ARTICLE IN PRESS

Engineering Science and Technology, an International Journal ■■ (2015) ■■–■■

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: http://www.elsevier.com/locate/jestch

Full Length Article

Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation

Basant Subba, Santosh Biswas *, Sushanta Karmakar

Department of Computer Science & Engineering, Indian Institute of Technology, Guwahati, Assam 781039, India

ARTICLE INFO

Article history:
Received 14 April 2015
Received in revised form
4 November 2015
Accepted 4 November 2015
Available online

Keywords:
Mobile Ad-hoc Network (MANET)
Intrusion detection system (IDS)
Game theory
Bayesian Nash equilibrium

ABSTRACT

Present Intrusion Detection Systems (IDSs) for MANETs require continuous monitoring which leads to rapid depletion of a node's battery life. To address this issue, we propose a new IDS scheme comprising a novel cluster leader election process and a hybrid IDS. The cluster leader election process uses the Vickrey–Clarke–Groves mechanism to elect the cluster leader which provides the intrusion detection service. The hybrid IDS comprises a threshold based lightweight module and a powerful anomaly based heavyweight module. Initially, only the lightweight module is activated. The decision to activate the heavyweight module is taken by modeling the intrusion detection process as an incomplete information non-cooperative game between the elected leader node and the potential malicious node. Simulation results show that the proposed scheme significantly reduces the IDS traffic and overall power consumption in addition to maintaining a high detection rate and accuracy.

1. Introduction

Mobile Ad-hoc Networks (MANETs) are a collection of heterogeneous, infrastructure less, self organizing and battery powered mobile nodes with different resources availability and computational capabilities. The dynamic and distributed nature of MANETs makes them suitable for deployment in extreme and volatile environmental conditions. They have found applications in diverse domains such as military operations, environmental monitoring, rescue operations etc. Each node in a MANET is equipped with a wireless transmitter and receiver, which enables it to communicate with other nodes within its wireless transmission range. However, due to limited wireless communication range and node mobility, nodes in MANET must cooperate with each other to provide networking services among themselves. Therefore, each node in a MANET acts both as a host and a router.

The dynamic and distributed nature of MANETs make them vulnerable to various types of attacks like black hole attack, traffic distortion, IP spoofing, DoS attack etc. Malicious nodes can launch attacks against other normal nodes and deteriorate the overall performance of the entire network [1–3]. Unlike in wired networks, there are no fixed checkpoints like router and switches in MANETs, where the Intrusion Detection System (IDS) can be deployed [4,5]. Therefore, nodes in MANETs must cooperate in many aspects including intrusion detection for their well being [6–8]. IDSs have been deployed with great degree of success across diverse domains like wireless Ad-hoc networks [5,9], MANETs [10–12], wireless sensor networks [13], cyber-physical system [14], cloud computing [15], large scale complex critical infrastructures [16] etc. In this paper, we focus on IDS for MANETs.

Due to absence of any centralized monitoring entity in MANETs, each node runs its own IDS and usually operates in a promiscuous mode. However, owing to limited battery life, it is not feasible to keep the IDS running continuously on MANET nodes. Most of the current MANET IDS schemes do not take into account the nature of the environment they are operating in and therefore they end up monitoring all nodes with equal probability, irrespective of whether or not the node being monitored has a history profile of being malicious. This results in a poor monitoring strategy wherein the node operating the IDS ends up wasting most of its energy monitoring the normal nodes. Another issue with many MANET IDS schemes [17–19] is that they generate heavy intrusion detection related traffic. Unlike the wired networks, MANETs have limited bandwidth and therefore, a large amount of intrusion detection related traffic can cause severe congestion in the network and limit the flow of normal traffic. In addition, heavy intrusion detection traffic also leads to more energy consumption among MANET nodes for processing them.

Designing a MANET IDS scheme that is energy efficient and generates a low IDS traffic, while at the same time maintaining a high accuracy and detection rate is an active area of research. In this paper, we model the intrusion detection process in MANETs using a game theoretical framework. Game theory based MANET IDSs [20–22] have been found to be energy efficient as well as generate low IDS traffic

* Corresponding author. Tel.: +91 9957561026, +91-361-2583000.
E-mail address: santosh_biswas@iitg.ernet.in (S. Biswas).
Peer review under responsibility of Karabuk University.

http://dx.doi.org/10.1016/j.jestch.2015.11.001

through application of dynamic and economical monitoring strategies. Game theory based IDS models the intrusion detection problem as a non-cooperative game between two competing players (attacker and defender), where the defender player (cluster leader node) tries to maximize its payoff by increasing its probability of successful intrusion detection while the attacker player (malicious node) tries to minimize its probability of being detected by the IDS.

Game theory based IDS scheme allows the IDS to assess the type of the node being monitored and adopt appropriate monitoring strategies. Nodes are assigned maliciousness values based on the history profile of their observed actions. Unlike most conventional IDSs that adopt promiscuous monitoring strategy and results in high IDS traffic generation, game theory based IDS uses a dynamic monitoring strategy wherein nodes with high maliciousness values are monitored more frequently compared to nodes with low maliciousness values. This helps the IDS to conserve its energy and minimize the overall IDS traffic generation. In a game theoretic IDS framework, a rigorous monitoring strategy is adopted by the IDS if the environment it is operating in is hostile. On the other hand, if the environment is less hostile, a less rigorous monitoring strategy is adopted by the IDS.

Most of the game theory based IDSs proposed in the literature [19–21,23] assume a complete information game, wherein all players (nodes) have complete information about the game, i.e., they make an implicit assumption that various network parameters like energy levels and types of network nodes (normal or malicious), accuracy and detection rate of IDS etc. are known to all nodes *a priori*. But, such assumptions have limitations, since in most of the real network settings each node only has a limited information about the network parameters. Therefore, to address this issue of incomplete information game, we propose a Bayesian game theory based MANET IDS scheme that models the interaction between the attacker (malicious node) and the defender (node operating IDS) in MANET as a two person multi-stage, non-cooperative and incomplete information game. The Bayesian model [19] allows the node operating the IDS to adopt the most efficient monitoring strategy in an incomplete information game settings by examining the maliciousness history profile of the node being monitored and by evaluating the Bayesian Nash equilibrium of the game.

In summary, this paper proposes a MANET IDS scheme with the following objectives:

1. Modeling the intrusion detection process in MANETs as an incomplete information Bayesian game as nodes in MANETs only have partial information about the network.
2. Minimization of power consumption for operating IDS in MANETs.
3. Minimization of intrusion detection related traffic in MANETs.
4. Developing a MANET IDS scheme with high accuracy and detection rate.

To achieve these objectives, we propose a new MANET IDS scheme consisting of the following two components:

1. *A MANET leader election mechanism*: This component elects the cluster leader node using the VCG mechanism [24] and entrusts it with the responsibility of providing intrusion detection services to all other cluster nodes for a predefined period of time. Cluster leader elections are held at regular intervals which ensures uniform energy consumption among various cluster nodes for operating the IDS.
2. *A hybrid MANET IDS*: This component comprises one lightweight module and one heavyweight module. The lightweight module is less powerful but requires less energy for its operation. On the other hand, the heavyweight module is more powerful than the lightweight module but requires more energy for its operation. Initially only the lightweight module is activated. If the action of the node being monitored by the lightweight module is determined to be malicious then the heavyweight module is activated, else the decision to activate the heavyweight module is determined by the Nash Equilibrium of the non-cooperative game played between the elected leader node and the node being monitored.

The elected leader node operates the *hybrid MANET IDS*. Initially, only the lightweight module of the hybrid MANET IDS is activated, which calculates the Packet Forwarding Rate (PFR) of the potential malicious node being monitored. The PFR of any given node is defined as the ratio of total number of packets received to the total number of packets forwarded by the node over a given period of time. If the PFR of the node being monitored is less than the threshold value, then its action is assumed to be malicious and the heavyweight module is activated for more rigorous analysis. However, if the action of the node is found to be normal then the decision to activate the heavyweight module is determined by modeling the intrusion detection process as a multi-stage Bayesian game between two competing players, where the players of the game are the cluster leader node and the potential malicious node.

The cluster leader node has incomplete information about the type of the opponent node (normal or malicious) and the following two strategies: *Monitor* and *Not Monitor*. Here, the strategy *Monitor* corresponds to the activation of the heavyweight module. Similarly, the attacker player has two strategies: *Attack* and *Not Attack*. The Bayesian Nash Equilibrium (BNE) of the game is the strategy pair of the players which corresponds to the probability of the leader node to play its strategy *Monitor*/*Not Monitor* and the probability of the attacker player to play its strategy *Attack*/*Not Attack*. Intrusion detection process in MANETs is usually an incomplete information game, where nodes only have partial information about network parameters. The Bayesian game model allows the cluster leader node to formulate its monitoring strategies based on its belief about the type of the node (malicious or normal) being monitored without requiring a complete information about that node. It also minimizes the overall IDS traffic by adopting a non-promiscuous monitoring strategy.

Simulation results in NS-2 [25] show that the proposed MANET IDS scheme significantly reduces the power consumption for operating the IDS among MANET nodes by 15–20% compared to a random model. Further, the proposed scheme also maintains a high level of detection rate against *route compromise*, *traffic distortion* and *black-hole* attacks without introducing any significant traffic.

The rest of the paper has been structured in the following way. Section 2 discusses about the background and related works on intrusion detection in MANETs. Section 3 presents the overall description of our proposed MANET IDS scheme. Bayesian Game model used for developing energy efficient IDS monitoring strategies is discussed in section 3.1. A distributed and energy efficient MANET leader election mechanism is discussed in section 3.2. A hybrid MANET IDS along with its main components are discussed in section 3.3. Experimental results and performance evaluation of the proposed hybrid MANET IDS and MANET leader election mechanism are provided in Section 4. Finally, Section 5 provides the conclusion and future work.

2. Background and related works

In this section, we provide a brief background study on different types of MANET IDS based on their detection mechanism and modes of operation. We then discuss about various intrusion detection issues in MANETs and analyze the related works which have been categorized into non-game theory based and game theory based. Finally, the drawbacks associated with the related works have

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

been listed out which provides us with the motivation for our work to address them.

Based on their mode of operations, IDS in MANETs can broadly be classified into anomaly based, signature based and specification based. The anomaly based IDSs consist of the training phase and the testing phase. The normal traffic profile of the network is developed during the training phase and then the learned model is used to analyze the current network traffic for sign of misbehavior during the testing phase. Numerous anomaly detection methods like statistical methods [26,27], data-mining methods [28] and machine learning based methods [29] have been developed. The main advantage of anomaly-based IDSs are their ability to detect previous unknown attacks not seen during the training phase. However, the main drawback of anomaly based IDSs is their high False Positive (FP) alarm rate. Signature-based IDSs [30] use a database of known attack signatures and raise an alarm wherever there is a malicious traffic that matches with one or more attack signatures in the database. They have high detection rate against known attacks but cannot detect new attacks. They require frequent updates to their signature database to detect new attacks. The specification-based IDSs [31] specify a set of constraints on the network traffic or protocols and any violations of these specifications are treated as intrusions. They provide detection against both known and unknown attacks with low false positive rate. However, the main drawback of specification-based IDSs is their requirement of detailed specifications for each program/protocol, which is a very time consuming and computationally expensive process.

Based on their modes of operations, IDSs in MANETs can be grouped into *Stand-alone IDS*, *Distributed IDS* and *Clustered IDS*. In the *Stand-alone IDS* architecture, each node independently runs its own IDS to determine intrusions. There is no cooperation between the nodes in the network and every intrusion decision made by the node is solely based on its own gathered information. Since partial information on each individual node might not be enough to detect attacks like network scans, this category of IDS is not suitable and generally not preferred for MANETs. In the *Distributed IDS* architecture, every node participates in the intrusion detection process by having an IDS agent running on them. The IDS agent collects local event data to detect and identify local network intrusions. However, neighboring IDS agents cooperate to perform a global intrusion detection, when the local intrusion detection evidence is inconclusive. In the *Clustered IDS* architecture, the network is divided into multiple clusters. Every cluster node runs its own IDS agent, which monitors and detects local intrusions for the given cluster node, while the cluster head runs the IDS agent both locally for its own node and globally for the entire set of cluster nodes.

The conventional IDSs used in wired networks are ineffective and inefficient for MANETs because of differences in their underlying characteristics and architectures. The major issues encountered while developing an IDS for MANETs are:

- *Lack of Central Monitoring Points*: Unlike in wired networks there are no centralized points like routers and gateways for monitoring network traffic in MANETs. IDS in MANETs needs to be distributed and cooperative. However, limited bandwidth, low energy levels, different computation capabilities of MANET nodes, presence of malicious nodes etc. put a serious constraint on cooperation among MANET nodes.
- *Mobility*: MANET topology may change frequently because of mobile nodes that can exit or join the network arbitrarily. This makes it difficult for the IDS to differentiate whether the node sending an out of date routing information is simply out of synchronization with other MANET nodes or whether the node has been compromised.
- *Wireless Links*: Wireless networks have limited bandwidth compared to wired networks. Heavy intrusion detection related traffic could cause network congestion and limit the flow of normal traffic. Therefore, MANET IDSs need to minimize their data flow to avoid network congestion. But constraining the IDS traffic flow may result in performance degradation of the IDSs and they may not be able to respond to intrusions in real time.
- *Limited Resources*: Mobile nodes in MANETs consist of various mobile devices with different computational capabilities and energy resources. Therefore, signature-based IDS for MANETs must take into account memory constraints for storing attack signatures, while the anomaly-based MANET IDS needs to be optimized to reduce energy usage for correlation of the network traffic with the learned IDS model.
- *Insecure Communication Link*: MANETs are vulnerable to various passive attacks like eavesdropping and interference. Therefore, IDS traffic needs to be encrypted to prevent the attacker from learning about the working principles of the IDS. However, employing cryptographic and authentication mechanism in MANETs is not feasible as they consume significant amount of energy and are computationally expensive.

2.1. Related works

Shakshuki et al. [18] proposed an IDS named Enhanced Adaptive Acknowledgment (EAACK) for MANETs. Their scheme requires all acknowledgment packets to be digitally signed by its sender and verified by its receiver. They used DSA and RSA as digital signatures and showed that their scheme is able to detect wide range of attacks. However, the drawback of their scheme is the requirement to digitally sign all the acknowledgments which increases computational overhead.

