Hit and Run: A Bayesian Game between Regular and Malicious Nodes in MANETs

Feng Li and Jie Wu
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431

***Abstract*—In mobile ad hoc networks (MANETs), nodes can move freely. Besides conducting attacks to maximize their utility and cooperating with regular nodes to deceive them, malicious nodes get better payoffs with the ability to move. In this paper, we propose a game theoretic framework to analyze the strategy profiles for regular and malicious nodes. We model the situation as a dynamic Bayesian signaling game, and analyze and present the underlining connection between nodes' best combination of actions and the cost and gain of the individual strategy. Regular nodes consistently update their beliefs based on the opponents' behavior, while malicious nodes evaluate their risk of being caught to decide when to flee. Some possible countermeasures for regular nodes that can impact malicious nodes' decisions are presented as well. An extensive analysis and simulation study shows that the proposed equilibrium strategy profile outperforms other pure or mixed strategies, and proves the importance of restricting malicious node's advantages brought by the flee option.**

I. Introduction

Mobile ad hoc networks (MANETs) rely on collaboration between the participants to achieve the aimed functionalities. Their topologies dynamically change because of node movement. Nodes in MANETs usually have no pre-defined trust between each other. Moreover, all nodes tend to maximize their own *utility* (also referred to as *payoff*) in activities. Among existing research, different mechanisms (e.g. virtual currency, barter economy) have been developed to stimulate cooperation and mitigate nodes' selfish behavior.

Besides regular nodes' selfish behavior, malicious nodes, which have different interests, also exist in the network. The common objective of malicious nodes is to maximize the damage to the network while avoiding being caught. Their utility comes from activities that disrupt the operation of the network and waste the resources of regular nodes.

To counter malicious nodes and stimulate cooperation, regular nodes monitor and continuously evaluate their neighbors. Certain criteria are set to distinguish a node's trust level towards others. Regular nodes will focus their resources on cooperating with neighbors they trust, decline requests from suspicious neighbors, and report when a neighbor is considered to be malicious. However, in this case, intelligent malicious nodes would elaborately choose a frequency at which they cooperate to deceive regular nodes.

Moreover, in MANETs, malicious nodes have the strategy of fleeing to avoid punishment. Therefore, a malicious node can start its malicious behavior all over again with a clean history in a new location by fleeing before being caught. However, this additional strategy does not imply that malicious nodes should continuously hit and run, since fleeing is also associated with a cost (e.g. the energy spent to move to the selected destination). We can instinctively describe malicious nodes' optimal strategy as follows: cooperate to deceive regular nodes' trust; attack to cause damage and maximize their own utility; flee before regular nodes accumulate enough evidence and decide to report. Now the critical questions become:

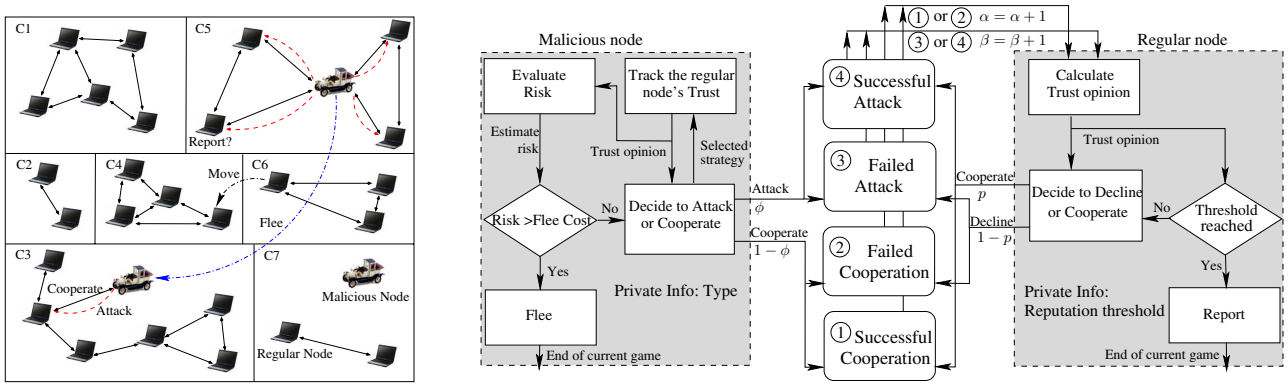
1) How to choose between acting good or bad?
2) When should a regular node report?
3) When should a malicious node flee?
4) What countermeasures are available to restrict the malicious node's advantages brought by the flee strategy?

We model the wrestling between the regular and malicious node as a *dynamic Bayesian game*, and provide answers to the aforementioned questions through analysis. In this game, nodes observe the result of each round of communication. Each node's type, regular or malicious, is its own private information. Its neighbor's actual type is the incomplete information in the game. Each node should form beliefs towards neighbors and update the beliefs according to the neighbors' actions as the game evolves.

Both regular and malicious nodes' best responses are guided by the threats about certain reactions from other players. Such threats are dependent on their current beliefs. The regular node sets a reputation threshold and judges other nodes' types based on the evaluated belief and this threshold. The malicious node continuously evaluates the risk, which is decided by the possibility that a regular node would choose to report under current conditions. On the basis of the risk and expected fleeing cost, the malicious node makes a decision on whether to flee.

The contributions of this paper are as follows:

1) We formulate a Bayesian game framework to study the strategy of regular and malicious nodes in MANETs.
2) We propose decision rules for regular nodes to report, which consider both the current belief and sufficiency of the evidence.
3) We develop decision rules for malicious nodes to flee, which balance the fleeing cost and risk of being caught.
4) We study the equilibrium strategy profiles for both parties based on the belief and expected payoff, and reveal the connection between nodes' best response and the cost and gain of each individual strategy.



(a) Example scenario

(b) Decision process

Fig. 1. The wrestling between regular and malicious nodes in MANETs.

5) We present several countermeasures to restrict malicious nodes' advantage from the flee strategy.

II. Basic Model and Assumptions

We consider a MANET which contains both regular and malicious nodes. We will not restrict malicious nodes' ability to coordinate. Hence, they would avoid playing the following game with each other because there is no gain in doing so. Regular nodes only know their own type. To simplify the analysis, time is divided into slots and players choose their strategies simultaneously at the beginning of each time slot.

