Modeling Energy Constrained Routing in Selfish Ad Hoc Networks

Petteri Nurmi
Helsinki Institute for Information Technology HIIT
Basic Research Unit P.O. Box 68, FI-00014, Finland
petteri.nurmi@cs.helsinki.fi

ABSTRACT

In static networks, game theory has long been used to model the routing decisions of network nodes. However, once we move to dynamic and resource constrained settings, such as ad hoc or sensor networks, traditional models are no longer sufficient. Instead, new models that capture the dynamic nature of the decisions and the resource constraints of the devices are needed. To date, several models that attempt to capture the dynamic nature of routing decisions have been proposed. However, the resource constraints of the devices and the uncertainty about the resources of other devices have been thus far ignored. To this end, we present a game theoretic model that formalizes the resources of the nodes and the beliefs the nodes have about the resources of other nodes. We also discuss the structure of strategies in the proposed model and make explicit the role the resources and beliefs of the nodes play in routing decisions. In addition to presenting a game theoretic model, we propose a method that allows the nodes to learn equilibrium strategies over time, and prove that the strategies suggested by the mechanism converge to a sequential equilibrium. Finally, we present simulations that give insights into the expected behavior of the devices under the proposed model.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*Data communications*; C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless communication*; I.6 [**Computing Methodologies**]: Simulation and Modeling

General Terms

Theory, Experimentation

Keywords

Game theory, ad hoc networks, sequential equilibrium, modeling, simulation

1. INTRODUCTION

In wireless ad hoc networks a collection of nodes, e.g., mobile hosts, forms a self-organizing network without any support from a pre-established infrastructure [9, 30]. The lack of infrastructure support forces the nodes to implement all networking tasks by themselves. In the context of routing, this means that packets must be routed using other nodes as intermediate relays. Simple routing schemes, such as the Dynamic Source Routing (DSR) protocol [14, 15], are sufficient only if all nodes are willing to participate in the forwarding. However, nodes are energy constrained and want to maximize their lifetime, which leads to potential selfishness as the nodes may refuse to forward packets for other nodes. Simulation studies have shown that when standard routing protocols are used and even a small fraction of the nodes act selfishly, the throughput of the network degrades significantly [22].

A routing decision is essentially a conflict situation that can be modeled using game theory [7, 34]. In static networks, game theory has long been used to model the routing decisions of network nodes. However, once we move to dynamic and resource constrained settings, such as ad hoc or sensor networks, traditional models are no longer sufficient. Instead, new models that capture the dynamic nature of the decisions and the resource constraints of the devices are needed. To date, several models that attempt to capture the dynamic nature of routing decisions have been proposed. However, the resource constraints of the devices and the uncertainty about the resources of other devices have been thus far ignored. To this end, we propose a game-theoretic model, *the energy constrained routing game*, that formalizes the resources of the nodes and the beliefs the nodes have about the resources of other nodes. We also discuss the structure of strategies in the proposed model and make explicit the role the resources and beliefs play in node decision making.

As our second contribution, we propose a learning rule that the agents can utilize to learn equilibrium strategies over time and prove that the resulting rule ensures convergence to a sequential equilibrium. Our proof is mainly based on results in [10]. As our third contribution, we present simulations that give insights about the expected behavior of the devices under the proposed model.

The rest of the paper is organized as follows: Section 2 introduces our routing model with the help of an example scenario. The energy constrained routing game is discussed in Section 3 and in Section 4 we discuss the strategies of the nodes and present a learning rule that the agents can use to learn equilibrium strategies. Section 5 proves that the model has at least one equilibrium point and that the strategies suggested by the learning rule converge to an equilibrium point. Section 6 presents experimental results and Section 7 introduces related work. Finally, Section 8 concludes the paper.

2. EXAMPLE SCENARIO

In this section we consider an example scenario that details what happens in the network nodes during the lifetime of a single set of packets. To begin with, we assume an arbitrary node i, called the *source*, has generated some packets that it wants to send to another node j. In order to send the packets, the source needs to know what routes can be used to reach the destination. To this end, the first phase of the routing process is to discover possible routes to the destination. We assume that a source routing protocol is used and in the following we focus on what happens once the source has information about possible routes. Contrary to the DSR protocol, we assume that whenever possible the protocol returns information about multiple paths to the destination.

Once the source has information about available routes, the next step is to decide which route to use. According to rationality arguments, the source wants to maximize the number of packets that arrive at the correct destination during its lifetime. Since the nodes are energy constrained, the path that is used for routing should be such that the source believes that the packets have a high probability of reaching the destination. In addition, the source should decide how many packets to send to the network based on its current beliefs and on its available resources. The packets that are not sent, as well as the packets that are dropped during the routing, remain in the buffer of the source and they are processed at a later stage. To this end, if the source is low on energy, it could send packets to the network only if it believes that the packets are likely to arrive at the correct destination. On the other hand, if the source has more energy, it may choose to "experiment" and send more packets even though it believes that the packets might not reach the correct destination.

We next assume that the source has sent some packets to the network and that the first intermediate node k_1 on the routing path has just received the packets and stored them in a buffer. When the node has resources to spare, it processes the packets by deciding whether to forward the packets or to discard them. When node k_1 makes its decision, it considers whether it still has resources to spare and how useful it is to spend resources on node i. In the former, the decision can be made purely based on available resources. However, for the latter, the node k_1 needs to look at its past experiences with the source node i and estimate how likely it is that the source will forward packets for it in the future.

As the resources of the intermediate nodes affect routing decisions, the source should have beliefs about their resources. It is clearly infeasible to assume that the source could know exactly the resources that the other nodes have available to them; especially since the other nodes do not have any incentives to reveal this information. However, based on, e.g., own observations and on prior knowledge, the source node can make crude estimates about the resources that other nodes have available to them. These estimates, however crude they may be, can then be used to refine the path estimates and to favor paths where nodes are both cooperative and have resources to spend. In terms of the entire network, this kind of a scheme enforces more balanced use of available resources.

