RESEARCH ARTICLE

Towards cross-layer approaches to coping with misbehavior in mobile ad hoc networks: an anatomy of reputation systems

Shuzhen Wang¹*, Zonghua Zhang² and Farid Naït-Abdesselam³

¹ School of Computer Science and Technology, Xidian University, Xi'an, China

² Institute Mines-Telecom/TELECOM Lille, CNRS UMR 5157 SAMOVAR, France

³ Paris Descartes University, Paris, France

ABSTRACT

In mobile ad hoc networks (MANETs), the nodes need to cooperate each other to establish multi-hop routes for out-of-range wireless communication. However, some of them may not always behave normally, either behaving selfishly for saving computational resource or maliciously for compromising communication protocols. Regardless of intents, such misbehavior would lead to the degradation of network performance. It is therefore important to design appropriate mechanisms to ensure that network performance could be maintained at an acceptable level in the presence of misbehaving nodes. But the open nature of MANETs makes such designs challenging. Reputation system has been widely recognized as an effective approach, which associates the behavior of nodes with its reputation, which is calculated by specifying and quantifying the observations of interest with respect to predefined performance metrics. More interestingly, the observations can be obtained and integrated from multiple layers, facilitating cross-layer analysis. This paper intends to take a deep look into several well-studied reputation systems and examine their operational characteristics in terms of modeling approaches and redemption techniques, with an objective to identify their capabilities in terms of misbehavior detection coverage and blind spots. Furthermore, such an anatomy allows us to better understand the failure curses of the deployment and operation of reputation systems in MANETs, so as to improve their performance by adopting effective countermeasures. Copyright © 2014 John Wiley & Sons, Ltd.

KEYWORDS

MANET security; misbehavior detection; reputation system

*Correspondence

Shuzhen Wang, School of Computer Science and Technology, Xidian University, Xi'an, China. E-mail: shuzhenwang@xidian.edu.cn

1. INTRODUCTION

A mobile ad hoc network (MANET) is a self-configuring wireless network that consists of a collection of mobile nodes, which serve as both router and host and cooperate each other to establish multi-hop routes for data packet delivery and out-of-range communication. As mobile nodes are autonomic entities, the network topology and connectivity may shift from time to time. A number of example networking scenarios are wireless sensor networks, vehicular ad hoc network, and even smartphone enabled applications. Regardless of the organization forms, the special characteristic of infrastructure and communication mode of MANET exposes the whole network to various abuses from accidental system faults to intentional malicious attacks, resulting in a lack of clear line of defense.

As the security design in wire-line networks, an in-depth defense line in MANET is constituted of intrusion prevention, detection, and response techniques. In particular, security prevention techniques rely on cryptography primitives, including authentication, authorization, and access control. For example, a large variety of secure routing protocols [1,2] have been proposed to prevent routes from being compromised by malicious attackers. Additionally, intrusion detection and response techniques are tightly integrated in order to achieve auto-defense functionality in MANET. For instance, a number of intrusion detection systems (IDS) [3,30] have been proposed to detect attacks by specifying routing protocol and analyzing packets feature. However, those systems usually require non-trivial computational overhead, impeding their applications in resource-constrained MANET.

In addition to IDS, a set of game theoretic models [4,5] have been used to enforce network nodes to cooperate with each other for achieving the maximum gain in terms of given metrics, such as network throughput and packet loss rate. A majority of those techniques, however, solely focus on selfish behavior of nodes for forwarding data packets. To make it more general, we introduce misbehavior to refer to the anomalies activities violating predefined network protocols and specifications, spanning from accidental system faults to incidental attacks, and ranging from physical layers to application layers. To describe the misbehavior resulting from various anomalies, reputation can be used to characterize the behavior of network nodes by quantifying the quality of services they provide, serving as a key metric to identify the deviation between regular behavior and misbehavior. In general, the objective of reputation system is to build and maintain reputation scores of each node in MANET, further encouraging them to trust and cooperate each other in normal manner and ultimately deter misbehavior. To achieve that, the behavior of network node needs to be characterized by effective observations, and the reputation of a node associated with its behavior must be quantified, stored, updated, and propagated in the network. In fact, reputation system has gained widespread application in open computing environments such as P2P network, social network, and e-commerce services.

This paper aims at providing an anatomy of reputation systems in MANETs and examine their potential to misbehavior diagnosis. In particular, this paper has three major contributions:

- We investigate the misbehavior in MANETs and classify them in terms of their consequence associated with high-level security properties. Such a classification allows us to gain insights into the effectiveness of reputation systems. For example, what kind of misbehavior can be possibly diagnosed, while what kind of misbehavior can never be diagnosed by any reputation system.
- We decompose a reputation system into several components in terms of functional role and discuss them independently, with an emphasis on the representation, quantification, and update of reputation value. We then put the components together as a unified framework by exploring their implicit relationships and specifically discuss the application of reputation system to misbehavior diagnosis in MANET based on a comparative study between a number of representative reputation systems.
- We address the design challenges of reputation system and identify the potential vulnerabilities attributing to their failure in misbehavior diagnosis. We also present a practical reputation-based misbehavior troubleshooting system by taking into account the countermeasures to those recognized vulnerabilities.

This paper starts with an analysis on misbehavior in MANETS in Section 2, along with some practical examples. Then the major functional components of reputation systems are discussed in Section 3, followed by the response issues analyzed in Section 4. Finally, we examine the root causes leading to the failure of reputation systems in Section 5, along with effective improvements.

2. MISBEHAVIOR IN MOBILE AD HOC NETWORKS

Typically, misbehavior of MANETs nodes can be caused by either accidental network faults or intentional attacks, and attacker intents can be either rational or malicious. For instance, a node may fail to relay data packets if its battery has been used out, and it probably drops the packets intentionally to save its limited power. More seriously, a node may manipulate the packets it forwards or violates routing protocols by some attack techniques. Because of the dynamic infrastructure, open medium, flexible topology, and unpredictable node mobility, as well as the signal noise, channel interference, and traffic congestion, it is extremely difficult to identify the root cause of misbehavior in MANETs. Misbehavior detection is therefore a significant issue attracting considerable research attention and effort, and a recent survey on trust management in MANET is given in [15].

It is commonly recognized that a comprehensive analysis on the root cause of misbehavior is a hard and meticulous process because of the intrinsic complexity of MANETs. An alternative way is to represent misbehavior in terms of consequence rather than specific techniques, covering a large class of low-level factors and significantly facilitating the analysis and design of models to characterize the behavior of nodes. In practice, although the network performance is also affected by the quality of communication links, the behavioral models usually integrate the factors associated with links into the relevant nodes instead of treating them independently.