Marti et al. [32] proposed an IDS scheme for MANET which consists of two different modules, *viz.* the Watchdog and the Pathrater. In this scheme, the Watchdog acts as an IDS for the MANET and detects malicious node behaviors in the network by promiscuously listening to its next hop's transmission. If the Watchdog notices that its immediate next node fails to forward the packet within a given period of time then it increments the node's failure counter. If the failure counter of the monitored node exceeds a threshold value then the Watchdog reports the node as misbehaving. The Pathrater is then employed to inform the routing protocol to avoid the reported nodes for further data transmission. The drawback of this scheme is that it requires continuous monitoring by the Watchdog for detecting intrusions.

Lui et al. [17] proposed a TWOACK MANET IDS scheme which requires every data packets transmitted over three consecutive nodes along the source to the destination path to be acknowledged. Every node along the route has to send back an acknowledgment packet to the node that is two hop counts away from it in the route. The arrival of TWOACK packet at first node X (in the three consecutive nodes along the route) indicates a successful transmission of packet from node X to node Z via the intermediate node Y. However, if this TWOACK packet is not received within a given predefined time interval, both nodes Y and Z are reported as malicious. The drawback of this scheme is that it introduces a routing overhead due to frequent TWOACK packet generation.

Misra et al. [33] proposed a distributed self-learning, energy-aware and low complexity protocol for intrusion detection in wireless sensor network. Their protocol uses the stochastic Learning Automata (LA) on packet sampling mechanism to obtain an energy efficient IDS. They showed that their approach was successful in detecting and removing malicious packets from the WSN. The drawback of this scheme is that the LA needs multiple rounds of learning before it becomes efficient. Haddadi and Sarram [34] proposed a hybrid IDS model for Wireless Local Area Network (WLAN) that uses both misuse and anomaly based IDS sub-modules to detect intrusion. The drawback of this approach is that the response times of the misuse based and anomaly based IDSs are different. It also introduces significant computational overhead due to processing of the same data traffic by two different IDSs.

A light weight, energy efficient and non-cryptographic intrusion detection solution against the gray hole attack in MANET is proposed in Reference [35] by Mohanapriya and Krishnamurthi. However, their scheme requires the IDS to operate in a promiscuous mode to detect intrusions, which results in high power consumption for operating the IDS.

A game-theoretic solution for Ad-hoc networks that models the cooperation and selfishness of the networks are discussed in References [36,37]. In these schemes, each node decides whether to forward or not forward a packet based on the trade-offs involved in cost (energy consumption) and benefits (network throughput) involved in collaborating with other nodes in the network. Therefore, enforcing a cooperation mechanism ensures that a selfish node that does not obey the network rules receives a low throughput. The drawback of this scheme is that it assumes the complete information game, where nodes have full knowledge about the network parameters.

Lui et al. [19] proposed a game theoretic framework to analyze the interactions between pairs of attacking/defending nodes using a Bayesian formulation in wireless Ad-hoc Networks. They suggested a Bayesian hybrid detection approach for the defender, in which a less powerful lightweight module is used to estimate the opponent's type, and a more powerful heavyweight module acts as a last line of defense. They analyzed the obtainable Nash Equilibrium (NE) for the attacker/defender Bayesian game in both static and dynamic settings and concluded that the dynamic approach is a more realistic model, since it allows the defender to consistently update its belief about the maliciousness of the opponent player as the game evolves. The drawback of their work is that it is difficult to determine a reasonable prior probability about the maliciousness of the attacker player.

Liu [38] proposed a general incentive-based method to model attacker's intent, objectives and strategies (AIOS) based on game theoretic formalization. The author developed an incentive-based conceptual framework for AIOS modeling which can capture the inherent inter-dependency between AIOS and defender objectives and strategies in such a way that AIOS can be automatically inferred. The AIOS modeling enables the defender to predict which kind of strategies are more likely to be taken by the attacker than the others, even before such an attack happens. The AIOS inferences lead to more precise risk assessment and harm prediction. The drawback of the scheme is that it assumes the complete information game.

Chen et al. [39] proposed a framework that applies two game theoretic schemes for economic deployment of intrusion detection agent. In the first scheme, the interaction between an attacker and the intrusion detection agent is modeled and analyzed within a non-cooperative game theory setting. The mixed strategy Nash Equilibrium solution is then used to derive the security risk value. The second scheme uses the security risk value derived by the first scheme to compute the Shapley value of the intrusion detection agent while considering the various threat levels. This allows the network administrator to quantitatively evaluate the security risk of each IDS agent and easily select the most critical and effective IDS agent deployment to meet the various threat levels to the network. The drawback of this scheme is the computational overhead involved for calculating the Shapley values of the intrusion detection agents.

A game theoretical framework to model the interaction between the service provider and the attacker as an intrusion detection game was proposed by Kodialam and Lakshman [23]. In this scheme, the game is represented as a two person zero-sum game, wherein the service provider tries to maximize its payoff by increasing its probability of successful detection while the attacker tries to minimize its probability of being detected by the IDS. The optimal solution for both players is to play the minmax strategy of the game. The drawback of this model is the assumption that both players (attacker and defender) have complete information about the topology of the network and all links in the network, which allows the players to choose the optimal path for playing the minmax strategy. However, this assumption is usually invalid in real networks where the players have an incomplete information about the network parameters.

Agah et al. [20] and Alpcan and Basar [21] addressed the attack–defense problem in a sensor network as a two-player non-cooperative, non-zero-sum game. In their model, the game is assumed to have a complete information and the payoff function of the opponent player decides each player's optimal strategy. The drawback of their work is the assumption that the players have complete information about the game.

In summary, we found that most of the non-game theory based IDS schemes proposed in the literature are computationally expensive and require continuous monitoring, thereby leading to more power consumption for operating the IDS. The game theory based IDSs proposed in the literature addresses this issue to some extent. However, most of the previous works on game theory based MANET IDS assumes a complete information game where both players (attacker and defender) have complete information about the game. But such an assumption is usually not valid in a real network, where each node only has a partial information about the network because all network parameters are not known *a priori*. We also found that most of the games are static in nature where the strategies and utilities of players are fixed and repeated over a period of time. This approach fails in a dynamic environment where players adopt different strategies at various stages of the game. We also found that most of IDSs proposed in literature for MANETs are specific to certain classes of attacks like blackhole attack, wormhole attack etc. [32,40]. All these drawbacks in the related works provide us with the motivation to propose a new MANET IDS scheme based on incomplete information game to address them.

In this paper, we propose a new IDS scheme for MANETs comprising of two different components *viz.* the MANET leader election mechanism and the hybrid MANET IDS. The former component minimizes the overall power consumption required for operating the IDS by distributing the task of intrusion detection among various cluster nodes. It elects the cluster leader node based on reputations and energy levels of nodes. The elected leader node is designated with the responsibility of providing intrusion detection services to all other cluster nodes for a predefined period of time.

The second component of the proposed IDS scheme is a game theory based hybrid MANET IDS, which performs the actual intrusion detection operation. The leader node elected by the election mechanism runs the hybrid MANET IDS. The hybrid MANET IDS comprises one lightweight module and one heavyweight module. The lightweight module is less powerful and uses simple analytical rules based on threshold values to detect intrusions. On the other hand, the heavyweight module is more powerful and uses complex association-mining rule techniques to detect anomalies. Initially, only the lightweight module is activated. The decision to activate the heavyweight module depends on the output of the lightweight module. If an intrusion is detected by the lightweight module, then it activates the heavyweight module for more rigorous analysis. However, if no malicious activity is detected by the lightweight module, then the network intrusion detection problem is modeled as a non-cooperative game between the elected leader node and the potential malicious node. In this case, the BNE of the game decides the probability of activating the heavyweight IDS module.

3. Proposed MANET IDS scheme

In this section, we describe various assumptions and aspects of our proposed MANET IDS scheme. First, the flowchart of the

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

proposed scheme is provided and then its components *viz.* the *MANET cluster leader election mechanism* and the *hybrid MANET IDS* are described. The *hybrid MANET IDS* comprises a lightweight IDS and a heavyweight IDS module. We make the following assumptions related to our proposed MANET IDS scheme:

- MANET is divided into a set of clusters using a standard cluster algorithm [41]. Every node in a given cluster is within the transmission range of each other.
- Each node n_i in a given MANET cluster has the following associated parameters: maliciousness value (p_i), reputation value (R_i) and energy value (E_i).
- The elected cluster node (C_L) provides the intrusion detection services to all other cluster nodes for a predefined period of time by operating the IDS.

Fig. 1 shows the flowchart of our proposed MANET IDS scheme. Initially, the cluster leader node C_L is elected using the VCG mechanism [24]. C_L is entrusted with the responsibility of providing intrusion detection services to the entire set of cluster nodes for a predefined period. The intrusion detection service provided by C_L to any given cluster node n_j depends on n_j's reputation value (R_j). Nodes with higher reputations are entitled to more service from C_L compared to nodes with lower reputations. The services provided by C_L to node n_j includes monitoring the incoming traffic received by n_j from its neighbors as well as monitoring the outgoing traffic of n_j. C_L may misbehave after being elected as a leader node by not providing intrusion detection services to other cluster nodes or by reporting the normal node as malicious. Therefore, a set of checker nodes are elected to monitor the operations of C_L. If C_L is found to be misbehaving by the checker nodes, then it is punished by lowering its reputation value. The detailed description of the MANET leader election and punishment mechanism is provided in section 3.2.

After being elected as the cluster leader, C_L assigns initial maliciousness belief value (p_i) to cluster node n_i being monitored and activates its lightweight IDS module to determine the action of n_i. The lightweight IDS module uses the packet forwarding rate (PFR) of n_i as a parameter to determine the action of n_i as *Attack* or *Normal*. The PFR of n_i is defined as the ratio of total number of packets received by n_i to the total number of packets forwarded by n_i over a given interval of time. If the PFR of n_i is less than the threshold value T_{PFR}, then the action of n_i is assumed to be *Attack*. The p_i value of n_i is then updated using the Bayes rule, and the heavyweight IDS module of C_L is activated for more rigorous analysis. However, if the PFR of n_i is greater than or equal to the threshold value T_{PFR}, then the action of n_i is assumed to be *Normal*. In this case too, the p_i value of n_i is updated using the Bayes rule but the decision to activate the heavyweight IDS module is determined by representing the interaction between C_L and n_i as a non-cooperative game between two competing players and calculating the Bayesian Nash Equilibrium (BNE) of the game. The BNE of this game corresponds to the strategy combination (q^*, p^*), where q^* is the probability of C_L to activate its heavyweight IDS module and p^* is the probability of n_i to play

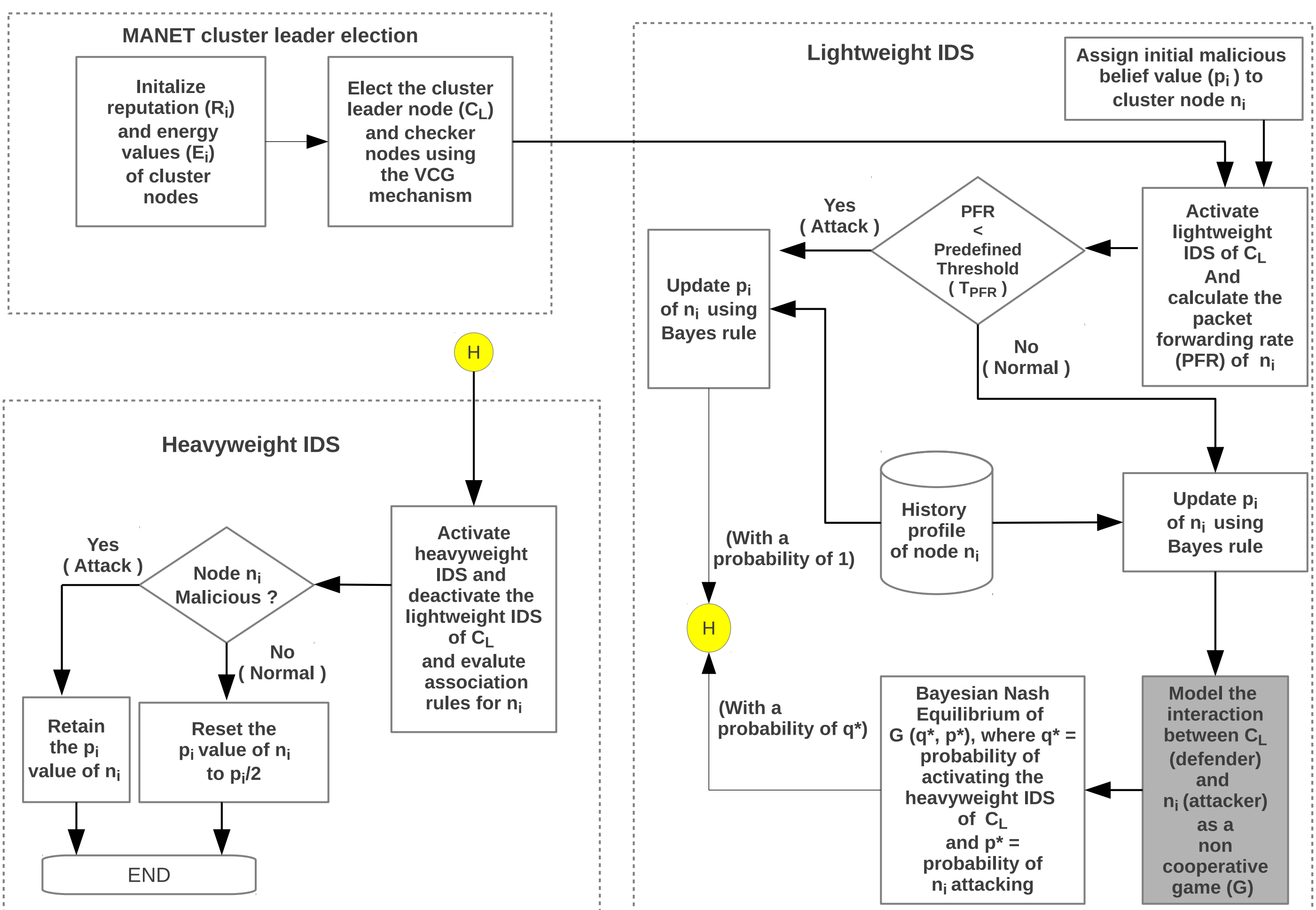

Fig. 1. Flowchart of the proposed MANET IDS scheme.