A. Community

Community denotes a logical region of the MANET, in which nodes are highly connected with each other. The grid in Fig. 1(a) indicates a community. A MANET can always be divided into communities. Nodes can dynamically leave or join a community during their movement. We assume an authentication method exists, e.g. the public-key-based authentication, and that the identity is bound with the physical node which cannot be changed or faked during the node's stay in the community. When a node newly joins a community, other nodes in the community authenticate the node and set their belief towards the newcomer to the initial value.

When a malicious node *flees* (F) into a community that it has never visited before, nodes in that community will treat it as a newcomer. When a malicious node rejoins a community, it can still appear to be a newcomer by using an identity that is different from the one which is used during its last stay since a node's behavior cannot be tracked and the identity binding cannot be monitored outside the community.

A community can either be organized or ad hoc. In the former case, a commonly trusted node is elected as the cluster head. When a regular node decides to *report* (R) one of its neighbors as a malicious node, it sends its report to the cluster head. In the latter case, the regular node will broadcast the report in its current community. If the report is considered to be true, the malicious node being reported will be punished. Otherwise, the reporting regular node's accountability will be affected for the false alarm. Methods to distinguish the false alarm and punish malicious nodes are out of the scope of this paper, and criteria in existing systems can be applied [1]. We use the expected gain of a correct report and expected loss of a false alarm to summarize the results of reporting in the game.

B. Neighbor monitoring

As shown in Fig. 1(b), the regular user can choose to cooperate or decline one round of communication while the malicious node can attack or cooperate. Here, *decline* (D) means a node simply rejects participation while *cooperate* (C) means a node makes itself available for communication. The packet can be forwarded through a link only when nodes on both endpoints of the link choose to cooperate. The regular node benefits from good network operations. However, each receiving and forwarding action also costs energy. If a regular node chooses to cooperate while the other node on the link chooses not to, the regular node wastes energy.

The malicious node *attacks* (A) to waste the resources and disrupt the operation of the network. Attacking leads to the failure of one round of communication between two neighbors. Malicious nodes can conduct a simple dropping-packet attack, which is in the same form as the decline strategy of regular nodes. However, malicious nodes get payoff from the attack, while regular nodes receive no gain from the decline. Malicious nodes can also conduct more sophisticated attacks, such as analyzing a received packet without further forwarding, or sending out a modified packet. To make the definition of an attack more general, we use the cost and gain metrics to summarize the characteristic of one type of attack in the game.

Exploiting the promiscuous nature of broadcast communication in wireless media, nodes track the outgoing packets of their one-hop neighbors through passive observation. However, a node cannot distinguish whether a failure in communication is caused by its opponent's A or D (A/D for short). Therefore, an observation is classified as either a detected C or a detected A/D. Accordingly, the corresponding discrete variable, α for detected C and β for detected A/D, is incremented as shown in Fig. 1(b). This mechanism is called *neighbor monitoring*.

C. Decision process

Fig. 1(b) shows the general decision process of regular and malicious nodes. The regular node obtains feedback from the neighbor monitoring, and evaluates the belief and sufficiency of evidence towards the opponent based on α, β. It follows a threshold policy to decide whether to report. If not, the regular node chooses C with a probability p, which is calculated based on its belief. The malicious node also tracks the regular node's trust opinion, evaluating the risk of being caught. It follows its rule to decide whether to flee. If not, the malicious node chooses A with a probability ϕ. The key issues in this decision process are the decision rules for both parties and the action profiles reflected by p and ϕ. We analyze the MANET to find the optimal decision rules and action profiles by using the dynamic Bayesian game framework.

We first explain some concepts [2] used in the framework. *Bayesian games* are the combination of game theory and probability theory that allow taking incomplete information into account. In Bayesian games, each player is allowed to have some private information that affects the progress of the game, e.g. the node's type in the regular/malicious node game. However, others are assumed to have beliefs about the private information. These beliefs are represented by probability distributions and updated using Bayes' rule whenever new information is available.

Signaling games are one specific category of Bayesian games with two players: the sender and the receiver. In the regular/malicious node game, the receiver is a regular node, and the sender is one of the regular node's neighbors that is being observed. The sender's type, which can be regular or malicious in our game, is its private information. Based on its own type, the sender chooses to send a message from a set of possible messages. The receiver observes the message but not the type of sender. The receiver chooses an action from a set of feasible actions as a response to the message.

Through analysis, we aim to find the *Perfect Bayesian Equilibrium* (PBE) of this game. Nodes' best response strategy profiles, which are based on the evolving beliefs towards the opponents' types, are described by the PBE.

III. Regular/Malicious node game

We model the regular/malicious node game as a multi-stage dynamic Bayesian signaling game to find the optimal strategy of regular and malicious nodes.

A. Game specification

In the regular/malicious node game, player i is the sender, and its type can be regular or malicious; player j is a regular node, and it is the receiver. In each time slot, each player chooses its action from its strategy space. The strategy space for regular nodes is $\{C, D, R\}$. For malicious nodes, the strategy space is $\{A, C, F\}$. After each time slot, each player gets a payoff that depends on its own action, its neighbors' actions, and its own type. The payoffs are listed in Table 1. Here, G denotes gain, C denotes cost, the subscript denotes the corresponding strategy, and L_F denotes loss of false alarm.

TABLE I
Strategic form of the regular/malicious node game

	C	D	R
A	$(G_A{-}C_A, -G_A{-}C_C)$	$(-C_A, 0)$	$(-G_R{-}C_A, G_R{-}C_R)$
C	$(-C_C, G_C{-}C_C)$	$(-C_C, 0)$	$(-G_R{-}C_C, G_R - C_R)$
F	$(-C_F, -C_C)$	$(-C_F, 0)$	$(-C_F, -C_R)$

(a) Node i is malicious: (i's utility, j's utility).

	C	D	R
C	$(G_C{-}C_C, G_C{-}C_C)$	$(-C_C, 0)$	$(-C_C, -L_F - C_R)$
D	$(0, -C_C)$	$(0, 0)$	$(0, -L_F - C_R)$
R	$(-L_F{-}C_R, -C_C)$	$(-L_F{-}C_R, 0)$	$(-L_F{-}C_R, -L_F{-}C_R)$

(b) Node i is regular: (i's utility, j's utility).

For both players, all possible strategies except D incur cost. The cost can be interpreted as the energy spent to conduct certain actions. For a malicious node, it gains G_A from a successful A. The success depends on its neighbor's strategy selection. Only when regular node j selects C will the attack succeed. The malicious node could also choose C to deceive node j. However, there is no gain for the malicious node if it chooses C in a one-shot game as it has a different objective compared to regular nodes. Regular nodes gain G_C from a successful C. They could also choose D which incurs zero gain and no cost even if the opponent chooses A in a stage game.