In this paper, our effort is limited to the misbehavior associated with high-level security properties. While the vulnerabilities, which cause misbehavior, tend to be cross layer and more destructive, we treat them independently for an easier layer-specific analysis and understanding. Generally, the misbehavior can be classified into two cases: rational ones and malicious ones. The rational misbehavior usually have clear objectives, for example, greedy behavior for achieving more bandwidth and selfish behavior for saving computational cost; the malicious misbehavior generally intend to disrupt network performance or functionality, such as denial-of-service (DoS) attacks consuming network bandwidth, sybil, blackhole, and wormhole attacks that tamper routing protocols. As such, we could classify a large variety of misbehavior into several categories by simply specifying their consequence in terms of security properties, potentially providing a baseline for examining the reputation systems that we will discuss.

On the basis of an extensive investigation on the misbehavior issues in MANET, we enumerate the most typical ones in Table I, with an objective to gain an understanding on the types of misbehavior (in terms of incentives and consequences), which can be detected and prevented by reputation systems. We have to claim here that in some cases, the incentives of misbehavior are not clear, so insisting in classifying them in a finer-grained category rather than suggested security properties is neither practical nor meaningful. In addition to the three security attributes shown in the table, another important property is *confidentiality*, which is usually preserved by cryptographic techniques and not listed here.

Specifically, Table I summarizes the potential attacks that may occur in MANET by exploiting vulnerabilities at different system layers. While the enumerated misbehavior can be further examined in terms of particular techniques, we concern more about the cross-layer variants via the combination of vulnerabilities simultaneously occurring at two or more protocol layers. Thus, by understanding attacker intent and examining attack consequence, we may cluster similar attacks. For instance, DoS is a generic term referring to denial of services that may happen at any layer: an attacker could use jamming signals to disrupt ongoing transmissions on the wireless channel (physical and MAC layer) [12,16], and the attacker could also send a huge amount of SYN packets to a victim to cause it too busy to respond the other legitimate node's connection requests. Another scenario is that a number of collusive attackers collude each other to prevent legitimate users from accessing particular network services (application layer). More seriously, an attacker could destroy or steal a mobile device in ad hoc network (device theft and tampering).

3. REPUTATION MANAGEMENT

Reputation management has gained widespread applications in e-commerce such as eBay [17], as well as other computing environments such as P2P systems [18]. As the definition in human social networks, the opinion a node has of another is called reputation, and the reputation management in MANETs is essentially a feedback process involving the monitor and tracking of a node's behavior and associated feedback from its witnesses. A survey regarding reputation management was given in [19], which generally discusses the creation and update of reputation models by exemplifying a number of reputation systems. Another survey on reputation systems was conducted by K. Hoffman et al. [20], with the focus on attacks and defense mechanisms in reputation systems, whereas the specific application in misbehavior diagnosis in MANET is out of concern. To complement those works, our aim is to present a framework for specifically analyzing the reputation system in MANET, to provide a theoretical basis for understanding their design rationale and operations and ultimately to suggest more effective designs by countermining the potential system vulnerabilities. Essentially, a reputation system deployed in MANET must consider following design concerns:

- What observable events can be monitored and selected for characterizing the behavior of mobile nodes.
- How to build models for the regular behaviors of mobile nodes using the selected observable events.
- How to update the records of reputation and compute them timely for misbehavior detection and response.
- How to enhance the security and dependability of reputation system for more reliable and accurate misbehavior detection.

3.1. Design principles of reputation system

This section aims at addressing the fundamental elements in reputation system, which can be decomposed into a number of functional modules including monitor, storage, computation, and propagation of the related

Layers	Misbehavior	Incentive		Consequence (security attributes)			
		Rational	Malicious	Availability	Integrity	Non-repudiation	
MAC	Spoofing[6]		\checkmark	o		0	
	Greedy [7,8]	\checkmark		0			
	Blackhole [9]		\checkmark	0			
	Sybil[10]		\checkmark	0	0	0	
Network	Wormhole [11]		\checkmark	0	0		
(routing protocols)	DDoS[12]		\checkmark	0			
	Selfish[13]	\checkmark		0			
Transport	Hijacking [14]	\checkmark		0	0		
Application	Repudiation[14]	\checkmark		0		0	
	DDoS[14]		\checkmark	0			

Table I. Lave	r-specific mist	ehavior ar	nd their	consequences.
---------------	-----------------	------------	----------	---------------

Note: this table does not intend to cover all the misbehavior variants, we only select the most representative ones as examples. DDoS, distributed denial of service.

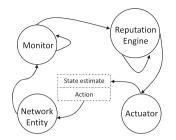


Figure 1. Reputation system and its components.

observations. Essentially, a reputation system is consisted of three components, which are shown in Figure 1, and their functions can be specified, respectively, as follows:

- Monitor: monitoring and gathering the behavioral traces from the interested events.
- Reputation manager: creating (and updating) reputation based on the observation of monitor.
- Actuator or responder: estimating node states with respect to reputation and taking appropriate actions, meanwhile sending feedback to reputation manager upon request.

Furthermore, the figure implies a working flow of reputation-based misbehavior detection system as follows:

- (1.) Collecting evidence by monitoring (first-hand) and recommendations (second-hand).
- (2.) Evaluating the evidence in terms of specified behavioral categories, for example, good and bad.
- (3.) Updating the new evaluation results with the previous reputation.
- (4.) Classifying the behavior of monitored node into trustable/non-trustable using a threshold.
- (5.) Setting response polices by using node reputation, for example, isolating non-trustable nodes out routing context.

In particular, observation in a reputation system generally refers to the observable events that can be monitored and gathered for representing the behavior of mobile nodes in MANETs. The events may occur at any protocol layer, and their activity is usually regulated by particular protocols, primarily routing protocols, and behaves as predefined rules. Because MANETs are dynamic, wireless-connected, and multi-hop networks, a node can only monitor the observations of its neighbors, which is called *first-hand information*. To collect those observations, each node must operate in a promiscuous mode and use enhanced passive acknowledgments to overhear the transmissions of its neighbors. Most of the existing reputation systems are developed from the basic assumption. A node can also gather the information of non-neighboring nodes, which are of interest through the intermediate nodes, and such information or evidence is viewed as second-hand information. The observations can

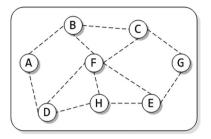


Figure 2. A network example.

be stored either locally or in a particular monitor module, depending on particular reputation systems. Moreover, in practical implementation, the reputation manager can be either centralized or distributed. In the case of fully distributed, the monitor and reputation manager may be coupled on a same node. It is also possible that the reputation manager is formed by a group of nodes that pool in the monitored evidence to compute the reputation.