We now assume that node k_1 has forwarded some of the packets and that the second intermediate node k_2 is currently processing those packets. Similarly as node k_1, the node k_2 decides whether to forward packets or not and then either sends some packets to the next node or discards the remaining packets. This process is continued until some packets arrive at the destination or all packets are dropped. After packets arrive at the destination, or are dropped at some intermediate node, the nodes that are on the route should update their beliefs, as well as other parameters of interest. As a consequence, nodes need to be able to observe at least partially the forwarding decisions of the intermediate nodes. To this end, we assume that each node is equipped with a neighborhood watch mechanism, such as the Watchdog [21]. We also assume that, whenever the destination receives some packets, it sends the source an acknowledgment message using the same path that was used to route the packets. Here we assume that the size of the acknowledgment messages is much smaller than the size of normal packets. This assumption ensures that the effect of the overhead caused by the acknowledgments is negligible. Finally, to cope with discarded packets, we require that packet timers are used so that, if a node does not acknowledge packets within a certain time period, they are considered to be dropped.

To conclude the example, a summary of the phases in a routing process is given in the list below.

1. Source node i generates a set of packets and performs the route discovery phase.
2. The source node evaluates available paths by estimating the probability that packets arrive at the correct destination.
3. After the path has been selected, the source decides based on the score of the path and its own resources how many packets, if any, to send to the network.
4. An intermediate node receives the packets and decides how many packets to send to the next node based on its own energy level and past experiences with the source node i.
5. While some packets are remaining and the current node is not the destination node, Steps 3 and 4 are repeated.
6. The destination node receives the packets, or a subset of them, and sends an acknowledgment message to the source node i.
7. The source node and the intermediate nodes receive the acknowledgment or packet timers go off and they update their beliefs based on this information.

3. THE ENERGY CONSTRAINED ROUTING GAME

In this section we formalize the interactions between the nodes of the network and define the game they are playing. The overall game model we consider is a repeated Bayesian game [7, 11, 12, 13] where each repetition is a stage. The stages also are games, more precisely Bayesian games in extensive form [7, 18]. The model is defined so that a new stage begins whenever some node generates packets to the network. In the rest of the section we focus on a single stage and fix an arbitrary node i to be the source of traffic. For ease of presentation, we assume that the destination of packets is the same during the stage.

Let $\langle N, E\rangle$ be a connected graph where N is the set of nodes. Each node $i \in N$ is assumed to have a discrete representation for time so that a new time step for node i begins when the node generates new traffic to the network. The time steps of node i are denoted t_k^i. As we are considering only a single game, we drop the indices i from all variables. We define a *packet* to be the smallest unit of communication and $g(t_k)$ to be the number of packets node i generates at time t_k. If there are packets from earlier time frames remaining, we assume they are included in the value of $g(t_k)$. By definition, the variable $g(t_k)$ is an integer value that is greater than zero, i.e., $g(t_k) \in \mathbb{Z}_+$. The number of packets node i actually sends to the network is denoted $s(t_k)$. The number of packets generated clearly serves as an upper bound for the number of packets sent, i.e., the relation $0 \leq s(t_k) \leq g(t_k)$ holds.

We assume that each node has a discrete representation for its remaining energy. We call the discretized energy the *energy class* of a node and use the variable θ_i to denote it. The set of all possible energy values is denoted Θ. For notational simplicity, we assume that the discretization mechanism is global so that the set Θ is the same for all nodes in the network. The energy class of a node is only known by the node itself and thus the energy class corresponds to the concept of type in Bayesian games. However, nodes are assumed to have beliefs about the energy classes of the other nodes and the beliefs node i holds about the energy class of some forwarding node j are denoted $\mu_{t_k}^j$. We assume that the beliefs are independent over the nodes, i.e., $\mu_{t_k}^j$ and $\mu_{t_k}^l$ are independent for all $j, l \in N\backslash\{i\}, j \neq l$. If the beliefs are not independent, we can map the game to a strategically equivalent game with independent beliefs [7]. We also assume that the beliefs assign a strictly positive value to all possible energy classes, i.e., for all $\theta' \in \Theta$, $\mu_{t_k}^j(\theta') > 0$. The easiest way to motivate this assumption is to assume that nodes do not know the true energy classes of other nodes beforehand, in which case they should assign strictly positive beliefs over all elements in the set Θ. Furthermore, since in reality the lifetime of a node is finite, the posterior probabilities derived from the prior beliefs are guaranteed to remain positive at all times. We also assume that the energy class of the source does not affect the decisions of the forwarders (see Section 8). Finally, we assume that forwarding nodes do no have beliefs about the energy classes of other forwarding nodes. As a consequence of the last two assumptions, in the formulation of the extensive game the forwarding nodes do not have beliefs[1].

We now assume that the route discovery phase of the routing protocol has been performed and that $J \subseteq N\backslash\{i\}$ is the collection of intermediate nodes j through which packets can be routed to the destination. Under the mentioned assumptions, the (active) set of players in the tk:th stage game is

[1]Note that we consider here a single game. All nodes have beliefs about the energy classes of other nodes, but the beliefs are used only when the node acts as the source of traffic.

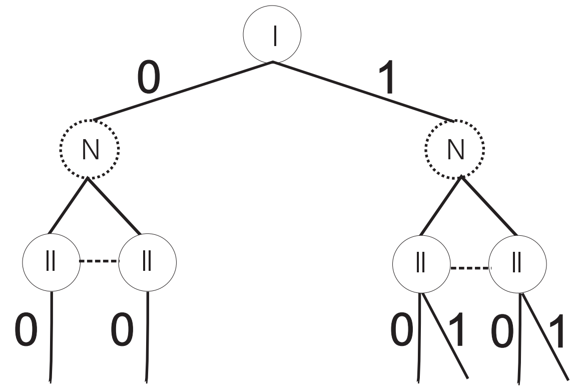

Figure 1: An example game tree, which has a single intermediate node (II) with two possible types, and where the source (I) sends either one or zero packets to the network. The nodes labeled with N represent moves by nature.

