SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks (2015)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/sec.1320

RESEARCH ARTICLE

On the evaluation of reputation and trust-based schemes in mobile ad hoc networks

Sohail Abbas[1]*, Madjid Merabti[2] and David Llewellyn-Jones[2]

[1] Department of Computer Science & IT, University of Malakand, Chakdara, KPK, Pakistan
[2] School of Computing and Mathematical Sciences, Liverpool John Moores University, Byrom St., Liverpool L3 3AF, U.K.

ABSTRACT

In multihop networks such as mobile ad hoc networks selfish or misbehaving nodes can degrade network performance. Reputation and trust-based schemes have been proposed in order to enforce cooperation and to discourage misbehaving nodes. These schemes detect and isolate selfish nodes and maintain network throughput by enabling nodes to construct paths that only include good nodes. In order to punish selfish nodes and restrict their services, selfish nodes are not provided with packet forwarding services anymore by the good nodes in the network. The aim of these schemes is to increase good nodes' benefit (throughput and utility) and to decrease malicious nodes' benefit. In this paper, we demonstrate that the commonly adopted simulation parameters used by the research community for the evaluation of reputation or trust-based schemes in mobile environments usually produce biased results. We further demonstrate this fact by adopting a widely used reputation-based scheme as our case study using the commonly used metrics, such as throughput, utility, delay, and overhead. The simulation parameters are analyzed analytically and via simulations in order to demonstrate their main reason of producing inaccurate and biased results. Finally, we compare simulation and analytical results to corroborate our claim and discuss how to mitigate the resulted inaccuracy and biasness. Copyright © 2015 John Wiley & Sons, Ltd.

KEYWORDS
selfish node; misbehavior detection; reputation; trust; mobile ad hoc networks

***Correspondence**
Sohail Abbas, Department of Computer Science & IT, University of Malakand, Chakdara, KPK, Pakistan.
E-mail: sabbas@uom.edu.pk

1. INTRODUCTION

Mobile ad hoc networks (MANETs) are constructed with the help of a large number of wireless nodes, such as laptops, tablets, personal digital assistants, and sensors, usually having limited transmission, battery, and computation powers. The key advantage of these networks is their capability to work without relying on any centralized architecture or control. These flexibilities make MANETs tempting for use in situations where there is no pre-deployed infrastructure or where it is costly to deploy one, such as in emergency situations, disaster relief scenarios, search and rescue operations, vehicular networks, casual meetings, campus networks, robot networks, and so on [1,2].

In ad hoc networks, each node provides packet forwarding services to other nodes in the network, because there are no dedicated routers present in the network for packet forwarding. However, due to a variety of reasons nodes may not always return such altruistic behavior. One reason for this might be that nodes do not belong to a single authority. This is quite a common scenario within civilian applications where users/nodes perform in the network for global connectivity according to their own interests and objectives. Another reason might be that nodes belong to a single authority (such as the military) and have common goals; however, they nonetheless operate in a physically insecure environment, where they are vulnerable to being captured or compromised. In any case, nodes show reluctance to spend their resources, such as battery and so on, for forwarding neighbors' packets; as a result, they may discard such packets. Such behavior of nodes is called selfish behavior[†] that can ultimately lead to network partitioning and performance degradation [3–6].

In the last decade, substantial research has been undertaken to promote cooperation and discourage selfish nodes in MANETs. Cooperation promotion schemes mainly include reputation, trust, and credit-based models. In reputation and trust-based models, usually nodes are

[†] Selfishness or packet dropping behavior is considered as passive attack, but this phenomenon can also be used as active attack, called blackhole attack.

characterized as cooperative and trustworthy based on their past behavior. As a result, nodes keeping high reputation or trust ratings are provided with services, whereas nodes keeping bad reputation or trust are separated from the network. On the other hand, in credit-based models, nodes are usually charged for services (usually packet forwarding service). Payments' transactions are carried out in a form of virtual currency. Nodes act as buyers or/and sellers of the services. Nodes can earn credit by forwarding each others' packets. The earned credit is then used to acquire packet forwarding service in the network. Due to the distributed nature of reputation and trust-based schemes, they are generally preferred as compared with the credit-based schemes that need credit protection mechanisms plus centralized virtual bank. Therefore, we will take reputation-based scheme as our case study in this paper; for the detailed discussion on these schemes, please consult [7–9].

The main objectives of reputation-based schemes are to detect selfish nodes, maintain network throughput by enabling nodes to construct paths that only include good nodes, and discourage selfish nodes by restricting their services, that is, selfish nodes are not provided with the packet forwarding services anymore by the good nodes acting as *intermediate* nodes. However, selfish nodes can still gain benefit (in the form of throughput and utility) from the network by interacting *directly* with the destination nodes, whether intentionally or unintentionally.[‡] It is due to the fact that there are no intermediate nodes along the path to restrict selfish nodes' services.

Direct interactions (DIs) are natural and may not be fully prevented in the presence of mobility; however, during the simulation lifetime, if sources and destinations interact directly for considerable amount of time, it will unfairly improve the overall throughput and utility of both the good nodes as well as the selfish nodes. This phenomenon can cause biased results when reputation or trust-based schemes are evaluated under certain metrics, such as throughput, utility, network delay, and overhead. We will demonstrate that the simulation parameters, especially simulation area and radio range, used in the research community for reputation and trust-based schemes' evaluation produce considerable amount of DIs that causes ambiguity and biasness in the results obtained. We will also discuss how to mitigate the confusion caused by DIs during evaluation. We will investigate and demonstrate this using cooperation of nodes and fairness in dynamic ad hoc networks (CONFIDANT) [4,10] as a benchmark scheme for our case study. A very basic version of this paper has been published in [11].

The paper is organized as follows. Section 2 covers relevant background material and the proposed related research work. In Section 3, the benchmark scheme is evaluated using a simulator. Furthermore, the problems caused by DIs will be discussed. In Section 4, we analyze the simulation parameters in order to demonstrate their main reason in causing ambiguities. Section 5 presents how the biasness in the results will be mitigated. The paper is concluded in Section 6.

2. BACKGROUND AND RELATED WORK

Reputation or trust schemes exploit the past interaction experiences to classify nodes as cooperative and uncooperative. The DI experiences are called firsthand ratings, whereas indirect (collected from neighbors) is called secondhand ratings. Nodes share experiences with neighbors in the form of firsthand ratings for better reputation or trust computations. Sharing ratings or interaction histories plays a vital role in maintaining cooperation in the network in many ways [12]. First, history indicates ability of an entity. Second, because the fear that the present actions will become history in the future will motivate each entity to act as to the best of its abilities. Finally, history can make help to distinguish between nodes having higher abilities from those having lower abilities.