Both players have one more option. When choosing F, the malicious node avoids the risk of being caught. Therefore, the expected gain for F is the value of risk. This risk is not static; it increases as the regular node j's evidence accumulates. If j chooses R, it gets the gain G_R if i is a malicious node. The malicious node is considered to be caught in this case. However, j should also consider the possible loss for false alarm. If i is a regular node and j reports i as a malicious node, node j needs to bear the loss L_F for this false alarm.

B. Stage game

Stage games are simple games played at individual time slots. In each stage game, the objective of both regular and malicious nodes is to maximize their expected payoff. This implies that both players are assumed to be *rational*.

The extensive form of the game is given in Fig. 2. Nature determines the type of node i, and this type is i's private information. Node j's current belief that i's type is malicious is represented by θ. Recall that α and β denote the number of detected C and detected A/D in previous stage games respectively. According to Bayes' rule, θ should be calculated as $\theta = \frac{\beta}{\alpha+\beta}$ while $1-\theta = \frac{\alpha}{\alpha+\beta}$. We assign the initial value $\alpha = \beta = 1$ at the beginning, which makes $\theta = 0.5$. This initial belief is in compliance with the no-evidence situation.

We analyze the possible *Bayesian Nash Equilibrium* (BNE), which is the Nash equilibrium for a single stage game given nodes' beliefs. Nash equilibrium refers to the situation where each player has chosen a strategy and no player can benefit by changing its strategy while the others keep theirs unchanged.

We discuss pure strategy BNE under two cases. In the first case, node i plays its pure strategy $\sigma_i = (A$ if malicious, C if regular), which means i always plays A if its type is malicious and C if regular. The expected payoffs $E_j(C)$ or $E_j(D)$ of j

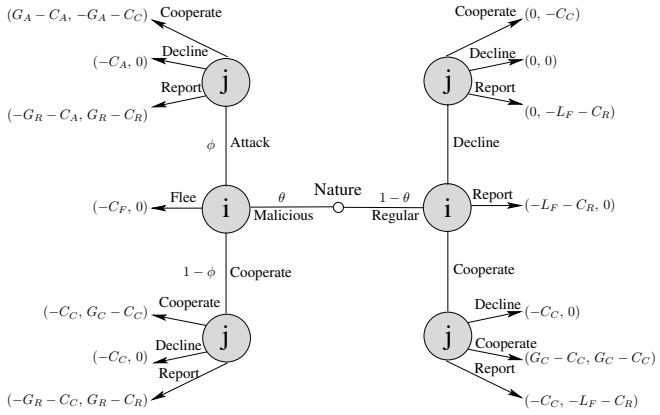


Fig. 2. Single stage of the flee game.

playing its pure strategy $\sigma_j = C$ or $\sigma_j = D$ are:

$$\begin{cases} E_j(C) = \theta \cdot (-G_A - C_C) + (1-\theta) \cdot (G_C - C_C) \\ E_j(D) = \theta \cdot 0 + (1-\theta) \cdot 0 \end{cases}$$

The formula of $E_j(C)$ enumerates two cases. One is that neighbor node i is a malicious node. According to j's current belief, this case appears with the probability θ. Since i will choose A, j's payoff in this case is $(-G_A - C_C)$. Another case is that i is a regular node, which appears with the probability $1-\theta$. j's payoff in this case is $(G_C - C_C)$. Other formulas in this section follow the same idea.

If $E_j(C) \geq E_j(D)$, node j's best response is to play C. That is, when the estimated probability $\theta \leq \frac{G_C - C_C}{G_C + G_A}$, the BNE strategy pair for i and j is $(\sigma_i, \sigma_j) = ((A$ if malicious, C if regular), $C)$. However, when $\theta > \frac{G_C - C_C}{G_C + G_A}$, there is no pure strategy BNE because when the malicious type node i plays A, the best response for j is to play D. However, if j plays D, it is possible that C is the best response for malicious type node i since C_A could be larger than C_C in some scenarios.

In the second case, malicious type node i plays pure strategy C. Then, j's best response is C, regardless of θ. However, if j plays C, malicious type node i's best response is A, which reduces to the previous case. In this case, $(\sigma_i, \sigma_j) = ((C$ if malicious, C if regular), $C)$ is not a BNE.

We now examine the mixed strategy BNE for situations without a pure strategy BNE. Recall that ϕ stands for the probability that the malicious type node i will play A, and p stands for the probability of that node j will play C. j's expected payoffs of C and D are:

$$\begin{cases} E_j(C) = \phi \cdot \theta \cdot (-G_A - C_C) + (1 - \phi \cdot \theta) \cdot (G_C - C_C) \\ E_j(D) = \phi \cdot \theta \cdot 0 + (((1-\phi) \cdot \theta) + (1-\theta)) \cdot 0 \end{cases}$$

To make C and D indifferent to j, i.e. $E_j(C) = E_j(D)$, the malicious type node i's equilibrium strategy is to play A with $\phi = \frac{G_C - C_C}{(G_C + G_A) \cdot \theta}$. i's expected payoffs of A and C are:

$$\begin{cases} E_i(A) = p \cdot G_A - C_A \\ E_i(C) = -C_C \end{cases}$$

By imposing $E_i(A) = E_i(C)$ to make A and C indifferent to malicious type node i, we get that j's equilibrium strategy is to play C with probability $p = \frac{C_A - C_C}{G_A}$ (when $C_A < C_C$, $p = 0$). Thus, mixed strategy pair $(\sigma_i, \sigma_j) = ((\phi$ if malicious, C if regular), $p)$ is a BNE for corresponding situations.

Therefore, the BNE of the stage game can be summarized as: when $\theta \leq \frac{G_C - C_C}{G_C + G_A}$, $(\sigma_i, \sigma_j) = ((A$ if malicious, C if regular), $C)$; after $\theta > \frac{G_C - C_C}{G_C + G_A}$, j becomes more conservative and $(\sigma_i, \sigma_j) = ((\phi$ if malicious, C if regular), $p)$.

When searching for the BNE of a single stage game, we only need to consider C and A/D. However, the regular node has an additional option R and the malicious node has an additional option F, which makes the sequential rationality more complicated.