In a reputation management system, a significant issue is to quantify the value of reputation of each node. To do that, a particular performance criterion regarding regular behaviors must be predetermined, which then serves as an evaluation metric for measuring the ongoing activities in terms of specific observable events. Assuming a MANET that works with some particular protocols at different layers, a node may rate its neighbors after each communication session (or transaction[†]). For easier understanding, let us consider Figure 2 as an illustrative example throughout this paper. Each time node A sends a packet to its neighbor B with destination G, it would rate B's behavior as good, if B did forward the packet; otherwise, A rates B as bad. Also, if the packet has successfully got to node G, node A then rate all the nodes, which have forwarded the packets as good. The sum of all the ratings that A assigned to B during a certain period Δt can be defined as *local trust* value LT(A, B) (or the first-hand information), while the ratings that A assigned to a non-neighboring node E can be regarded as global trust value GT(A, E) (or the secondhand information). The reputation is therefore a general concept, which essentially integrates both local and global trust values, represented as R = AG(LT, GT).

While the local trust value in terms of first-hand information is the basis for computing reputation, the aggregation of global trust value in terms of second-hand information can provide more evidence for the computation and accelerate misbehavior detection and response. However, it is necessary to adopt some measures to avoid node deception and ensure the quality of second-hand information. Especially, a group of nodes may collude each to raise a node's reputation by arbitrarily assigning its global trust values. We will give further discussion on this issue in the following section.

[†] We use transaction here to generally refer to the communication that may occur at any level.

3.2. Computation engine of reputation system

A core component of reputation system is reputation computation engine, which represents, models, quantifies, and updates the reputation value of each node. In [19], reputation engines are classified into six types: simple summation or average of ratings, Bayesian systems, discrete trust models, belief models, fuzzy models, and flow models. However, our discussion is based on two genera categories as in [21]: probabilistic estimation techniques and social network theory-based techniques. Formally, the first class of techniques extracts the feedback from the nodes of interest and uses Bayes theorem [22] as Equation (1) to create reputation models, where the probability of a random event is obtained given available observations. The second cluster techniques rely on the aggregation of the entire available weighted feedback in the network and can be expressed as shown in Equation (2), where W(i) is node i's witness set, R_i is the trustworthiness of node i, and w_j is the feedback of node j. Considering node F in Figure 2, its witness set should be W_F = $\{B, C, D, H, E\}$ for calculating local trust value:

$$Pr(E_i|O) = \frac{Pr(O|E_i)Pr(E_i)}{\sum_{i=1}^{n} Pr(O|E_i)Pr(E_i)}$$
(1)

$$R_i = \sum_{j \in W(i)} w_j \cdot \frac{R_j}{\sum_{k \in W(i)} R_k}$$
(2)

The most challenging issue for modeling and representing reputation is to incorporate the second-hand information that cannot be directly observed. The nodes maintain and publish their local observations and in parallel take information from others so as to gradually gain a global view of the whole network. The reputation aggregation basically involves two concerns, the trustworthiness of information provider and the information itself. In another word, the second-hand information source must be reliable, and the provided information must be authentic.

3.2.1. Probabilistic models.

The fundamental aspect of modeling reputation is to characterize a node's behavior based on the observation samples that have been monitored and to estimate a hidden probability of misbehavior with the updated observation. A predefined threshold is used to determine whether a node's behavior deviates from the historic record, or it behaves abnormally. For example, a Bayesian approach is used in [23,24] to model and represent reputation, where $\theta \in [0, 1]$ is an unknown probability denoting the occurrence of misbehavior. A pair of parameters (α, β) of Beta function is used to estimate θ as Equation (3). Starting with a prior distribution, for example, uniform distribution $f_0(\theta) \sim Beta(1, 1), f_{k-1}(\theta) \sim Beta(\alpha, \beta)$ is updated to $f_k(\theta) \sim Beta(u\alpha + s, u\beta + (1 - s))$ with new observations, where s = 1 is misbehavior, and the weight u is a discount factor for past observations. Especially, in [24], the opinion

(trustworthiness) of node *A* on *B*, or R_{AB} , is derived from a pair of parameter (t, c) (where *t* is the mean value and *c* is the standard deviation of of *Beta* (α, β)) by mapping them to a point on an elliptic curve centered at the point (1, 1).

$$f(\theta) = \frac{\theta^{\alpha-1} \cdot (1-\theta)^{\beta-1}}{\int_0^1 \theta^{\alpha-1} \cdot (1-\theta)^{\beta-1} d\theta}$$
(3)

The effect of evidence spreading in reputation systems has been discussed in [22,25], while Bayesian approach was taken as the analytical basis. In [26], the second-hand information is simply aggregated by using a significance factor after a compatibility test. For example, in Figure 2, if A wants to get reputation rating on F, it would incorporate the second-hand information LT(B, F) or LT(D, F)or both of them to its original reputation rating R_{AF} provided their deviation is less than a threshold, that is, R_{AF} := $R_{AF} + w_b \cdot LT(B, F) + w_d \cdot LT(D, F)$. The incompatible second-hand information will be discarded. In [24], a node A's opinion on non-neighboring node E is computed using a shortest path algorithm in order to achieve a route that has the maximum trustworthiness value by simply multiplying local trust values, for example, $R_{AE} = LT(A, B)$. $LT(B, F) \cdot LT(F, E)$. The objective is to simply find the most dependable path to delivery data packets. They employed a sliding averaging window to update observations and calculated reputation ratings using a simple linear weighted averaging scheme.

We have also seen some attempts to apply machine learning algorithms such as support vector machine to create reputation [27]. But such methods suffer same problem as Bayesian approach about reputation update. They even take longer time to create reliable and accurate reputation profile by training effective classifiers.

3.2.2. Social models.

The social models do not necessarily rely on probabilistic models, and the reputation score can be calculated and aggregated using the evidence collected by a node either directly or indirectly [28]. For instance, CORE [13] represents reputation values (ranges from -1 to 1) in a reputation table by integrating subjective, indirect, and functional information using a function.