In MANET, reputation has basically evolved locally through monitoring cooperation or packet forwarding activities using passive acknowledgements (which can be accomplished using promiscuous listening mode of network interface cards). This monitoring is called Watchdog mechanism [3]. Watchdog is a monitoring process working on each entity, responsible for maintaining and recording of reputation ratings into a reputation table. Each time when a source node S sends a packet to the destination node D via the intermediate nodes along the path, S holds a copy of the packet in memory until it overhears (i.e., in promiscuous mode) the same packet forwarded by the intermediate node before time T expires. Node S augments the reputation of the next node after it is confirmed that it has forwarded its packet and decreases its reputation otherwise after a time-out period. If a packet dropper exceeds a prescribed threshold, the node is considered to be bad reputed.

Cooperation promotion schemes such as CONFIDENT [4], CORE [5], and others [3,13–15] use Watchdog mechanism for local reputation built up. In order to collaboratively detect and isolate selfish or misbehaving nodes, these schemes disseminate firsthand ratings in the network. However, to avoid the problem of badmouthing and to reduce communication overhead, some authors proposed schemes, such as locally aware reputation system [16] and OCEAN [17], that exploit only firsthand ratings for selfish nodes' detection. Passive acknowledgements, being a lightweight solution, suffer from ambiguous and receiver collision problems [3,18]. To overcome these problems, some authors propose *explicit* acknowledgement-based monitoring [19] but at the cost of increased

[‡] Direct interactions can also be used by selfish nodes to increase their benefits after being detected by strategically selecting their locations in order to interact directly with their sources or destinations. In this paper, we will not consider this scenario.

communication overhead. To reduce this overhead, Zhao and Delgado-Frias [20] proposed a scheme based on multipath routing with an end-to-end feedback mechanism. However, the technique only detects and isolates misbehaving *paths* that reduce good nodes utilization in the network.

Li and Wu [13] demonstrate that effect of mobility on the scope of interactions and secondhand reputation/trust ratings' dissemination. The authors come up with the conclusion that increased mobility increases the overall trust convergence. Pirzada *et al.* [21] show that throughput in the network is increased with the increase in mobility even in the presence of selfish nodes. This increase in the throughput is due to the increased interactions among nodes. However, they do not explore this fact that the increase in throughput is actually caused by the DI of nodes. The work, they propose, differs from ours in that they show the effect of mobility on throughput; however, they do not explore the fact that this increase in throughput is actually caused by DIs among nodes.

In Table I, various reputation and trust-based schemes have been summarized and compared. We have identified various metrics that will be affected by the DIs and the factors (area and radio range) that usually produce these DIs. In the subsequent sections, we will discuss these issues in detail.

3. CASE STUDY

Buchegger and Boudec [4] proposed a cooperation promotion scheme for MANETs, which they named as CONFIDANT. The authors enhanced their scheme in [10] by using a Bayesian model in order to address a problem, called rumor spreading. Due to space limitations, these schemes cannot be explained here; interested readers are referred to [4] and [10].

3.1. Simulation setup

We simulate the CONFIDANT scheme in the network simulator NS-2.30 and evaluate it using the simulation parameters listed in Table II. These parameters are commonly used in the research community, especially simulation area and radio range. In this simulation study, our aim is to find out how often a source node interacts directly with a destination node and to what extent itaffects the throughput, utility, delay, and routing overhead. We categorize the throughput and utility of the network into direct and indirect ones in order to improve the clarity of the results obtained from the experiments. The results are obtained from the average of 20 random seeds.

3.2. Metrics

The following metrics have been used for the CONFIDANT scheme evaluation.

(1) *Throughput*. The ratio between the total number of data packets successfully received by destination nodes to the total number of data packets sent by source nodes at the application layer.
(2) *Good throughput*. The throughput available to good nodes.
(3) *Evil throughput*. The throughput available to selfish nodes.
(4) *Direct good throughput*. The throughput achieved by good nodes through DIs, with no intermediate nodes involved in the path.
(5) *Direct evil throughput*. The throughput gained by selfish nodes through DIs.
(6) *Real or indirect throughput*. The throughput achieved without DIs.
(7) *Utility*. Buchegger *et al.* [4] define the benefit or utility of a node as follows.

$$u_i = b_r \sum P_{received} + b_s \sum P_{sent} - c_f \sum P_{forwarded} \quad (1)$$

where u_i is the utility of a node i, b_r is the benefit when i receives a packet as the destination node, b_s is the benefit acquired by i when its packet successfully reaches its destination, and c_f is the cost (in terms of memory, bandwidth, and central processing unit (CPU) usage) incurred by node i of forwarding a packet for others. The exact values of benefit and cost variables (b_r, b_s, and c_f) depend on energy models, CPU, and memory usage per packet, but the value of each one is in the range [0, 1]. We used the value of b_r, b_s, and c_f to be 1 in order to measure the utility in packets only. A negative utility value means that a node incurs an overall loss.

(8) *Evil utility*. The utility gained by evil nodes.
(9) *Good utility*. The utility achieved by good nodes.
(10) *Direct evil utility*. The sum of $P_{received}$ and P_{sent} where the source, destination, or both are selfish and happen to be interacting directly. Please note that in our sce nario, selfish nodes do not forward any data packets; hence, $P_{forwarded}$ does not apply here.
(11) *Direct good utility*. The sum of $P_{received}$ and P_{sent} where the source, destination, or both are good nodes that happen to be interacting directly. Please note that our aim here is to show that a considerable amount of sent and received packets is gained from DIs; hence, we do not plot $P_{forwarded}$.
(12) *Delay*. The total time (in seconds) taken by data packets to reach their destinations divided by the total number of packets.
(13) *Routing overhead*. The total number of sent as well as forwarded events on the network layer.

3.3. Analysis

In order to first demonstrate that the DIs really exist in the simulation results, we show the effect of selfish nodes on

Table I. Reputation and trust-based schemes with their features and affected metrics by the direct interactions.

