Information Systems 71 (2017) 68–77

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/is

The Social Relation Key: A new paradigm for security

Sihyun Jeong, Jaehoon Lee, Junhyun Park, Chong-kwon Kim*

Dept. of Computer Science and Engineering, Seoul National University, Gwanak-gu, Seoul 151–744, Republic of Korea

ARTICLE INFO

Article history:
Received 23 June 2017
Revised 7 July 2017
Accepted 7 July 2017
Available online 18 July 2017

Keywords:
Online social network
Security key
SMS
Twitter
Spam
Authentication

ABSTRACT

For the last decade, online social networking services have consistently shown explosive annual growth, and have become some of the most widely used applications and services. Large amounts of social relation information accumulate on these platforms, and advanced services, such as targeted advertising and viral marketing, have been introduced to exploit this social information. Although many prior social relation-based services have been commerce oriented, we propose employing social relations to improve online security. Specifically, we propose that real social networks possess unique characteristics that are difficult to imitate through random or artificial networks. Also, the social relations of each individual are unique, like a fingerprint or an iris. These observations thus lead to the development of the Social Relation Key (SRK) concept. We applied the SRK concept in different use cases in the real world, including in the detection of spam SMSes, and another in pinpointing fraud in Twitter followers. Since spammers multicast the same SMS to multiple, randomly-selected receivers and normal users multicast an SMS to friends or acquaintances who know each other, we devise a detection scheme that makes use of a clustering coefficient. We conducted a large scale experiment using an SMS log obtained from a major cellular network operator in Korea, and observed that the proposed scheme performs significantly better than the conventional content-based Naive Bayesian Filtering (NBF). To detect fraud in Twitter followers, we use different social network signatures, namely isomorphic triadic counts, and the property of social status. The experiment based on a Twitter dataset again confirmed the feasibility of the SRK. Our codes are available on a website[1].

1. Introduction

Over the last decade, we have witnessed explosive growth in online social networking services (SNS), such as Facebook, Twitter, Ren-Ren, and Sina Weibo. These started as microblogs or online replicas of off-line social interaction; and since then, social networking services and social media have rapidly changed the landscape of Internet use. Social media services are now a fast artery of news propagation. Furthermore, many online systems, such as audio/video contents providers, question-and-answer sites, review and rating systems, and e-commerce systems, increasingly support social media to provide product information, or to improve customer retention. Before the proliferation of social media, telephone and cellular companies were the primary points at which social relation information would accumulate, but now many companies that support online social interactions can collect an abundance of social relation information that can open unprecedented opportunity to develop novel services.

The everyday lives of people increasingly depend on online access to information and services. Many first hear news through friends (or followers) rather than through mass media [1], and we frequently consult online reviews and ratings before making purchase decisions [2]. Even though online systems certainly make our lives more efficient and convenient, they are not without drawbacks. One drawback is that of online security. Online security can be susceptible to a number of problems, and as a rule of thumb, attacks evolve as new online services are introduced. Since we first hear news from un-authorized sources, such as Twitter, attackers propagate false rumors, some of which are maliciously fabricated. In addition, many users consult and trust online information including reviews and ratings, so systematic campaigns with the intention to promote or undermine the reputations of certain products, affiliations, and public figures have been increasingly carried out. In contrast to earlier attacks that were performed rather naively for leisure, recent attacks and information contortion have been motivated by monetary gain, and have been carried out by adept and vicious offenders, who overwhelm average security operators.

* Corresponding author.
E-mail addresses: sihyunj@snu.ac.kr (S. Jeong), jhlee@popeye.snu.ac.kr (J. Lee), jhpark041@popeye.snu.ac.kr (J. Park), ckim@snu.ac.kr, ckkim@popeye.snu.ac.kr (C.-k. Kim).

[1] https://github.com/sihyunj/TSP-SS-SRK

http://dx.doi.org/10.1016/j.is.2017.07.003
0306-4379/

A plethora of methods have been proposed to detect and prevent attacks. Taking spam e-mail detection as an example, content-based methods including Nave Bayesian Filtering (NBF) and SVM have been introduced and widely deployed in production e-mail systems. Also, network operators eagerly contribute to lists of spammers' IP address for blacklisting [3]. However, attackers can evade these security measures by avoiding using prohibited words, or by utilizing IP address spoofing. Therefore, security operators and attackers engage in an endless arms race: once a security person devises a new defense, attackers come up with new attacks that can elude the security checks. In addition to the lack of generality, many security solutions are optimized for a specific service. Content-based spam e-mail detection methods are not very effective for spam SMS detection because SMS messages, mostly less than or equal to 140 characters, use the same corpus as non-spam SMS messages and also contain characters in an image form and URLs. To terminate the arms race between operators and attackers, we aim to devise security measures that are difficult to circumvent and have general applicability.

