Ad Hoc Networks 10 (2012) 435–457

Contents lists available at SciVerse ScienceDirect

Ad Hoc Networks

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

# EMLTrust: An enhanced Machine Learning based Reputation System for MANETs ☆

Rehan Akbani [a], Turgay Korkmaz [b,*], G.V. Raju [c]

[a] Bioinformatics and Computational Biology, M.D. Anderson, Houston TX, United States
[b] Department of Computer Science, University of Texas at San Antonio, TX, United States
[c] Electrical and Computer Eng., University of Texas at San Antonio, TX, United States

ARTICLE INFO

*Article history:*
Received 28 November 2010
Received in revised form 20 July 2011
Accepted 15 August 2011
Available online 24 August 2011

*Keywords:*
Trust management
Machine Learning
Digital credentials
Fading Memories

## ABSTRACT

Many mission critical networks including MANETs for military communications and disaster relief communications rely on node cooperation. If malicious nodes gain access to such networks they can easily launch attacks, such as spreading viruses or spam, or attacking known vulnerabilities. One way to defend against malicious nodes is to use Reputation Systems (RS) that try to predict future behavior of nodes by observing their past behavior. In this paper, we propose a Machine Learning (ML) based RS that defends against many patterns of attacks. We specifically consider the proposed RS in the context of MANETs.

After introducing a basic RS, we propose further enhancements to it to improve its performance and to deal with some of the more challenging aspects of MANETs. For instance, we consider digital signature based mechanisms that do not require trusted third parties, or servers that are always online. Another enhancement uses an algorithm called Fading Memories that allows us to look back at longer histories using fewer features. Finally, we introduce a new technique, called Dynamic Thresholds, to improve accuracies even further. We compare the performance of our RS with another RS found in the literature, called TrustGuard, and perform detailed evaluations against a variety of attacks. The results show that our RS significantly outperforms TrustGuard, even when the proportion of malicious nodes in the network is high. We also show that our scheme has very low bandwidth and computation overhead. In contrast to existing RSs designed to detect specific attacks, ML based RSs can be retrained to detect new attack patterns as well.

## 1. Introduction

Let us consider a military MANET in a battle field where several vehicles and soldiers are using wireless communications to exchange mission critical information and to provide various services to each other. In such a realistic MANET, it is deemed necessary to make sure that only the authorized/legitimate users/nodes can access the network resources and the services provided by the other nodes. For example, a typical MANET may have several resources including file servers, databases, web servers, etc. In addition, many nodes may provide different services as part of a larger Service Oriented Architecture (SOA) [1] approach. In SOA, large applications are modularized into smaller services which run on heterogeneous devices. It especially makes sense to use SOA in MANETs so that large, computationally expensive applications can be implemented on resource constrained devices in a distributed fashion. But from a security standpoint, we need a mechanism to regulate access to those resources and services so that we can guard them against malicious transactions from malevolent or compromised nodes.

☆ An abridged version of this paper appeared in Proc. of MILCOM 2008 under the title, "Defending against malicious nodes using an SVM based Reputation System." That paper only described the core SVM RS without any of the enhancements proposed in this paper.

* Corresponding author. Tel.: +1 210 458 7346; fax: +1 210 458 4437.

*E-mail addresses:* rehan01@hotmail.com (R. Akbani), korkmaz@cs.utsa.edu (T. Korkmaz), GVS.Raju@utsa.edu (G.V. Raju).

1570-8705/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2011.08.003

MANETs can be classified as open or closed. In an open MANET anyone is free to enter or leave the network (e.g., in airports and university campuses), whereas in a closed MANET only designated nodes are allowed to access the network (e.g., in a military setting). In general, it is more difficult to provide security in an open MANET since there is no restriction on who may access the network. Fortunately, the security requirements of such networks are also not very demanding since users expect public networks to be insecure. By contrast, closed networks may have very strict security requirements (e.g., in the military or in the police department). Therefore, we specifically focus on *closed* MANETs. Various attacks can be launched against such MANETs by outsiders and/or insiders. Often different mechanisms are needed to defend the underlying network against insiders or outsiders. In this paper, we study how to defend MANETs against malicious transactions from malevolent or compromised (*insider*) nodes. Defending against outsiders is out of the scope of this paper, but has been extensively investigated in our related work [2,3] and references therein.