Schemes	Features	Metrics used	Area	Radio range	Affected metrics
Watchdog and Pathrater [3]	The scheme is proposed to mitigate the effect of selfish nodes by ranking paths based on selfish nodes in them; hence, selfish nodes are bypassed.	Throughput, overhead, and false positive	670 m × 670 m	Not given	Throughput and overhead
CONFIDANT [4]	Selfish nodes are isolated, and trusted recommendations are taken into account.	Throughput, goodput, packet drop, overhead, and utility	1 km × 1 km	250 m	Throughput, goodput, overhead, and utility
LARS [16]	The mechanism is proposed to use only locally aware observations that are further used to detect selfish nodes.	Goodput at cooperative nodes	670 m × 670 m	250 m	Goodput at cooperative nodes
2ACK [19]	Active/explicit acknowledgements for two hops are used to remedy the problems occurred in Watchdog mechanism.	PDR, routing overhead, and false alarms	700 m × 700 m	250 m	PDR and routing overhead
Comparison of trust-based protocols [21]	Various trust-based reactive routing protocols are compared in different scenarios and malicious nodes percentages.	Pkt loss, Pkt forwarded, overhead, latency, path optimality, and probability of detection	1 km × 1 km	250 m	Pkt forwarded, overhead, and latency
RTBD [15]	A trust-based model is used to counteract the issues of both selfish and congestion drop rate in the network.	PDR, drop rate, detection ratio, average latency, false positive rate, and overhead	100 m × 100 m	250 m	PDR, average latency, and overhead
EAACK [18]	The problems of Watchdog mechanism are addressed, that is, ambiguous collisions, receiver collisions, and so on.	PDR and routing overhead	670 m × 670 m	Not given	PDR and routing overhead
Detection and accusation [14]	This mechanism is developed for detecting and accusing nodes that exhibit packet forwarding misbehavior.	Percentage of detections, average network throughput, and network overhead	200 m × 200 m 283 m × 283 m 346 m × 346 m	100 m	Average network throughput and network overhead

CONFIDANT, cooperation of nodes and fairness in dynamic ad hoc networks; LARS, locally aware reputation system for mobile ad hoc networks; PDR, Pkt delivery ratio; RTBD, record-based and trust-based detection; EAACK, enhanced adaptive acknowledgement.

Table II. Simulation parameters.

Parameter	Level
Area	1000 m × 1000 m
Maximum speed	10 m/s
Pause time	0 to 1000 s
Radio transmission range	250 m
Number of nodes	50
Application	CBR
Packet size	64 B
Simulation time	900 s
Mobility model	Random waypoint model
Placement	Uniform
Selfish population	20%

CBR, constant bit rate.

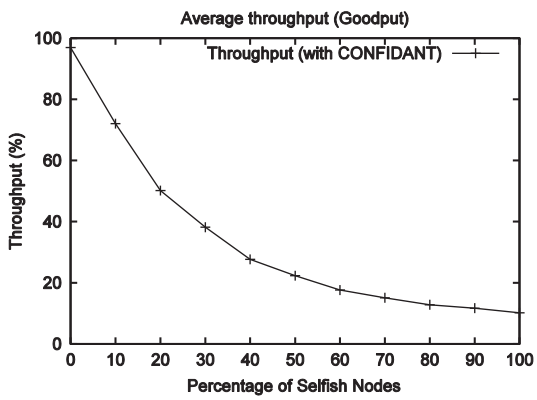

Figure 1. Average throughput versus selfish nodes. CONFIDANT, cooperation of nodes and fairness in dynamic ad hoc networks.

network throughput, with 60 s pause time, as shown in Figure 1. Here, it is evident that if all nodes are selfish, throughput still exists in the network because of the DIs. This throughput will be increased when the pause time reduces, because in high mobility environments, DIs tend to increase, as we will show in the later results.

Reputation or trust-based schemes are usually evaluated based on measurements of the utility and throughput [4,5,10] of nodes. Other metrics have also been considered [19,22] such as delay or the routing overhead of the network. These metrics are affected when the schemes are evaluated in environments where there are high levels of DIs among nodes in the network, as we will discuss this fact in the succeeding texts.

As shown in Figure 2, both the *normal*[§] evil and good throughputs are increased due to mobility where the increase in evil throughput is more than that of the good throughput. Moreover, when the mobility increases, the direct evil and direct good throughputs also increase, and

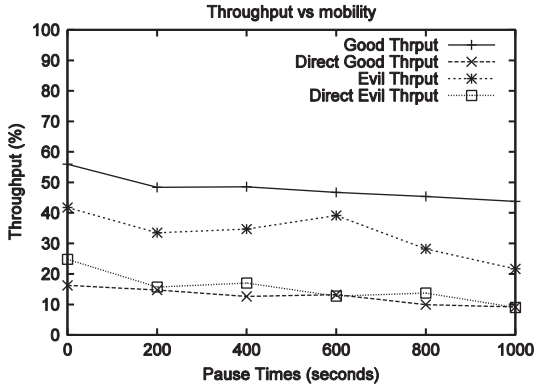

Figure 2. Direct and indirect throughputs versus mobility.

here again, the increase in direct evil throughput is affected more than the direct good throughput. This is because the number of good nodes is greater than the number of selfish nodes. In continuous mobile environment (zero second pause time), 59% of the evil throughput is gained through DIs, and 29% of the good throughput is achieved through DIs. In order to clarify this fact further, we found out that in 41.75% evil throughput, 24.75% is attained through DIs, while in 55.96% good throughput, 16.24% is achieved through DIs.

The evil throughput is high at a 600 s pause time. One of the reasons for this is that the evil throughput is also affected by the actual act of the node pausing. For example, during a pause, when a selfish node is detected by its neighbors, it then moves to a new neighborhood after the pause where it is likely to be treated as a new node. And hence, it starts receiving services again. However, more than one pause throughout the simulation lifetime can increase the chance of repeating interactions among nodes and hence faster detection. For example, a selfish node detected in one neighborhood (after having paused there) may interact with the same nodes again without obtaining any services a second time. The 600 s pause may be the optimal time for a selfish node to stay static along with other nodes, after which it then starts moving for the remaining period of 300 s until the simulation ends.

So the real benefit to a selfish node when a cooperation promotion scheme is active is considerably lower than the results suggest, after the DIs are filtered out, as shown in Figure 3. Likewise, the good throughput is also increased because of the DIs of nodes. This difference suggests that in work related to the effectiveness of reputation or trust-based schemes, there is a need to clarify, whether the increased good throughput is due to the DIs or the cooperation promotion scheme itself. Likewise, this also helps to establish how much evil throughput is reduced due to the reputation-based scheme. In short, categorizing throughput and utility clarifies that (i) the actual evil throughput is much lower and (ii) the good throughput is also lower than that claimed.

[§] We use the term "normal" for the throughput usually used in the literature that contains both direct and indirect throughputs.

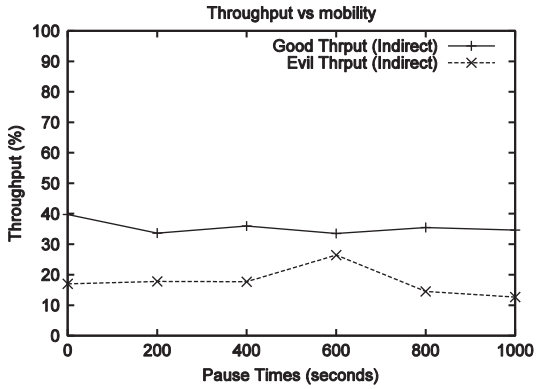

Figure 3. Indirect good and evil throughputs versus mobility.

