MATHEMATICAL MODEL FOR THE DETECTION OF SELFISH NODE IN MANETS USING K-NEAREST NEIGHBOR TECHNIQUE

Md. AMIR KHUSRU AKHTAR & G. SAHOO

ABSTRACT: Mobile Ad-Hoc Networks special categories of wireless networks that are attracting attention of industry and academia for quite some time due to its Low cost, low power, low bandwidth, multi-functional networks, and design and implementation issues. New algorithms for media access and routing are being designed to optimize the performance. MANETS use multi-hop radio relaying and are capable of operating without the support of any fixed infrastructure. Security is a challenge for implementing Ad-hoc networks. In this paper we are proposing mathematical model for the detection of selfish node in MANETs using k-nearest neighbor technique that classify the nodes into two classes and improves the efficiency.

1. INTRODUCTION

Mobile Ad-Hoc Networks (referred to as MANETs), are impromptu wireless communication networks increasingly appearing in the Commercial, Military, and Private sector as portable wireless computers become more and more ubiquitous. Mobile Ad-Hoc Networks allow users to access and exchange information regardless of their geographic position or proximity to infrastructure. In contrast to the infrastructure networks, all nodes in MANETs are mobile and their connections are dynamic. Unlike other mobile networks, MANETs do not require a fixed infrastructure. This offers an advantageous decentralized character to the network. Decentralization makes the networks more flexible and more robust. Applications for MANETs are wide ranging and have use in many critical situations: An ideal application is for search and rescue operations. Such scenarios are characterized by the lack of installed communications infrastructure. This may be because all of the equipment was destroyed, or perhaps because the region is too remote. Rescuers must be able to communicate in order to make the best use of their energy, but also to maintain safety. By automatically establishing a data network with the communications equipment that the rescuers are already carrying, their job made easier.

Every node in an ad hoc network participates in the routing process of the packets which uses any algorithms presented in [1-6, 10]. Attacks in ad hoc routing protocols

Keywords: MANETs, Regular node, Selfish node, Prior Probability, Density function, K-NN.
are a major factor of concern. A variety of attacks like modification, fabrication, wormhole attack (tunneling), blackhole attack, denial of service attack, invisible node attack, sybil attack, rushing attack and non-cooperation reduce the reliability of these routing protocols. To overcome these security threats the concept of secured routing protocols came into existence. Some popular secured routing algorithms are SRP (Secure Routing Protocol) [5], ARAN (Authenticated Routing for Ad-hoc Networks) [2, 3], Ariadne [1], SEAD (Secure Efficient Ad hoc Distance vector routing) [4] etc. Where as most of the attacks based on manipulations of routing data can be detected by the use of a secure routing protocol like ARAN [2, 3], Ariadne [1] and others [3-6]. But when nodes simply drop packets, or show its selfishness all of the secure routing protocols fail, as they focus only on the detection of modifications to routing data but not on the concealment of existing links. In this paper we proposed a mathematical model for the detection of selfish node. This paper is organized as follows. Section 2 describers the background and related work. Section 3. defines the proposed mathematical model. Section 4 presents the experimental analysis. In this work we have used OPNET to find the simulation results. Section 5 concludes the paper.

2. BACKGROUND AND RELATED WORK

2.1 Background

Many attacks in MANET had broken routing protocols. This is due to developing routing services without defining security issues. In this section, we describe the security threats, advantage and disadvantage of some common routing protocols.

The Ad-hoc On-demand Distance Vector (AODV) [10] routing algorithm is a reactive algorithm that routes data across wireless mesh networks. The advantage of AODV is that it is simple, requires less memory and does not create extra traffic for communication along existing links. In AODV, the attacker may advertise a route with a smaller distance metric than the original distance or advertise a routing update with a large sequence number and invalidate all routing updates from other nodes.