In this paper, we propose using social relation information to design security solutions. As mentioned before, the availability of abundant social relation information has opened opportunities to create new services, and several applications that exploit social relation information have already appeared. Most of these applications use social relation to improve monetary gain. One application is Influence Maximization (IM) [4], which identifies most of the important users who may maximize the sales of newly-introduced goods along established social relations. Viral marketing also utilizes social relations to expedite and to expand the propagation of advertisements. Various types of link prediction schemes have been proposed to recommend plausible friends [5,6]. More recently, researchers have devised clever recommender systems that extract social relation features to identify similar users [7,8].

In contrast to previous methods that utilize social relation information for commercial purposes, we aim to devise security schemes that use social relations. Like fingerprints or irises that are unique for each person, every person has a unique set of social relations. Due to this uniqueness, we can use social relations as an authentication key. We call the set of social relations for each individual a Social Relation Key (SRK). Several systems have already used social relations as a supplementary authentication tool. For example, Facebook requests users to list several friends, and asks if the user is indeed their friend [9]. This is a use case with direct adoption of social relation information for security purposes. However, a request for explicit social relations can be cumbersome, and can be easily thwarted. Instead of implementing an explicit request for social relations, we propose using social relations naturally, and in unobtrusive ways. Ideally, the security mechanisms should detect attackers and attack-related activity during normal use, without the need for requests for additional information or interruptions.

As far as we are aware, the idea of actively using accumulated social relation information to design security solutions has not been proposed before. Several researchers proposed a scheme that detects attack-related activity according to abnormal behavior, such as plural review postings in a short time span [10]. However, behavior-based schemes do not use individual social relations, but rather depend on the average habits of individuals. Sybil detection methods based on a random walk implicitly use clustering in a social network. However, these methods do not use individual social relations, but use the global property of clustering. Of course, global social network properties, including clustering and small world phenomena, are characteristics useful in the design of security solutions. However, we believe that social network properties that are observable at the level of the individual or ego-network, including homophily, clustering coefficient, balancedness, status, etc., provide stronger and more robust grounds to develop security solutions.

This paper proposes the overarching idea of using the SRK in various security solutions. We devised two schemes based on the SRK and applied them to security problems in two different security research areas to test the feasibility and effectiveness of the proposed idea. The first problem involves detecting SMS spam, and the second is to detect spam in Twitter accounts. The spam SMS detection experiment used real voice calls and SMS data obtained from one of the major cellular operators in Korea. We devised a spam-detection mechanism based on the clustering coefficient of the recipients. Even though we use a constrained social network derived from a two-week long communication history, the proposed scheme performs significantly better than a conventional content-based spam detection scheme. We also designed a Twitter spam detection method. We suggest the use of two social related features, triad counts and social status, that are easily obtainable from each individual's ego-network. Ego-networks of a spammer and a legitimate user present discriminating power on these two features. When compared to a previous study [11,12], our method based on integrated social features is superior in terms of a notably high true positive and low false positive.

The contributions of this paper are as follows.

- We propose using social relations for security purposes. We develop the concept of the SRK, which may be able to spawn various applications.
- To show the feasibility and effectiveness of the proposed SRK concept, we implemented a spam SMS detection scheme with the SRK. In addition, we developed a scheme that detects Twitter follower spam using two social network features: the triad counts and social status.
- We conducted experiments using real-world large-scale social network data. In particular, the cellular network data is one of the largest and most recent sets of data, and our experimental results indicated the feasibility of the umbrella idea and the effectiveness for both SMS detection and Twitter follower spam detection.

The rest of this paper is organized as follows. Section 2 provides an overview of several security solutions for SMS and Twitter. Section 3 introduces the basic SRK model, while Section 4 illustrates the proposed approach for SMS Spam detection, and presents the results of the performance obtained with a rigorous experiment using real-world data. Section 5 delineates the details of the proposed approaches for Twitter Spam detection using the triad count and social status. Section 6 concludes the paper, and suggests interesting topics for future research.

2. Related work

2.1. SMS Spam filtering

2.1.1. Content-based SMS spam filtering

Since Email and SMS share fundamental characteristics, in that communication consists of exchanging text messages and various attachments, existing methods can be easily adopted for SMS without extensive modifications. The Nave Bayesian filter [13,14] and Support Vector Machine (SVM) [13,15–17] are well-known approaches in this category. The Nave Bayesian filter determines a bag of words that occur in spam and non-spam messages in advance. Then by computing the Bayesian inference, it classifies a message as spam or non-spam, according to the probability that a message is spam. SVM is a famous method for machine learning, and it is also introduced here to clear out SMS Spam. Although it shows a high classification performance when compared to that of Bayesian filters, it has a fundamental limitation due to its high