As shown in Figures 4 and 5, there is considerable amount of evil and good utility gained due to DIs that can cause confusion for the developers as well as for the readers. In the scheme evaluation, a developer may be frustrated about the increase in evil utility (which is against the aim of the reputation-based systems); however, in reality, this might not be the case. Because the evil utility is partly gained using direct interactions which a reputation-based system does not have the mechanisms to prevent or even mitigate it. In the case of good utility, the actual indirect good utility might not be so much to claim the scheme to be efficient.

The reason for the greater received packets than sent packets in Figure 5 is that the sent packets are recorded for good nodes only, that is, for 80% of the population (because it is good utility) while the received packets are recorded for the good as well as evil nodes, that is, all population.

The DIs also affect the delay and routing overhead of the network. Because, during the times the nodes interact directly, there is no packet forwarding activity occurs along the path that reduces the overall forwarding events and routing overhead. Similarly, no intermediate nodes are involved in the path will reduce the overall delay. The aforementioned facts can be seen in Figures 6 and 7.

It becomes harder to establish from the results that whether the reduced delay and routing overhead are caused by the DIs or by the beauty of the designed reputation-based scheme.

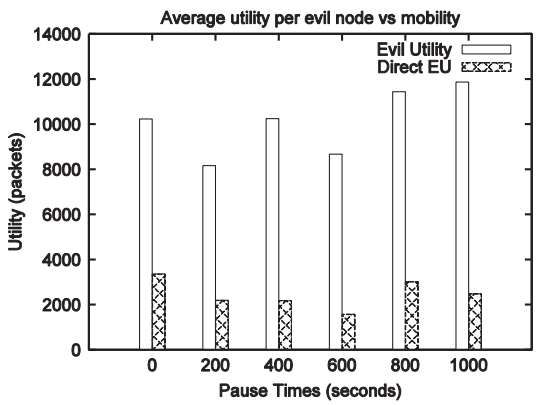

Figure 4. Direct and indirect evil utilities versus mobility. EU, evil utility.

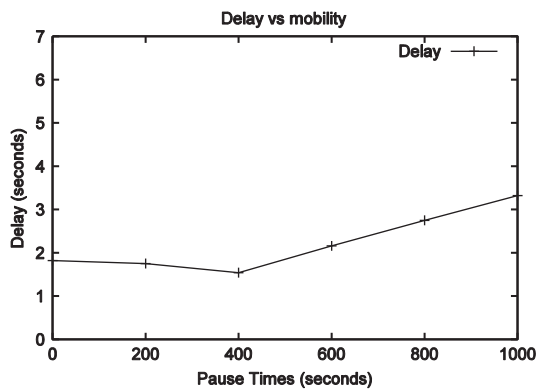

Figure 6. Delay versus mobility.

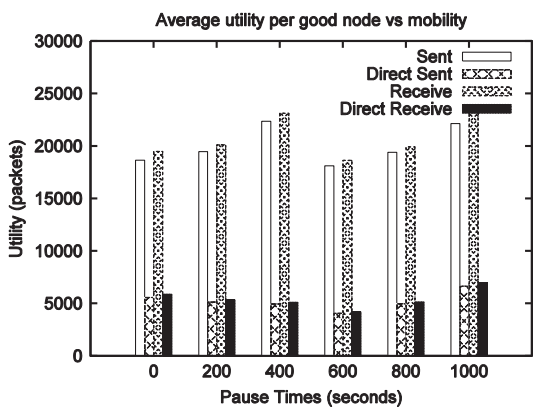

Figure 5. Direct and indirect good utilities versus mobility.

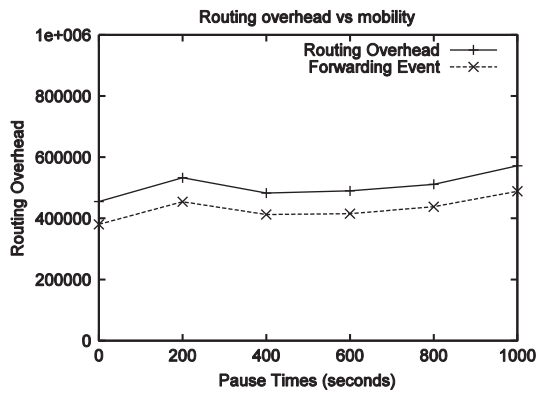

Figure 7. Routing overhead versus mobility.

4. ANALYZING SIMULATION PARAMETERS

As discussed previously, the simulation parameters normally used for reputation-based scheme evaluation in mobile environment can cause substantial DIs in the network. It is because the reputation and trust-based schemes are not evaluated in a good multihop environment. The examples of different simulation settings have been presented in Table I. The paths, established during routing in such simulation settings, are hardly few hops long. A node can have too many nodes in its direct communication range when the network nodes having larger radio ranges are deployed into a smaller simulation area. And after the deployment, when the traffic sources and sinks/destinations are randomly selected during simulation setup (such as using cbr.tcl in NS2), most of the sources and sinks fall directly in the range of each other; hence, high DIs are produced in the network. Those sources and sinks that do not fall directly in each others' ranges, even a low mobility (in any pattern, hence using any mobility model) will make them interact directly. In Section 3, while evaluating our benchmark scheme, we have shown that even 250 m radio range in 1000 m squared area (in mobile environment) produced considerable DIs. We believe that 250 m radio range in 1000 m or less will produce more DIs, hence more biased results.

Because, intuitively mobility increases the DIs, we take a worst case scenario where the possibility of DIs is low, that is, the scenario where nodes are randomly deployed and after deployment, they do not move. In order to simplify the scenario, we try to calculate the likelihood of any two nodes in a MANET that are able to communicate directly with one another, without the need to perform multihop communication using other nodes.

We start by considering the territory that the nodes inhabit. For the sake of simplicity (and because it is an assumption invariably made for MANET simulations or tests), we suppose that the territory is rectangular. Suppose it has dimensions of X and Y along its two edges as shown in Figure 8(a).

We must also consider the characteristics of the MANET nodes. We suppose each as a radio range of r and that their centers must lie within the territory. So if the node's location is at (x, y), we must have $0 \leq x \leq X$ and $0 \leq y \leq Y$. Figure 8(a) indicates two nodes with centers (x_1, y_1) and (x_2, y_2). The two nodes are able to communicate if each falls within the other's radio range. That is, the distance between the two nodes is less than r, or equivalently one node lies within a circle of radius r drawn around the other node.