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

its strategy *Attack*. Any unilateral deviation by either players (C_L or n_i) from the BNE strategy reduces the payoff (increase in monitoring cost for C_L or increased probability of getting caught for n_i) of the deviating player. Therefore, in this case, the decision to activate the heavyweight IDS is probabilistic and depends on the BNE of the game. The probabilistic activation of the heavyweight IDS module is achieved by using a random number generator that generates a random number between 0 and 1. If the generated number is greater than or equal to the value of q^* then the heavyweight IDS module is activated; otherwise, it is not activated. The heavyweight module is an anomaly based IDS that uses association-rule mining technique to determine the action of n_i as attack or normal. If the action of n_i is found to be normal by the heavyweight module then the p_i value of n_i is reset to $p_i/2$; otherwise, the p_i value of n_i is retained.

The basic philosophy of the proposed hybrid IDS scheme is that, data packets in MANETs can be dropped due to various reasons like network congestion, depletion of node's resources, presence of malicious nodes etc. Nevertheless, excessive packet dropping is a strong indication of presence of malicious node in the network. Therefore, the calculation of node n_i's PFR value by the lightweight IDS module provides a strong insight into n_i being malicious or not. So, if a node n_i is ascertained to be malicious by the lightweight module, executing the heavyweight module is justified. However, a node can be malicious but still maintain its PDR above the threshold value by carrying out sniffing and probe types of attacks. Therefore, probabilistic activation of the heavyweight IDS module ensures the monitoring of such malicious nodes. Since the energy required for operating the heavyweight IDS module is comparatively higher than that required for operating the lightweight IDS module, using the lightweight IDS module as a precursor before activating the heavyweight IDS module reduces the overall power consumption required for operating the IDS. More elaborate details about the proposed hybrid MANET IDS is provided in section 3.3.

In the next section, we introduce the preliminaries of the game theory which is a prerequisite for developing monitoring strategies of the proposed hybrid MANET IDS.

3.1. Bayesian game model for proposed MANET IDS

Game theory allows us to study events of conflict and cooperation between two or more rational decision makers (players) with different set of objectives and competing for the same set of resources. Therefore, game theory is concerned with finding the best actions for individual decision makers in such situations and recognizing stable outcomes.

The interaction between the monitoring node and the potential malicious node in a MANET can be represented as a two player static Bayesian game in which one of the player P_i is a potential attacker and the other player P_j is a defender. The private information of player P_i is its type θ_i (*normal or malicious*). The type $\theta_i = 1$ if the player P_i is normal and $\theta_i = 0$ if it is malicious. This private information regarding the type of player P_i is unknown to the defender player P_j. The type of the defender player is always normal and denoted by $\theta_j = 1$, which is a common knowledge known to both the players. The attacker player of type $\theta_i = 0$ has two pure strategies: {*Attack, Not attack*} while the normal player of type $\theta_i = 1$ has only one pure strategy: {*Not attack*}. Similarly, the defender player P_j has two pure strategies: {*Monitor, Not monitor*}.

Both the players simultaneously choose their strategies at the beginning of the game with prior knowledge about the costs involved in monitoring and attacking any given node in the network along with the beliefs about the types of their opponents. This non-cooperative incomplete information game between the two players P_i and P_j can be represented as a triplet $G = \langle N, S, U \rangle$, where

Table 1
Payoff matrix when player P_i is malicious.

	Monitor	Not Monitor
Attack	$(1-2\alpha)w_k - C_{a_k},\ (2\alpha-1)w_k - C_{m_k}$	$w_k - C_{a_k}, -w_k$
Not Attack	$0,\ -\gamma w_k - C_{m_k}$	0, 0

- $N = \{P_i, P_j\}$ are the two players of the game.
- $S = S_i \times S_j$ is the strategy space of the game with S_i and S_j being the strategy space of players P_i and P_j, respectively.
- $U = U_i \times U_j$ is the payoff utility corresponding to the strategy space S. U_i and U_j are the payoffs of players P_i and P_j corresponding to their strategy spaces S_i and S_j, respectively.

In the subsequent sections, the terms player and node refer to the same entity and we use them interchangeably. Let $C = \{n_1, n_2, \ldots, n_t\}$ be a set of t nodes in a given MANET cluster. Consider any given node $n_k \in C$, where k $(1 \le k \le t)$ is the index of n_k and the asset value of n_k is w_k. Therefore, the symbol k in n_k refers to the index number of the k^{th} node in the given cluster and w_k refers to the associated asset value of the node n_k. The loss of asset when the attacker player P_i successfully exploits the node n_k represents a loss, whose value is equivalent to degree of damage such as loss of reputation, compromise of data integrity, cost of controlling damages etc. The defender player P_j is the cluster leader node. P_j is equipped with an IDS and is entrusted with the responsibility of providing intrusion detecting services to all other cluster nodes. Let the detection rate and the false alarm rate (*FP rate*) of P_j's IDS be denoted by α and γ, respectively where $\alpha, \gamma \in [0, 1]$. Let the cost involved in attacking the node n_k by P_i be denoted by C_{a_k} and the cost involved in monitoring the node n_k by P_j be denoted by C_{m_k}.

Tables 1 and 2 show the payoff matrices corresponding to the interaction between players P_i and P_j over the node n_k whose asset value is worth w_k, when the type of P_i is malicious and normal, respectively. These tables define various payoffs obtained by the defender and the attacker/normal players when interacting over a node n_k. The following conclusions can be drawn from Table 1, when the type of player P_i is malicious.

- When the malicious player P_i attacks and the defender player P_j monitors, i.e., for strategy combination $S_1 = $ (*Attack, Not Monitor*), the defender player P_j gets a payoff

  $$U_j(S_1) = -w_k$$

  which represents the loss of asset worth w_k. On the other hand, for this strategy, the malicious player P_i receives a payoff which is its gain from the successful exploitation of node n_k minus the cost involved in attacking the node n_k (C_{a_k}). Therefore, the payoff utility of player P_i with strategy S_1 is

  $$U_i(S_1) = w_k - C_{a_k}$$

- For strategy combination $S_2 = $ (*Attack, Monitor*), the defender player P_j's payoff is the gain from successful attack detection against node n_k minus the monitoring cost C_{m_k}. However, successful attack detection against node n_k depends on the detection rate (α) of the IDS monitoring the node n_k. Therefore, the payoff utility of defender player P_j playing strategy S_2 is

Table 2
Payoff matrix when player P_i is normal.

	Monitor	Not Monitor
Not Attack	$0,\ -\gamma w_k - C_{m_k}$	0, 0

$$U_j(S_2) = \alpha w_k - (1-\alpha) w_k - C_{m_k}$$
$$= (2\alpha - 1) w_k - C_{m_k}$$

where $(1 - \alpha)$ represents the false negative rate of the IDS. On the other hand, the malicious player P_i's loss after being caught is equal to player P_j's gain minus the attacking cost C_{a_k}. Therefore, player P_i's payoff utility with strategy S_2 is

$$U_i(S_2) = (1-2\alpha) w_k - C_{a_k}$$

- For the strategy $S_3 =$ (*Not Attack, Monitor*), the defender P_j's expected loss is $-\gamma w_k$ due to false alarm of IDS plus the monitoring cost C_{m_k}, while the payoff of malicious player P_i is 0. Therefore, the payoff utilities of players P_j and P_i with strategy S_3 are

$$U_j(S_3) = -\gamma w_k - C_{m_k}$$
$$U_i(S_3) = 0$$

- For the strategy $S_4 =$ (*Not Attack, Not Monitor*) the payoffs of both the players are 0, i.e., $U_j(S_4) = U_i(S_4) = 0$.

Similarly from Table 2, we observe that when the type of player P_i is normal, the payoff of player P_i is always 0. The payoff of defender player P_j is 0 if it plays its pure strategy (*Not Monitor*). On the other hand, if it plays its pure strategy (*Monitor*) its payoff utility is $-\gamma w_k - C_{m_k}$, which is the cost incurred due to false IDS alarms and the monitoring cost.

3.1.1. Bayesian Nash Equilibrium (BNE) analysis

Fig. 2 shows the extensive form of the Bayesian Game described in the preceding section. This game is also an imperfect information game since the defender player P_j is not aware about the type (*Normal, Malicious*) and action (*Attack, Not Attack*) of the player P_i while choosing its own action (*Monitor, Not Monitor*). In Fig. 2, **N** is the nature node that determines the type of player P_i. Let p_o be the prior probability of player P_i being malicious. We make an implicit assumption that both players are rational and their main objective is to maximize their respective payoffs. The attacker would want to play a strategy that minimizes its probability of being detected by the IDS while the defender would like to play a strategy that maximizes its probability of successfully detecting the attack.

In the subsequent section, we analyze the BNE of the game assuming that player P_j's prior belief (p_o) about player P_i being malicious is a common prior, i.e., player P_i (attacker) knows player P_j's (defender) belief about player P_i being malicious. We make the following observations about the Bayesian game described by Tables 1 and 2 and Fig. 2.

- If the type of player P_i is malicious and if it plays its pure strategy *Attack* then the expected payoff of player P_j playing its pure strategy *Monitor* is:

$$U_j(Monitor) = p_o((2\alpha - 1) w_k - C_{m_k}) - (1 - p_o)(\gamma w_k + C_{m_k})$$

and when it plays its pure strategy *Not Monitor*, its expected payoff is:

$$U_j(Not\ Monitor) = -p_o w_k$$

- When the defender player P_j plays its pure strategy *Monitor*, the expected payoffs of malicious player P_i playing its pure strategies *Attack* and *Not Attack* are:

$$U_i(Attack) = p_o((1-2\alpha) w_k - C_{a_k}) \quad and$$
$$U_i(Not\ Attack) = 0, \quad respectively.$$

- Therefore, if $U_j(Monitor) > U_j(Not\ Monitor)$, i.e., if $p_o > \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma) w_k}$, the best response of the player P_j is to play its pure strategy *Monitor*. However, when player P_j plays its pure strategy *Monitor*, the best response of player P_i would be to play its pure strategy *Not Attack*. Hence the strategy ((*Attack* if malicious, *Not Attack* if normal), *Monitor*, p_o) is not a BNE, when $p_o > \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma) w_k}$. Similarly, if $U_j(Monitor) < U_j(Not\ Monitor)$ i.e., if $p_o < \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma) w_k}$, the best response of player P_j is to play *Not Monitor*, since in this case the payoff obtained by playing strategy *Monitor* is less than the payoff obtained by playing strategy *Not Monitor*. Therefore, ((*Attack* if

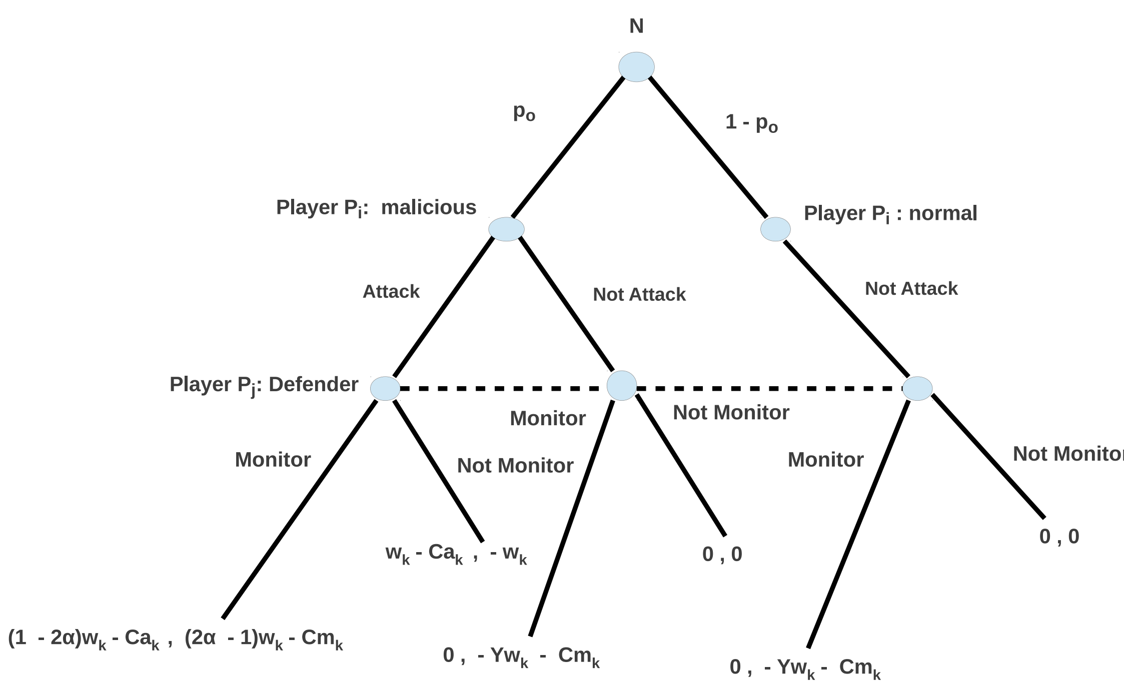

Fig. 2. Extensive form of the Bayesian game.

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

malicious, *Not Attack* if normal), *Not Monitor*, p_o) is a pure strategy BNE, when $p_o < \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma)w_k}$.

- If the player P_i plays its pure strategy *Not Attack*, then the player P_j's dominant strategy is to play *Not Monitor* regardless of the value of p_o. However, if the player P_j plays *Not Monitor*, the best response of player P_i if its type is *malicious* is to play *Attack*. Therefore, the strategy ((*Not Attack* if malicious, *Not Attack* if normal), *Not Monitor*) is not a BNE.

From our previous discussions we have shown that when $p_o > \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma)w_k}$, then there does not exist any pure-strategy BNE. But any game with a finite set of players and finite set of strategies has a Nash equilibrium of mixed strategies. Therefore, in such case where no pure strategy BNE exists, we derive a mixed strategy BNE for the game.