$\{i\} \cup J$. Respectively, the type space of the stage game is the product space $\Theta = \theta_i \times \Theta^{|J|}$, where Θ_i is a singleton set (identity) and Θ_j is the set of possible energy classes for node j. Since we assumed that the discretization is global, the sets Θ_j are the same for all intermediate nodes[2]. The belief system of the stage game is $\mathcal{B} := \{\mu_{t_k}^j\}_{j\in J}$. The action space of source i is the set $A(t_k) = \{s(t_k)|s(t_k) \leq g(t_k)\}$, i.e., the possible amounts of packets node i can send to node j at time step t_k. Let $f_j(t_k)$ denote the number of packets node j forwards for node i. The action space of a forwarding node j is the set $A_j(t_k) = \{f_j(t_k)|0 \leq f_j(t_k) \leq g(t_k)\}$. Note that at the beginning of the stage the intermediate nodes do not know how many packets node i sends to the network and thus the action space of node j consists of all alternatives between zero and $g(t_k)$. However, the actions of node i clearly restrict the actions of node j within the game tree. As an example of this kind of situation, Figure 1 contains a game tree with a single intermediate node j that has two possible types and where node i has sent a single packet to the network.

From Figure 1, we can deduce that a forwarder has $|A(t_k)|$ information sets[3]. The size of each information set is $s(t_k)$, i.e., the number of packets on the route to that particular information set. The probabilities for the different vertices of the forwarder are determined by a common objective distribution [13] and thus are known to the forwarder. The information sets of the source are singleton sets and thus the probabilities assigned over the vertices of the source are trivial (one everywhere).

In order to finalize the definition of a stage game, we still need to define the utility functions of the nodes. The utility functions are discussed later in more detail and, at this point, we only assume that the utility functions of the players are functions from $\mathbb{Z}$ to $\mathbb{R}$. The variables u_i and u_j are used to denote the utility functions of the source and a for-

[2]If we assume the forwarding nodes have beliefs about the energy class of the source, the set Θ_i equals Θ_j.

[3]As the intermediate nodes do not have beliefs, the only thing that matters to node j is its own energy class and the action of node i. Thus the game tree resembles that of Figure 1 everywhere along the routing path.

warder. Using these definitions, we define the t_k:th stage game of node i formally as follows:

DEFINITION 1. *The t_k:th stage game of source i is the 8-tuple $\langle \{i\} \cup J, \mathcal{T}, u, \mathcal{A}, \mathcal{I}, p, \Theta, \mathcal{B} \rangle$, where u is the vector function $u = (u_i, u_{j_1}, \cdots, u_{j_n})$, $\mathcal{A}$ is the collection $\{A_i(t_k^i)\} \times \prod_{j \in J} \{A_j(t_k^i)\}$, Θ is the type space of the game, $\mathcal{B}$ is the belief system of the game and $\mathcal{T}$ is the game tree, which is constructed in a similar fashion as in Figure 1. The information space $\mathcal{I}$ is derived from the game tree in a trivial fashion and the probabilities for the vertices within information sets are given by a common objective distribution p, which is possibly unknown to the nodes.*

4. STRATEGIES AND OVERALL GAME

In this section we first detail the strategies of the source and the intermediate nodes, after which we finalize the definition of the energy constrained routing game by fixing what takes place between individual iterations.

4.1 Behavior of the intermediate nodes

In order to define optimality in the stage games, we must define the utility functions of the nodes more precisely. We start by considering the intermediate nodes. We fix j to be an arbitrary intermediate node and assume θ_j to be its current energy class. Let $\gamma(\theta_j)$ be the probability that indicates the willingness to forward packets given the energy class θ_j. We assume that the function $\gamma(\cdot)$ maintains ordering of the energy classes, i.e., if $\theta_{j1} < \theta_{j2}$, the willingness $\gamma(\theta_{j1})$ is smaller than $\gamma(\theta_{j2})$. Intuitively this simply means that the more energy the node has, the more likely it is that the node is willing to forward packets for other nodes.

Next we assume that a cooperation mechanism that gives probabilistic estimates about the cooperation of another node is employed. Such mechanisms are presented, e.g., in [4, 5, 16, 23, 32]. Let $\hat{\gamma}_{j,i}$ denote node j's estimate about the willingness of node i to cooperate as given by the employed cooperation mechanism. The probability $\sigma_{j,i}$ with which node j forwards packets for node i is assumed to be a combination of the energy dependent probabilities and the probabilities given by the cooperation mechanism, i.e.,

$$\sigma_{j,i} = \alpha_j \hat{\gamma}_{j,i} + \beta_j \gamma(\theta_j), \quad (1)$$

where α_j and β_j are importance parameters that indicate how important energy and cooperation rate are for a particular node. We require that the weights sum up to one, i.e., $\alpha_j + \beta_j = 1$, in which case the Equation 1 can be seen as an instance of Bayesian model averaging. Thus, given the action s_i of node i, on average the node j forwards $\sigma_{j,i} s_i$ packets, where the probability $\sigma_{j,i}$ changes as a function of available energy and actions of node i.

4.2 Behavior of the source node

The source node is assumed to have theories that guide its decisions on how many packets to send to the network. These theories are about the types of the other players and about their behavior strategies. We define $\phi_j(t_k)$ to be the theory of the source node about the quantities of interest of a forwarding node j, i.e.,

$$\phi_j(t_k) = \{\hat{\sigma}_{i,j}, \hat{\theta}_j\}, \quad (2)$$

where $\hat{\sigma}_{i,j}$ is node i's theory about the probability distribution of $\sigma_{j,i}$ (behavior strategy of node j) and $\hat{\theta}_j$ is node i's theory about the [probability distribution of] energy classes (type) of node j.

Intuitively, the actions of the source node should balance throughput of own messages and energy usage. To this end, let $\mathcal{P}$ be a collection of routing paths from the source node i to the destination. In a similar fashion as in the Watchdog approach [21], we assume that the source node evaluates the different paths in $\mathcal{P}$ so that each path is calculated a probability with which the packets are routed to the destination. The path with the highest score is usually select for routing, but to guarantee that all alternatives are properly explored, we assume a suboptimal path (path with lower score) is selected with a small probability ϵ_k that depends on the current time and that converges to zero as a function of time. In case a suboptimal path is selected, we require that the probability distribution that is used to select the alternative path has full support, i.e., all (available) suboptimal paths have a strictly positive probability of being selected, conditioned on the event that the node will experiment in its next decision. More details about the sequence ϵ_k are given in the next section.