Dynamic Source Routing (DSR) [6] protocol is similar to AODV in that it also forms route on-demand. But the main difference is that it uses source routing instead of relying on the routing table at each intermediate node. It also provides functionality so that packets can be forwarded on a hop-by-hop basis. In DSR, it is possible to modify the source route listed in the RREQ or RREP packets by the attacker. Deleting a node from the list, switching the order or appending a new node into the list is also the potential dangers in DSR.

Authenticated Routing for Ad-hoc Networks (ARAN) [2, 3] is an on-demand routing protocol that detects and protects against malicious actions carried out by third parties.
and peers in particular Ad-hoc environment. This protocol introduces authentication, message integrity and non-repudiation as a part of a minimal security policy. Though ARAN is designed to enhance ad-hoc security, still it is immune to rushing attack and cannot identify selfish nodes.

ARIADNE [1] is an on-demand secure Ad-hoc routing protocol based on DSR that implements highly efficient symmetric cryptography. It provides point-to-point authentication of a routing message using a message authentication code (MAC) and a shared key between the two communicating parties. Although ARIADNE is free from a flood of RREQ packets and cache poisoning attack, but it is immune to the wormhole attack and rushing attack. Specifically, SEAD[4] builds on the DSDV-SQ version of the DSDV (Destination Sequenced Distance Vector) protocol. It deals with attackers that modify routing information and also with replay attacks and makes use of one-way hash chains rather than implementing expensive asymmetric cryptography operations. Two different approaches are used for message authentication to prevent the attackers. SEAD does not cope with wormhole attacks. The ARAN protocol was observed to defend almost against all security attacks in MANETs. However, by doing more research in the field of MANETs, one major flaw in any of the existing secure routing protocols was discovered. This is that all of these secure routing protocols do not account for selfish nodes whether by detecting or isolating them from the network.

2.2 Related Work to identify Selfish Nodes

A Mobile Ad-Hoc Network is a collection of autonomous mobile nodes that communicate with each other over wireless links without any central administration. In Ad-hoc networks, each host has to act as a router for itself to communicate with hosts outside its transmission range due to the limited range of each host’s wireless transmission. An Ad-hoc routing protocol runs on every host and is subject to the limit of the resources at each mobile host. Therefore, a traditional routing protocol, which is used in IP network, is not suitable to the Ad-hoc network. Also, it is not appropriate to use conventional protocols, which are proposed or used in IP networks, such as MAC, QoS, Multicast, and TCP. So in an adhoc network we have adhoc routing protocols but these protocols have security issues. This is due to developing routing services without defining Selfishness is a major security issues. There are various methods to detect selfish nodes. These methods are categorized in incentive-based methods or reputation-based methods. In the first method it discourages a node to become selfish by giving virtual money or credits when a node forward packets of others because to send or receive its own packets the node requires enough credit. Buttayan and Hubaux [16] method uses virtual currency, called nuglets to
detect selfish node. In this method a nuglet counter is incremented monotonically when it forwards a packet for others. When a node wants to send its own packet, it requires enough credit because if it is less than certain threshold it is not allowed to send packets. But this method requires tamper proof hardware to maintain the nuglet. In reputation-based method it detects a selfish node and performs proper action by using a reputation system that detect and rate a selfish node. The reputation is defined by the participation seen by others [17]. When node’s reputation is good it participate in network activity otherwise it is marked as selfish.

3. Proposed Work

3.1 K-nearest Neighbor Technique

In this paper, we will show via a simple model and with existing routing protocols that in an adhoc network we have to detect selfish nodes using the $K$-nearest Neighbor technique [15].

