication: Somebody you know," in *Proc. 13th ACM Conf. Comput. Commun. Security (CCS)*, 2006.

[5] J. Podd, J. Bunnell, and R. Henderson, "Cost-effective computer security: Cognitive and associative passwords," in *Proc. 6th Australian Conf. Comput.-Human Interact.*, 1996.

[6] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[7] (2013, May). *Facebook's Trusted Contacts* [Online]. Available: goo.gl/xHmVHA

[8] (2011, Oct.). *Facebook's Trusted Friends* [Online]. Available: goo.gl/KdyYXJ

[9] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Zhao, "Detecting and characterizing social spam campaigns," in *Proc. Internet Meas. Conf. (IMC)*, 2010.

[10] E. Gilbert and K. Karahalios, "Predicting tie strength with social media," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2009.

[11] N. Z. Gong *et al.*, "Evolution of social-attribute networks: Measurements, modeling, and implications using Google+," in *Proc. ACM Conf. Internet Meas. Conf. (IMC)*, 2012.

[12] P. Jaccard, "Étude comparative de la distribution florale dans une portion des Alpes et des Jura," *Bulletin Soc. Vaudoise Sci. Naturelles*, vol. 37, no. 1, pp. 547–579, 1901.

[13] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2003.

[14] H. Kim, J. Tang, and R. Anderson, "Social authentication: Harder than it looks," in *Proc. Financial Cryptography (FC)*, 2012.

[15] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *Proc. 19th Int. Conf. World Wide Web (WWW)*, 2010.