The probabilities assigned to the different paths in $\mathcal{P}$ are assumed to utilize the theories so that the assigned probability is a combination of the observed behavior history of a node and the estimated energy class. Thus, the probability of a routing path $P \in \mathcal{P}$ is

$$\mu(P) = \prod_{j \in P} \left[\alpha_i \hat{\gamma}_{i,j} + \beta_i \gamma(\hat{\theta}_j) \right], \quad (3)$$

where α_i and β_i are importance parameters of node i (as with the intermediate nodes), $\hat{\gamma}_{i,j}$ and $\hat{\theta}_j$ are the current theory and $\gamma(\cdot)$ is a function that assigns a probability to an energy class. Thus the evaluation scheme of the source node is assumed to be similar as the decision mechanism of a forwarding node (see Equation 1).

After the source node has decided which path to use for the packets, it has still to decide how many packets it should send to the network. To this end, we define the utility of an action $s(t_k)$ and a path P as

$$u(s(t_k), P) = \mu(P) \left[h(s(t_k)) - h_\gamma(s(t_k)) \right], \quad (4)$$

where $h(\cdot)$ is a function that tell how much the packets are worth to the source node, whereas $h_\gamma(\cdot)$ tells how much the energy needed to forward the packets is worth to the source node. We require the functions $h(\cdot)$ and $h_\gamma(\cdot)$ to be of von Neumann-Morgenstern type, i.e., the differences between utilities (or score values) reflect the agents sensitivity to take, or respectively to avoid, risks, and the magnitudes of the values are induced by some underlying preference relation. In accordance with the adaptive learning framework, we assume that node i selects the action $\hat{s}(t_k)$ that maximizes Equation 4 with a probability $1 - \epsilon_t$, where ϵ_t is a sequence of small errors that decreases as a function of time, i.e., $\lim_{t \to \infty} \epsilon_t = 0$. In our experiments we use the sequence $\epsilon_k = 1/k$, but also other sequences are possible (see next section).

4.3 The overall game

From the definition of the overall game, we are still lacking the definition of how and when the updates are carried out. In the single hop case, it suffices that the source node is equipped with a watchdog-like mechanism and with a packet timer mechanism. Namely, these assumption allow the source to observe the actions of the forwarders and thus the source can update its theories when it observes that node j has forwarded packets (positive update) or when a packet counter goes out (negative update). Thus we can define the *routing game* between a source node i and the network $N\backslash\{i\}$ to be a repeated game where each stage game is as in Definition 1 and where the theories are as discussed above. When the routing path contains more intermediate nodes, additional technical solutions are needed for gathering feedback information. For our purposes it suffices to assume that some mechanism exists with which the actions of the individual nodes can be tracked, but for concreteness, in the following we present one such scheme.

4.4 The binary propagation scheme

At the end of a stage, the source needs to update its beliefs about all the intermediate nodes. The way this can be implemented is to consider each packet separately in which case the actions are binary – either forward all or forward none. At this point we need to make some assumptions about the underlying protocol and require that the protocol implements packet numbering, timers and acknowledgments. Thus each packet can be uniquely identified by the combination of the source address and the packet number. Moreover, each node is assumed to maintain a timer for the packets it sends or forwards. The timer mechanism should work so that once the packet is sent, the node starts the timer. Once an acknowledgment message is received with the same packet number and source address, the timer is stopped. Because the nodes are assumed to be rational, the mechanism works if the size of the acknowledgment messages is much smaller than the size of packets. If a node has already forwarded a packet for the source node, it is reasonable to assume that, because it already has spent energy for forwarding, it wants to maintain a good status in the eyes of the source. To achieve this, the source needs to know that the node actually forwarded (some) packets, which is achieved via the acknowledge scheme just discussed. A problem with this scheme is that nodes may be able to manipulate the packet header in which case the information is not reliable anymore. However, the encryption schemes that are needed to avoid this problem are out of scope for the paper.

The theoretical model that we introduced assumed that each new time step considers a single set of packets instead of a single packet. Thus we still need to modify the binary propagation scheme discussed above. They way to modify the scheme is to consider a continuous probability distribution defined over the interval $[0, 1]$ and where the value $x \in [0, 1]$ is the fraction of messages that were forwarded by some intermediate node. By adding a time frame number to each packet, we can separate between packets belonging to a particular time frame. Assuming that at time step t_k node i sent $s_i(t_k)$ packets, the unbiased estimator of the forwarding rate of the first relay node is $\sum_b f_j(b)/s_i(t_k)$, where each $f_j(b)$ is a binary indicator that tells whether packet b was forwarded by node j or not. Now the source node can update the probability distributions of the different relay nodes by considering the value of the estimator as the new evidence in the calculation of the posterior beliefs. The exact calculations depend on the used probability distribution. In our case we use a multinomial distribution and information on how the updates can be carried out can be found from any standard statistics book (e.g. [3]).

5. EQUILIBRIUM AND OPTIMALITY PROOFS

In this section we first prove that the game-theoretic model proposed in Sections 3 – 4 has at least one sequential equilibrium point. After this, we prove that the strategies of the nodes, as presented in Section 4, converge to a sequential equilibrium point. We start by stating an equilibrium result, which is important for our proof.

THEOREM 1. *Every perfect equilibrium point is a sequential equilibrium*

PROOF OF THEOREM 1. *[17, p.882].* ■

We now show that the game-theoretic model admits at least one perfect (Bayesian) equilibrium, after which we can use Theorem 1 to conclude that there exists at least one sequential equilibrium for the model. To show that the game admits a perfect Bayesian equilibrium, we must show that it satisfies certain conditions, which are given in **B1** – **B4** below.

- **B1** For each information set, the players must have beliefs about the node the game play has reached.
- **B2** Whenever it is a player's turn to move, her actions must be optimal from that point onward given her beliefs.
- **B3** The player's beliefs about reachable (on-the-path) nodes, must be determined using the Bayes' rule.
- **B4** The player's beliefs about unreachable (off-the-path) nodes must be determined using the Bayes' rule, whenever possible (whenever probabilities are positive).

