Future Generation Computer Systems ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

# Privacy-preserving trust management for unwanted traffic control

Lifang Zhang [a], Zheng Yan [b,a,*], Raimo Kantola [a]

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

[b] State Key Laboratory on Integrated Services Networks, School of Cyber Engineering, Xidian University, China

## HIGHLIGHTS

- A trust management system controls unwanted traffic with privacy protection.
- The system complies with legal requirements and overcomes internal attacks.
- Evaluation on extra costs introduced by privacy protection shows system practicality.

## ARTICLE INFO

*Article history:*
Received 15 November 2015
Received in revised form
26 April 2016
Accepted 29 June 2016
Available online xxxx

*Keywords:*
Privacy preservation
Trust management
Unwanted traffic control
Intrusion detection

## ABSTRACT

The pervasive use of the Internet has caused an incredible growth of unwanted traffic, such as spam, malware and malicious intrusions. Controlling unwanted traffic based on trust and reputation mechanisms has invited significant and rigorous research in recent years. However, few of existing solutions concern and preserve the privacy of Internet hosts that report suspicious attacks. They cannot fulfill legal requirements, and are therefore impractical. In this paper, we propose a privacy-preserving trust management system for unwanted traffic control by applying homomorphic cryptosystem. The proposed system protects privacy, which is proved to be one-way and semantically secure against chosen-plaintext (IND-CPA) attacks if Computational Composite Residuosity Assumption (CCRA) holds. The system is implemented and its performance is extensively examined in terms of computation complexity, communication overhead and storage consumption. The result shows the effectiveness and practicality of our system to preserve the privacy of Internet hosts in the detection and control of unwanted traffic.



## 1. Introduction

The Internet has witnessed an incredible growth in its pervasive use. People are enjoying unprecedented convenience brought by it. However, when users benefit tremendously from the Internet, increasing volumes of unwanted traffic, such as spam, malware and vicious intrusions, also trouble them. Such incidents caused by unwanted traffic undoubtedly increase public worries on the Internet security. Thus, developing an efficient solution to control the unwanted traffic in the Internet, as well as the next generation wireless networks has become a crucial task that brooks no delay.

### 1.1. Motivations

To block unwanted traffic, techniques such as firewalls, network monitoring and intrusion detection systems (IDS) have been widely used, achieving certain positive effects. Moreover, a number of new approaches have been put forward in order to control rapidly evolving spam, malware, and DDoS attacks in the Internet. A very promising approach among them is trust and reputation management [1]. As a result, different variants of trust and reputation mechanisms have been proposed to control different types of unwanted traffic, such as email spam [2–11], Instant Messaging Spam (i.e., spim) [12–15], Spam over Internet Telephony (SPIT) [16–19] and web page spam [20–23]. However, none of them preserved the privacy of spam reporting hosts, which prevents existing solutions from being able to be legally deployed in practice due to user privacy concerns and privacy demands from legal policies.

In our previous work, we proposed a Global Trust Management (GTM) system, which executes accurate, effective and robust Unwanted Traffic Control (UTC) based on trust evaluation and management on each network entity [1,24,25]. However, without privacy protection, unwanted traffic reports in the GTM system can be disclosed to Internet Service Providers (ISPs), thus malicious hosts can easily collude with ISPs to block or modify legitimate and useful reports, degrading the performance of the GTM system.

* Corresponding author at: State Key Laboratory on Integrated Services Networks, School of Cyber Engineering, Xidian University, China.
*E-mail addresses:* lifang.zhang@aalto.fi (L. Zhang), zheng.yan@aalto.fi, zyan@xidian.edu.cn (Z. Yan), raimo.kantola@aalto.fi (R. Kantola).

http://dx.doi.org/10.1016/j.future.2016.06.036
0167-739X/

Moreover, trust management incorporates intrusion reporting, which could impact the privacy of reporters. An incident report is an accusation against a remote party. In many cases such an accusation may cause a quarrel, which is not in the interests of the reporter. The privacy of spam reporters herein concerns their legal rights to have the reporting contents protected from being accessed and modified by any unauthorized network entities.

In addition, ISPs cannot perform traffic monitoring and controlling without sufficient evidence according to current privacy laws. For example, the European privacy law allows an ISP (same goes with any other parties) to monitor traffic only when the ISP network itself is under an attack or the attack is so severe that a part of ISP network is under imminent threats. The ISP cannot help a customer network by providing network monitoring (in plain text) without sufficient evidence to support this action. Moreover, the privacy of network hosts has gained increasing attention in many countries recently. However, existing work [2–12,20–23] and our previous work [1,25] for UTC based on trust management cannot fulfill these legal requirements. In short, user privacy concerns and privacy laws necessitate privacy-preserving trust management for UTC in reality.