C. Sequential rationality: Report

For a multi-stage game, a strategy profile is *sequentially rational* if and only if the player's expected payoff is maximal in the subsequent play given the strategies played by its opponent, no matter what type the opponent is.

With the strategy R, regular node j faces the following question in each stage: should I report node i as a malicious node? Supposing it decides to report in this stage, there are two possible results: 1) Node i is malicious, and the report is correct. 2) Node i is regular, and the report is a false alarm.

The second result may occur since regular nodes also play D in some stage games to maximize their utility. Such a false alarm would draw unnecessary attention and reduce regular nodes' sensitivity to real attacks. Therefore, regular nodes should estimate the loss L_F for the event of the false alarm. L_F is a subjective value that reflects the regular node's characteristic. Larger L_F indicates a more conservative characteristic. L_F is the private information of the regular node.

The regular node's decision about whether it should report depends on the comparison between the expected correct report gain and the expected false alarm cost. Besides the formed belief θ, regular node j also needs to evaluate the sufficiency of the evidence before making decisions.

We use uncertainty u, which is calculated on the basis of α and β to measure the sufficiency of evidence. The uncertainty metric u could be defined as follows [3]:

$$u = \frac{12 \cdot \alpha \cdot \beta}{(\alpha + \beta)^2 \cdot (\alpha + \beta + 1)} \tag{1}$$

Two important attributes make the uncertainty metric suitable for measuring the sufficiency of evidence. First, when there is more evidence, u will consequently be lower. Second, when the evidence for detected C or A/D dominates, there will be less u when compared to the equal evidence situation.

Regular node j decides whether to report in the current stage game by checking whether a threshold T has been reached. The threshold T should reflect the combined requirement on both the proportion of detected A/D in the evidence and the sufficiency of the evidence. Consider a case where $\alpha = 1$ and $\beta = 2$. Although $\theta = 0.67$ is high, the sufficiency of evidence is low. If we use a T that only focuses on θ, it is highly possible that a regular type node i is falsely reported in this case. Since $1 - u$ can be regarded as regular node j's certainty towards

the current evidence, $\theta \cdot (1-u)$ is the proportion of certainty which supports the proposition that node i is a malicious node. The threshold T should be imposed on $\theta \cdot (1-u)$ to reflect both requirements.

To satisfy the sequential rationality, node j should decide to report only when $E_j(R) > max\{E_j(C), E_j(D)\}$, where $E_j(R) = \theta \cdot (1-u) \cdot (G_R - C_R) - ((1-\theta) \cdot (1-u) + u) \cdot (L_F + C_R)$. j should not choose R when $E_j(C) > 0$, as it should not end the game when it still expects to gain in the following stage games. Therefore, T should be calculated as the condition that makes $E_j(R) > 0$. We get $T = \frac{L_F + C_R}{G_R + L_F}$. When $\theta \cdot (1-u) > \frac{L_F + C_R}{G_R + L_F}$, regular node j will choose R.

Assume $T = 0.42$, when $\alpha = \beta = 10$, j should report as $\theta \cdot (1-u) = 0.43 > T$. Consider the case where $\alpha = \beta = 2$. j should not report as $\theta \cdot (1-u) = 0.2 < T$, although θ is the same. As the evidence is insufficient, R has a good chance of leading to a false alarm in the latter case.

As regular node j follows the above rules to R, malicious nodes have two choices: 1) Keep $\phi < T = \frac{L_F + C_R}{G_R + L_F}$. Hence, no matter how the uncertainty is reduced, node j will not report. 2) Choose a higher ϕ to attack, and flee before being reported.

D. Sequential rationality: Flee

When a malicious node decides to flee, the expected gain of such an action is to avoid the risk of being caught. But what is the definition for the risk? Since i's attack frequency ϕ depends on node j's belief θ, and j's reporting rule depends on belief and uncertainty, the malicious type node i's risk should be calculated based on opponent j's current opinion and threshold. The risk is defined as the expected loss of being reported $Risk = P(catch) \cdot G_R$, where $P(catch)$ denotes the probability of being caught. The malicious node should check whether $E_i(F) = Risk - C_F > max\{E_i(A), E_i(C)\}$. If this condition is satisfied, the malicious node should flee.

As the malicious type node i has perfect information about the transaction history between itself and regular node j, it can precisely estimate j's belief towards it. Since L_F is a subjective cost for the false alarm of node j, node i cannot know the exact value of the L_F. However, node i would have enough knowledge about the network and know the distribution of L_F. If the number of nodes is large enough in the network, L_F should comply to the normal distribution. Node i could know the standard deviation $VAR(L_F)$ and the expected value $E(L_F)$. $P(catch)$ is equal to the probability that the current $\theta \cdot (1-u)$ will pass j's threshold T, and $P(\theta \cdot (1-u) > T) = P(L_F < \frac{\theta \cdot (1-u) \cdot G_R - C_R}{1 - \theta \cdot (1-u)})$. Therefore, we have:

$$P(catch) = \Phi(\frac{\frac{\theta \cdot (1-u) \cdot G_R - C_R}{1 - \theta \cdot (1-u)} - E(L_F)}{VAR(L_F)}) \tag{2}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{u^2}{2}) du$.

From the analysis of the F strategy we can see the malicious node enjoys its advantage of choosing its optimal ϕ to attack and escapes punishment with the option to flee. It needs to keep evaluating the risk of staying and playing, and find a tradeoff between risk and C_F to maximize its payoff.

Strategy Profile 1 Player j's PBE strategy σ_j^*

```
while θ · (1 − u) < T do
    if θ ≤ (G_C − C_C)/(G_C + G_A) then
        Choose C with p = 1;
    else
        Choose C with p = (C_A − C_C)/G_A;
    end if;
    Updated α, β, get θ and calculate u;
end while
Report node i as a malicious node;
```

Strategy Profile 2 Malicious type player i's PBE strategy σ_i^*

```
while E_i(F) < max{E_i(A), E_i(C)} do
    if θ ≤ (G_C − C_C)/(G_C + G_A) then
        Choose A with φ = 1;
    else
        Choose A with φ = (G_C − C_C)/((G_C + G_A) · θ);
    end if;
    Track j's θ, estimate risk of being caught and E_i(F);
end while
Flee to a remote area and attack again;
```

E. Perfect Bayesian Equilibrium (PBE)

PBE is a refinement of BNE in multi-stage dynamic games since it demands sequential rationality. The PBE of this game describes the optimal decision rules for both regular and malicious nodes, and reveals the connection between the best strategy profile and the cost and gain of individual strategies.