LEMMA 1. *The model satisfies the postulates* **B1** – **B4***:*

PROOF OF LEMMA 1. *The condition* **B1** *is trivially satisfied as all information sets are singleton sets and we can assign probability one to each node. The condition* **B2** *follows directly from the definition of the game. We assumed that the support of the action space of an arbitrary node equals the common action space which implies that there are no off-the-path nodes and thus condition* **B4** *is satisfied. The last condition,* **B3***, follows directly from the way the belief system is constructed.* ■

Next we give the necessary conditions that a learning process must satisfy for it to converge to a sequential equilibrium point. These conditions are given in Conditions 1 – 4 below.

CONDITION 1. *Every information set is reached infinitely many times. More formally, let h be some information set and $\mathbf{a}^t(h)$ the set of information sets that are reached at repetition t. We require that for each information set h_0 the set $\{n \in N | h_0 \in \mathbf{a}^n(h)\}$ is infinite.*

CONDITION 2. *Suboptimal play vanishes in the long run. More precisely, let h_0 be an information set which is reached infinitely many times. The empirical frequency of reaching* *a suboptimal information set h_1 from information set h_0 approaches zero as the number of iterations grows, i.e.,*

$$\frac{\#\{n \in N | h_0, h_1 \in \mathbf{a}^n(h)\}}{\#\{n \in N | h_0 \in \mathbf{a}^n(h)\}} \rightarrow 0, \text{ when } n \rightarrow \infty \quad (5)$$

CONDITION 3. *If the empirical frequencies converge to some point, the theory of each player j must converge to that point, i.e., if the frequencies of the actions over the information sets σ converge to σ^* and if the frequencies of the probabilities defined of the information sets μ converge to μ^*, the theory (σ^t, μ^t) converges to (σ^*, μ^*).*

CONDITION 4. *The frequencies of natures moves converge to the objective distribution p.*

Groes et al. prove the following Theorem, which is important for our proof:

THEOREM 2. *A learning process satisfying conditions 1 - 4 converges to a sequential equilibrium.*

PROOF OF THEOREM 2. *[10, p.136].* ■

We now prove in Lemmas 2 – 4 that the adaptive learning process described in the previous section satisfies the Conditions 1 – 4, after which we can use Theorem 2 to conclude that the strategies converge to a sequential equilibrium point.

LEMMA 2. *The learning sequence $\epsilon_k = \frac{1}{k}$, satisfies Condition 1, i.e., all information sets are visited infinitely many times.*

PROOF OF LEMMA 2. *First of all, since we assume beliefs to be uncorrelated and strictly positive over all information sets, there is a strictly positive probability for any possible action sequence. Secondly, with the learning sequence, a suboptimal action is selected with probability $\frac{1}{k}$. We define a random variable $\chi_I(k)$ such that $\chi_I(k) = 1$, if information set I is visited on the k^{th} iteration and $\chi_I(k) = 0$ otherwise. The iterations for which $\chi_I(k) = 1$ form a harmonic series, and, as a consequence, $\sum \epsilon_k \chi_I(k) = \infty$ since ϵ_k has a divergent subsequence* ■.

An alternative way to prove Lemma 2 is to resort to stochastic approximation theory, according to which, Condition 1 is satisfied, if $\sum_k \epsilon_k = \infty$ and $\sum_k \epsilon_k^2 < \infty$ [20]. In our case, these conditions are clearly satisfied. The learning sequence $\epsilon = 1/k$ is by no means the only possible choice as the learning sequence and especially in the field of reinforcement learning alternative sequences, such as the Bolzmann rule, have been considered. Discussion about different learning sequences is, however, out of scope for the paper and we refer to [31] for more information.

LEMMA 3. *The model satisfies Condition 2, i.e., suboptimal play vanishes in the long run.*

PROOF OF LEMMA 3. *As $\lim \epsilon_k = 0$, suboptimal play clearly vanishes in the long run* ■.

LEMMA 4. *The proposed model satisfies Conditions 3 and 4, i.e., if the empirical frequencies converge at some point, the theory of each player converges to that point and the frequencies of natures moves converge to the objective distribution.*

PROOF OF LEMMA 4. *These are immediate consequences of using the Bayes rule as the Bayesian updates ensure asymptotic converge to the true distribution* ■.

THEOREM 3. *The proposed model has at least one sequential equilibrium point and the learning sequence converges to a sequential equilibrium. If there is only one sequential equilibrium, the sequence converges to that point. Otherwise the learning sequence converges to some sequential equilibrium point.*

PROOF OF THEOREM 3. *Existence of a sequential equilibrium follows from Theorems 1 and 2 and the fact that a sequential equilibrium is reached follows from Lemmas 1 - 4 and Theorem 3* ■.

6. EXPERIMENTS

In order to give insights into node performance under our model, we have conducted a set of simulation experiments. The goal of the experiments is to show how an energy constrained network might function and to highlight what kind of effects we can expect to see in the network. In the following subsections we first detail our simulation setup, after which we present results from the simulations.

6.1 Simulation setup

In our simulations, each experiment consisted of 1000 iterations (seconds). The number of nodes was fixed (100) and the topology of the network was randomly generated. The value of the importance parameter β was varied between zero and one in steps of 0.1. In a single experiment, all nodes used the same value for β and the value of α was set to $1.0 - \beta$. For each value of β we repeated the experiment 30 times.

The nodes were deployed into a 125×125 m^2 area and the range of the nodes' transmitters was set to 25 meters. In the initial deployment we required that the minimum distance between two nodes is 10 meters. This way the network topologies were initially fairly uniform.

The movements of the nodes were modeled using a random waypoint model with constant speed (4 m/s) and Poisson distributed waiting times. The mean of the waiting times was set to 10 seconds (= iterations). A Poisson distribution was used also for traffic generation. The mean of the traffic generating distribution was set to five and hence the nodes would generate about 500 packets per iteration.