Let the player P_i play its strategy *Attack* with probability p if its type is *malicious* and play its pure strategy *Not Attack* if its type is *Normal*. In this case, the expected payoff of the defender player P_j playing its pure strategy *Monitor* is:

$$U_j(Monitor) = pp_o((2\alpha-1)w_k - C_{m_k}) - (1-p)p_o(\gamma w_k + C_{m_k}) \\ - (1-p_o)(\gamma w_k + C_{m_k})$$

and the expected payoff of the defender player P_j playing its pure strategy *Not Monitor* is:

$$U_j(Not\ Monitor) = -pp_o w_k$$

Similarly, the expected payoffs of attacker player P_i playing its pure strategies *Attack* and *Not Attack* when the defender player P_j plays its strategy *Monitor* with probability q and *Not Monitor* with probability $(1 - q)$ are:

$$U_i(Attack) = p_o(q((1-2\alpha)w_k - C_{a_k}) + (1-q)(w_k - C_{a_k})) \quad and \\ U_i(Not\ Attack) = 0, \quad respectively.$$

By equating $U_j(Monitor) = U_j(Not\ Monitor)$, we get $p = \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma)w_k p_o}$, which is the equilibrium strategy probability of malicious player P_i to play its pure strategy *Attack*. Similarly, by equating $U_i(Attack) = U_i(Not\ Attack)$, the player P_j's equilibrium strategy probability to play *Monitor* is $q = \frac{w_k - C_{a_k}}{2\alpha w_k}$. Therefore, when the prior probability of player P_i being malicious i.e., $p_o > \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma)w_k}$, no pure strategy BNE exists. But there exists a mixed-strategy BNE which corresponds to the strategy pair ((*Attack* with probability p if malicious, *Not Attack* if normal), *Monitor* with probability q, p_o), where $p = \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma)w_k p_o}$ and $q = \frac{w_k - C_{a_k}}{2\alpha w_k}$.

From the BNE strategy obtained above, we observe that the monitoring probability (q) of the defender does not depend on the current maliciousness belief of the opponent (attacker) player, but rather influences the attacker's behavior, as the probability of attack (p) is inversely proportional to the defender's maliciousness belief about the attacker player. A high maliciousness belief of the defender on its opponent results in the attacker drastically reducing its attack. This is a result of the fact that both the attacker and the defender are rational players and the cost and maliciousness beliefs are common knowledge known to both players.

The static Bayesian game approach described above can be used to model most types of attacks in MANETs like Denial of Service (DoS) attacks, network routing protocol disruption attacks like blackhole attack [42] and wormhole attack [43] etc. The proposed Bayesian game model enables the defender to implement its monitoring strategy based on its BNE solution that maximizes its expected payoff without requiring the IDS to be running all the time. However, the drawback of the scheme is that it is not always easy to determine the prior malicious belief (p_o) about the type of the opponent player in dynamic and distributed networks. Therefore, depending on the nature of the environment it is operating in, the defender may assign an appropriate value for p_o. If the environment is hostile, a high value of p_o should be assigned.

3.2. Energy efficient MANET IDS leader election mechanism

MANET nodes are essentially selfish in nature to preserve their energies. Taking this fact into account, Mohammed et al. [44] proposed a secure leader election mechanism for MANET. They simply treated IDS as a service and developed a computational cost metric for electing the leader node without considering metrics such as detection rate and false positive rate. In this section, we build on their work and develop a secure MANET leader election mechanism. We then integrate this mechanism with the dynamic hybrid IDS model proposed in section 3.3 and eventually evaluate the performance of the overall IDS scheme.

We model MANET as a set of clusters. Nodes in each cluster elect a leader node which carry out intrusion detection services for the entire set of cluster nodes for a predefined period of time (one slot period). Re-election is conducted to elect a new leader node after the timer expires. In most conventional schemes, the IDS operates in a promiscuous mode in all cluster nodes with a predefined sampling rate. This can have an adverse impact on the overall lifetime of the network as most of the node's energy is consumed for operating the IDS irrespective of whether intrusions take place or not. Contrary to this, in our proposed scheme, only the elected leader node operates the IDS and provides intrusion detection services to all other cluster nodes. This ensures that the power consumption required for operating the IDS in each individual cluster node is minimized through distribution of intrusion detection task among various MANET nodes.

The mechanism that elects a random node as a cluster leader [22] without considering energy level of nodes causes faster death of nodes with low energy levels. Therefore, the election mechanism must take into consideration the energy level of nodes while electing the leader node. Moreover, there are some selfish nodes in the cluster that are unwilling to participate in the intrusion detection process to preserve their resources (CPU time, energy etc). To address these issues, we propose a reputation based leader node election mechanism to encourage all cluster nodes including the selfish ones to participate in the leader node election process by truthfully revealing their energy levels. The elected leader node is provided with a payment in the form of reputation gain. Nodes with higher reputations are considered as more trusted nodes and given higher priorities in the cluster's services.

The sampling budget allotted by the leader node to any given node in the cluster is proportional to its reputation. The sampling budget (SB_{n_i}) of the i^{th} node (n_i) in the cluster denotes the amount of service it is entitled to receive from the leader node at the current game stage and is given as:

$$SB_{n_i} = (R_i) \Big/ \sum_{j=1}^{N} R_j$$

where N is the total number of nodes in the cluster under consideration and R_i is the reputation value of node n_i.

Every time a given node is elected as a leader its reputation value increases. This motivates the cluster nodes to truthfully reveal their private information (energy levels) during the leader-node election process. A default reputation value of R_o is assigned to all nodes during the cluster formation period, which gets updated when the node is elected as a cluster leader.

Let the energy required by the cluster leader node to operate the IDS for the elected period of time be denoted by E_{ids} and its cost for intrusion detection analysis during this period be denoted by Cst_i. We divide the N nodes in the cluster into k energy classes $\{Class_1, Class_2, \ldots, Class_k\}$ based on their power factor denoted by

$PF_i = E_i/NT_i$, where $1 \le i \le k$, E_i is the energy level of node n_i and T_i is the user defined scaling factor.

$$Class\ of\ n_i = \begin{cases} Class_1, & \text{if } PF_i < \rho_1 \\ Class_d, & \text{if } \rho_{d-1} \le PF_i < \rho_d \\ Class_k, & \text{if } PF_i \ge \rho_{k-1} \end{cases}$$

where $\rho = \{\rho_1, \rho_2, \ldots, \rho_{k-1}\}$ is a set of $(k-1)$ threshold values. The cost analysis value of node $n_i \in Class_i$ for analyzing data packets for specified period of time is given as:

$$Cst_i = \begin{cases} \left(\dfrac{\lambda * SB_{n_i}}{PF_i}\right) = \lambda * \left(\dfrac{R_i}{\sum_{i=1}^{N}(R_i)}\right) * \dfrac{NT_i}{E_i}, & \text{if } E_{n_i} \ge E_{ids} \\ \infty, & \text{if } E_{n_i} < E_{ids} \end{cases}$$

where $\lambda \in [0,1]$ is the sampling budget weighing factor. If the energy level of any node n_j is less than the threshold energy required for carrying out intrusion detection analysis i.e., if $E_{n_j} < E_{ids}$, then node n_j cannot be elected as a cluster leader since its cost of analysis would be infinite.

To motivate all nodes in the cluster including the selfish ones for cooperation, we model the leader node election problem as a game with mobile nodes as its players. Each node n_i holds a confidential information θ_i about its type. The type of θ_i can be either *Normal* or *Selfish*. The payoff utility function of player (node) n_i is given by:

$$U_i(\theta_i, \theta_{-i}) = P_i - W_i(\theta_i, O(\theta_i, \theta_{-i})) \quad (1)$$

where

- θ_{-i} represents the types of all other cluster nodes except node n_i
- $O(\theta_i, \theta_{-i}) = O(\theta_1, \ldots, \theta_i, \ldots, \theta_N)$ is the output corresponding to the types chosen by the players of the game.
- W_i is the cost of analysis (Cst_i) incurred by node n_i for providing intrusion detection services. However, if n_i is not elected as a leader, then W_i is 0 since no cost will be incurred to run the IDS.
- $P_i \in \mathbb{R}$ is the payment provided in the form of reputation to the elected leader node.

Each node n_i seeks to maximize its utility U_i. It signifies the amount of gain obtained by the player n_i if it follows the type θ_i. Player n_i might deviate from revealing its true cost analysis value Cst_i by either under-valuing or exaggerating its Cst_i value if doing so leads to better payoff. Therefore, we need to develop a mechanism with truth-telling as its dominant strategy.

The game begins with every node selecting its type θ_i and evaluating its cost of analysis value W_i. The objective of our mechanism design is to elect a node n_i with the least cost analysis value (Cst_i) as a cluster leader. Since $Cst_i \propto 1/E_i$, electing node with least cost analysis value is equivalent to electing a node with highest energy level. We refer to this objective as a Social Choice Function (SCF) and is defined as:

$$SCF = Min\{W_i(\theta_i, O(\theta_i, \theta_{-i})) \quad i = 1, 2, \ldots, N\} \quad (2)$$

If two or more nodes in the cluster have the same cost analysis value, then the node having the highest reputation among them will be elected as the cluster leader by the *SCF*. Payment in the form of reputation is made to the elected leader node using a VCG mechanism [24]. The amount of service provided by the elected leader node to any given node n_k is proportional to its reputation (R_k). The payment P_i received by the leader node n_i in the form of reputation (R_i) is equal to the second least cost analysis value C_j excluding the cost analysis value of the leader node n_i and is given by Equation (3).

$$P_i = R_i = Min\{W_j(\theta_j, O(\theta_j, \theta_{-j}) \quad j \ne i\} \quad (3)$$

$n_i \rightarrow cluster^{I}_{-n_i}$: $Begin\_Election\Big(\mathrm{H}(ID_{n_i}, Cst_i, TS_i), T_1\Big)$
$n_i \rightarrow cluster^{I}_{-n_i}$: $Election\Big(ID_{n_i}, Cst_i, TS_i\Big)$
if $Leader_{IDS} \ne n_i$; **then**
  $n_i \xrightarrow{Elected} Leader_{IDS}$
  $Leader_{IDS} \xrightarrow{Confirmation} n_i$
  $n_i \xrightarrow{Payment(R_{Leader_{IDS}})} Leader_{IDS}$
else
  After time T_2
  $n_i \xrightarrow{Confirmation} cluster^{I}_{-n_i}$
  $cluster^{I}_{-n_i} \xrightarrow{Payment(R_{n_i})} n_i$
end

Algorithm 1: Distributed MANET leader node election algorithm

We model MANET as a set of clusters as shown in Fig. 3. Based on the cost analysis value of different nodes, the leader election mechanism computes the SCF in a distributed manner which ensures that all nodes in the cluster elects the same leader. Algorithm 1 illustrates our proposed distributed leader election algorithm in a MANET cluster. Initially, a random node n_i initiates the election process by sending a *Begin_Election* message to all the other nodes in the cluster. The *Begin_Election* message contains the hash value H() corresponding to *Election* message to be sent by the leader node n_i later on. The receiving nodes use this hash value to authenticate and verify the *Election* messages received from node n_i. The time T_1 specifies the duration of the election process. All the participating nodes should interchange the *Begin_Election* messages within time T_1 after the node n_i has started the election process. Those nodes that do not participate in the exchange of *Begin_Election* messages are excluded from cluster's services.

After the completion of exchanges of *Begin_Election* messages the node n_i broadcasts the *Election* message containing its identity ID_{n_i}, its cost analysis value (Cst_i), and the time stamp TS_i to other nodes in its cluster. The receiver nodes then verify that the *Election* message indeed came from node n_i by generating a hash value $\mathbf{H^*()}$ of the received *Election* message. This generated hash value is then compared with the hash value $\mathbf{H()}$ received in *Begin_Election* message earlier. Upon successful verification, each node in the cluster computes the SCF, which is the least cost analysis value as defined in Equation (2).

After the completion of exchanges of *Begin_Election* messages between the nodes, if the elected leader node as per the SCF is different from node n_i, then the node n_i sends an *Elected* message to the chosen leader node. The elected leader node on receiving the *Elected* message sends back the *Confirmation* message to node n_i. The node n_i then calculates the payment $Payment(R_{Leader_{IDS}})$ for leader node using the VCG mechanism as described in Equation (3). The node n_i increases the reputation of the elected leader node ($Leader_{IDS}$) by value $Payment(R_{Leader_{IDS}})$ in its reputation table. However, if the node n_i finds itself to be the elected leader after calculating the SCF, then it sets the timer T_2 and starts verifying all the *Elected* messages from other nodes. If the timer T_2 expires without receiving *Elected* messages from all the nodes, then those nodes that did not participate in the leader election process are debarred from cluster's services. The node n_i then sends the *Confirmation* messages back to the nodes from which it received the *Elected* messages. Upon receiving the *Confirmation* message, other cluster nodes calculate the payment for node n_i and update their reputation tables.

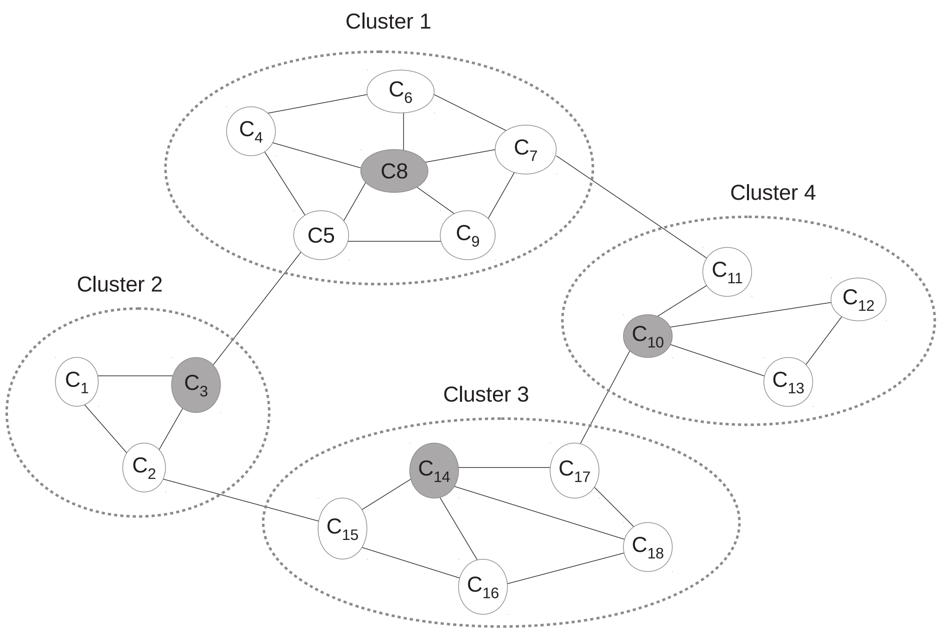

Fig. 3. MANET topology with leader IDS.