Similar to the probabilistic models, the ratio between the number of packets requested for forwarding and that of have been actually forwarded is also used by social models to compute reputation rating among neighbors, while the approach to aggregate the indirect information is different. An effective algorithm named EigenTrust was proposed in [29] for computing global trust values for P2P systems using power iteration. The key concept of this algorithm is *transitive trust*, that is, $GT(i,j) = \sum_j LT(i,k) \cdot LT(k,j)$. A global trust vector \overrightarrow{GT} (GT_j) is used to quantify the trust that the whole system places on node *j*. When the number of nodes is very large, the trust vector \overrightarrow{GT}_i converges to the same vector for every node *i*. Because the size of nodes in MANET is not as large as that of P2P system, the convergence of the global vector cannot be guaranteed. The transitive trust value, however, can be used to propagate reputation ratings.

For instance, the *local evaluation record* proposed in [28] (which we refer as LT(i, j),) contains two entries: one is the ratio for evaluating forwarding behavior, represented as RF(i, j); another is the absolute number of packets that monitoring node requests monitored node to forward, noted as NF(i, j), which is used to quantify the confidence on reputation rating. Each node then periodically updates its local evaluation record of each of its neighbors and calculates an overall evaluation record as follows:

$$R_{ij} = \frac{\sum_{k \in W_i \cup \{N\}, k \neq j} \lambda_{ik} \cdot NF(k, j) \cdot RF(k, j)}{\sum_{l \in W_i \cup \{N\}, l \neq j} \lambda_{il} \cdot NF(l, j)}$$
(4)

In Equation (4), the witness set W_i limits to the node *i*'s neighbors, and λ_{ik} is node *k*'s credibility (or trustworthiness) from the view of node *i*. It is set as RF(i, k) in the original literature. As such, a node's neighbors can share the reputation information of other nodes. One limitation in Equation (4) is that *i*'s neighbors are not necessarily *j*'s neighbor, so NF(k, j) and NF(k, j) are not always available. Also, the reputation of node *j* from the perspective of node *i* is combined by both local trust values and global trust values, that is, $R_{ij} = \beta \cdot LT(i,j) + (1-\beta) \cdot GT(i,j)$. Note that LT(i, j) is updated periodically upon the record (a reputation table) of a new observation RF(i, j), and GT(i, j) is updated on the basis of the compatible second-hand information, which is given as $GT(i, j) = \sum_k R_{ik} \cdot R_{kj} / \sum_k R_{ik}$. Here, R_{ik} is essentially the referral k's trustworthiness from the perspective of *i*, and incompatible R_{kj} would cause the decrease of R_{ik} and discarding of R_{ki} .

4. REPUTATION SYSTEMS BASED MISBEHAVIOR DIAGNOSIS

While the goal of reputation systems in MANETs is to enforce the nodes to cooperate with each other to forward packets and enhance network throughput, they can be also applied to detect and countermine other types of misbehavior. The operational flows of a misbehavior detection system are illustrated in Figure 3, containing two ingredients: misbehavior detection and response.

4.1. Detecting misbehavior based on reputation

As all the reputation systems we have analyzed mainly operate with routing protocols, our discussion exemplifies network layer in this section. In essence, the reputation of nodes has two implications supporting the detection of misbehavior: firstly, each node is encouraged to behave regularly to maintain and increase its reputation; secondly, misbehavior would decrease a node's reputation and cause it to be punished. Thus, the principle of detecting misbe-

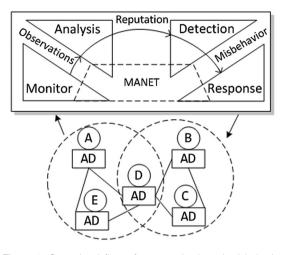


Figure 3. Operational flow of a reputation-based misbehavior detection system.

havior by reputation system is straightforward: updating reputation ratings periodically and comparing them with a predefined threshold, the nodes whose reputation falling lower than a tolerant range is regarded as misbehaving.

It is generally granted that the fundamental understanding of misbehavior would significantly facilitate their detection. Regardless of specific attack techniques, attacks in MANETs can be categorized according to their consequences: blackhole, routing loop, network partition, selfishness, sleep deprivation, and DoS. As for routing protocols, the specific misbehavior can be either direct packet dropping or routing table manipulation using modification, fabrication, injection, and rushing attacks. Referring to Table I, we see that the misbehavior associated with packet dropping would lead to the degradation of network performance with respect to availability. Alternatively, the performance degradation associated with the other security attributes is usually caused by routing manipulation. Thus, we believe the examination of misbehavior effects may cover a large class of specific techniques and thus facilitate our analysis and understanding on the detection capability of reputation systems.

As a matter of fact, there is no such an reputation system, which can detect all, even a majority of, the misbehavior introduced in Section 2. Actually, most of their detection coverage are limited to some specific misbehavior variants. We will take a number of system examples in Section to further illustrate this point. Additionally, another major counterpart for detecting misbehavior is anomaly-based IDS, which are usually specified to particular protocols with fine-grained analysis. In [30], a distributed anomaly detection architecture is proposed, which employs statistical approaches and data mining techniques for analyzing information from several network layers. The cross feature of routing packets are then analyzed and correlated for creating normal profiles. In addition, Tseng et al. [3] proposed a distributed intrusion detection model by capturing the message exchanges among mobile nodes

that use OLSR routing protocol, and specification-based techniques are used to characterize detection models.

4.2. Coping with misbehavior

As Figure 1 shows, a reputation system does not necessarily have an actual detection engine, which is mainly responsible for evaluating node's behavior and estimating its ongoing states. Its subsequent component, namely *actuator*, plays a significant role. Because MANET is a self-policing network, self-healing property is one of the most desirable properties. Thus, the response of an ideal reputation system generally contains two phases: (i) punishing the misbehaving node and (ii) recovering the node, which does not misbehave any longer.

As the reputation system is essentially an enforcement scheme (another category is credit-based schemes like Sprite [5]), the reaction effect tends to be *soft*. In another word, the misbehavior will be mitigated gradually rather than being countermined swiftly. A more direct reaction of a reputation system can be simply described as a decision process: when the actuator receives an ALARM message from the reputation manager, as shown in Figure 1, it would avoid to involve the misbehaving nodes in its routing table and thereby refusing to forward their packets. As a result, the misbehaving node would be eventually isolated from the network. In [28], the misbehaving node is published by its neighbors, which probabilistically (the probability is determined by both reputation value and link quality) drop the packets originated from the misbehaving nodes. A similar scheme was applied in [31], where a primary reputation rating ranging from 0 to 100 was used to represent the willingness of the node for forwarding the packets of its neighbors, and the misbehavior of a node would cause the decrease of the willingness of its neighbors (the forwarding probability is proportional to the willingness value).