Energy consumption was modeled using a squared loss function, according to which the energy C that is needed to send a packet is given by

$$C = \lambda d^2, \quad (6)$$

where d is the distance over which the packets are sent and λ is a scaling parameter (see Table 1). The coefficient and the energy of the nodes was set so that first nodes would start to die out approximately at around 100 iterations.

An important factor in simulating our model is to specify the utility function of the nodes. First of all, we set the utility function of the source to be

$$u(s(t_k), P) = \mu(P) \left[h(s(t_k)) - h_\gamma(s(t_k)) \right] \quad (7)$$
$$= \mu(P) \left(c \cdot s(t_k) - \exp\left(s(t_k) \cdot C\right) \right), \quad (8)$$

where c is a predefined constant and $\mu(P)$ is the probability of the path as given by Equation 3. As the function $\gamma(\theta_j)$

of the forwarding nodes we use the exponential function

$$\gamma(\theta_j) = \exp\left(\frac{-(1.0 - \theta_j)}{\nu}\right), \quad (9)$$

where ν is a constant and θ_j is the fraction of battery that is remaining. We do not claim that our choice of utility functions is realistic, but we merely wanted to give insights into node performance when the utility functions are more complex and non-linear. A summary of the simulation parameters is given in Table 1.

Parameter	**Value**
α	1.0 - β
β	0.0 to 1.0
c	12.5
λ	0.0008525
ν	1.15
$\gamma(\theta_j)$	$\exp(-(1.0 - \theta_j)/\nu)$
$h(s(t_k))$	$c \cdot s(t_k)$
$h_\gamma(s(t_k))$	$\exp(s(t_k) \cdot C)$
Area	$125 \times 125\ m^2$
Energy consumption	$C = \lambda d^2$
Initial energy	$\sim$ Uniform$(8000, 9000)$
Iterations	1000
Movement	random waypoint
Number of nodes	100
Range	25 m
Repeats	30
Speed	constant (4 m/s)
Traffic generation	$\sim$ Poisson(5)
Waiting times	$\sim$ Poisson(10)

Table 1: Summary of the used simulation parameters.

6.2 Results

In this section we present the results of the simulations. All statistics are averaged over the 30 repetitions. While calculating the different statistics, we have taken into account only those nodes that still participate in the network. Thus, this means that only nodes that have energy left and that are reachable at some point of time are considered in the calculations.

First of all, we measured the average success rate of the nodes, which measures the ratio of packets arriving at the correct destination to the overall number of packets sent. The success rate in our simulations is shown in Fig. 2 for the varying parameter values of β. As the results indicate, when $\beta = 0.0$, i.e., when the nodes do not consider their energy level at all, all packets arrive at the correct destination. However, as we will later see, this has the drawback that nodes run faster out of resources and the resulting resource usage of the network is highly unbalanced. On the other hand, once the nodes start to consider also their energy level, the success starts to decrease exponentially, but the resource usage is better balanced.

The second aspect that we measured was the number of nodes alive as a function of iterations. As Figure 3 indicates, this was fairly similar over the different parameter choices. Moreover, as the figure indicates, most simulations ended with around 80 alive nodes from the overall 100 nodes. An exception from this is the value $\beta = 0.0$, i.e., when the nodes did not consider energy at all. Namely, in this case nodes would start to die out sooner and actually more nodes died out, leaving us with only around 70 nodes alive.

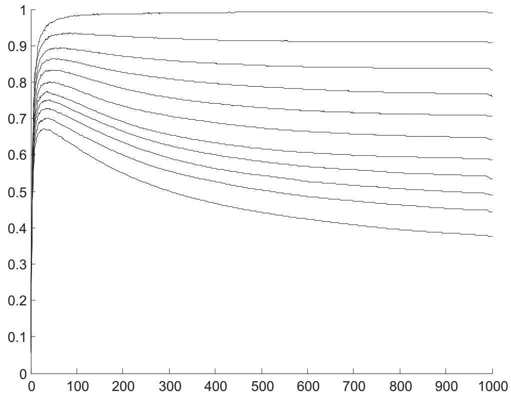

Figure 2: The success rate of nodes as a function of time when the value of the parameter β is varied between 0.0 and 1.0. The topmost plot corresponds to the case $\beta = 0.0$ and the consequent plots (from top to bottom) correspond to increases in the value of β by 0.1.

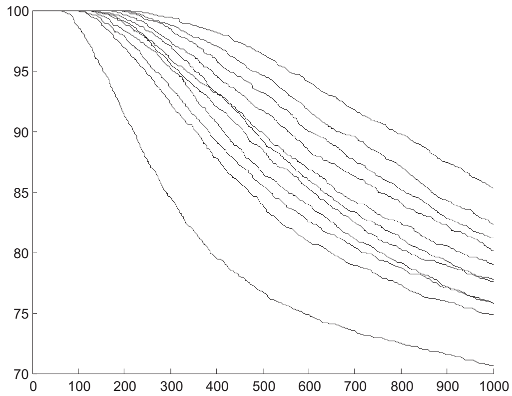

Figure 3: The amount of nodes that are alive as a function of iterations. The topmost plot corresponds to the case $\beta = 1.0$ and the consequent plots (from top to bottom) correspond to decreases in the value of β by 0.1.

Finally, we also measured the number of packets sent and number of packets forwarded during a single iteration. The results are given in Figures 4 and 5. As was expected, both the number of sent packets and the number of forwarded packets decreases significantly over time. What is interesting to see is that when the nodes consider also their energy, the decrease in sent and forwarded packets is nearly linear whereas in the case when the nodes do not consider their energy ($\beta = 0.0$) the decrease is exponential. This is due to the fact that there are less nodes available in the network and thus less routes for use. On the other hand, the closer the value of β is to one, the better balanced the number of sent (and forwarded) packets is.