Table 3
Leader IDS election example.

Nodes	N_1	N_2	N_3	N_4	N_5	N_6
i^{th} round reputation	**7**	9	2	4	5	3
i^{th} round energy	5	6	4	5	10	7
i^{th} round sampling (%)	23.33	30	6.66	13.33	16.66	10
i^{th} round cost valuation (Cst_i)	0.28	0.30	0.1	0.16	0.1	0.09
$(i+1)^{th}$ round reputation	7	9	2	4	5	3.1
$(i+1)^{th}$ round energy	5	6	4	5	10	6.8

The election process is repeated after every T_{elect} time interval. If the cluster has not changed after the time interval T_{elect}, then the cluster formation step is skipped and only the leader election process is carried out. Re-election is also conducted when the elected leader node quits the cluster before the completion of T_{elect} time interval.

We illustrate the proposed leader election scheme with an example as shown in Table 3. The reputations of different nodes at i^{th} round are shown in the 1^{st} row of table with node N_1 elected as a leader node. The 2^{nd} row gives the energy level of different nodes at the i^{th} round. The leader node's sampling budget for different nodes (in terms of percentage) is shown in the 3^{rd} row.

The election of new leader node for the $(i+1)^{th}$ round involves every node to compute its corresponding cost analysis value Cst_i as shown in 4^{th} row using the following Equation:

$$Cst_i = \left(\frac{\lambda * SB_{n_i}}{PF_i}\right) = \lambda * \left(\frac{R_i}{\sum_{i=1}^{N}(R_i)}\right) \times \frac{NT_i}{E_i}$$

For Table 3 the value of N is 6. The values of λ and T_i are assumed to be 0.1 and 10, respectively. Similarly, the energy required for operating the IDS is assumed to be 0.2 units. Since node N_6 has the least cost analysis value (0.09), it is elected as a new leader node. Nodes then calculate the payment for the new elected node N_6, which is equal to the 2^{nd} least cost analysis value i.e., $P_i = 0.1$ unit. All the nodes increase the reputation of the elected node N_6 by 0.1. The new reputations of different nodes at $(i+1)^{th}$ round are shown in the 5^{th} row. The payoff utility of node N_6 calculated using Equation (1) is $0.1 - 0.09 = 0.01$, which represents the benefit gained by the node N_6.

3.2.1. *Mechanism analysis*

The primary objective of our mechanism design is to encourage players (nodes) into truthfully revealing their private information by providing them incentives for doing so. In this section, we validate our mechanism design to ensure that our proposed model meets the cost-efficiency and truthfulness properties even in the presence of malicious and selfish nodes in the cluster. This is validated by demonstrating that truth-telling is the dominant strategy of our mechanism.

We consider two untruthful revelations of selfish node n_i *viz. under-declaration* and *over-declaration* of its cost analysis value Cst_i, and show that in both cases it is never better off compared to when it truthfully reveals its cost analysis value.

Node n_i may under-declare its cost analysis value by revealing a false value W_i^*, where $W_i^* < W_i$. By declaring a false cheaper cost analysis value, node n_i wins the cluster leader election. However, under-declaring its cost analysis value will not benefit the node n_i for the following two reasons. In the 1^{st} case, if the real cost analysis value W_i of the node is already least among all the nodes, then under-valuing its cost analysis value to W_i^* does not increase its payment, since payments are made on the basis of second least cost analysis value. Therefore, its utility function U_i remains unchanged since it is calculated with respect to its real cost analysis value W_i. On the other hand, if the node n_i does not have the least cost analysis value but wins the election by declaring a fake under-valued cost analysis value W_i^* then it leads to negative utility function U_i. This is because the payment P_i received by node n_i is less than the real cost analysis value W_i. Therefore, in this case, the work done by node n_i exceeds the amount of payment P_i that it receives.

Similarly, in case of *over-declaration*, if a node n_i over-declares its cost analysis value by declaring a fake W_i^*, where $W_i^* > W_i$, then such a strategy would never increase the payoff utility U_i for the following two reasons. First, if node n_i indeed has the least cost analysis value W_i, then pursuing this strategy leads to node n_i not being elected as the leader node and hence it loses the payment. Second,

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

if the real cost analysis value W_i of node n_i is not the least among all the nodes, then this strategy would never increase its payoff utility U_i as the node n_i would not be chosen as a leader node.

3.2.2. Cooperative catch and punish model

The leader node may misbehave after being elected. Therefore, we need a mechanism to detect misbehaving leader node and take appropriate measures. A leader node is said to be misbehaving if it does not provide intrusion detection services to cluster nodes proportional to their reputations. Checker nodes are employed to monitor the misbehaving leader node. The checker nodes perform a small part of the computation executed by the leader node to determine the misbehavior of the leader node. Let Chk_{cost} be the cost incurred by any given checker to monitor the leader node for elected period of time. Incentives in the form of checker reputation payments P_{chk} are provided to the checker nodes for monitoring the leader node such that $P_{chk} - Chk_{cost} > 0$.

```
k-checkers ← k random nodes excluding the Leader_IDS with least cost analysis values.
(N − k) ← Non-Leader IDS node + Non-checker nodes.
for i = 1, i++, i < N
if n_i ≠ k-checkers ‖ Leader_IDS; then
  |  n_i --Chk_ele--> k-checkers
  |  k-checkers --Chk_conf--> n_i
  |  n_i --P_chk--> k-checkers
else
  |  After time T_2
  |  n_i --Chk_conf--> (N − k) non checker nodes /* if n_i is checker */
  |  (N − k) non checker nodes --P_chk--> n_i
end
```

Algorithm 2: Distributed checkers election algorithm

Algorithm 2 illustrates the proposed mechanism for election of checker nodes. Initially, k nodes in the cluster with least cost analysis Cst_i excluding the leader node are chosen as checkers. Each node n_i in a cluster then verifies whether it is one of the k checkers. If it is not a checker then it sends a Chk_{ele} message to all the k checkers to inform them that they have been elected as a checker. The k checkers then send back a checker confirmation message Chk_{conf} to node n_i. Upon receiving the confirmation messages from the checker, the node n_i increments the reputation of the k checker nodes in its reputation table by P_{chk} reputation units. After time interval T_2, if the checker node n_i has not yet received a Chk_{ele} message from any of the non-checker nodes then it sends Chk_{conf} messages to all the non-checker nodes from which it has not yet received the Chk_{ele} message. Upon receiving the Chk_{conf} message from the checker node n_i, the non-checker node n_j verifies that the Chk_{conf} message indeed came from one of the checkers by referring to its cost analysis table. Upon successful verification, the receiver node updates its reputation table by incrementing the reputation of checker node by P_{chk} reputation units.

If the leader node n_i is found to be misbehaving by the checker nodes, the mechanism punishes the leader node by lowering its reputation and paying it a negative payment value $-p_j$ i.e., the mechanism instructs all the cluster nodes to decrement the reputation of leader node in their reputation table by value R_i as calculated in Equation (3). Leader node election is then conducted to elect a new leader.

To detect a misbehaving leader node, we propose a set of detection level given by $DL = \{dl_1, dl_2, \ldots, dl_j\}$. The proposed catch and punish model comprises j detection-levels with each level representing the severity of the misbehaving leader node. We define a threshold set $T = \{t_1, t_2, \ldots, t_{j-1}\}$ to categorize the misbehaving detection levels. Setting the threshold value above which the leader node is considered to be misbehaving is crucial. Setting this threshold value too high increases the false positive (FP) rate wherein even the sincere leader nodes are penalized whereas setting it too low increases the false negative (FN) rate wherein the mechanism fails to catch the misbehaving leader node. Therefore, this value must be set appropriately so as to balance and maintain a good trade-off between the FP and the FN rates.

Let $Chk_{set} = (Chk_1, Chk_2, \ldots, Chk_x)$ be the set of checker nodes and $S_{set} = (n_a, n_b, \ldots, n_x)$ be the set of nodes monitored by the checkers such that $|Chk_{set}| = |S_{set}|$. Each $Chk_i \in Chk_{set}$ monitors one of the nodes $n_j \in S_{set}$. We then define an aggregate function of checkers as:

$$T(n) = \sum_{i \in Chk_{set} \,\&\, j \in S_{set}} (R_i) * f(j) \tag{4}$$

where R_i is the reputation of the checker node Chk_i and $f(j)$ is the *catch function* defined as the ratio of actual number of data packets analyzed by the leader node for node n_j ($n_j \in S_{set}$) to the actual sampling budget allocation of node n_j as observed by the checker node Chk_i. We then classify the detection-levels as follows:

$$DL = \begin{cases} dl_1, & \text{if } T(n) < t_1 \\ dl_f, & \text{if } t_{f-1} \le T(n) < t_f;\ f \in [2, j-1] \\ dl_j, & \text{if } T(n) \ge t_{j-1} \end{cases}$$

Grouping the severity of misbehaving leader node into j different levels minimizes the FP rate while determining the misbehaving leader node. The leader node found misbehaving with detection level (DL) lower than the threshold level (dl_{th}) is penalized by computing its payment in negative and temporarily debarring it from cluster services. This acts as a deterrence and discourages the leader node from misbehaving. Hence a malicious node has no valid motivation to become a leader node since it has a high probability of being caught and punished by checker nodes.

3.3. Hybrid MANET IDS

In section 3.1, we discussed about static Bayesian game where the player P_j (defender) has a fixed prior belief (p_o) about the opponent player P_i being malicious. However, determining this prior belief is usually difficult and depends on the nature of the environment the IDS is operating on. Nodes in MANETs are usually energy constrained and may become less responsive as their energy levels drain out. In addition, some trustworthy nodes may be compromised over a period of time and made to act maliciously. Taking all these factors into account, the IDS needs to re-evaluate the malicious beliefs of MANET nodes at regular intervals. In this section, we extend the static Bayesian game to a multi-stage dynamic Bayesian game, wherein the defender player updates its maliciousness belief about the opponent player as the game evolves.

In the multi-stage Bayesian game, the game is played repeatedly after every time interval t_k. However, the payoffs of the game and the identities of the players remain the same throughout each iteration of the game. The strategies of players in the dynamic game depends on the history profile of the game. At any stage t_k of the game, the optimal strategy of the attacker player P_i depends on the maliciousness belief of the defender player P_j about P_i. The defender player P_j's initial belief about player P_i being malicious at the first stage (t_0) of the game is given by the prior probability p_o. The defender player P_j then updates its malicious belief about the opponent player P_i at the k^{th} stage of the game by evaluating its posterior belief $p_j(\theta_i|a_i(t_k), a_i(t_{k-1}))$, where $a_i(t_k)$ and $a_i(t_{k-1})$ represent the actions taken by the player P_i at the k^{th} and $(k-1)^{th}$ stage of the game. The player P_j evaluates its posterior belief about player P_i using the following Bayes' rule.

$$p_j(\theta_i|a_i(t_k), a_i(t_{k-1})) = \frac{p_j(\theta_i|a_i(t_{k-1}))P(a_i(t_k)|\theta_i, a_i(t_{k-1}))}{\sum_{\tilde{\theta}_i} p_j(\tilde{\theta}_i|a_i(t_{k-1}))P(a_i(t_k)|\tilde{\theta}_i, a_i(t_{k-1}))} \tag{5}$$

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

where $P(a_i(t_k)|\theta_i, a_i(t_{k-1}))$ is the probability that the player P_i plays the action $a_i(t_k)$ at the k^{th} stage, given the type of player P_i is θ_i and its action at the $(k-1)^{th}$ stage was $a_i(t_{k-1})$.

From Equation (5), it can be observed that the defender player needs to continuously monitor the opponent player at every game stage to update its belief. However, operating IDS in an always-on promiscuous mode is not an energy-efficient monitoring strategy. Therefore, to minimize the energy spent on operating the IDS, we propose a two layered hybrid IDS detection model. The proposed hybrid model consists of one lightweight module and one heavyweight module. The former module is less powerful but requires less energy for its operation, while the latter module is more powerful but requires more energy to operate. By default, only the lightweight module is activated.

In Fig. 1, we have shown the proposed two layered hybrid IDS framework. The malicious belief of node n_i is updated using the input from the lightweight IDS module and the history profile of n_i's actions. The lightweight module calculates two parameters of n_i *viz.* its packet reception rate (PRR) and the packet forwarding rate (PFR). (However, in Fig. 1 only the PFR calculation is shown.) The details about these parameters are discussed in section 3.3.2. The lightweight IDS module updates the malicious belief of n_i using the observed behavior of n_i in the current and previous stage of the game by employing the Bayes rule. If the PRR or PFR values of n_i exceeds or falls below the threshold value, then the action of n_i is assumed to be attack and the heavyweight module is activated in the next stage of the game for more rigorous analysis. The maliciousness value of n_i can be unilaterally reset to a lower value by the heavyweight IDS module if n_i has not acted maliciously for a pre-defined period of time. After the maliciousness value of n_i is reset to lower value, the heavyweight IDS module is turned off and the lightweight IDS module is turned on. This process is repeated over the period of time and only one of the IDS module is activated at any given time.

3.3.1. Heavyweight intrusion detection system (HIDS)

The HIDS uses an unsupervised association-rule mining technique [45,46] on a set of packet-level transmission events to find the association patterns. The extracted association rules are then used to build the normal profile of the network. There is a trade-off between effectiveness and efficiency while selecting the feature set for IDS analysis. A higher number of features can help the IDS to detect various types of attacks; however, it also results in a higher power consumption and computational overhead. Considering the energy constrained MANET nodes, we select a minimum number of features for developing HIDS normal profile. The transmission events consist of features listed in Table 4 that are extracted from the MAC and network layer at a pre-defined sampling rate. A brief description about each of these features is provided below:

- *Packet event type*: This feature represents the type of the transmission event taking place.
- *Sender Address*: This feature represents the MAC address of the sender node.
- *Destination Address*: This feature represents the MAC address of the destination node.
- *MACFrameType*: This feature represents the type of MAC frame observed in the transmission event.
- *RoutPktType*: This feature represents routing control packets (routingCtrlPkt) like Route Request, Route Reply, Route Error etc. and data packets (routingDataPkt) from network layer.
- *Route change percentage*: It is defined as $(|S_2 - S_1| + |S_1 - S_2|)/|S_1|$, where $(S_2 - S_1)$ indicates the newly increased routing entries and $(S_1 - S_2)$ indicates the deleted routing entries during the time interval $(t_2 - t_1)$.