To calculate the probability that two nodes can communicate, we therefore simply have to calculate the likelihood that, given two randomly placed points in the territory, the second lies within the area representing the radio range of the first.

If the territory is wrapped at the edges (i.e., a torus geometry), the calculation would be straightforward. In such a world, the area of the radio range around a point in the territory could be calculated as πr^2, the area of a circle with radius r. For a circle like this within the territory, the probability that any other point would fall within this radio range would be $\frac{\pi r^2}{X \times Y}$, because there are $X \times Y$ possible points in the territory to choose from, and πr^2 of them fall within radio range of the node.

However, for most simulations, the territory will be bounded, and so this calculation will no longer hold. The reason is that if a node is close to the edge of the territory, some of its radio range will fall outside of the area. As a result, the probability of a second node falling in to the radio range of the first node will depend on the precise location of the first. If the first node is sufficiently far from the edge that its radio range does not cross it, then the aforementioned calculation will hold. However, if it is close enough to the edge that some of the radio range is out of bounds, the probability will be lower.

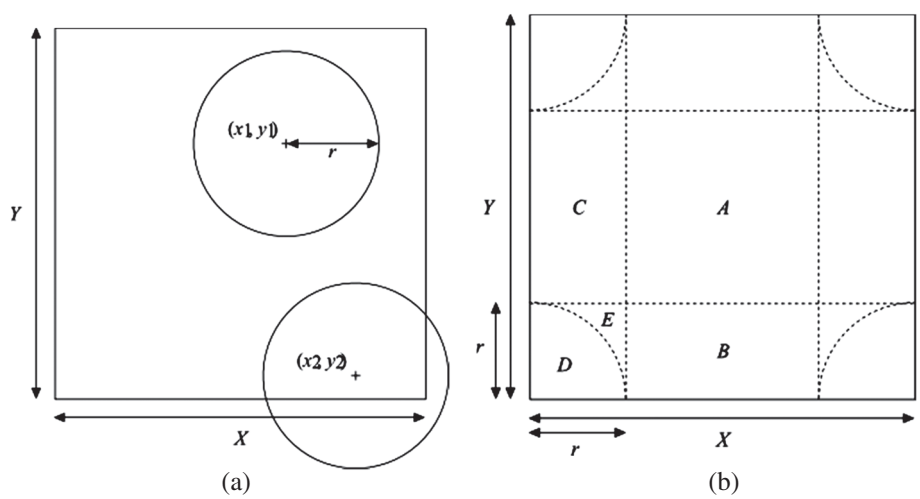

Figure 8. (a) Radio range of two nodes within a rectangular territory and (b) expected radio range area depends on where a node is located within the territory.

In order to calculate the probability, in general, we therefore calculate the *expected* size of the area covered by the radio range of the first node. We can then perform a similar division to that previously to calculate the probability that a second node will fall within this area. For the expected area, we break down the territory into a number of regions as shown in Figure 8(b).

Area A represents every point in the territory that is a distance of more than r away from the boundary. The radio range of a point within this part of the territory will not be cut off by the boundary. For the other areas, we need to perform some more complex calculations in order to figure out the expected area. For example, in area B, the radio range area will be cut off by the bottom edge of the territory, as shown in Figure A-1(a) in Appendix.

Suppose we have the expected radio range areas for each of the areas A, B, C, D, and E in Figure 8(b). By symmetry, we see that the calculation for each of the four corners is the same and that areas B and C have equivalent areas on each side of the territory. Let R be the area of the radio range around node n inside the territory. The expected area is based not only on the size of the radio range but also the relative size of the territory within which that area can occur, denoted by $E(R)$ in Table III.

The size of area A is $(X-2r)(Y-2r)$. The size of area B is $r(X-2r)$. For area C, it is $r(Y-2r)$; for D, it is $\frac{1}{4}\pi r^2$; and for E, it is $r^2-\frac{1}{4}\pi r^2 = r^2\left(1-\frac{\pi}{4}\right)$.

We can also calculate the expected area of the radio range for each of these portions of the territory by integrating all of the possible radio range areas across the possible locations of the node (x, y). We work through these calculations in Appendix. The results of all of these calculations are summarized in Table III.

Considering the values in this table, we get the following.

$$\begin{aligned} E(R) = {} & 1\times\frac{(X-2r)(Y-2r)}{XY}\times E(R:n\in A) \\ & + 2\times\frac{r(X-2r)}{XY}\times E(R:n\in B) \\ & + 2\times\frac{r(Y-2r)}{XY}\times E(R:n\in C) \\ & + 4\times\frac{\frac{1}{4}\pi r^2}{XY}\times E(R:n\in D) \\ & + 4\times\frac{r^2\left(1-\frac{\pi}{4}\right)}{XY}\times E(R:n\in E) \end{aligned}$$

Table III. Properties of the various portions of the territory.

Area	Size	Relative size	Occurrences	$E(R)$
A	$(X-2r)(Y-2r)$	$\frac{(X-2r)(Y-2r)}{XY}$	1	πr^2
B	$r(X-2r)$	$\frac{r(X-2r)}{XY}$	2	$r^3\left(\pi-\frac{2}{3}\right)$
C	$r(Y-2r)$	$\frac{r(Y-2r)}{XY}$	2	$r^3\left(\pi-\frac{2}{3}\right)$
D	$\frac{1}{4}\pi r^2$	$\frac{\frac{1}{4}\pi r^2}{XY}$	4	$\frac{r^4}{8}(\pi^2+1)$
E	$r^2\left(1-\frac{\pi}{4}\right)$	$\frac{r^2\left(1-\frac{\pi}{4}\right)}{XY}$	4	$r^4\left(\pi-\frac{\pi^2}{8}-\frac{4}{3}\right)$

Substituting the calculations from Appendix and performing some simplification, we therefore obtain the following.

$$\begin{aligned} E(R) = {} & \frac{1}{XY}\Big(\pi r^2(X-2r)(Y-2r) \\ & + 2r^4(X-2r)\left(\pi-\frac{2}{3}\right) \\ & + 2r^4(Y-2r)\left(\pi-\frac{2}{3}\right) \\ & + \frac{\pi r^6}{8}(\pi^2+1) \\ & + 4r^6\left(1-\frac{\pi}{4}\right)\left(\pi-\frac{\pi^2}{8}-\frac{4}{3}\right)\Big) \end{aligned} \tag{2}$$