The secondary response, or nodes redemption, is the next reaction to punishment. This means that the misbehaving node should be allowed to return the network if it does not misbehave any longer. We classify the existing techniques into following two categories:

- Timeout-based approach. A timeout parameter is introduced, and the reputation ratings of a node, which has been viewed as misbehaving, would be reset when the timeout expires. The timeout parameter is usually predefined and fixed over time.
- Observation-decay approach. In contrast to the first approach, this approach is more flexible. The negative reputation rating is assumed to be alleviated smoothly over time as some statistical models, such as exponential averaging windows, and the misbehaving would be allowed to rejoin the network once its recent reputation rating surpasses a tolerant threshold.

Obviously, these two approaches do not contradict each other, and it is possible to integrate them together. No matter which approach is used, an absolvent node would be kept in the record and punished more seriously and quickly than those nodes that do not have a track record for misbehaving. For example, the *direct interaction recovery* proposed in [31] is essentially a timeout-based approach, while the *timeout*, or the delay, is defined by two reputation ratings (a primary one and a secondary one). The misbehaving node would not be recovered until the secondary reputation rating grows larger than the primary one.

The typical examples of the second approach are CONFIDANT [26] and Hermes [24]. A so-called fading technique was used in CONFIDANT to impel nodes to discount all reputation ratings periodically and update observations by exponential decay. As such, the misbehaving nodes can alleviate their negative reputation ratings and return to the network automatically given the sufficient time (no direct observation is available for misbehaving node before its redemption). Although the approach adopted in Hermes was specifically designed for selecting trustworthy nodes, its property enables it to be extended for secondary response as well. In particular, a sliding window was used to expire old observation data, and only the observations that exist in the current time window were used for computing reputation ratings. However, as observations were only obtained between neighbors, and the second-hand information was excluded, its application to secondary response is impeded. Therefore, the recovery of a misbehaving node may take a long time.

In addition, there are some reputation systems that aim to secure the routing protocols by solely relying on the trustworthy nodes [28], the secondary response is therefore not a strict requirement.

4.3. Comparative studies

Here, we examine six well-studied reputation systems for a comparison, and the purpose is to provide a guide for more effective designs by studying their limitations and advantages.

Because most of the reputation systems take *Watchdog* and *path rather* as the design basis, we treat it as a simple reputation system. Also, we select the following criteria for this comparative study.

- Detection performance, which is the key performance metric of reputation systems and measured by detection coverage, blind spot, and false positive. But here, we mainly examine the misbehavior variants that can (or cannot) be detected by a particular system.
- Observation, which generally refers to the observable evidence/subjects collected from the network. Most of the existing reputation systems operate with reactive routing protocols such as Dynamic Source Routing, and we extend it to other similar routing protocols.
- Models, which are built for characterizing the reputation of network nodes in formal ways.
- Information, which is an abstract term referring to observations that are monitored, processed, and fed to

the reputation management system. Particularly, local information is captured among neighbors, and the global information is obtained from non-neighbors.

- Complexity, which measures the implementation cost in terms of the total number of messages exchanged among the nodes for maintaining reputation ratings, by excluding the message complexity associated with misbehavior response. Because the topology of MANETs is dynamic, we consider the worst case of this property for each reputation system.
- Scalability, which determines whether a system is scalable with the network size. In other words, whether a system is fully distributed, its complexity does not increase dramatically with the increasing number of network nodes.
- Extendability, which implies whether a system can be easily extended to detect a larger variety of misbehavior. It also implies that a system can work on multiple layers and has potential to detect cross-layer misbehavior.

In terms of detection coverage, Watchdog, OCEAN, and SORI are aware of "non-forwarding" behavior, but they are blind to malicious attacks. CORE has the potential to detect some particular DoS attacks. CONFIDANT has the largest detection coverage, which can detect out both selfish behavior and wormhole, blackhole, and distributed DoS attacks [32]. The objective of Hermes [24] is not to detect misbehavior but establish trust relationships using reputation among the nodes between source and destination so as to achieve secure and reliable data packet delivery.

A detailed comparison is summarized in Table II, along with the following discussions:

- All the reputation systems assume wireless interfaces to support *promiscuous* mode operation, which essentially impedes their applications in security-sensitive networks, which unlikely allow such mode. While acknowledgement-based (or feedback) scheme [31] is an alternative for collecting evidence and quantifying the reputation of a group of nodes that cooperate with each other for forwarding packets to the destination, it tends to trigger extra message complexity and response latency.
- Because Watchdog is adopted by all the reputation systems as the monitoring scheme, its complexity is

set to be O(1), even though in practice no actual messages are exchanged for querying and maintaining reputation ratings. There is no significant difference between the message complexity of probabilistic models and that of social models.

- For probabilistic models, the cost associated with exchanging messages is O(E), where *E* is the number of communication links. If we assume the number of network nodes is *N* and each link is bidirectional, we obtain E = N(N - 1)/2, so the complexity in terms of network node is $O(N^2)$.
- Because social models work in a flooding-like form, the overhead is almost same to that of probabilistic models. However, in practice, each node only maintains a reputation table with limited length (a small portion of network), so the complexity is reduced to $O(SizeofTable(T) \cdot N) = O(N)$. This enables them to be scalable with the network size. As for the computational overhead, SORI and Hermes cost more than the rest of the systems.
- None of them is essentially deception-resilient, which means the system is vulnerable to intentional manipulation of reputation ratings. We will further discuss this issue in Section 5.

5. FAILURE CURSES OF REPUTATION SYSTEMS AND REMEDIES

While the reputation system can serve as a basis for fulfilling self-healing functionality in MANETs, a number of crucial challenges must be tackled for achieving secure and dependable operations. Otherwise, the resulted vulnerabilities may lead to the inefficiency or failure of the reputation system. A number of design challenges and vulnerabilities are discussed in details in this section and outlined by Figure 4, which is extended from Figure 1.