Table 4
HIDS feature set.

Features	Values
Packet event type (Event)	SEND, RECV, DROP, FWD
Sender Address (SA)	$SrcMAC_i$
Destination Address (DA)	$DestMAC_i$
MACFrameType	RTS, CTS, DATA, ACK
RoutPktType	routingCtrlPkt, routingDataPkt
Route change percentage	PCR

The HIDS uses multiple segments of training data set to extract the association rules. These rules are then aggregated to build the normal profile. The association rule describes the association of attributes within transaction records of an audit data set. Let $T = \{T_1, T_2, \ldots, T_n\}$ be the set of n transaction records and $F = \{F_1, F_2, \ldots, F_k\}$ be a k feature set defined over T. A transaction record T_i is a collection of k-tuple features i.e., $T_i = \{f_1, f_2, \ldots, f_k\}$, where f_k represents a value from the k^{th} feature F_k.

Let A and B denote two disjointed item subsets in T_i. The support of item subset A denoted by *sup(A)* represents the percentage of transactions containing A in T and the support of A and B denoted by $sup(A \cup B)$ represents the percentage of transactions containing both A and B. The association rule between A and B is given as $A \Rightarrow B, (s, c)$, where $s = sup(A \cup B)$ and $c = \frac{sup(A \cup B)}{sup(A)}$ are defined as the support value and confidence value of the association rule, respectively. The rule holds good if $s \geq minsup$ and $c \geq minconf$, where *minsup* and *minconf* denote the predefined minimum support threshold and minimum confidence threshold values, respectively.

Apriori algorithm [45] was used to build the association rules for the normal profile. The algorithm mines the frequent itemsets from the transactional dataset and uses an iterative approach to find itemsets of larger size at each iteration. The algorithm works on the principle that any subset of a frequent itemset must also be a frequent itemset. Therefore, the algorithm reduces the number of item candidates being considered by only exploring the itemsets whose support count is greater than the minimum support count. For our analysis, we have used *minsup* and *minconf* values as 15% and 70%, respectively.

A transaction record is a packet level event with the following format $\langle Event, SA, DA, MACFrameType, RoutPktType \rangle$. An example association rule is $(SrcMAC_6, routingCtrlPkt \rightarrow DestMAC_{15}, RECV),(0.35,1)$, which describes an event pattern related to the RECV flows of the monitoring node i.e., 35% of transaction records match the event of "node 6 sends data packets to node 15", and when node 15 receives data packets, they are 100% of the time from node 6. Another example is $(SrcMAC_3, routingCtrlPkt \rightarrow DestMAC_7, PCR),(0.20,0.80)$, which indicates that route change between node 3 and node 7 constitutes 20% of total route change in the network, and 80% of changes in node 7's route is related with change in node 3's route.

The association rules extracted from the test data (real time data) are then correlated with the normal profile and any deviation of the test association rules from the normal profile is considered as an anomaly by the HIDS.

3.3.2. Lightweight intrusion detection system (LIDS)

It is not efficient to operate the association-rule based HIDS in an always-on mode since it uses massive packet-level transmissions of network and MAC layers to detect intrusions. Therefore, we propose an alternative lightweight monitoring system (LIDS) to update the malicious belief of the defender node about the opponent node n_i on every stage of the game. The LIDS being a lightweight module uses simple rules and methods to detect intrusions. It uses two different approaches for detecting the inbound and outbound attacks. The following inbound attacks are considered in our study: *Sleep deprivation, Flooding, DoS* and *Forging attack*. The outbound attacks considered are *Black hole attack* and *packet dropping attack*. Let N_j represent the set of neighboring nodes of defender node P_j and let the potential

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

attacker node $P_i \in N_j$. Let $R_j^i(t_k)$ denote the number of data packets received by node P_j from node P_i during the game stage t_k.

We define the following two terminologies to determine the outbound and inbound attacks: Packet Reception Rate *(PRR)* and Packet Forwarding Rate *(PFR)*. The *PRR* of node P_j from node P_i for game stage t_k is defined as the rate of inbound data traffic from node P_i to node P_j with respect to the total data traffic rate in the vicinity of node P_j. It is given as:

$$\phi_j^i(t_k) = \frac{R_j^i(t_k)}{\sum_{a \neq b} R_{a \in N_j}^{b \in N_j}(t_k) + R_j^{b \in N_j}(t_k)} \tag{6}$$

If the value of *PRR* is greater than the threshold value τ, the action of the player P_i is assumed to be an inbound attack. Therefore, the action of node P_i i.e., $a_i(t_k)$ = inbound attack, if $\phi_j^i(t_k) > \tau$.

The *PFR* of node P_i for game stage t_k is defined as the ratio of number of packets received by the node P_i from its neighboring nodes to the number of packets forwarded by node P_i to its neighboring nodes (N_i) and is given by:

$$\psi_i(t_k) = \frac{R_i^{j \in N_i}(t_k)}{R_{k \in N_i}^i(t_k)} \tag{7}$$

The action of node P_i is implied to be an outbound attack if the value of $\psi_i(t_k)$ is less than the threshold value Θ. In other words, the action of node P_i i.e., $a_i(t_k)$ = outbound attack if $\psi_i(t_k) < \Theta$.

The choices of *PRR* and *PFR* threshold values τ and Θ influence the performance of the LIDS. These threshold values can be calculated experimentally from the normal data traffic patterns. Employing this simple analysis rule of LIDS as a precursor before applying the association-rules of the HIDS can significantly lower the FP rate of the overall IDS.

Let the detection rate and FP rate of LIDS be α_L and γ_L, respectively. Let $P(a_i(t_k)|\theta_i, a_i(t_{k-1}))$ be the conditional probability of player P_i playing action $a_i(t_k)$ at k^{th} stage of game, given its type θ_i and its action at the $(k-1)^{th}$ stage was $a_i(t_{k-1})$. This conditional probability can be updated as follows:

$$\begin{aligned} P(a_i(t_k) = Attack|\theta_i = 1, a_i(t_{k-1})) \\ = p\alpha_L + (1-p)\gamma_L \end{aligned} \tag{8}$$

$$\begin{aligned} P(a_i(t_k) = Not\ Attack|\theta_i = 1, a_i(t_{k-1})) \\ = p(1-\alpha_L) + (1-p)(1-\gamma_L) \end{aligned} \tag{9}$$

$$P(a_i(t_k) = Attack|\theta_i = 0, a_i(t_{k-1})) = \gamma_L \tag{10}$$

$$P(a_i(t_k) = Not\ Attack|\theta_i = 0, a_i(t_{k-1})) = 1-\gamma_L \tag{11}$$

In above equations, p represents the probability of the malicious player P_i to play its strategy *Attack* under Nash Equilibrium (NE). Similarly, $(1 - \alpha_L)$ and $(1 - \gamma_L)$ represent the false negative (FN) rate and the true negative (TN) rate of the LIDS, respectively. The LIDS can determine the action of the node P_i using Equation (6) and Equation (7). It then updates the maliciousness value of the player P_i using Equation (5) along with Equation (8) through Equation (11).

3.4. Numerical Example

Continuing with our standard notation, let α and γ be the detection rate and FP rate of the heavyweight IDS, respectively. Similarly, let α_L and γ_L be the detection rate and FP rate of the lightweight IDS, respectively. Consider a defender attacker game interacting over a node n_k. Let C_{m_k} and C_{a_k} be the cost associated with monitoring and attacking node n_k. Let the asset value of n_k be w_k. In previous sections, we have developed the BNE of the game, which corresponds to the strategy combination $(p^*, q^*, p(\theta_i))$, where $p^* = \frac{\gamma w_k + C_{m_k}}{(2\alpha+\gamma) w_k p(\theta_i)}$ is the attacking probability of the attacker player (P_i), $q^* = \frac{w_k - C_{a_k}}{2\alpha w_k}$ is the monitoring probability of the defender player P_j and $p(\theta_i)$ is the maliciousness belief of P_j about P_i, which is given by Equation (5). Consider a heavyweight and a lightweight module with the following values, $\alpha = 0.9178$, $\gamma = 0.0025$, $\alpha_L = 0.833$ and $\gamma_L = 0.0029$. Let $w_k = 9.45$ and $C_{a_k} = C_{m_k} = w_k/1000$. Assume that the initial belief of P_j about P_i being malicious is 0.5, i.e. initial value of $p(\theta_i) = 0.5$. Therefore, the probability of player P_i playing its strategy attack for the 1^{st} stage of the game is $p^* = \frac{0.0019}{p(\theta_i)} = \frac{0.0019}{0.5} = 0.0038$. Similarly, the monitoring probability $q^* = 0.5442$. Next, we update the malicious belief of player P_i under following conditions:

Case 1: The observed action of P_i by the lightweight module of P_j is *Attack*:

$$p(\theta_i = 1)(t_1) = \frac{p(\theta_i = 1)(t_0)\, P(a_i(t_1) = Attack|\theta_i = 1, a_i(t_0))}{\sum_{\tilde{\theta}} p(\tilde{\theta}_i)(t_0)\, P(a_i(t_k) = Attack|\tilde{\theta}_i, a_i(t_0))} = 0.6756$$

Case 2: The observed action of P_i by the lightweight module of P_j is *Not Attack*:

$$\begin{aligned} p(\theta_i = 1)(t_1) &= \frac{p(\theta_i = 1)(t_0)\, P(a_i(t_1) = Not\ Attack|\theta_i = 1, a_i(t_0))}{\sum_{\tilde{\theta}} p(\tilde{\theta}_i)(t_0)\, P(a_i(t_k) = Not\ Attack|\tilde{\theta}_i, a_i(t_0))} \\ &= 0.49920 \end{aligned}$$

From the above results, it can be observed that when the action of P_i is detected as an *Attack* by P_j (defender) then the maliciousness belief of P_j about P_i increases, which in turn decreases the probability of P_i to play its strategy *Attack* in the next game stage. On the other hand, when the action of P_i is detected as *Not Attack* by P_j, then P_j's malicious belief about P_i decreases, which increases the probability of P_i to play its strategy *Attack* in the next stage of the game. It can also be observed that the proposed hybrid MANET IDS reduces the power consumption by activating the heavyweight IDS module 54.42% of the time instead of turning it on 100% of the time.

Summarizing the above results and discussion, we conclude that the monitoring probability of the P_j does not depend on its current maliciousness belief about P_i, but rather influences the P_i's behavior. A high maliciousness belief results in P_i drastically reducing its attack. This is result of the fact that both P_i and P_j are rational players, and the cost and maliciousness beliefs are common knowledge for both the players.

4. Experimental results

Since our work comprises two different components, we classify our analysis into following two subsections:

- Analysis of MANET leader election mechanism.
- Analysis of the hybrid MANET IDS.
  1. Evaluate the detection rate and the FP rates of the lightweight module and the heavyweight module of the proposed hybrid MANET IDS.
  2. Evaluate the payoff utilities of the attacker and defender nodes under different BNE strategies.
  3. Analysis of reduction in IDS traffic generation achieved by the proposed MANET IDS scheme.
  4. Performance analysis comparison of the proposed MANET IDS scheme with other well known schemes.

We have implemented our proposed model in the network simulator NS2 [25] on Ubuntu 12.04 running gcc version 4.6.3. We restrict the movements of the mobile nodes to a predefined flat-grid area of 15×15 m^2. Table 5 lists the various parameters used in our simulation.

Table 5
Parameters used for simulation.

Parameters	Value
Simulation time	900–3000 s
Number of nodes	12–30
Simulation area	600×600 m^2
Transmission range	150 m
Mobility	Random way point
Routing protocol	DSR
MAC layer	DCF of IEEE 802.11
Max. node movement speed	20 m/s
Pause time	500 s
Traffic type	CBR/UDP
Election period	60 s
Data rate	20k bps
Packet size	512 Bytes
MAC protocol	IEEE 802.11b
Sampling interval	3 s

4.1. MANET leader election mechanism analysis

We analyze our proposed model to study the impact of our scheme (leader IDS election) on the average life span of nodes. Initially, nodes in the cluster are assigned energy levels between 5 and 50 Joules. The energy consumed by the leader IDS for elected period of time (15 s) is assumed to be 4 Joules. The energy required by nodes for their normal operations and transmissions has been ignored to simplify the analysis.

We analyze our proposed model in a cluster consisting of 12 nodes, with 25% i.e., 3 malicious nodes. Figs. 4 and 5 show the energy levels of different nodes using the random leader election model and the VCG leader election model, respectively. It can be observed that in the random model some of the nodes die out over a period of time, while the energy levels of other nodes remain constant or decrease marginally. On the other hand, the VCG mechanism based leader election model balances the energy levels of all nodes by always electing the most cost-efficient node (high-energy level node) as cluster leader. In general, it was found that the proposed leader election model increases the average lifetime of the cluster node by 15–20% compared to a random model that does not employ leader election mechanism.

Fig. 6 shows the percentage of normal alive nodes *versus* percentage of malicious nodes in a cluster consisting of 20 nodes after 2400 s. A malicious node avoids being elected as a leader node by exaggerating its cost analysis value. It can be observed from the figure

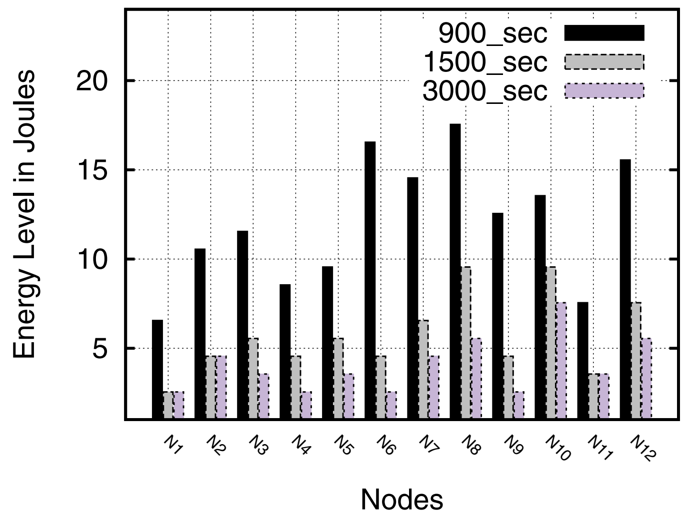

Fig. 5. Energy consumption using VCG model.