For any particular size of area and radio range, we will know X, Y, and r and so can resolve this equation. To calculate the probability, we then have

$$P(\text{nodes are within range}) = \frac{E(R)}{XY}$$

We note that as X and Y become large, area A becomes larger relative to the other areas, assuming that r remains constant. We would therefore expect that

$$\lim_{X,Y\to\infty} E(R) = \pi r^2$$

As $X, Y\to\infty$, we can see that $\frac{1}{XY}\to 0$, $\frac{X-2r}{XY}\to 0$, $\frac{Y-2r}{XY}\to 0$, and $\frac{(X-2r)(Y-2r)}{XY}\to 1$. Because with r constant, the other terms in Equation (2) are also constant; we therefore obtain the following.

$$\begin{aligned} \lim_{X,Y\to\infty} E(R) = {} & (\pi r^2\times 1 \\ & + 2r^4\times 0\times\left(\pi-\frac{2}{3}\right) \\ & + 2r^4\times 0\times\left(\pi-\frac{2}{3}\right) \\ & + 0\times\frac{\pi r^6}{8}(\pi^2+1) \\ & + 0\times 4r^6\left(1-\frac{\pi^2}{4}\right)\left(\pi-\frac{\pi^2}{8}-\frac{4}{3}\right)\Big) = \pi r^2 \end{aligned}$$

as required.

5. DISCUSSION AND RESULTS COMPARISON

In order to verify the results of this equation, we performed a simulation that placed a pair of nodes within a given territory area. By running this process over several thousand iterations, we calculated the proportion of times that the nodes were within range of one another.

The results of these simulations are shown in Figure 9, along with the calculated value from the equation. The lower and upper bounds shown in this figure correspond to probabilities of $\frac{\pi r^2}{XY}$ and $\frac{\pi r^2}{(X-2r)(Y-2r)}$, respectively, because the expected area should fall within this range. In Figure 9, we can see that the simulated values match the theoretical expected values very closely, hence verifying the results.

In order to compare our empirical results with our case study results, we simulate our benchmark scheme in NS-2.30 based on the simulation parameters shown in Table II. We perform the simulation for various areas and use about 25 random seeds for different scenarios while observing network throughput. The result is shown in Figure 10. It is evident from both Figures 9 and 10 that the DIs tend to increase when the simulation area is reduced for a given (constant) radio range. It is further observed from Figure 10 that for small areas, the throughput is increased, but it is actually due to the increase in DIs occurred in the territory.

It is evident from Figures 9 and 10 that our empirical results closely match the case study simulation results. We have shown that even in a worst case scenario, that is, in the absence of mobility, our claim is valid. However, intuitively, in mobile environments, the DIs will be increased; hence, we believe that in such an environment, the results will be more biased and confusing.

Because DIs are natural and cannot be completely avoided or vanished even though the area is considerably expanded, as shown by Figure 10, it is suggested to categorize the throughput and utility based on direct and indirect entities in order to reduce the ambiguity and confusion caused by DIs. We have shown in Section 3.3 that this method is suitable for reducing the prevalence of confusing results. Because high levels of DIs produce high evil and good throughput in a network, categorizing utility and throughput into direct and indirect forms will help clarify the evaluation of a scheme. In other words, when evaluating the throughput of a reputation-based scheme, it may be better to isolate the regular (indirect) throughput from the direct throughput (based on DIs) to establish whether the throughput is increased because of the sophistication and efficiency of the designed scheme or as an effect of the DIs.

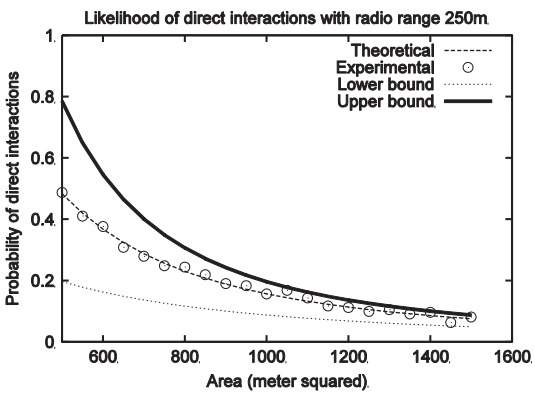

Figure 9. Area versus direct interactions.

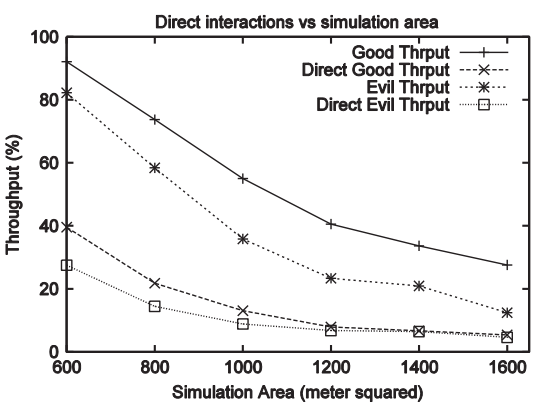

Figure 10. Direct interactions (DIs) with different simulation areas.

6. CONCLUSION AND FUTURE WORK

We have established that commonly used simulation parameters produce a considerable amount of DIs as a by-product, which could affect the interpretation and evaluation of any results obtained from the reputation or trust-based models. We demonstrated this fact through simulating a widely used scheme as our case study in which we have shown that during evaluation of cooperation promotion schemes in mobile environments, DIs can affect certain metrics, such as throughput, utility, delay, and routing overhead. Hence, the results obtained caused confusion due to biasness in it. We investigated this problem and empirically analyzed the simulation parameters used. Based on our simulations and empirical analysis, we suggest that cooperation promotion schemes should be evaluated using wider areas for constant radio ranges in order to reduce the effect of DIs. Alternatively, the induced gain or loss resulted from DIs should be highlighted in the results obtained when evaluating reputation or trust-based schemes, especially in mobile environments. In our future work, we will investigate the effect of different mobility models on the evaluation of reputation and trust-based schemes.

ACKNOWLEDGEMENTS

The authors are thankful to Shanshan Song and Sonja Buchegger from Deutsche Telekom Laboratories, Berlin, Germany, for contributing their NS-2 source code in support of the simulations experiments.