From the discission, we can summarize player j's PBE strategy σ_j^* as strategy profile 1. The regular type player i has the same PBE strategy profile as j, and the PBE strategy σ_i^* of malicious type player i is listed as strategy profile 2.

IV. Analysis

In this section, a detailed analysis is given to prove some of the claims made in Section III.

A. Rationality of the reporting rule

We examine the $(k+1)th$ stage game. Node j's current belief $\theta = \frac{\beta}{\alpha + \beta}$ is the prior probability of the decision. We assume that the average probability that malicious type node i chooses A is ϕ, and the average probability that regular type node i will choose D is $1-p$. The results of the previous k stage games' should comply to the binomial distribution, and $k = \alpha + \beta - 2$. So, $P((\alpha, \beta)|r) = \binom{k}{\beta - 1}(1-p)^{\beta - 1} p^{\alpha - 1}$, and $P((\alpha, \beta)|m) = \binom{k}{\beta - 1} \phi^{\beta - 1} (1-\phi)^{\alpha - 1}$, where r denotes regular and m denotes malicious.

Hence, the probability that the regular node's decision to report under current conditions leads to a false alarm is:

$$P(r|(\alpha, \beta)) = \frac{(1-\theta)\binom{k}{\beta} p^{\beta} (1-p)^{k-\beta}}{(1-\theta)\binom{k}{\beta} p^{\beta} (1-p)^{k-\beta} + \theta \binom{k}{\beta} \phi^{\beta} (1-\phi)^{k-\beta}}$$

and $P(m|(\alpha, \beta)) = 1 - P(r|(\alpha, \beta))$. As the sequential rationality condition for R is: $P(r|(\alpha, \beta)) \cdot (L_F + C_R) \leq P(m|(\alpha, \beta)) \cdot (G_R - C_R)$, we can derive:

$$\frac{k - \beta}{\beta} (\frac{1-p}{\phi})^{\beta} (\frac{p}{1-\phi})^{k-\beta} \leq \frac{(G_R - C_R)}{(L_F + C_R)} \tag{3}$$

Inequality 3 reflects that both β's proportion of k and the value of k must be considered. Hence, we get lemma 1:

Lemma 1: Based only on the belief θ, a threshold-based reporting policy cannot guarantee the sequential rationality for regular nodes.

Proof: If we take only belief into account when imposing the threshold without considering uncertainty, the decision rule becomes: $\frac{\beta}{\alpha+\beta} > T$, and the regular node should report.

The threshold T in the decision rule regulates only the relationship between α and β, where $\beta \geq \frac{T}{1-T} \cdot \alpha$. T cannot regulate the value of k, which reflects the sufficiency of observations. It violates the sequential rationality requirement reflected in Inequality 3. ■

Therefore, only measuring belief is not enough for the sequential rationality requirement. Theorem 1 states that a combined threshold on uncertainty u and belief θ is necessary.

Theorem 1: By imposing a threshold T on $\theta \cdot (1-u)$, the sequential rationality for regular nodes can be guaranteed.

Proof: By imposing the threshold T on disbelief d, we actually impose a combined requirement on both u and $\frac{\beta}{\alpha+\beta}$. As $\frac{\beta}{\alpha+\beta} \cdot (1-u) > T$, we have: $u = \frac{12 \cdot \alpha \cdot \beta}{(\alpha+\beta)^2 \cdot (\alpha+\beta+1)} \leq 1-T$ according to Equation 1. Hence $k > 12 \cdot T - 3$ as $\beta \geq \frac{T}{1-T} \cdot \alpha$. So we now have a combined requirement for both the proportion of evidence for *A*/*D* and the sufficiency of observations. ■

B. Optimality of strategy profiles (σ_i^*, σ_j^*)

We first prove that the regular/malicious node game has a PBE, since the game satisfies the Bayesian postulates. After that, we prove that both proposed PBE strategy profiles 1 and 2 satisfy the sequential rationality condition.

Lemma 2: (Bayesian postulates) The described regular/malicious node game satisfies four Bayesian conditions [2]:

B1: Posterior beliefs are independent. All types of receivers have the same beliefs.

B2: Bayes' rule is used to update beliefs (θ) from stage game k to stage game $k+1$ whenever possible.

B3: The players do not signal what they do not know.

B4: All players must have the same belief about the type of another player.

Proof: B1 is satisfied because receiver j has only one type which is regular. Since $\theta = \frac{\beta}{\alpha+\beta}$, the updated θ satisfies Bayes' rule when α or β is incremented. Hence, B2 is satisfied. The regular and malicious nodes select their signal (*A*/*D* or C) based on their own payoff respectively, which fulfills B3. Because there are only two players in the game and no other players influence the belief updates, B4 is satisfied. ■

Lemma 3: (Sequential rationality) For each player x, given any alternative strategy σ_x of x, σ_x^* satisfies: $E_x(\sigma_x^*) \geq E_x(\sigma_x)$. Here, $E_x(\sigma_x)$ denotes the expected payoff of x's strategy σ_x when other players' play best response to σ_x.

Proof: As the receiver of the game, regular node j plays R only when $E_j(R) > max\{E_j(C), E_j(R)\}$. Otherwise, it will play C with the optimal probability p. The actions of the receiver maximize its expected payoffs given its beliefs.

Similarly, the sender i, which could be a malicious or regular node, chooses to *A*/*D* or C depending on which action will maximize its payoff given j's strategy and its own type. ■

As stated in [2], Theorem 2 can be derived from Lemmas 2 and 3, which indicates that (σ_i^*, σ_j^*) are the optimal decision rules for both parties in this game.

Theorem 2: The (σ_i^*, σ_j^*) described by strategy profiles 1 and 2 is a perfect Bayesian equilibrium of the regular/malicious node game.