The figure implies that reputation is calculated by quantifying the observations of interest in accordance with quality of service performance metrics. The detection engine monitors the nodes behavior characterized by the reputation and sends out warning messages once anomalous node is detected. A reaction module is attached to

			Information				
Systems	Observations	Models	Local	Global	Complexity	Scalability	Extendability
Watchdog[33]	DSR	SM	Y	Ν	<i>O</i> (1)	Y	Ν
CONFIDANT[26]	DSR	PM	Y	Y	$O(N^2)$	Y	Ν
SORI[28]	DSR	SM	Y	Y	$O(N^2)$	Ν	Y
CORE[13]	DSR	SM	Y	Y	$O(N^2)$	Y	Y
OCEAN[34]	DSR-MAC	SM	Y	Y	$O(N^2)$	Y	Y
Hermes[24]	Data packets	PM	Y	Y	$O(N^2)$	Y	Ν

Table II. A comparison of six representative reputation systems.

N is the number of network nodes; SM, social models; PM, probabilistic models; Y, Yes; N, No; DSR, dynamic source routing

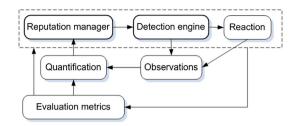


Figure 4. Functional modules of reputation system and its operational flowchart.

the detection engine, taking appropriate action to handle anomalous observations.

5.1. Observation collection and quantification

The observation collection process must be reliable (or trustworthy), and their quantification must be authentic. Intuitively, the more observation collected the more accurate of the calculated reputation. In particular, a wide range of observations from different network layers (or we call cross-layer observations), as well as their interconsequence, must be characterized sufficiently well, providing accurate reference for quantifying node reputation.

Given the observation, a set of models need to be particularly developed for data processing for obtaining the numerical values of reputation. While the information extracted from direct observations can be fed to the reputation models, most of reputation systems also use secondhand information (we defined as global trust value) to accelerate detection speed and enhance detection accuracy. For instance, in [22], the propagation effect of secondhand information in reputation systems was discussed, and another analysis on the negative effect of fake second-hand information was given in [25], where a critical probability model was given for identifying a liar's behavior. However, both of the two analyses were specified to Bayesian approach-based reputation system. In such system, a node (or a group of collusive nodes) can manage to subvert a reputation system by brainwashing, intoxication, identity spoofing [23], and other schemes. In general, we assume network nodes fall into two categories in terms of reputation, that is, liars and honest nodes. As discussed in [21], there are four types of deception misbehavior in P2P systems (as shown in Figure 5) systems, there are four types of deception behavior may occur (as illustrated by Figure 5):

- Individual liar. Node F always reports spurious information to the remaining nodes.
- Simple collusion. Nodes A and B report low reputation ratings for the other nodes and report high reputation ratings for each other.
- Collusive chain. The liars form a chain to recommend each other by escalating reputation ratings. One of the most serious case is the collusive loop, given by *A*, *B*, *F*, and *D* in Figure 5 (c).

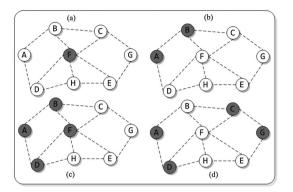


Figure 5. Deception behavior of propagating second-hand information.

 Collusive groups. Nodes A and D (C and G) always report honestly for the other nodes except for C and G (A and D), so they gain the high *credibility* for each other.

5.2. Vulnerabilities and system enhancements

A comprehensive analysis on the modeling of reputation system has been given in Section 3. This section discusses on some specific vulnerabilities they may encounter, primarily including reputation deception and identity spoofing. The existence of those vulnerabilities may dramatically affect the detection and reaction of reputation systems.

Generally, the design of reputation systems are based on a strong assumption, that is, the node's identity is unique and persistent, which may impede the applications of some anonymous secure routing protocols [35]. Also, the identity spoofing attack may lead to the impersonation of nodes and the manipulation of reputation ratings, eventually causing the reputation management to be intractable. More specifically, the majority of existing reputation systems in MANETs do not support identity management policies as in the other network forms, while simply assume the existence of some identity authorities relying on public key systems [5,36]. While authentication schemes pertaining to pseudonyms, digital signature, and key management is an alternative to tackle this issue, their effective combination and efficient implementation with reputation systems is a non-trivial issue.

It is obvious that the two major vulnerabilities may serve as failure curses of the reputation systems. To make the system more dependable, we may employ a fault-tolerant mechanism as a computing basis for reputation quantification. In doing so, the system may allow *Byzantine failure* of any witness of a particular network node, and the achieved *consensus* can be used to evaluate the trust values instead of by simple aggregation. In particular, for a node with *n* witness, the mechanism allows at most *f* of them to fail, where 3f + 1 = n.

In order to enhance the secure reputation computing in the presence of a group of malicious nodes (based on the scenarios shown in Figure 5), we may assume the existence of *a prior* trust values on some trustworthy network nodes during network deployment stage. This is a feasible assumption considering the application of tamper-proof hardware, which has very small probability of being compromised. The pre-trusted nodes are then severed as the supervisor to monitor the network and participate in the computing of other node's reputation. In other words, a node's reputation has to involve one of the pre-trusted node's opinions. By requiring each node to place some trust on trustworthy nodes, the potential malicious collectives can be broken. The security of reputation systems can also be hardened by integrating with lightweight cryptographic techniques [36], for example, hash chain [28] and certificateless signatures [37]. The initial version of CONFIDANT [26] employed a predetermined trust mechanism similar to the trust management adopted in pretty good privacy (PGP) for trust accumulation during routing and forwarding. The EigenTrust algorithm for reputation management in P2P networks [29] uses a distributed hash table to assign reputation managers. This is then located by hashing a unique Identifier (ID) of the peer, for example, IP address and transmission control protocol port, (TCP port) into a point in the distributed hash space. Such scheme can also be extended to MANET considering the similarities between MANET and P2P networks.

5.3. Performance evaluation

Performance evaluation of reputation system remains as a challenging issue and attracts much less attention than it deserves. On the one hand, as the reputation systems are usually based on theoretical models, their soundness can be verified. On the other hand, as reputation systems are specific to observations and scenarios, their performance can vary differently in different environments.

To date, we have observed that most of reputation systems were evaluated using simulation tools such as NS2 [38], OPNET [39], and QUALNET [40], by giving diverse assumptions, configurations, and parameters. So the results cannot be hardly analyzed and compared in terms of commonly granted evaluation metrics. Although simulations can help us to gain a certain amount of understanding on the designed systems, their performance applied in real networks is still unpredictable and cannot be guaranteed because of the dynamic characteristics of MANET [41].