REFERENCES

1. Conti M, Giordano S. Mobile ad hoc networking: milestones, challenges, and new research directions. *IEEE Communications Magazine* 2014; **52**:85–96.
2. Biagioni E, Giordano S. Ad hoc and sensor networks [series editorial]. *IEEE Communications Magazine* 2015; **53**:248–248.
3. Marti S, Giuli TJ, Lai K, Baker M. Mitigating Routing Misbehaviour in Mobile Ad hoc Networks, presented at the Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, Boston, Massachusetts, United States, 2000.
4. Buchegger S, Boudec J-YL. Performance Analysis of the CONFIDANT Protocol, presented at the *Proceedings of the 3rd ACM International Symposium on Mobile Ad hoc Networking & Computing*, Lausanne, Switzerland, 2002.
5. Michiardi P, Molva R. Core: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks, in *Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security*, 2002.
6. Thorat SA, Kulkarni P. Opportunistic routing in presence of selfish nodes for MANET. *Wireless Personal Communications* 2015:1–20.
7. Abbas S, Merabti M, Llewellyn-Jones D. A Survey of Reputation Based Schemes for MANET, in *The 11th Annual Conference on the Convergence of Telecommunications, Networking & Broadcasting* (*PGNet 2010*), Liverpool, UK, 2010.
8. Djahel S, Nait-abdesselam F, Zhang Z. Mitigating packet dropping problem in mobile ad hoc networks: proposals and challenges. *IEEE Communications Surveys & Tutorials* 2011; **13**:658–672.
9. Jesudoss A, Raja SK, Sulaiman A. Stimulating truth-telling and cooperation among nodes in VANETs through payment and punishment scheme. *Ad Hoc Networks* 2015; **24**:250–263.
10. Buchegger S, Boudec J-YL. A robust reputation system for P2P and mobile ad-hoc networks, in Proceedings of P2PEcon, Harvard University, USA, 2004.
11. Abbas S, Merabti M, Llewellyn-Jones D. The Effect of Direct Interactions on Reputation Based Schemes in Mobile Ad Hoc Networks, in *IEEE Consumer Communications and Networking Conference* (*CCNC*) 2011; 297–302.
12. Friedman E, Resnick P, Sami R. Ch: 27- manipulation-resistant reputation systems. In *Algorithmic Game Theory*, Nisan N, *et al*. (eds). Cambridge University Press: New York, 2007; 677–697.
13. Feng L, Jie W. Uncertainty modeling and reduction in MANETs. *IEEE Transactions on Mobile Computing* 2010; **9**:1035–1048.
14. Gonzalez OF, Ansa G, Howarth M, Pavlou G. Detection and accusation of packet forwarding misbehaviour in mobile ad-hoc networks. *Journal of Internet Engineering* 2008; **2**:181–192.
15. Senthilkumar S, William J, Karthikeyan S. A distributed framework for detecting selfish nodes in MANET using record-and trust-based detection (RTBD) technique. *EURASIP Journal on Wireless Communications and Networking* 2014; **1**:1–10.
16. Hu J, Burmester M. LARS: A Locally Aware Reputation System for Mobile Ad hoc Networks, presented at the Proceedings of the 44th Annual Southeast Regional Conference, Melbourne, Florida, 2006.
17. Bansal S, Baker M. Observation-based cooperation enforcement in ad hoc networks, Stanford University Technical Report, July 2003. Available at: http://arxiv.org/pdf/cs.NI/0307012.
18. Kang N, Shakshuki EM, Sheltami TR. Detecting Misbehaving Nodes in MANETs, in *Proceedings of the ACM 12th International Conference on Information Integration and Web-based Applications & Services*, 2010.
19. Kejun L, Jing D, Pramod KV, Kashyap B. An acknowledgement-based approach for the detection of routing misbehavior in MANETs. *IEEE Transactions on Mobile Computing* 2007; **6**:488–502.
20. Zhao L, Delgado-Frias JG. MARS: Misbehavior Detection in Ad Hoc Networks, presented at the *IEEE Global Telecommunications Conference* (*GLOBECOM*) 2007.
21. Pirzada AA, McDonald C, Datta A. Performance Comparison of Trust-Based Reactive Routing Protocols, IEEE transactions on mobile computing, vol. **5**, 2006.
22. Yang H, Shu J, Meng X, Lu S. SCAN: self-organized network-layer security in mobile ad hoc networks. *IEEE Journal on Selected Areas in Communications* 2006; **24**:261–273.

APPENDIX

It remains to show that the values of $E(R)$ shown in Table III are indeed correct. We go through each of the five expected areas for the five regions.

Area *A*

Area A is the easiest one. Because a circle of radius r in this portion of the terrain will always have area πr^2, the average area is the same: πr^2.

Area *B*

A circle drawn in area B will always be cut along the bottom edge, as shown in Figure A-1(a). If the center of t

circle is chosen entirely at random and assuming we restrict our consideration solely to area B, every possible value of x is equally likely where $0 \le x \le r$. In order to calculate the average area of the circle in this part of the environment, we should therefore integrate the area of the cut circle across all of the possible values of x.

The area GHI can be calculated as

$$\begin{aligned} GHI &= FGH - FGI \\ &= \frac{1}{2}r^2\cos^{-1}\left(\frac{x}{r}\right) - \frac{1}{2}x\sqrt{r^2 - x^2} \end{aligned} \tag{A1}$$

To calculate the area we are interested in, we can subtract the shaded area in the diagram from the area of the circle. In other words

$$\begin{aligned} \text{Area } A &= \pi r^2 - 2\times(FGH - FGI) \\ &= \pi r^2 - \left(r^2\cos^{-1}\left(\frac{x}{r}\right) - \sqrt{r^2 - x^2}\right) \end{aligned}$$

For area B, we therefore get $r^3\left(\pi - \frac{2}{3}\right)$, as given in the succeeding texts,

$$\int_0^r \pi r^2 - \left(r^2\cos^{-1}\left(\frac{x}{r}\right) - \sqrt{r^2 - x^2}\right) dx = r^3\left(\pi - \frac{2}{3}\right)$$

Area C

For area C, we get the same as for $B\, r^3\left(\pi - \frac{2}{3}\right)$

Area D

For area D, the circle is cut across two edges with the corner of the territory inside the circle, shown as point M in Figure A-1(b). The area FHJ is simply a quarter of the circle and therefore has area

$$FHJ = \frac{1}{4}\pi r^2$$

And the area $FIML$ is a rectangle with width x and height y and therefore has area

$$FIML = xy$$

Combining these, we therefore find the area of the circle inside the territory GMK is equal to

$$\begin{aligned} GMK = &\frac{3}{4}\pi r^2 + \frac{1}{2}r^2\left(\cos^{-1}\left(\frac{x}{r}\right) + \cos^{-1}\left(\frac{y}{r}\right)\right) \\ &-\frac{1}{2}\left(x\sqrt{r^2 - x^2} + y\sqrt{r^2 - y^2}\right) - xy \end{aligned}$$

This provides us with a formula for the area for a given x and y, but in order to determine the *average* area for all possible locations of the center of the circle within area D, we must integrate across the possible values of x and y. We must therefore integrate across x where $0 \le x \le r$ and y where $0 \le y \le \sqrt{r^2 - x^2}$:

$$\text{Area } D = \int_0^r \int_0^{\sqrt{r^2 - x^2}} GMK\, dy\, dx$$

We consider each part of the sum separately.