Proof: Since the described regular/malicious node game satisfies the Bayesian conditions B1-B4 (Lemma 1) and (σ_i^*, σ_j^*) described by strategy profiles 1 and 2 satisfies the Sequential rationality condition (Lemma 2), (σ_i^*, σ_j^*) is a PBE. ■

C. Possible equilibrium of never-fleeing

When node j chooses the pure strategy C, the malicious type node i can also follow the mixed strategy of C and A with a fixed low ϕ, where $\phi < \frac{G_C - C_C}{G_C + G_A}$. Since, in this case $E_j(C) > E_j(D) = 0$, node j should always follow C and never choose R. However, whether the malicious type node i wants to keep such a low attack frequency and benefits from this mixed strategy depends on the parameters. More specifically, if there exists a ϕ, which makes $E_j(C) > 0$ and $E_i(\phi) \geq 0$ at the same time, regular and malicious nodes should follow (σ_i^*, σ_j^*) =((ϕ if malicious, C if regular), C), which turns out to be another possible equilibrium. The F and R strategies won't be used in this case, so the conditions are:

$$\begin{cases} E_j(C) = \phi \cdot (-G_A - C_C) + (1-\phi) \cdot (G_C - C_C) > 0 \\ E_i(\phi) = \phi \cdot (G_A - C_A) + (1-\phi) \cdot (-C_C) \geq 0 \end{cases}$$

Hence, $\phi \in [\frac{C_C}{G_A - C_A + C_C}, \frac{G_C - C_C}{G_A + G_C}) \neq \emptyset$, this equilibrium exists. The intuitive explanation for this equilibrium is: when the cost for A and C is small enough and the gain is high, malicious nodes would like to afford small C costs to persuade the regular nodes to C most of the time, and only A occasionally. The game repeats infinitely and the malicious nodes won't flee. However, this equilibrium only exists when $\frac{C_C}{G_A - C_A + C_C} < \frac{G_C - C_C}{G_A + G_C}$. The PBE in this paper is more general and could be applied when this condition does not hold.

V. Countermeasures

The regular node needs to balance the possible loss for false alarm and gain in order to yield a correct report. It needs an evidence accumulation process to make a confident reporting decision. The malicious node clearly gains advantages by fleeing before the end of this process. Therefore, shorten the length of this process and make it less predictable becomes the networks' main countermeasure against malicious nodes.

A. Dynamic threshold

Regular nodes can use a dynamic threshold to mitigate malicious nodes' threats. However, regular nodes cannot define their threshold T arbitrarily since this would violate the sequential rationality requirement. A regular node will have a number of neighbors when it stays in one community. Through communicating with these nodes, it becomes more familiar

Algorithm 1 Belief dissemination

1: Each community elects a commonly-trusted node to be the CH;
2: **if** a node decides to move out of the community **then**
3: This node informs the CH about its destination;
4: **end if**;
5: CH selects some moving nodes it trusts to be its messengers;
6: Messengers carry the beliefs and blacklists, move and pass the belief information to CHs of their destination communities;
7: CHs of the destination communities broadcast the information;

with this community. The aforementioned L_F, which is the evaluated cost for the false alarm, decreases as it gains more confidence about its current community.

The decrease of L_F leads to the decrease of T. This indicates that a regular node tends to be more aggressive in reporting as it learns more about the community it stays in. Malicious nodes, on the other hand, move around more frequently then regular nodes because they flee when the risk is high. When the malicious node enters a new community, regular nodes in the community would have a low L_F already. The malicious node cannot cause much damage before being caught or fleeing again.

B. Inter-community belief dissemination

In the above game, the flee strategy leads to a reputation reset with a 100% success probability. However, if this probability can be reduced, malicious nodes are forced to be more conservative. Malicious nodes tend to flee earlier and the damage to the current community is reduced. If the node's identity binding cannot be changed in the MANET, and the belief is disseminated among communities as well, the above probability will be reduced.

To enforce identity binding in the MANET, a network-wide single PKI with a strict identity policy should be used. When a node applies for the public key certificate, the CA will verify the node with certain out-of-band mechanisms, assign a unique ID to the node, and make sure the node cannot obtain multiple public key certificates. Therefore, the identity binding cannot be changed during fleeing. Methods to thwart the sybil attack should also be employed to prevent faked identities.

The belief towards moving nodes can be disseminated following Algo. 1. CH denotes the cluster head. The black list records nodes that have been reported as malicious. In step 1 of Algo.1, CH could also move. However, it should select one successor before moving. In step 2, if a node fails to inform the CH, it will be reported as malicious. So even a malicious node will inform the CH when it prepares to flee.

We skip some details in this scheme, such as building trust among communities. Let us assume community level trust does exist and the communities use the above steps to securely share the beliefs. When a malicious node flees into one of the communities that shares beliefs with its original community, it will face the same situation as that in its original community.

C. Intra-community belief sharing

In the game, regular nodes build their beliefs exclusively based on first-hand observations. This increases the detection time and makes their decisions more predictable. Since regular nodes also observe each other and buildup trust, they can utilize this trust and share their belief. With this mechanism, which is usually called second-hand information integration, each attack of malicious nodes leads to multiple regular nodes' θ increase in a community. Hence, the number of attacks that a malicious node can conduct before being reported is reduced.

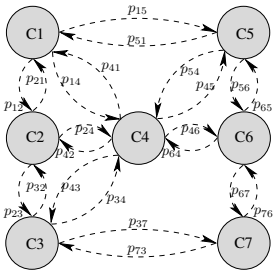


Fig. 3. Community-based mobility pattern for nodes in Fig. 1(a).

VI. Simulation

We conduct the simulation to evaluate the regular and malicious nodes' pure, mixed, and PBE strategies.

A. Simulation setup

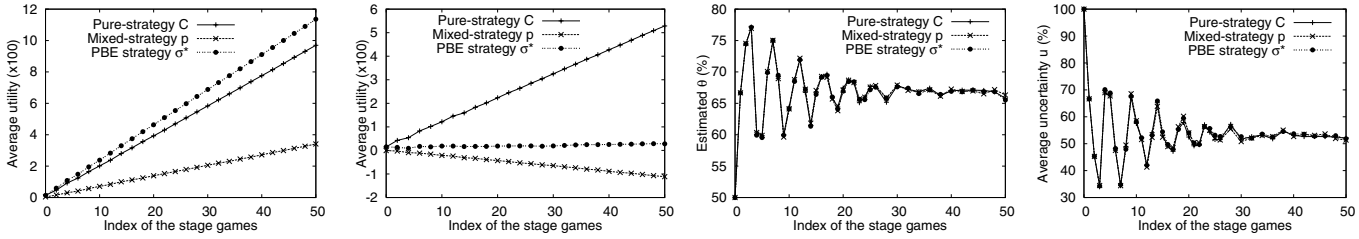
All proposed strategies have been implemented on a custom simulator ds (see http://sourceforge.net/projects/wrss/ for details). All simulations are conducted in randomly generated MANETs. We assume an ideal situation where the regular node can track its neighbor's outgoing packets.