An elemental yet most essential step for evaluating reputation system is to define a suit of performance metrics such as accuracy for long-term performance, impact of current behavior, robustness, and update smoothness. In Section 4.3, we defined a number of fine-grained criteria to enrich the evaluation metrics, because an ideal metric set is always desirable, which can measure each dimensionality of the behavior of reputation system. In addition, it is a compelling need to develop a benchmark dataset and a pool of application scenarios, which cover different misbehavior variants and their consequence in term of node/network behavior characterized by cross features. An easier approach to partially achieve that is to create a set of standard scenarios files and shell scripts as reference, which are publicly available in research community. More practically, a testbed, which encompasses a wide range of physical factors in MANET, for example, signal noise/interference, link error, and traffic congestion, is always desirable for testing and evaluating the prototypes of reputation systems. A preliminary version of testbed was proposed in [32], but the implementation details and subsequent work have not appeared yet.

As a matter of fact, the evaluation of reputation systems shares many similarities with that of IDS in traditional computer networks. We envision such an evaluation framework, which provides a formal way for theoretical analysis of reputation models and also a rich set of experimental/simulation data and synthetic data for building standard scenarios and hopefully creating practical factors in the real world. The ultimate goal is to yield a set of numerical scores/degrades for a particular reputation system by mapping its long-term results to each predefined evaluation metric. Such a framework may also significantly facilitate the implementation and evaluation of other designs in MANET, including MAC mechanisms, routing algorithms, and applications.

5.4. A practical reputation system for misbehavior diagnosis

As the objective of reputation systems is to enforce the nodes to cooperate each other, the observations are mainly extracted from the mobile nodes. In practical environment, however, the packet collisions and signal interference may cause cooperative nodes to appear as selfish ones. Thus, a reputation system is expected to take into account the properties of communication links as well [42].

In addition, as previously discussed, the observations extracted from protocol layers other than routing layer should be incorporated into the reputation systems for cross-layer misbehavior detection. To do that, we have designed a more practical reputation-based anomaly detection system called RADAR for wireless mesh networks [43], and it can be readily extended to MANET scenarios and other autonomic computing networks [44]. The design principle is illustrated in Figure 6, and its salient features are highlighted as follows,

- Each network node may serve as a host of an anomaly detector, playing double functional roles, that is, the agent and manager, collecting evidence and query trust values, respectively.
- Observation extracted from different network layers is used to calculate and quantify global view and local view for defining trust values and reputation.

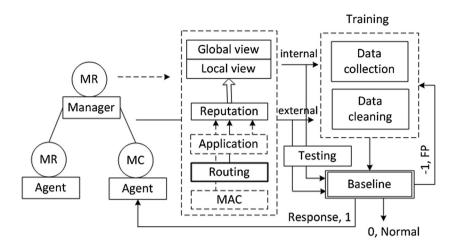


Figure 6. A practical reputation-based misbehavior troubleshooting system in mobile ad hoc network.

 The reputation management system considers both internal properties of network nodes and external factors regarding link failure, signal interference, traffic congestion, and so on.

This system takes reputation as a key metric for evaluating the behavior of network nodes and employs anomaly detection algorithms to spot anomalous nodes whose reputation is lower than a predefined threshold. The reputation may incorporate observations drawn from different protocol layers, and the system architecture allows the application of cryptographic primitives, as well as security and dependability mechanisms. Specifically, as Figure 6 shows, it involves both internal (nodes) and external (communication links) factors, and a data preprocessing module (data collection and data cleaning) is used before creating normal profiles. As such, the vulnerabilities presented in Section 5.2 can be partially fixed, and a friendly interface is enabled to support inter-operability with other counterpart systems.

6. CONCLUDING REMARKS

This paper provides an anatomy of the reputation systems in MANETs. Specifically, starting from the understanding and analysis of misbehavior in MANET, we examined the capabilities of reputation systems by identifying misbehavior consequence. We observed that most of reputation systems can only detect a subset of misbehavior, and each reputation system has its own detection coverage and blind spot. Further, we addressed the key components of a generic reputation system model, with emphasis on reputation manager, which encompasses observation monitor, reputation modeling, representation, and update. Moreover, a comparative study between several existing reputation systems was conducted. We extensively discussed the applications of reputation systems for misbehavior detection in MANET, especially the detection schemes and response for mitigating misbehavior, as well as the node redemption issue. Finally, we identified the challenging issues that may undermine the effectiveness of current reputation systems and proposed the corresponding enhancements for achieving dependent and secure reputation computation.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China under grant nos.61100156 and 61070143.

REFERENCES

- Hu Y-C, Perrig A. Ariadne: a secure on-demand routing protocol for ad hoc networks. *Wireless Networks* 2005; 11: 21–38.
- Sanzgiri K, Dahill B, Levine BN, Shields C, Belding-Royer EM. A secure routing protocol for ad hoc networks, *Proceedings of 10th IEEE International Conference on Network Procotols*, Paris, France, 2002; 78–89.
- Tseng CH, Wang S-H, Ko C, Levitt K. DEMEM: distributed evidence-driven message exchange intrusion detection model for MANET, *Proceedings of* 9th International Symposium on Recent Advances in Intrusion Detection (RAID2006), Hamburg, Germany, 2006; 249–271.
- Srivastava V, Neel J, MacKenzie AB, Menon R, DaSilva LA, Hicks JE, Reed JH, Gilles RP. Using game theory to analyze wireless ad hoc networks. *IEEE Communications Surveys and Tutorials* 2005; 7(1–4): 46–56.

