Future Generation Computer Systems 49 (2015) 77–93

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

TruSMS: A trustworthy SMS spam control system based on trust management

Liang Chen[a], Zheng Yan[a,b,*], Weidong Zhang[a], Raimo Kantola[b]

[a] The State Key Lab of ISN, School of Telecommunications Engineering, Xidian University, Xi'an, China

[b] Department of Communications and Networking, Aalto University, Espoo, Finland

HIGHLIGHTS

- TruSMS, a trustworthy SMS spam control system design and implementation.
- A convincing performance evaluation to show TruSMS's trustworthiness.
- A TruSMS business model for practical deployment.

ARTICLE INFO

Article history:
Received 29 November 2013
Received in revised form
9 June 2014
Accepted 27 June 2014
Available online 9 July 2014

Keywords:
Trust management
Spam control
Traffic analysis
Trust mining
Performance evaluation

ABSTRACT

The fast growth of mobile networks has greatly enriched our life by disseminating information and providing communications at any time and anywhere. However, at the same time, when people gather and exchange useful information, they also receive unwanted data and contents, such as spam and unsolicited commercial advertisements. SMS (Short Message Service) spam is one typical example of unwanted contents, which has caused a serious problem to mobile users by intruding their devices, occupying device memories and irritating the users. More critically, some of these fraudulent messages deceive users and cause them incalculable loss. SMS spam control has become a crucial issue that impacts the further success of mobile networks. A number of researches have been conducted to control unwanted contents or traffic, some are based on trust and reputation mechanisms. But the literature still lacks an effective solution for SMS spam control. In this paper, we present the design and implementation of an SMS spam control system named TruSMS based on trust management. It can control SMS spam from its source to destinations according to trust evaluation by analyzing spam detection behaviors and SMS traffic data. We evaluate TruSMS performance under a variety of intrusions and attacks with a prototype system implementation. The result shows that TruSMS is effective with regard to accuracy, efficiency and robustness, which imply its trustworthiness.

1. Introduction

Mobile networks have been playing an important role for information interaction, resource sharing, and social communications. They have become a global platform for the provision of various applications and services, such as mobile Email, Short Message Service (SMS), mobile commerce, multimedia communications and mobile social networking. All of these bring us a great convenience and become more and more indispensable in our modern life. On the basis of mobile cellular network and local connection technologies, billions of mobile devices can connect the Internet and access its services and applications, which satisfy people's needs and enrich their life.

However, the mobile network has also become a channel to transmit a lot of unwanted contents. The unwanted content is malicious, harmful or unexpected digital information for its receivers, for example, malware, virus, spam, intrusion, and unsolicited commercial advertisement. Thus, it could greatly annoy mobile users and intrude or infect their devices. The unwanted content only benefits its source, but burdens both users and network service providers by adding extra load into the network, which greatly increases the possibility of normal traffic congestion. It consumes network and computing resources in a way that does not benefit its receivers, thus it should be controlled, filtered or discarded

* Corresponding author at: The State Key Lab of ISN, School of Telecommunications Engineering, Xidian University, Xi'an, China. Tel.: +86 18691958048.
E-mail addresses: 772029466@qq.com (L. Chen), zyan@xidian.edu.cn (Z. Yan), wdzhang@xidian.edu.cn (W. Zhang), raimo.kantola@aalto.fi (R. Kantola).

http://dx.doi.org/10.1016/j.future.2014.06.010
0167-739X/© 2014 Elsevier B.V. All rights reserved.

during transmission from its source to destinations. Recently, mobile networks have been suffering from rampant unwanted contents, such as SMS spam, and the problem is worsening. In China, statistics show that each mobile user received 10.7 SMS spam messages on average per week. Most of the SMS spam is advertisements, which occupy 75.2% of the total SMS spam [1]. In addition, 66.3% of mobile users have received fraud SMS, and the numbers are still increasing. Taking measures to control SMS spam becomes crucial.

In the literature, quite a number of solutions have been proposed for unwanted traffic or content control. Trust management is one of them. It is feasible and effective to collect evidence and through trust and reputation evaluation to figure out a strategy for unwanted traffic or content filtering and control. Trust management can also utilize symbolic representation of social trust to aid decision-making. It enables users to make their own judgment for the purpose of unwanted traffic control (UTC), instead of just relying on passive software detection. In the literature, we can find a number of mechanisms based on trust and reputation management that control spam [2–9], spim (i.e., Instant Messaging spam) [10], SPIT (Spam over Internet Telephony) [11] and web page spam [12–15]. Seldom, people apply trust management to control SMS spam. Specifically, a robust and effective, in short trustworthy SMS spam control system is still lacking. On the other hand, existing anti-SMS-spam solutions are not very effective in practice. The approaches based on a blacklist suffer either from inaccurate filtering of useful SMS messages or ignoring real spam if corresponding keywords are not listed in the blacklist. The methods based on text classification are not effective for filtering SMS spam due to the specific nature of SMS messages, e.g., using non-standard words.