In an adhoc Network, the K-nearest Neighbor classifies an unknown sample on the “votes” of $k$ of its nearest neighbors. The $K$-nearest Neighbor classification procedure is denoted by $K$-NN. If the cost of error is equal for each class, the estimated class of an unknown sample is chosen to be the class that is most commonly represented in the collection of its $k$-nearest neighbors. The error rate for $k$-nearest neighbors is similar to bayesian technique as both $k$ and $n/k$ tends to infinity, where $n$ is the arbitrary taken samples of known classification in the data set. Therefore from the set of classes Nodes $X$ should be chosen for the test in comparison to Nodes $Y$ when

$$
\frac{N(X) - L(X)}{P(X)_{\text{reference set}}} > \frac{N(Y) - L(Y)}{P(Y)_{\text{reference set}}}
$$

where $N(X)$ is the number of $X$s found in the $K$-nearest neighbor and $L(X)$ is the loss of misclassifying a sample from Nodes $Y$ and $N(Y)$ is the number of $Y$s found in the $K$-nearest neighbor and $L(Y)$ is the loss of misclassifying a sample from Nodes $X$. Where $X$ and $Y$ denotes Regular nodes and Selfish nodes respectively. If there are two classes, $N(X) + N(Y) = k$, the total number of neighbors used.

3.2 Scale Factors

In the nearest neighbor technique the difference is taken in terms of nodes features. If we change the feature parameter it affects the decision. So we have to apply the scale factor for calculating the distance. If the features should be taken without any
Mathematical Model for the Detection of Selfish Node in Manets Using...

problem then each of the features has equal mean deviation, range and other measures. Features $X_i$ with mean (or median) 0 and standard deviation (or $AMD$) 1 for all data set. $AMD_i$ represent the absolute deviation of $X_{ij}$ from median $M_i$.

$$AMD = \frac{1}{n} \sum_{j=1}^{n} |x_{ij} - M_i|$$

Normalization replaces each feature $x_i$ by

$$Z_i = \frac{(x_i - \mu_i)}{\sigma_i} \text{ (or } Z_i = \frac{(x_i - M_i)}{AMD_i} \text{ )}$$

before computation of distance.

Another technique used for making the scale factor of each feature be proportional to power $a$ of this probability of being correct when the is taken self. Using MinKowski distance for this nonlinear accuracy produces distance metric with two parameters

$$d_{ar} = \sum_{i=1}^{n} \left[ P(C|i)^a |x_{iu} - X_{icj} \right]^r$$

where $P(C|i)$ is the probability of correctness when only feature $i$ is used which is used to identify on the basis of single feature. If $r = 1$ and $a = 0$ value is given to the parameters gives unweighted city blocks distance and if $r = 2$ and $a = 0$ value is given to the parameters gives unweighted Euclidean distance squared. We can take other values for the parameters also.

4. Experimental Analysis

Let us consider a free space that contains three nodes from class $X$ and two nodes from class $Y$ as given in Fig. 1. Suppose a node of unknown class is located at (1, 1).
Using the Euclidean distance metric, the closest point of known classification is a sample from class $X$ located at $(1, 3)$. Using this method the unknown node is classified and belonging to Class $X$. But Using $K$-NN if three neighbors are to be used as given in Fig. 1, the unknown sample at $(1, 1)$ would be classified and belongs to class $Y$, since the three nearest neighbors consist of the sample from class $X$ at $(1, 3)$ and the two samples of class $Y$.

5. Conclusion

As we have seen that protocols fails to detect selfish or malicious nodes in Ad-hoc network because it is a complex task. In this paper we have presented the $K$-Nearest Neighbor technique that can detect selfish nodes with a good confidence as shown by our simulation results. In an adhoc Network, the $K$-nearest Neighbor classifies an unknown sample on the “votes” of $k$ of its nearest neighbors. It classifies the network into classes and a selfish node is detected easily, then this is a good indication for excluding the node from the network. This mathematical model is verified by experimentation and gives acceptable accuracy and provides a solution for secured routing in an Ad-hoc environment. So, this model gives more accurate class information to identify selfish node using the defined mathematical model.

References


Mathematical Model for the Detection of Selfish Node in Manets Using ...


Md. Amir Khusru Akhtar
ICFAI University, Ranchi, Jharkhand, India.
E-mail: akru2008@gmail.com

G. Sahoo
Department of Information Technology,
Birla Institute of Technology,
Mesra, Ranchi, India.
E-mail: gsahoo@bitmesra.ac.in