- Zhong S, Chen J, Yang R. Sprite: A simple, cheat-proff, credit-based system for mobile adhoc networks, *Proceedings of IEEE INFOCOM*, San Francisco, USA, 2003; 3: 1987–1997.
- Bellardo J, Savage S. Savage 802.11 denial-of-service attacks: real vulnerabilities and practical solutions, *Proceedings of USENIX Security*, Washington, DC, 2003; 15–28.
- Raya M, Aad I, Hubaux J-P, Fawal AE. DOMINO: detecting MAC layer greedy behavior in IEEE 802.11 hotspots. *IEEE Transactions on Mobile Computing* 2006; 5(12): 1691–1705.
- Radosavac S, Moustakides G, Baras JS, Koutsopoulos I. An analytic framework for modeling and detecting access layer misbehavior in wireless networks. *ACM Transactions on Information and System Security* (*TISSEC*) 2008; **11**(4): 19:1–19:28.
- Deng H. Routing security in wireless ad hoc networks. *IEEE Communications Magazine* 2002; 40(10): 70–75.
- Piro C, Shields C, Levine BN. Detecting the sybil attack in ad hoc networks, *Proceedings of IEEE/ACM International Conference on Security and Privacy in Communication Networks (SecureComm)*, Baltimore, MD, USA, August 2006; 1–11.
- Hu Y, Perrig A, Johson DB. Packet leashes: a defense against wormhole attacks in wireless networks, *Proceedings of IEEE INFOCOM*, New York, USA, 2002; 12–23.
- Aad I, Hubaux J-P, Knightly EW. Impact of denial of service attacks on ad hoc networks. *EEE/ACM Transactions on Networking (TON)* 2008; 16(4): 791–802.
- Michiardi P, Molva R. CORE: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks, *Proceedings of the 6th IFIP Conference on Security Communications, and Multimedia*, Melbourne, Australia, 2002; 107–121.
- Yang H, Luo H, Ye F, Lu S, Zhang L. Security in mobile ad hoc networks: challenges and solutions. *IEEE Wireless Communication* 2004; 11(1): 38–47.
- Cho J-H. A survey on trust management for mobile ad hoc networks. *IEEE Communications Surveys & Tutorials* 2008; **13**(4): 72–75.
- Liu Z, Liu H, Xu W, Chen Y. Extracting jamming signals to locate radio interferers and jammers, *Proceedings of MobiHoc*, South Carolina, USA, 2012; 257–258.
- 17. http://www.ebay.com.
- Marti S, Molina HG. Taxonomy of trust: categorizing P2P reputation systems. *Computer Networks* 2006; 50: 472–484.

- Josang A, Ismail R, Boyd C. A survey of trust and reputation systems for online service provision. *Decision Support Systems* 2007; 43(2): 618–644.
- Hoffman K, Zage D, Nita-Rotaru C. A survey of attack and defense techniques for reputation systems. ACM Computing Surveys 2009; 42(1): 1–31.
- Despotovic Z, Aberer K. P2P reputation management: probabilistic estimation vs. social networks. *Computer Networks* 2006; 50: 485–500.
- 22. Buchegger S, Le Boudec J-Y. The effect of rumor spreading in reputation systems for mobile ad-hoc networks, *Proceedings of WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, France, March 2003.
- Buchegger S, Le Boudec J-Y. Self-policing mobile ad hoc networks by reputation systems. *IEEE Communications Magazine* 2005; 43(7): 101–107.
- Zouridaki C, Mark BL, Hejmo M, Thomas RK. Hermes: a qauantitative trust establishment framework for reliable data packet delivery in MANETs. *Journal of Computer Security* 2007; 15(1): 3–38.
- Mundinger J, Le Boudec J-Y. Analysis of a reputation system for mobile ad-hoc networks with liars, *Proceedings of The 3rd International Sympo*sium on Modling and Optimization, Trento, Italy, April 2005.
- Buchegger S, Le Boudec J-Y. Performance analysis of the CONFIDANT protocol, *Proceedings of ACM MobiHoc*, Lausanne, Switzerland, 2002; 226–236.
- 27. Akbania R, Korkmazb T, Raju GV. EMLTrust: an enhanced machine learning based reputation system for MANETs. *Ad Hoc Networks* 2012; **10**(3): 435–457.
- He Q, Wu D, Khosla P. SORI: A secure and objective reputation-based incentive scheme for ad hoc networks, *Proc. of Wireless Communications and Networking Conference*, Atlanta, American, 2004; 825–830.
- 29. Kamvar SD, Schlosser MT, -Molina HG. The Eigen-Trust algorithm for reputation management in P2P networks, *Proceedings of the 12th International Conference on World Wide Web*, Banff Alberta, Canada, 2003; 640–651.
- Zhang Y, Lee W, Huang Y. Intrusion detection techniques for mobile wireless networks. ACM Wireless Networks Journal 2003; 9(5): 545–556.
- Vilela JP, Barros J. A feedback reputation mechanism to secure the optimized link state routing protocol, *Proceedings of 3rd IEEE International Conference* on Security and Privacy in Communication Networks (SecureComm 2007), Nice, France, September 2007; 294–303.

- 32. Buchegger S, Tissieres C, Le Boudec J-Y. A testbed for misbehavior detection in mobile ad-hoc networks—how much can Watchdogs really do? *Proceedings of IEEE WMCSA 2004*, English Lake District, UK, December 2004; 102–111.
- 33. Marti S, Giuli TJ, Lai K, Baker M. Mitigating Routing Misbehavior in Mobile Ad Hoc Networks, Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, ACM Press, New York, USA, 2000; 255–265.
- Bansal S, Baker M. Observation-based cooperation enforcement in ad hoc networks. *Technical Report*, Stanford University, 2003.
- Seys S, Preneel B. ARM: anonymous routing protocol for mobile ad hoc networks. *International Journal of Wireless and Mobile Computing* 2010; 3(3): 145–155.
- Zhou L, Haas Z. Securing ad hoc networks. *IEEE Network* 1999; **13**(6): 24–30.
- 37. Zhang Z, Liu J, Kadobayashi Y. STARS: a simple and efficient scheme for providing transparent traceability and anonymity to reputation systems,

Proceedings of DPM/SETOP, Athens, Greece, 2010; 170–187.

- 38. Nework Simulator. http://www.isi.edu/nsnam/ns/.
- 39. http://www.opnet.com/.
- 40. The QualNet simulator from Scalable Networks Inc. http://www.scalable-networks.com/.
- Konorski J. Effective data-centric reputation systems for MANETs: a novel evaluation framework, *Proceed*ings of IEEE ICC, Kyoto, Japan, 2011; 1–6.
- Jaramillo JJ, Srikant R. DARWIN: distributed and adaptive reputation mechanism for wireless ad-hoc networks, *Proceedings of ACM MobiCom*'07, Quebec, Canada, 2007; 87–98.
- Zhang Z, Ho P-H, Nait-Abdesselam F. RADAR: a reputation-driven anomaly detection system for wireless mesh networks. *ACM Wireless Networks* 2010; 16(8): 2221–2236.
- 44. Zhang Z, Kadobayashi Y, Nait-Abdesselam F. Towards an evaluation framework for reputation systems in autonomic computing networks, *Proc. of ChinaCom'09*, Xi'an, China, 2009; 1–5.