$$\begin{aligned} &\int_0^r \int_0^{\sqrt{r^2 - x^2}} \frac{3}{4}\pi r^2\, dy\, dx = \frac{3\pi^2 r^4}{16} \\ &\int_0^r \int_0^{\sqrt{r^2 - x^2}} \cos^{-1}\frac{x}{r}\, dy\, dx = \frac{r^2}{4}\left(\frac{\pi^2}{4} + 1\right) \\ &\int_0^r \int_0^{\sqrt{r^2 - x^2}} x\sqrt{r^2 - x^2}\, dy\, dx = \frac{r^4}{4} \\ &\int_0^r \int_0^{\sqrt{r^2 - x^2}} y\sqrt{r^2 - y^2}\, dy\, dx = \frac{r^4}{4} \\ &\int_0^r \int_0^{\sqrt{r^2 - x^2}} xy\, dy\, dx = \frac{r^4}{8} \end{aligned}$$

Combining each of these parts together, we get the following.

$$\begin{aligned} \text{Area } D = &\frac{3\pi^2 r^4}{16} - \frac{r^2}{2}\left(\frac{r^2}{4}\left(\frac{\pi^2}{4} + 1\right) + \frac{r^2}{4}\left(\frac{\pi^2}{4} + 1\right)\right) \\ &+ \frac{1}{2}\left(\frac{r^4}{4} + \frac{r^4}{4}\right) + \frac{r^4}{9} = \frac{r^4}{8}(\pi^2 + 1) \end{aligned}$$

For area D, we therefore get $\frac{r^4}{8}(\pi^2 + 1)$.

Area E

For area E, the circle is also cut across two edges as shown in Figure A-2; however, in this case, the corner M lies outside the circle. From Equation (A1), we know that

$$KJN = \frac{1}{2}r^2\cos^{-1}\left(\frac{x}{r}\right) - \frac{1}{2}x\sqrt{r^2 - x^2}$$

and

$$GHL = \frac{1}{2}r^2\cos^{-1}\left(\frac{y}{r}\right) - \frac{1}{2}x\sqrt{r^2 - y^2}$$

Assuming that x and y are randomly placed within this area, we can calculate the average area by integrating across x and y within these constraints. We therefore need to calculate

$$\int_0^r \int_{\sqrt{r^2 - x^2}}^r \pi r^2 - KJN - GHL\, dy\, dx$$

We consider each of the parts of the sum separately.

$$\int_0^r\int_{\sqrt{r^2-x^2}}^r \pi r^2\,dy\,dx = \pi r^4\left(1-\frac{\pi}{4}\right)$$

$$\int_0^r\int_{\sqrt{r^2-x^2}}^r \cos^{-1}\left(\frac{x}{r}\right)dy\,dx = \frac{r^2}{4}\left(3-\frac{\pi^2}{4}\right)$$

$$\int_0^r\int_{\sqrt{r^2-x^2}}^r \cos^{-1}\left(\frac{y}{r}\right)dy\,dx = \frac{r^2}{4}\left(3-\frac{\pi^2}{4}\right)$$

$$\int_0^r\int_{\sqrt{r^2-x^2}}^r x\sqrt{r^2-x^2}\,dy\,dx = \frac{r^4}{12}$$

$$\int_0^r\int_{\sqrt{r^2-x^2}}^r y\sqrt{r^2-y^2}\,dy\,dx = \frac{r^4}{12}$$

Combining each of these parts together, we get the following.

$$\begin{aligned}\text{Area}\,E = \pi r^4\left(1-\frac{\pi}{4}\right)\\ -r^2\left(r^2-\left(\frac{r^2}{4}\left(\frac{\pi^2}{4}+1\right)\right)+\frac{r^2}{4}\left(3-\frac{\pi^2}{4}\right)\right)\\ +\left(\frac{r^3}{3}-\frac{r^4}{6}+\frac{r^4}{12}\right) = r^4\left(\pi-\frac{\pi^2}{8}-\frac{4}{3}\right)\end{aligned}$$

For area E, we therefore get $r^4\left(\pi-\frac{\pi^2}{8}-\frac{4}{3}\right)$.

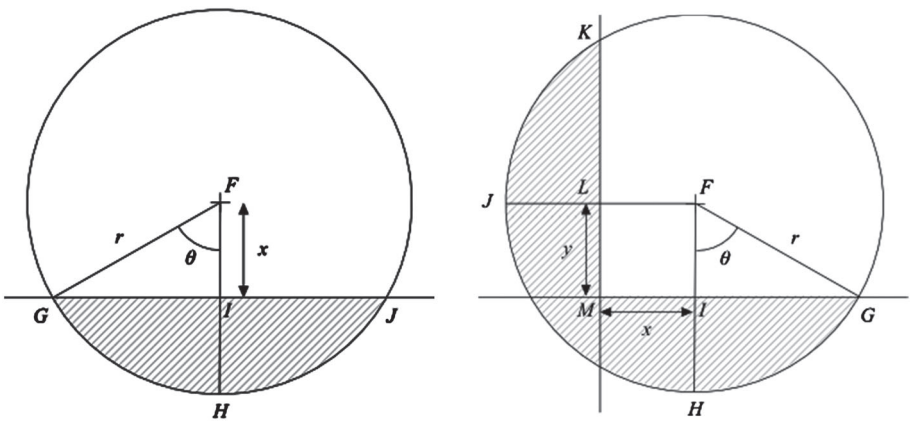

Figure A-1. (a) Radio range area bounded by the bottom edge of the territory and (b) radio range area bounded by a corner of the territory, where the corner lies inside the radio range area.

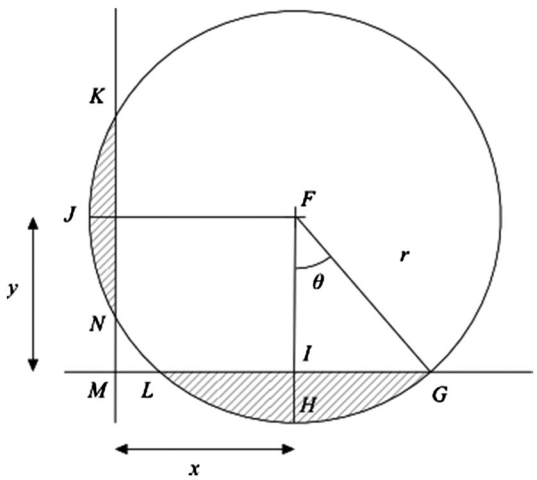

Figure A-2. Radio range area bounded by a corner of the territory, where the corner lies outside the radio range area.