100 nodes are randomly placed in a $900m \times 900m$ region which is evenly divided into 9 communities. The transmission range is 250m. Any two nodes within the same community are considered neighbors. Nodes follow the community-based mobility model [4]. Fig. 3 illustrates this mobility model for nodes in Fig. 1(a). The p_{xy} in Fig. 3 is the probability that regular nodes in community Cx will move to community Cy.

Each simulation is repeated 500 times, and the average data is used as the final result. The default number of malicious nodes is 40. The amount of energy for C_C is regarded as the unit cost/gain. We select the drop-packet attack as the sample attack in the simulation. The default value for the expected gain and cost parameters are $G_A = 20$, $G_C = 30$, and $G_R = 80$. L_F complies to normal distribution with $E(L_F) = 100$ and $VAR(L_F) = 20$. The utility in the following figures shows the actual average payoff of nodes.

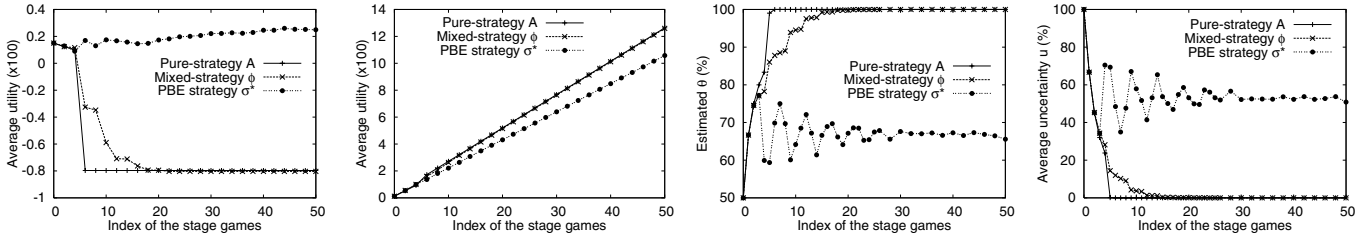
B. Simulation results

In Figs. 4(a) to 4(d), malicious nodes always follow their PBE strategy. We record the results of different stage games to compare regular nodes' different strategy profiles. In Figs. 4(a) and 4(b), regular nodes' PBE strategy outperforms the other two strategies. From Fig. 4(a) we can see that when regular nodes follow pure strategy C, their utility is high. This is due to the fact that regular nodes hold all the opportunities



(a) Regular nodes' utility (b) Malicious nodes' utility (c) Belief (d) Uncertainty

Fig. 4. Regular nodes' strategies comparison when malicious nodes follow their PBE strategy.



(a) Malicious nodes' utility (b) Regular nodes' utility (c) Belief (d) Uncertainty

Fig. 5. Malicious nodes' strategies comparison when regular nodes follow their PBE strategy.

to cooperate with other regular nodes. However, it will surely stimulate the malicious nodes to attack. As shown in Fig. 4(b), the utility of malicious nodes is the highest in this case.

Regular nodes can choose the mixed strategy σ_j: $\{p = \frac{C_A - C_C}{G_A}\}$ which makes $E_i(A) = E_i(C)$ for malicious nodes. This method greatly reduces malicious nodes' payoff. The corresponding curve in Fig. 4(b) shows that the utility for malicious nodes is negative. However, this mixed strategy is too conservative. While greatly reducing malicious nodes' utility, regular nodes' average utility is the lowest in Fig. 4(a).

Fig. 4(c) illustrates the convergence process of the estimated θ. As the estimated θ is mainly decided by the malicious nodes' strategy, the curves for three cases are very close to each other. θ intensely vibrates in the earlier stages, and converges in the later stages. Malicious nodes' periodic fleeing causes the vibration. When a malicious node attacks continuously at one location, θ goes up quickly. After it flees to a new destination, it attacks again with a clean history. As the malicious nodes' strategy selection becomes more diverse in later stages, the regular nodes' belief converges.

In Figs. 5(a) to 5(d), regular nodes always follow their PBE strategy. We compare malicious nodes' different strategies and the PBE strategy outperforms the others. In Fig. 5(a) and Fig. 5(b), when malicious nodes exploit pure strategy A or mixed strategy ϕ, they can only affect regular nodes' utility in the first several stages and their utility drops dramatically.

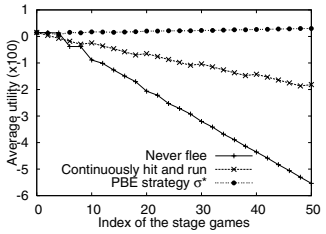
Fig. 5(c) illustrates the connection between the malicious nodes' strategy and the variation of the regular nodes' belief. When the malicious node follows pure strategy A, the estimated θ should converge to 100% in the first few stages. When the malicious node exploits the mixed strategy ϕ, it is more deceptive. We can see that the curve for the mixed strategy has a ladder shape. However, the malicious node will still be reported. When applying the PBE strategy, the malicious node has a good chance of escaping being reported by fleeing.

Figs. 4(d) and 5(d) show the uncertainty u in the above two cases. The tendency of the curves are just opposite to those in Figs. 4(c) and 5(c). This proves that uncertainty is the determinant element of regular nodes' decision.

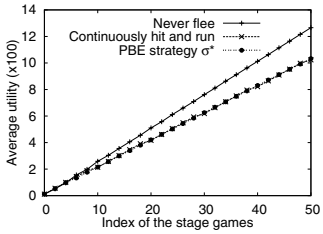
In Figs. 6(a) and 6(b), we compare different methods of fleeing. The first is to never flee. Malicious nodes could only select $\phi < \frac{G_C - C_C}{G_C + G_A}$ or increase ϕ but bear the loss of being caught. The latter case is shown in Fig. 5. For the former case, the average utility of regular nodes is the highest in Fig. 6(b) as malicious nodes can only choose a low ϕ to attack. Besides, the average utility of the malicious node is the lowest in this case. The average utility of regular nodes are similar when malicious nodes follow the PBE strategy or continuously hit and run. However, the malicious nodes' utility is much higher in the PBE strategy case. Therefore, following the PBE strategy outperforms other flee options.