As a summary, our simulations indicate that when nodes consider also energy in their decisions, we can expect the network performance to be initially somewhat lower. How-

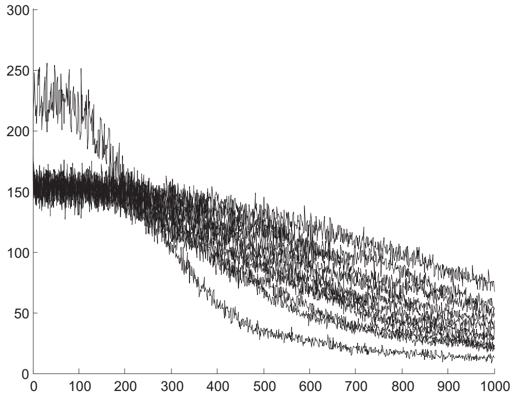

Figure 4: The amount of packets send by the nodes during a single iteration. The plot with the steepest descent corresponds to the case $\beta = 0.0$ and the consequent plots (in order of steepness of the plot) correspond to increases in the value of β by 0.1.

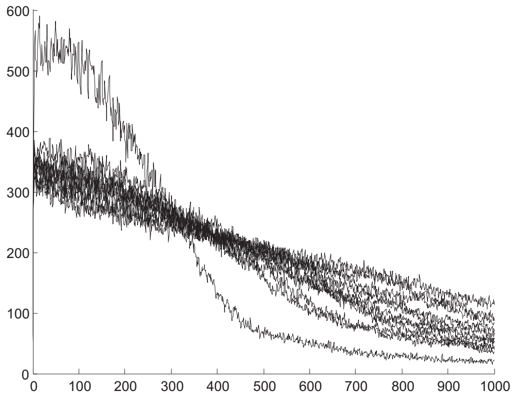

Figure 5: The amount of packets forwarded by the nodes during a single iteration. The plot with the steepest descent corresponds to the case $\beta = 0.0$ and the consequent plots (in order of steepness of the plot) correspond to increases in the value of β by 0.1.

ever, when the resources are considered in the routing decisions, the resources are spent in a more balanced way and the network will remain operational for a longer period of time.

7. RELATED WORK

Although the idea to model routing in ad hoc networks using game theory is relatively new, other types of communication networks have been analyzed with tools from game theory. Especially, game theoretic analysis of Internet congestion control has been an active research area (see, e.g., [8, 24, 26, 28]). However, due to node mobility and resource constraints, ad hoc networks have more inherent sources for uncertainty and, as a consequence, models designed specifically for ad hoc networks are needed. In this section we introduce earlier work that has applied either evolutionary game theory [34, 35] or static Bayesian games, i.e., Bayesian games with a non-dynamic structure.

Chronologically the first approach is by Srinivasan, Nuggehalli, Chiasserini and Rao who model the routing situation using the repeated prisoners' dilemma [29]. In their model, a node has two possible modes of behavior when it makes a routing decision: to cooperate or to defect. According to evolutionary game theory, an effective strategy in this kind of setting is the so-called *TIT-FOR-TAT* strategy [2, 27], in which the player cooperates on the first move and thereafter imitates the other player [36]. The authors argue that this kind of behavior is too severe for ad hoc networks and propose a modification, called the *generous TIT-FOR-TAT* (GTFT), in which the behavior of the nodes is stochastic so that node i forwards packets for node j with a probability μ_i^j. The probabilities are estimated over time using

$$\hat{\mu}_i^j = \frac{C_i^j}{D_i^j}, \tag{10}$$

where C_i^j is the number of packets node j has forwarded for node i and D_i^j is the total number of messages node i has sent to node j to be forwarded. The authors prove that with GTFT the total throughput rate of the network is *Pareto-optimal*, i.e., all resources of the network are utilized optimally. Our model differs from this model on three aspects: we consider exploration of actions, we model explicitly the resources of the nodes and we model the uncertainty the nodes have about the resources of other nodes.

The next approach that we consider is by Altman, Kherani, Michiardi and Molva [1]. Also in their model each node j is assumed to forward packets with some probability μ_j. The probabilities are assumed to be independent of the source. When a packet is sent to the network, each node computes the current equilibrium strategy and uses the probability indicated by the equilibrium strategy for forwarding. However, without a punishment or a reward mechanism, the equilibrium strategy is $\mu_j = 0$ for each node j. To solve this, the authors propose a punishment mechanism where other nodes decrease their forwarding probabilities, if someone deviates from the equilibrium strategy. The authors prove that, under this framework, the probability indicated by the equilibrium strategy of the nodes increases as the number of intermediate nodes increases. Furthermore, the authors give an algorithm for calculating the equilibrium strategies (in this setting) in a distributed manner.

The main problem in this approach is that the routing model is not generic. First of all, it is unlikely that the forwarding probabilities of the nodes are independent of the source – at least in the long run. Secondly, the complexity of calculating the equilibrium strategy is not feasible in terms of storage and computational effort. Finally, also this approach ignores the exploration of actions, the modeling of resources and the uncertainty about the resources of other devices.

Closest to our work is the work of Urpi, Bonucelli and Giordano [33] who use static Bayesian games and infinitely repeated games to model the routing situation. In their model, the nodes are divided into classes so that nodes in the same class generate traffic according to the same stochastic process. Once assigned, the energy class of a node stays fixed for the entire lifetime of the node[4]. A node can only have

[4]In the traditional approach to Bayesian games, Nature draws the type of each player from a fixed probability distribution before each repetition / stage game and thus the types are not fixed. However, the model can be seen to be-

information about nodes that belong to its neighborhood Γ_i and nodes that belong to Γ_i are assumed to have beliefs about the class of node i and vice versa. In the model, time is assumed to be discrete and during each time step a Bayesian game is played. In each instance of the game a node has to decide how many of its own packets to send to the network and how many packets to forward for a neighboring node. Because the overall game is a repeated game, the utility of a player depends on the discount factor, which is used to weight future utilities. In the model each node has a discount factor γ_j for each node j in Γ_i, which is interpreted as the probability that node j belongs to the neighboring of node i in the next time step.

The model has the same problems as the others we have discussed. That is, the uncertainty about resources of others is ignored. Furthermore, the model utilizes only local information and thus no guarantees can be given that packets are actually routed to the destination. The model has also a game theoretic flaw. Namely, the players' information increases over time and thus the beliefs of the players should be updated. However, in the proposed model the beliefs stay the same.