Suppose an adversary is somehow able to join a closed MANET as a legitimate user. Such a compromise may occur in many ways. For example, an adversary may hack into and gain access to a legitimate node, or obtain the secret key of a legitimate node and assume its identity. Due to the wireless nature of MANETs, those types of attacks would be common and easier to launch. Unfortunately, there is no litmus test to enable one to verify whether an insider node is malicious or benign. We can only predict future behavior of a node by observing and analyzing its past behavior. That is the basic idea behind Reputation Systems (RSs) that have been extensively investigated in the literature [4–9]. For example, eBay uses that form of Reputation System where users leave feedback about other users [10] and Google uses PageRank where web pages are ranked for relevance by other pages [11]. RSs are also used in the context of P2P networks, large scale distributed networks, and the Internet [12,6,13–15]. In general, any network where nodes frequently transact with each other can benefit from RSs. RSs are especially warranted for mission critical networks that rely on node cooperation, such as closed MANETs for the military, emergency and disaster relief networks, and corporate networks [4–6,13].

Fig. 1 illustrates the basic steps in a typical RS. In general, a node that needs to decide whether to transact with another node or not must first gather historical data about that node (e.g., the proportion of good vs. bad transactions in the last $x$ minutes). Then it applies a customized mathematical equation (or statistical model) to the data to produce an output score. For example, the RS in [6] is based on Eigen values from Linear Algebra, the one in [5] is based on using derivatives and integrals, and the one in [8] is based on Bayesian systems utilizing the Beta distribution. Depending on the output of the equation or model, the system then decides how to respond. In most cases, the equation or model is customized to detect specific types of malicious behavior only. For instance, the algorithm in [5] is designed to detect malicious behavior that alternates with good behavior and varies over time.

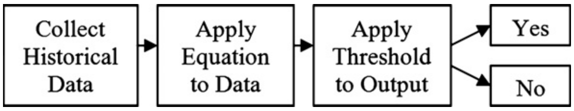

**Fig. 1.** General framework of a Reputation System that decides whether to transact with a given node or not.

Rather than developing a separate module for each attack pattern manually, we propose the use of Machine Learning (ML) to build more flexible and dynamic RSs that can be retrained to thwart a multitude of attack patterns easily and efficiently. Specifically, we consider Support Vector Machines (SVM) and discuss how they can be used for designing RSs. The basic form of the proposed ML-based RS can be used as a general model in wired networks like the Internet. However, to be able to use it effectively in MANETs, we need to deal with several challenges unique to MANETs. For example, eBay has dedicated and trusted centralized reputation servers to collect and store reputation scores for buyers and sellers. Users can trust that (i) the scores being reported by eBay are genuine, (ii) the transactions actually did occur between the buyers and sellers, (iii) unfair scores can be challenged by users and arbitrated by eBay, and (iv) the scores have not been tampered with by Internet routers en-route from eBay to the user. All of those assumptions do not necessarily hold for MANETs due to various reasons. For instance, there is no online central authority, many nodes are limited in their computational resources, nodes may go offline at any time, and nodes are not guaranteed to be completely trustworthy [16].

To deal with the challenges of MANETs while further improving performance, we enhance our core SVM based RS with various mechanisms. To guard against fake transactions and dishonest/incorrect feedback, we propose a digital signature based scheme that does not need online trusted third parties. Using extensive simulations, we demonstrate the efficiency and effectiveness of the proposed core SVM approach, and compare it against two other algorithms found in the literature, namely TrustGuard Naive and TrustGuard TVM [5]. We consider TrustGuard because it has been shown to perform very well compared to eBay's Reputation System. We simulate five different attack scenarios and show that our approach outperforms TrustGuard in all five scenarios, including when there is oscillating or steady behavior, collusive or non-collusive behavior. Our scheme can achieve high accuracy and correctly predict good vs. malicious nodes, even when the proportion of malicious nodes in the network is very high. The ROC curves show that the improvement of SVM over TrustGuard is statistically significant, as their 95% confidence intervals do not overlap each other. We also show that SVM has the same bandwidth overhead as TrustGuard Naive while having much less overhead than TrustGuard TVM.

We propose two further enhancements to improve the performance of our core SVM based RS. First, we consider how to look back at longer histories using only a few features. That enhancement forces the adversary to behave well for longer periods of time in order to boost its reputation score. We evaluated its performance and showed that it was much better at detecting malicious behavior that varied over longer periods. Second, we introduce an algorithm called Dynamic Thresholds that further improves the