Though researchers have proposed many methods to control unwanted traffic and contents, there is still little in depth research on trustworthy SMS spam control based on trust management [16]. In this paper, we design and develop TruSMS, a trustworthy SMS spam control system based on trust management. We design a trust scheme for controlling SMS spam according to trust evaluation by analyzing spam detection behaviors and SMS traffic data at both mobile devices and network Service Providers (SP). The system can detect malicious or indifferent detection behaviors, thus encourage good behaviors of mobile users and SPs and punish their bad behaviors with regard to SMS spam control. Herein, malicious behavior refers to the behaviors that are harmful to the correct control of SMS spam. Indifferent behavior refers to selfish behaviors that hide evidence, ignore social responsibility or always show a neutral attitude and that can therefore influence evidence collection for SMS spam control and filtering. We implement the system using Android phones as mobile devices to send and receive normal SMS or SMS spam. The prototype system applies a centralized server to evaluate trust of each system entity by collecting, aggregating and processing SMS spam detection reports from mobile devices and SMS traffic analysis reports from SPs. Based on the trust evaluation, the system can identify the sources of SMS spam and malicious or indifferent mobile devices. We further evaluate TruSMS performance under a number of generally applied malicious attacks based on a real SMS dataset, such as hide evidence attack, bad mouthing attack, on–off attack and conflict behavior attack in order to demonstrate its efficiency, accuracy and robustness [17,18]. The performance evaluation based on a close-to-real prototype system makes our evaluation results convincing and reliable. Therefore, the contributions of this paper can be summarized as three folds: (a) TruSMS, a trustworthy SMS spam control system design and implementation; (b) a convincing performance evaluation method applied to show TruSMS's trustworthiness by evaluating the main aspects of trust based on a prototype system; (c) discussions on TruSMS adoption and its deployment business model from an economic point of view.

The rest of the paper is organized as follows. Section 2 gives a brief review on SMS spam control, unwanted traffic control through trust management, and evaluation methods for testing the performance of trust management solutions. In Section 3, we introduce TruSMS system design and algorithms. Section 4 describes TruSMS implementation and its principal functionalities. Section 5 reports TruSMS performance evaluation results. We further discuss privacy and social acceptance issues of TruSMS in Section 6. Finally, conclusions and future work are presented in Section 7.

2. Related work

2.1. SMS spam control

There are mainly two types of techniques currently applied to filter SMS spam: Black and White List and Text Classification [19]. Black and White List allows mobile users to make a blacklist and a whitelist that contain the phone numbers or keywords used by an SMS spam filtering system. Thereby, every SMS coming from those phone numbers or containing the keywords in the blacklist will be put into a spam folder. One way in this approach is to filter SMS spam by comparing the words in SMS messages with the keywords saved in the blacklist. This technique has been practically deployed in mobile phones, e.g., an SMS Spam Manager running in Nokia Symbian phones and a Spam SMS Blocker running in Android phones. However, this technique does not perform very well in many situations because it depends on the keywords listed in the blacklist. Obviously, accuracy of SMS spam filtering and control is an issue. Some useful SMS messages could be filtered and some real SMS spam cannot be controlled if the corresponding keywords are not listed in the blacklist. On the other hand, Text Classification distinguishes SMS spam from other messages based on message content. It can be applied at a mobile cellular operator to block SMS spam or on a smart phone to filter SMS spam. Text Classification relies on the patterns of SMS spam. Some examples of patterns are word occurrences, length and frequency of words in messages. Text Classification uses pattern recognition algorithms such as Naive Bayes, Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Tree, *k*-Nearest Neighbor (*k*NN), and Hidden Markov Model (HMM) for SMS spam filtering. Deng and Peng proposed a method to filter Chinese SMS spam with a naïve Bayesian classification algorithm by introducing such attributes as the length of the SMS and rules found by statistics into an attribute set [20].

Although there are a number of existing anti-SMS-spam solutions as described above, they may not be very effective in practice. This is due to the nature of SMS messages. For example, the SMS messages have significantly less characters than e-mails. A standard SMS message can only contain 160 characters. So people tend to use non-standard words in their messages, such as "how r u (how are you)", and "asap (as soon as possible)", which make it difficult for the above mentioned methods based on words or semantic contents of messages to accurately filter SMS spam. However, a trust management based solution, such as TruSMS can overcome such a shortcoming because user subjective feedback plays an essential role in the trust evaluation on SMS sources. In addition, the anti-SMS spam toolkit installed in the mobile devices (as described above) can greatly help TruSMS automatically generate detection reports.