8. CONCLUSIONS

In this paper we have presented a novel game theoretic model for modeling energy constrained routing in ad hoc networks and we have discussed the strategies of the participants. Furthermore, we have introduced an adaptive learning framework that allows the participants to learn the optimal strategies over time and we have proven that the mechanism ensures convergence to the true strategies. A minor flaw in our model currently is that the forwarding nodes do not have beliefs about the energy class of the source node and in the future we plan to extend the model to take also this aspect into account.

9. ACKNOWLEDGMENTS

The author wishes to thank Jukka Suomela and Patrik Floréen for reading and commenting earlier versions of the paper. This work was supported in part by the IST Programme of the European Community, under the PASCAL network of excellence, IST-2002-506778. The publication only reflects the authors' views.

10. REFERENCES

[1] E. Altman, A. A. Kherani, P. Michiardi, and R. Molva. Non-cooperative forwarding in ad-hoc networks. In *Proceedings of the 4th International IFIP-TC6 Networking Conference*, volume 3462 of *Lecture Notes in Computer Science (LNCS)*, pages 486 – 498. Springer-Verlag, 2005.

[2] R. Axelrod. *The Evolution of Cooperation.* Basic Books, 1984.

[3] J. M. Bernardo and A. F. M. Smith. *Bayesian Theory.* Wiley, 2000.

[4] S. Buchegger and J.-Y. L. Boudec. A robust reputation system for mobile ad hoc networks. Technical Report IC/2003/50, EPFL-IC, CH-1015 Lausanne, July 2003.

long to a more general class of games, stochastic Bayesian games, where the probability distribution is allowed to change over time [6].

[5] S. Buchegger and J.-Y. L. Boudec. A robust reputation system for P2P and mobile ad-hoc networks. In *Proceedings of the 2nd Workshop on the Economics of Peer-to-Peer Systems*, 2004.

[6] E. Dekel, D. Fudenberg, and D. K. Levine. Learning to play Bayesian games. *Games and Economic Behaviour*, 46:282 – 303, 2004.

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

[8] R. Garg, A. Kamra, and V. Khurana. A game-theoretic approach towards congestion control in communication networks. *ACM SIGCOMM Computer Communications Review*, 32(3):47–61, 2002.

[9] S. Giordano. Mobile ad hoc networks. In *[30]*, chapter 15, pages 325 – 346. 2002.

[10] E. Groes, H. J. Jacobsen, and B. Sloth. Adaptive learning in extensive form games and sequential equilibrium. *Economic Theory*, 13:125 – 142, 1999.

[11] J. C. Harsanyi. Games with incomplete information played by Bayesian players i-iii. part i: The basic model. *Management Science*, 14(3):159–182, November 1967.

[12] J. C. Harsanyi. Games with incomplete information played by Bayesian players i-iii. part ii: Bayesian equilibrium points. *Management Science*, 14(5):320 – 334, January 1968.

[13] J. C. Harsanyi. Games with incomplete information played by Bayesian players i-iii. part iii: The basic probability distribution of the game. *Management Science*, 14(7):486 – 502, March 1968.

[14] D. B. Johnson, D. A. Maltz, and J. Broch. DSR: The dynamic source routing protocol for multi-hop wireless ad hoc networks. In *[25]*, chapter 5, pages 139 – 172. 2001.

[15] D. B. Johnson, D. A. Maltz, and Y.-C. Hu. The dynamic source routing protocol. IETF-MANET Draft, July 2004.

[16] A. Josang and R. Ismail. The Beta reputation system. In *Proceedings of the 15th Bled Conference on Electronic Commerce*, 2002.

[17] D. M. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, 50(4):863 – 894, July 1982.

[18] H. W. Kuhn. Extensive games and the problem of information. In *[19]*, pages 46 – 68. 1953.

[19] H. W. Kuhn, editor. *Classics in game theory.* Princeton University Press, Princeton, New Jersey, 1997.

[20] H. J. Kushner and G. G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications.* Springer-Verlag, 2nd ed. edition, 2003.

[21] S. Marti, T. J. Giuli, K. Lai, and M. Baker. Mitigating routing misbehavior in mobile ad hoc networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 255 – 265, August 2000.

[22] P. Michiardi and R. Molva. Simulation-based analysis of security exposures in mobile ad hoc networks. In *European Wireless Conference*, 2002.

[23] P. Nurmi. Bayesian game theory in practice: A framework for online reputation systems. In *Proceedings of the International Conference on*

Internet and Web Services. IEEE Computer Society, 2006.

[24] C. Papadimitriou. Algorithms, games, and the internet. In *Proceedings of the 33rd annual ACM symposium on Theory of computing (STOC)*, pages 749–753. ACM Press, 2001.

[25] C. E. Perkins, editor. *Ad Hoc Networking*. Addison-Wesley, New York, 2001.

[26] T. Roughgarden. *Selfish routing*. PhD thesis, Cornell University, May 2002.

[27] R. Selten. The chain-store paradox. *Theory and Decision*, 9:127 – 159, 1978.

[28] S. Shenker. Making greed work in networks: a game-theoretic analysis of switch service disciplines. *ACM/IEEE Transactions on Networking*, 3(6):819–831, December 1995.

[29] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao. Cooperation in wireless ad hoc networks. In *Proceedings of the 22nd INFOCOM*, volume 2, pages 808 – 817. IEEE, March 2003.

[30] I. Stojmenović, editor. *Handbook of Wireless Networks and Mobile Computing*. John Wiley & Sons, New York, 2002.

[31] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[32] W. L. Teacy, J. Patel, N. R. Jennings, and M. Luck. Coping with inaccurate reputation sources: experimental analysis of a probabilistic trust model. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 997 – 1004, 2005.

[33] A. Urpi, M. Bonuccelli, and S. Giordano. Modelling cooperation in mobile ad hoc networks: a formal description of selfishness. In *Proceedings of Modelling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, March 2003.

[34] F. Vega-Redondo. *Economics and the Theory of Games*. Cambridge University Press, 2003.

[35] J. W. Weibull. *Evolutionary Game Theory*. MIT Press, 1997.

[36] E. W. Weisstein. Tit-for-tat. From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/Tit-for-Tat.html`, April 2005.