Figs. 7(a) and 7(b) demonstrate the effectiveness of the proposed countermeasures. Using the inter-community belief dissemination or the dynamic threshold method (combined with intra-community belief sharing), the utility of the malicious node is reduced. However, both methods rely on the regular nodes' mobility model and organization.

Simulation results can be summarized as follows: 1) The PBE strategies for both parties are better than other pure or mixed strategies; 2) Regular nodes' decision rules, which consider the evidence sufficiency, balance the possible gains from cooperation with regular nodes and the threats from malicious nodes; 3) The flee strategy is one key point for the malicious nodes. It greatly increases the malicious nodes' utility.

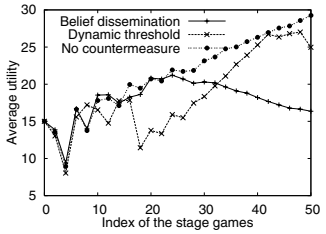


(a) Malicious nodes' utility

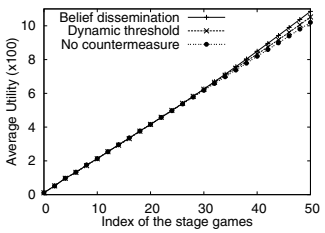


(b) Regular nodes' utility

Fig. 6. Flee strategies comparison.



(a) Malicious nodes' utility



(b) Regular nodes' utility

Fig. 7. Countermeasures comparison.

VII. Related Work

The incentives for nodes to cooperate are analyzed and presented in [5], [6], [7]. However, in these works, malicious nodes are modeled as never cooperative, without any further sophistication, since their main focus was discouraging selfish nodes. There is no degree of selfishness that can approximate the behavior of the malicious nodes.

Some recent works study the incentives for malicious nodes and model their behavior more rationally. In [8], Liu et al present a general incentive-based method to model the attackers' intents, objectives, and strategies. In [9], Theodorakopoulos et al further study the payoff of the malicious nodes and identify the influence of the network topology. However, the good nodes' behavior in [9] is simple and it fails to consider the possibility that an attacker might choose different attack frequencies towards different opponents.

Game theory [2] is a powerful tool in modeling interactions among self-interested nodes and predicting their choice of strategies. Therefore, it is widely employed to study wireless networks [10], [11], [12]. The equilibria of the contention window game is studied in [12]. The results of the analysis show that selfishness does not always lead to network collapse and may help the network to operate at an efficient Nash equilibrium. In [10], a mixed strategy equilibrium is studied to counter the jamming attack. A Bayesian game is studied in [11] to save energy in distributed intrusion detection systems.

We use a monitoring and reputation system as the basic setting for regular nodes. Many related works also use reputation systems [3], [13], [14] and a game theory model to analyze the problem. Srinivasan et al [15] analyze a modified Tit-for-Tat, where each node compares its own frequency to the aggregate frequency of cooperation of the network. Altman et al [16] propose a scheme for punishing users whose frequency of cooperation is below the one dictated by the Nash equilibrium.

VIII. Conclusion

In this paper, we use a dynamic Bayesian game framework to analyze the wrestling between regular and malicious nodes in mobile networks. The regular node forms belief and measures uncertainty to evaluate the type of its opponent. It chooses the probability to cooperate with its opponent based on its belief, and follows a decision rule to report which balances between the loss for false alarm and the gain for accurate reporting. The malicious node keeps evaluating the risk of being caught and exploits its flee strategy to avoid punishment. We analyze the perfect Bayesian equilibrium in this game and emphasize the advantages that malicious nodes would gain from the flee strategy. Our future work will focus on multi-user scenarios for the regular/malicious node game and analyze the topology's influence on equilibrium.

Acknowledgments

This work was supported in part by NSF grants CCR 0329741, CNS 0422762, CNS 0434533, CNS 0531410, and CNS 0626240. Email: fli4@fau.edu, jie@cse.fau.edu.

References

[1] S. Buchegger and J. Boudec. A robust reputation system for p2p and mobile ad-hoc networks. In *Proc. of the Second Workshop on the Economics of Peer-to-Peer Systems*, 2004.

[2] D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, Cambridge, Massachusetts, 1991.

[3] F. Li and J. Wu. Mobility reduces uncertainty in MANETs. In *Proc. of IEEE INFOCOM*, 2007.

[4] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy. Modeling time-variant user mobility in wireless mobile networks. In *Proc. of IEEE INFOCOM*, 2007.

[5] A. Blanc, Y. Liu, and A. Vahdat. Designing incentives for peer-to peer routing. In *Proc. of IEEE INFOCOM*, 2005.

[6] L. Buttyan and J. Hubaux. Stimulating cooperation in self-organizing mobile ad hoc networks. *ACM Mobile Networks and Applications*, 8(5), 2003.

[7] M. Felegyhazi, J. Hubaux, and L. Buttyan. Nash equilibria of packet forwarding strategies in wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 5(5), 2006.

[8] P. Liu, W. Zang, and M. Yu. Incentive-based modeling and inference of attacker intent, objectives and strategies. *ACM Transactions on Information and Systems Security*, 56(3), 2005.

[9] G. Theodorakopoulos and J. Baras. Malicious users in unstructured networks. In *Proc. of IEEE INFOCOM*, 2007.

[10] X. Liu, G. Noubir, R. Sundaram, and S. Tan. Spread: Foiling smart jammers using multi-layer agility. In *Proc. of IEEE INFOCOM (Mini-Symposiums)*, 2007.

[11] Y. Liu, C. Comaniciu, and H. Man. A bayesian game approach for intrusion detection in wireless ad hoc networks. In *Proc. of ACM GameNets*, 2006.

[12] L. Chen and J. Leneutre. Selfishness, not always a nightmare: Modeling selfish mac behaviors in wireless mobile ad hoc networks. In *Proc. of IEEE ICDCS*, 2007.

[13] A. Josang, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644, 2007.

[14] F. Li, A. Srinivasan, M. Lu, and J. Wu. Uncertainty mitigation for utility-oriented routing in MANETs. In *Proc. of IEEE GLOBECOM*, 2007.

[15] V. Srinivasan, P. Nuggehalli, C. Chiasseriniand, and R. Rao. Cooperation in wireless ad hoc networks. In *Proc. of IEEE INFOCOM*, 2003.

[16] E. Altman, A. Kherani, P. Michiardi, and R. Molva. Non-cooperative forwarding in ad hoc networks. *INRIA, Tech. Rep. RR-5116*, 2004.