### 1.2. Main contributions

This paper aims to solve the privacy issues mentioned above by designing a UTC system based on trust management with privacy preservation. The system is based on our previous GTM system [1,24,25], which neglected the privacy concerns for spam reporting hosts. Our system evaluates trust according to the analysis of traffic data and its processing behaviors at both hosts and ISPs. In addition, a centralized Global Trust Operator (GTO) is applied to evaluate trust of each system entity by collecting, aggregating and processing detection/monitor reports from hosts and ISPs based on unique anonymous IDs. On the basis of trust evaluation, the system can identify the sources of unwanted traffic and malicious or indifferent hosts in terms of intrusion detection.

The system ensures the privacy of spam reporting hosts by applying Homomorphic Encryption (HE) on host reports sent to ISPs. In ISPs, the encrypted reports are aggregated without decryption. Then, the aggregation result is sent to GTO to decide which entity should be monitored at ISPs. During this process, ISPs have no idea of the plain values of host reports, thus guaranteeing the privacy of reporting hosts. To decide which entity is suspicious and should be monitored, GTO performs decryption on the aggregation result of ISPs without knowing the real IDs (i.e. IP addresses) of hosts and ISPs since only anonymous IDs are disclosed to GTO in the designed system. GTO then issues the anonymous IDs of suspicious hosts to ISPs and authorizes them to monitor the traffic of the suspicious hosts at ISPs, where the real IDs of the suspicious hosts can be recovered. If the result of ISP monitoring is serious, the ISP reports it to GTO along with host detection reports regarding the same suspicious host. Those reports are then processed at GTO for trust evaluation in order to figure out real unwanted traffic sources. According to an announced greylist of unwanted traffic sources by GTO, ISPs perform unwanted traffic control accordingly. In the trust evaluation at GTO, we introduce a concept named detection trust of each host or ISP, which is evolved based on detection performance. Thus GTO can detect malicious or indifferent detection behaviors according to detection trust values. In this way, the proposed system encourages good behaviors of hosts and ISPs and helps GTO kick out malicious reports during trust evaluation in order to ensure accurate unwanted traffic control and at the same time preserve the privacy of system entities to fulfill legal requirements. Specifically, the contributions of the paper can be summarized as below:

(1) Design and implementation of a trust management system for UTC with privacy preservation, which strives to comply with European laws on privacy and overcome potential system attacks raised by ISPs and hosts.
(2) Performance evaluation on extra computation, communication and storage costs introduced by privacy protection in order to show the applicability and effectiveness of the proposed system.

The rest of the paper is organized as follows. Section 2 briefly reviews related work. Section 3 defines a system model and our design goal from the points of European law and Internet hosts. In Section 4, we describe the detailed design of a privacy-preserving trust management system for UTC. Security and privacy are analyzed, followed by system performance evaluation in Section 5. Finally, a conclusion is presented in the last section.

## 2. Related work

Recent years have witnessed an increasing number of advanced techniques in improving network security, such as Internet traffic classification with machine learning [26,27] and unwanted traffic control via trust and reputation mechanisms [1,24,25]. Among them, controlling unwanted traffic via trust and reputation mechanisms has attracted extensive research interest recently due to its effectiveness and high performance. A number of trust based solutions have been proposed with regard to dealing with different types of unwanted traffic, e.g., email spam [2–11,28,29], spim [12–15], SPIT [16–19] and web page spam [20–23,30–33].

### 2.1. Solutions for email spam (SPAM)

Email spam has long been a hot issue of research. Moreover, it is expected to be capricious due to the fact that spammers are adopting increasingly sophisticated methods and applying new strategies to spread Email spam [34]. Thus, most existing UTC solutions based on trust and reputation mechanisms target email spam. A reactive spam-filtering system based on reporter reputation was proposed in [2]. It kept records of spam-reporting behaviors using a trust-maintenance component and conducts spam filtering according to the feedback from trustworthy users. A layered trust management framework was proposed in order to help email receivers eliminate their unwitting trust and provide them with accountability support [4]. A distributed architecture and a Trust Overlay Protocol for Anti Spam (TOPAS) protocol for establishing and maintaining trust between mail servers was proposed in [5]. It tunes the threshold of filters using trust information to improve spam filtering. IPGroupRep clustered the senders into different groups based on their IP addresses, computed the reputation value of each group according to the feedback sent from the group, and used the reputation value to indicate incoming spam [6]. A predictive approach introduced two classification methods, Support Vector Machines (SVM) and Random Forests (RF) to analyze the behaviors of email servers to figure out malicious ones [7]. A multi-level reputation-based greylisting system adopted a centralized reputation system embedded in the Simple Mail Transfer Protocol (SMTP) at a receiver side to improve the efficiency of traditional greylisting anti-spam methods [8]. MailTrust filtered out dishonest feedback in order to obtain an accurate trust value of each mail server to perform spam control [9]. A system named "Ostra" thwarted unwanted communications by limiting the total amount of communications according to social trust relationships [10]. SocialFilter, a trust-aware collaborative spam filtering system, derived the system belief that a host is a spammer by weighting spam reports according to the trust of report senders accessed from online social networks [11]. LENS designed an Email

Please cite this article in press as: L. Zhang, et al., Privacy-preserving trust management for unwanted traffic control, Future Generation Computer Systems (2016), http://dx.doi.org/10.1016/j.future.2016.06.036