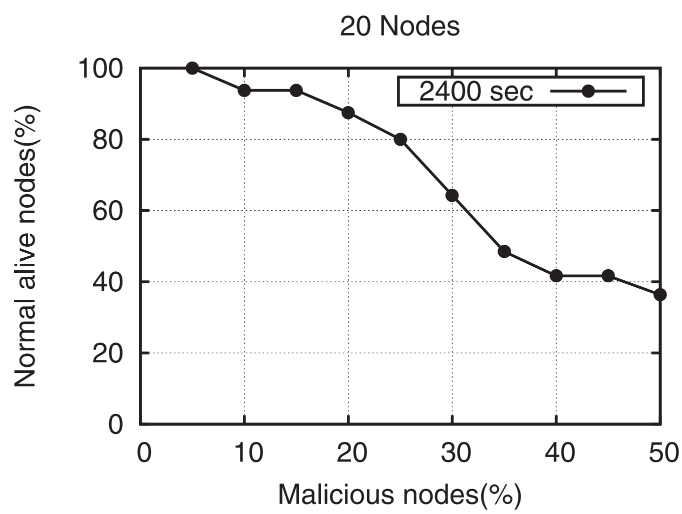

Fig. 6. Percentage of normal alive nodes *versus* percentage of malicious nodes.

that as the number of malicious node increases in the network, the number of alive normal nodes decreases. This shows that the normal nodes carry out more intrusion detection services and die out faster as the number of selfish nodes increase in the cluster.

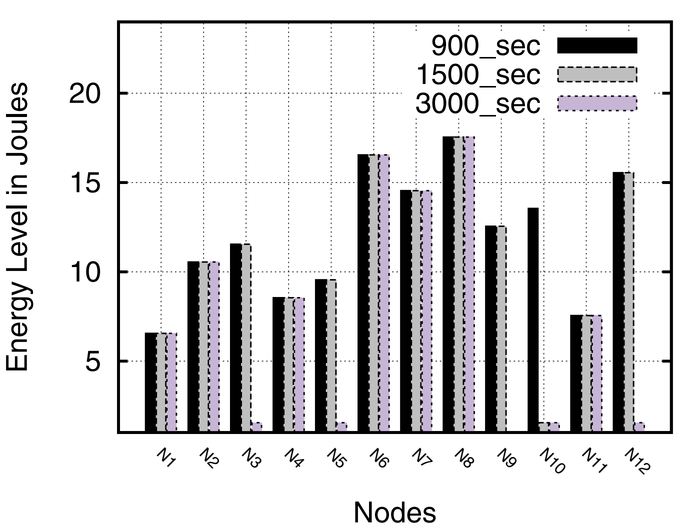

Fig. 4. Energy consumption using random model.

4.2. Hybrid MANET IDS analysis

For analyzing the proposed hybrid MANET IDS, the *Packet Reception Rate (PRR)* threshold (τ) and *Packet Forwarding Rate (PFR)* threshold (Θ) values of the lightweight module are taken as 0.5 and 0.3, respectively. The observed detection rate (α_L) and false positive rate (γ_L) of the lightweight module against different types of attacks like *DoS, Packet dropping, Packet distortion, Route compromise, Black-hole* etc. using the above *(PRR)* and *(PFR)* threshold values were found to be 81.33% and 0.61%, respectively.

The features listed in Table 4 are used to build the association rules for the heavyweight IDS module. We considered different sampling intervals for creating a training dataset, with each training instance containing a summary statistics of network activities for the specified time interval. The values of minimum support threshold (*minsup*) and minimum confidence threshold (*minconf*) are taken as 15% and 65%, respectively.

The performance analysis of association-rule based *HIDS* is carried out under different traffic conditions and against different types of attacks. Two different test scripts are used to generate training traces. 8k Trace and 5k Trace are normal training traces without any intrusions and with running time of 8000 s and 5000 s, respectively. The sampling rate of 3 s is used to record the feature values. The association rules extracted from these traces are then used to build the normal profile of the network.

Larger test traces with execution time from 10,000 (10k) seconds to 15,000 (15k) seconds were then generated. The association rules extracted from the test data (*real-time monitoring data*) were then compared against the normal profile. Any deviation of test association rules from the normal profile are considered as an anomaly, which triggers an intrusion alert. These test traces contain various types of attacks like *Route compromise*, *Traffic distortion* and *Black-hole attacks*. A brief description of these attack types is provided below:

- *Route compromise*: This type of attack involves either forwarding a packet to an incorrect node or propagating false route updates.
- *Traffic distortion*: These attacks change the normal traffic behavior by randomly dropping packets, generating packets with faked source address, reporting false misbehavior against normal node, corrupting the packet contents and denial of service.
- *Black-hole attack*: In this attack, a malicious node advertises spurious routing information, thus receiving packets and dropping them instead of forwarding them.

Table 6 shows the performance of the proposed unsupervised association-rule based *HIDS* against different types of attacks. It can be seen that the *HIDS* effectively detects the simulated attacks with

Table 6
Performance of association-rule based heavyweight IDS for different classes of attacks.

Attack Type	Detection rate	False alarm rate
Route compromise	91.4%	0.45%
Traffic distortion	95.3%	0.87%
Black-hole	99.5%	0.35%

Table 7
Performance of association-rule based heavyweight IDS.

Test trace	Detection rate	False alarm rate
10k	92.39%	0.45%
12k	91.68%	0.52%
15k	91.28%	0.53%

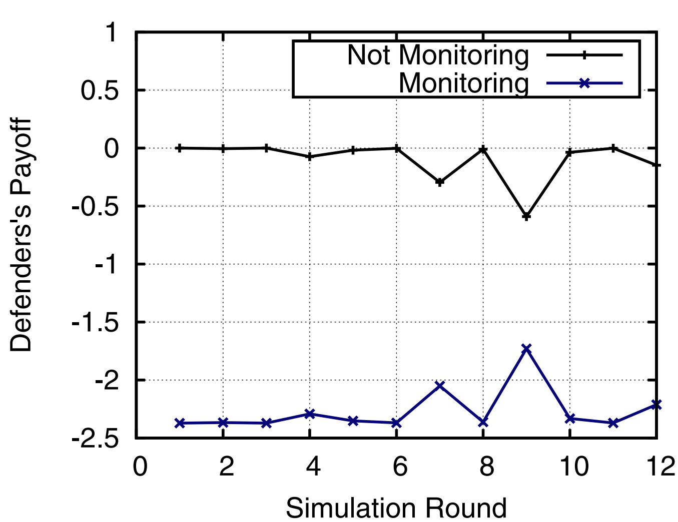

Fig. 7. Defender's Payoff when $p_o < p_{th}$ and Attacker is playing pure strategy *Attack*.

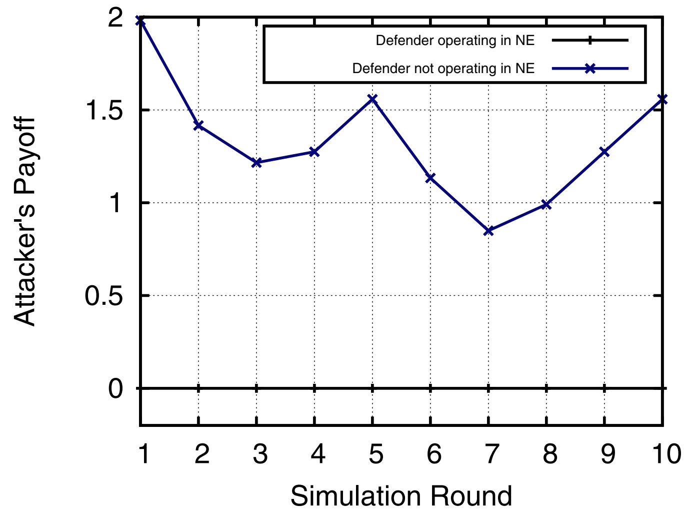

Fig. 8. Attacker's Payoff corresponding to different strategies of Defender.

relatively low FP rate. Table 7 shows the detection rate and FP rate of the *HIDS* on the test traces. The average detection rate and false alarm rate of the *HIDS* on these test traces are 91.78% and 0.5%, respectively.

Fig. 7 shows the defender's payoff playing its pure strategies *Monitor* and *Not Monitor* when the defender's maliciousness belief about opponent player is less than the malicious threshold (p_{th}), i.e., $p_o < p_{th}$. It can be observed from the figure that the defender is always better of playing its pure strategy *Not Monitor* when $p_o < p_{th}$.

The game under consideration is strictly non-cooperative. Therefore, each player tries to minimize the opponent's payoff. Fig. 8 shows the attacker's payoff corresponding to two different pure strategies of the defender. Similarly, Fig. 9 shows the defender's payoff corresponding to two different pure strategies of the attacker. It can be observed from these figures that the payoff of the opponent player increases when the player deviates from its BNE strategy. Fig. 10 shows the attacker's payoff under static and dynamic Bayesian games. It can be observed from the figure that in the static Bayesian game, the attacker gets a higher payoff.

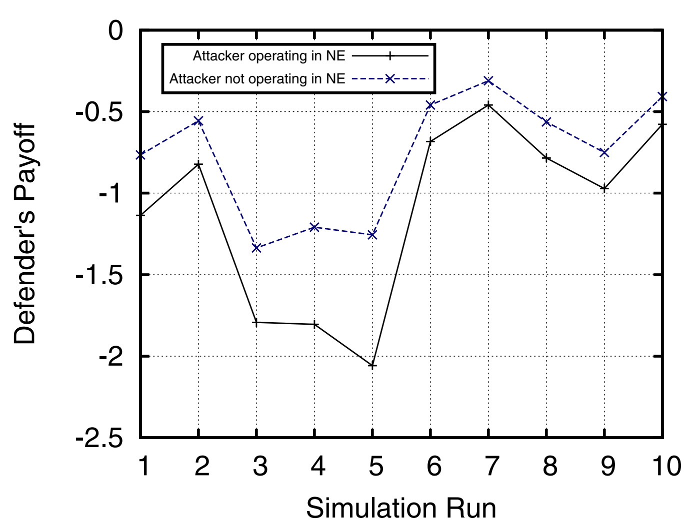

Fig. 9. Defender's Payoff corresponding to different strategies of Attacker.

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

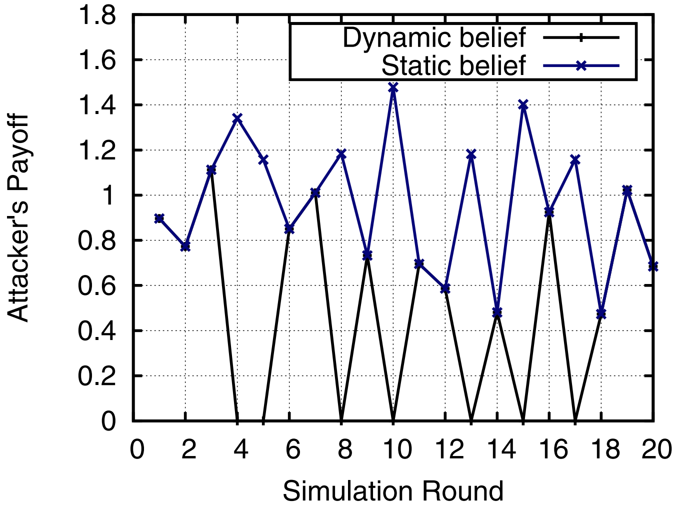

Fig. 10. Attacker's Payoff with static and dynamic maliciousness beliefs.

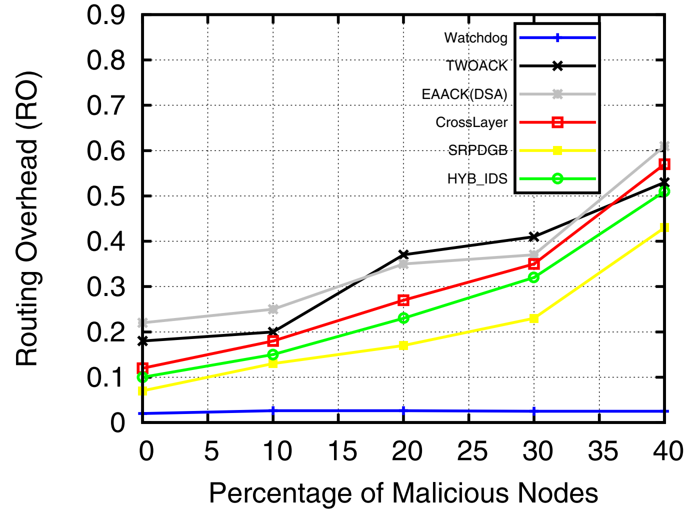

Fig. 12. Routing Overhead.

4.2.1. Comparison of proposed MANET IDS scheme with other methods

We have evaluated the performance of our proposed hybrid MANET IDS scheme with various other models like SRPDBG [47], CrossLayer [48], SPF [49], Watchdog [32], TWOACK [17] and EAACK [18]. These models were chosen for comparison since they represent a spectrum of MANET IDS schemes based on game theory (SRPDBG), data mining (CrossLayer), specification (SPF) and rules (Watchdog, TWOACK and EAACK). The following metrics were used for evaluation of the proposed hybrid MANET IDS scheme with other IDS schemes:

- *Packet delivery ratio (PDR)* refers to the ratio of the number of packets delivered to the destination node against the number of packets generated by the source node.
- *Routing overhead (RO)* refers to the overhead involved in transmission due to introduction of additional routing control packets like Route Request (RREQ), Route Reply (RREP), Route Error (RERR), ACK etc.

Figs. 11 and 12 show the *PDR* and *RO* of the various IDS schemes under varying percentage of malicious nodes. It can be observed from these figures that all the four schemes (TWOACK, EAACK, SRPDBG and proposed IDS) have higher *PDR* than the simple WatchDog scheme. The *PDR* of our proposed IDS scheme is comparable to that of EAACK and CrossLayer schemes, while it outperforms the TWOACK and SRPDBG schemes. On the other hand, the Watchdog scheme has the least *RO*, as it does not use any acknowledgment scheme to detect misbehaving nodes. The *RO* of the proposed IDS is less than the TWOACK, EAACK and CrossLayer schemes but higher than the SRPDBG scheme. The *RO* of the proposed IDS scheme is primarily due to exchanges of election messages for electing the MANET leader node and checker nodes.

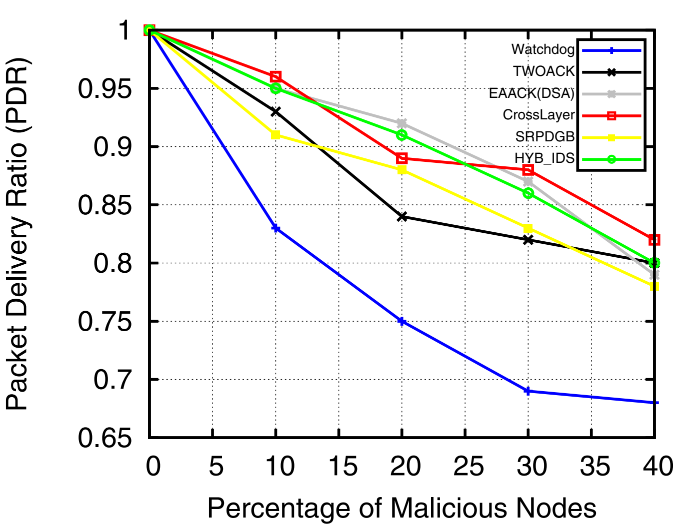

Fig. 11. Packet Delivery Ratio.

Table 8
Performance comparison of various IDS models.

IDS Models	Attack Type	Detection Rate	False Alarm rate
SPF	Route Compromise	47.56%	0.57%
	Traffic Distortion	43.24%	0.49%
	Black Hole	81.23%	0.51%
CrossLayer	Route Compromise	92.36%	0.38%
	Traffic Distortion	97.33%	0.93%
	Black Hole	99.7%	0.53%
SRPDGB	Route Compromise	65.43%	0.36%
	Traffic Distortion	51.56%	0.55%
	Black Hole	99.42%	0.37%
HYB_IDS	Route Compromise	91.4%	0.45%
	Traffic Distortion	95.3%	0.87%
	Black Hole	99.5%	0.35%

Table 8 shows the detection rate and false alarm rate of various IDS models on different classes of attacks. It can be observed from the table that our proposed HYB_IDS achieves high detection rate against all categories of attacks while producing a minimal amount of false alarms. The performance analysis comparison of various IDS models has been provided in Table 9.

From Tables 8 and 9, it can be summarized that the proposed hybrid scheme achieves high detection rate against different classes of attacks, while at the same time minimizes the overall false alarm rate and the computational overhead required for operating the IDS. However, the drawback of the proposed scheme is that it incurs a marginal overhead due to its cluster leader election process.

5. Conclusion and future work

In this paper, we proposed a new IDS scheme for MANETs which comprises a cluster leader node election mechanism and a hybrid

Table 9
Comparison of various MANET IDS models.

IDS Models	Proposed HYB_IDS	CrossLayer [48]	SRPDGB [47]	SPF [49]
Detection rate	High	High	Low	Low
False alarm	Low	Low	High	High
Detection method	Game theory based hybrid approach	Data mining anomaly based	Game theory and trust based	Specification based
Attack types addressed	Routing attacks, DoS attacks, Packet dropping, Packet spoofing	Routing attacks, Packet dropping, Packet spoofing	Routing attacks, Packet dropping	Routing attacks, Packet dropping, Packet spoofing
Advantage	High detection rate, Low false alarm rate, Low power consumption	High detection rate, Low false alarm rate	Low power consumption	Detect routing attacks with high accuracy
Disadvantage	Marginal overhead incurred in cluster leader node election	High power consumption, Overhead in training the IDS model	Low detection rate, High false alarm rate	Low detection rate, High false alarm rate, High power consumption

IDS. The main contributions of our proposed hybrid IDS scheme to the field of intrusion detection in MANETs is development of an IDS model that minimizes the power consumption and achieves a high detection rate across a wide range of attacks along with reduced false alarm rate. The proposed scheme minimizes the power consumption required for operating the IDS in MANETs through distribution of intrusion detection task among various nodes by employing a VCG mechanism based cluster leader election process. On the other hand, high detection rate and reduced false alarm rate are achieved by the hybrid IDS which comprises a threshold based lightweight module and a powerful anomaly based heavyweight module.

Our future work will be focused on improving the detection rate and decreasing the false positive rate of both the lightweight and the heavyweight modules of the hybrid MANET IDS. At present, the detection rate of the lightweight and the heavyweight modules are 91.78% and 81.33%, respectively. We also plan to investigate application of other equilibrium concepts like Pareto Equilibrium, Subgame Perfect Nash Equilibrium and Correlated Equilibrium in our future work. The refinement of the MANET leader election mechanism to address various issues like identification of selfish nodes in MANETs with greater accuracy, minimizing the computational overhead involved in execution of cluster leader node election mechanism, etc. are other possible potential research directions.

References

[1] A. Mishra, K. Nadkarni, A. Patcha, Intrusion detection in wireless Ad-hoc networks, IEEE Wirel. Commun. 11 (1) (2004) 48–60.
[2] Y. Zhang, W. Lee, Y.-A. Huang, Intrusion detection techniques for mobile wireless networks, Wirel. Netw. 9 (5) (2003) 545–556.
[3] M. La Polla, F. Martinelli, D. Sgandurra, A survey on security for mobile devices, IEEE Commun. Surv. Tutor. 15 (1) (2013) 446–471.
[4] F. Anjum, P. Mouchtaris, Intrusion Detection Systems, John Wiley & Sons, Inc., 2006.
[5] P. Brutch, C. Ko, Challenges in intrusion detection for wireless ad-hoc networks, in: Proceedings of Symposium on Applications and the Internet Workshops, 2003, pp. 368–373.
[6] Y.-C. Hu, A. Perrig, D. Johnson, Ariadne: a secure on-demand routing protocol for ad-hoc networks, Wirel. Netw. 11 (1–2) (2005) 21–38.
[7] S. Bu, F. Yu, X. Liu, P. Mason, H. Tang, Distributed combined authentication and intrusion detection with data fusion in high-security mobile ad hoc networks, IEEE Trans. Veh. Technol. 60 (3) (2011) 1025–1036.
[8] Z. Fadlullah, H. Nishiyama, N. Kato, M. Fouda, Intrusion detection system (IDS) for combating attacks against cognitive radio networks, IEEE Netw. 27 (3) (2013) 51–56.
[9] Y. Zhang, W. Lee, Intrusion detection in wireless ad-hoc networks, in: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, ACM, 2000, pp. 275–283.
[10] T. Anantvalee, J. Wu, A survey on intrusion detection in Mobile Ad Hoc Networks, in: Wireless Network Security, Signals and Communication Technology, Springer, 2007, pp. 159–180.
[11] A. Mitrokotsa, C. Dimitrakakis, Intrusion detection in MANET using classification algorithms: the effects of cost and model selection, Ad Hoc Netw. 11 (1) (2013) 226–237.
[12] C. Xenakis, C. Panos, I. Stavrakakis, A comparative evaluation of intrusion detection architectures for mobile ad hoc networks, Comput. Secur. 30 (1) (2011) 63–80.
[13] I. Butun, S. Morgera, R. Sankar, A survey of intrusion detection systems in wireless sensor networks, IEEE Commun. Surv. Tutor. 16 (1) (2014) 266–282.
[14] R. Mitchell, I. Chen, Effect of intrusion detection and response on reliability of cyber physical systems, IEEE Trans. Reliab. 62 (1) (2013) 199–210.
[15] A. Patel, M. Taghavi, K. Bakhtiyari, J.C. Jnior, An intrusion detection and prevention system in cloud computing: a systematic review, J. Netw. Comput. Appl. 36 (1) (2013) 25–41.
[16] M. Ficco, L. Romano, A generic intrusion detection and diagnoser system based on complex event processing, in: First International Conference on Data Compression, Communications and Processing, 2011, pp. 275–284.
[17] K. Liu, J. Deng, P.K. Varshney, K. Balakrishnan, An acknowledgment-based approach for the detection of routing misbehavior in MANETs, IEEE Trans. Mob. Comput. 6 (5) (2007) 536–550.
[18] E.M. Shakshuki, N. Kang, T.R. Sheltami, EAACK – a secure intrusion-detection system for MANETs, IEEE Trans. Ind. Electron. 60 (3) (2013) 1089–1098.
[19] Y. Liu, C. Comaniciu, H. Man, A Bayesian game approach for intrusion detection in wireless ad hoc networks, in: Proceedings of the 2006 Workshop on Game Theory for Communications and Networks, ACM, 2006.
[20] A. Agah, S. Das, K. Basu, M. Asadi, Intrusion detection in sensor networks: a non-cooperative game approach, in: Proceedings of Third IEEE International Symposium on Network Computing and Applications, 2004, pp. 343–346.
[21] T. Alpcan, T. Basar, A game theoretic approach to decision and analysis in network intrusion detection, in: Proceedings of 42nd IEEE Conference on Decision and Control, 2003, pp. 2595–2600.
[22] Y. Huang, W. Lee, A cooperative intrusion detection system for ad hoc networks, in: Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks, 2003, pp. 135–147.
[23] M. Kodialam, T. Lakshman, Detecting network intrusions via sampling: a game theoretic approach, in: Twenty-Second Annual Joint Conference of the IEEE Computer and Communications, vol. 3, 2003, pp. 1880–1889.
[24] A. Mas-Colell, M. Whinston, J. Green, Microeconomic Theory, New York, Oxford University Press, 1995.
[25] T. Issariyakul, E. Hossain, Introduction to Network Simulator NS2, 1st ed., Springer Publishing Company, Incorporated, 2008.
[26] P. Barford, J. Kline, D. Plonka, A. Ron, A signal analysis of network traffic anomalies, in: Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurement, 2002, pp. 71–82.
[27] A. Lakhina, M. Crovella, C. Diot, Mining anomalies using traffic feature distributions, Comput. Commun. Rev. 35 (4) (2005) 217–228.
[28] J. Dickerson, J. Dickerson, Fuzzy network profiling for intrusion detection, in: 19th International Conference of the North American Fuzzy Information Processing Society, 2000, pp. 301–306.
[29] A. Valdes, K. Skinner, Adaptive, model-based monitoring for cyber attack detection, in: Recent Advances in Intrusion Detection, vol. 1907, 2000, pp. 80–93.
[30] M. Roesch, Snort – lightweight intrusion detection for networks, in: Proceedings of the 13th USENIX Conference on System Administration, 1999, pp. 229–238.
[31] R. Sekar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang, et al., Specification-based anomaly detection: a new approach for detecting network intrusions, in: Proceedings of the 9th ACM Conference on Computer and Communications Security, 2002, pp. 265–274.
[32] S. Marti, T.J. Giuli, K. Lai, M. Baker, Mitigating routing misbehavior in mobile ad hoc networks, in: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, ACM, 2000, pp. 255–265.
[33] S. Misra, P. Krishna, K. Abraham, Energy efficient learning solution for intrusion detection in Wireless Sensor Networks, in: Second International Conference on Communication Systems and Networks, 2010, pp. 1–6.
[34] F. Haddadi, M. Sarram, Wireless intrusion detection system using a lightweight agent, in: Second International Conference on Computer and Network Technology, 2010, pp. 84–87.
[35] M. Mohanapriya, I. Krishnamurthi, Modified DSR protocol for detection and removal of selective black hole attack in MANET, Comput. Electr. Eng. 40 (2) (2014) 530–538.
[36] J. Cai, U. Pooch, Allocate fair payoff for cooperation in wireless ad hoc networks using Shapley Value, in: Proceedings of the 18th International Parallel and Distributed Processing Symposium, 2004, p. 219.

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001

[37] A. Urpi, M. Bonuccelli, S. Giodano, Modelling cooperation in mobile ad hoc networks: a formal description of selfishness, in: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2003, pp. 3–5.
[38] P. Liu, Incentive-based modeling and inference of attacker intent, objectives, and strategies, in: Proceeding of the 10th ACM Computer and Communications Security Conference, 2003, pp. 179–189.
[39] Y.-M. Chen, D. Wu, C.-K. Wu, A game theoretic framework for multi-agent deployment in intrusion detection systems, in: Security Informatics, vol. 9, Annals of Information Systems, Springer, 2010, pp. 117–133.
[40] S. Buchegger, J.-Y. Le Boudec, Performance analysis of the CONFIDANT protocol, in: Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing, ACM, 2002, pp. 226–236.
[41] P. Krishna, N.H. Vaidya, M. Chatterjee, D.K. Pradhan, A cluster-based approach for routing in dynamic networks, Comput. Commun. Rev. 27 (2) (1997) 49–64.
[42] H. Yih-Chun, A. Perrig, A survey of secure wireless ad hoc routing, IEEE Secur. Priv. 2 (3) (2004) 28–39.
[43] Y.C. Hu, A. Perrig, D.B. Johnson, Wormhole attacks in wireless networks, IEEE J. Sel. Areas Commun. 24 (2006) 370–380.
[44] N. Mohammed, H. Otrok, L. Wang, M. Debbabi, P. Bhattacharya, Mechanism design-based secure leader election model for intrusion detection in MANET, IEEE Trans. Dependable Secure Comput. 8 (1) (2011) 89–103.
[45] R. Agrawal, R. Srikant, Fast algorithms for mining association rules in large databases, in: Proceedings of the 20th International Conference on Very Large Data Bases, 1994, pp. 487–499.
[46] A. Savasere, E. Omiecinski, S.B. Navathe, An efficient algorithm for mining association rules in large databases, in: Proceedings of the 21th International Conference on Very Large Data Bases, 1995, pp. 432–444.
[47] M. Kaliappan, B. Paramasivan, Enhancing secure routing in Mobile Ad Hoc Networks using a Dynamic Bayesian Signalling Game model, Comput. Electr. Eng. 41 (2015) 301–313.
[48] R. Shrestha, K.-H. Han, D.-Y. Choi, S.J. Han, A novel cross layer intrusion detection system in MANET, in: 24th IEEE International Conference on Advanced Information Networking and Applications (AINA), 2010, pp. 647–654.
[49] C.-Y. Tseng, P. Balasubramanyam, C. Ko, R. Limprasittiporn, J. Rowe, K. Levitt, A specification-based intrusion detection system for AODV, in: Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks, 2003, pp. 125–134.

Please cite this article in press as: Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation, Engineering Science and Technology, an International Journal (2015), doi: 10.1016/j.jestch.2015.11.001