ABSTRACT

Ad hoc networks are self-configuring networks of mobile nodes, connected by wireless links. If a destination node is beyond the transmission range of an origin node, then the nodes must cooperate to provide a multi-hop route. Any node can act as a sender, receiver or transit node. It is in a node’s interest to be a sender or receiver, but it is less clear what the value is of forwarding traffic on behalf of other nodes. The nodes should therefore be given incentives to act as transit nodes, because otherwise the network would fail to function. A way to do so is by introducing for each node a credit balance, where nodes use credits to pay for the costs of sending their own traffic, and earn credits by forwarding traffic from other nodes.

The incentive scheme requires that the information needed to compute the credit balance at each node be locally available at the node. In this paper we present a model of a signalling protocol that obtains and distributes the prices that the nodes charge for processing flows so that the credit balances can be locally computed at each node. Simulation experiments show that the protocol can efficiently gather and distribute the control information such that effective flow allocation takes place for reasonable values of the signalling rate and signal processing delays.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Wireless communication

General Terms
Performance

Keywords
Congestion prices, distributed control, incentives for collaboration, mobile ad hoc networks, scalability, self-organization

Modelling Incentives and Protocols for Collaboration in Mobile Ad Hoc Networks

Johannes Göbel
Department of Informatics
University of Hamburg
22527 Hamburg, Germany
goebel@informatik.uni-hamburg.de

Anthony E. Krzesinski
Department of Mathematical Sciences
University of Stellenbosch
7600 Stellenbosch, South Africa
aek1@cs.sun.ac.za

1. INTRODUCTION

Ad hoc networks are self-configuring networks of mobile nodes, connected by wireless links. They enable infrastructure-free communication: no fixed equipment is needed, but instead each node acts as a router. When information needs to be transmitted across the network, it is sent from the sender to the receiver by relaying the information along intermediate nodes. An excellent survey on ad hoc networks, with special emphasis on Quality-of-Service aspects, is given in [7].

Any node in an ad hoc network can act as a sender, receiver or transit node. It is in a node’s interest to be a sender or receiver, but it is less clear what the value is of forwarding traffic on behalf of other nodes. However, if nodes would not be willing to act as transit nodes, the ad hoc network would fail to function. Therefore a crucial question is: how can nodes be given incentives to act as transit nodes?

Crowcroft et al. [2] rely on earlier work on pricing and rate control [4, 5] to devise such an incentive scheme. In their framework, each node has a credit balance that determines how much the node can spend on transmission resources in the next time interval. For each node there are two resources: bandwidth and power, each with its own price. The price of the resource increases when the resource is scarce, and decreases when the resource is abundant. When a call arrives, given the current prices, the node selects the least cost route to send its traffic on. At the same time the node earns credit when acting as a source, destination or transit node. The interplay between the prices, flow allocations, and credit balances is such that global stability of the system is achieved. Importantly, such a scheme is decentralized: no central controller is needed, and the scheme therefore has favourable scalability properties.

The results presented in [2, 3] assume that the congestion prices and the route costs are instantaneously available at all nodes. This is unrealistic: we therefore present a model of a signalling protocol to gather the route costs and make them locally available at the originating nodes where the route costs are needed to adjust the flow rates.

The remainder of this paper is organized as follows. Section 2 presents a summary of the mobile ad hoc network model. Section 3 describes a signalling protocol to gather and distribute the information required for the distributed operation of the credit-based incentive scheme. Section 4 describes the simulation model and presents the results of several simulation experiments. Conclusions are given in Section 5.
2. A SUMMARY OF THE MODEL

Consider a set of mobile nodes located on a 2-dimensional surface. An origin node transmits to a destination node. If the destination node is beyond the transmission range of the origin node, then the other nodes must cooperate to provide a multi-hop route. This raises the question as to why an intermediate node on the route would expend bandwidth and power in forwarding transit traffic without being compensated.

The setup of [2] can be summarized as follows. There are two scarce resources: bandwidth and power. For each node there is a limit on the bandwidth and power that can be consumed. When a node acts as a source, transit or destination node, it obtains compensation in the form of credits for the congestion costs of the bandwidth and power resources consumed. A node uses its credits to pay for the bandwidth and power congestion costs incurred when it sends its own traffic. New calls are connected on the least cost routes. The bandwidth and power congestion prices are periodically updated, and are meant to reflect the level of congestion at any specific node along any specific route.

2.1 Bandwidth and power congestion prices

The power congestion price $\mu_j^p(t)$ expressed in credits per unit power (credits/W) satisfies the differential equation (DE)

$$\frac{d}{dt} \mu_j^p(t) = \frac{\kappa \mu_j^p(t)}{\Gamma_j} (\gamma_j(t) - \Gamma_j)$$

with initial value $\mu_j^p(0) = 1$ where $\kappa$ is a constant of dimension seconds$^{-1}$, $\gamma_j(t)$ is the power in use at node $j$ at time $t$ and $\Gamma_j$ is the power available at node $j$. Likewise, the bandwidth congestion price $\mu_j^B(t)$ expressed in credits per unit flow (credits/Mbit) satisfies the DE

$$\frac{d}{dt} \mu_j^B(t) = \frac{\kappa \mu_j^B(t)}{C_j} (c_j(t) - C_j)$$

with initial value $\mu_j^B(0) = 1$ where $c_j(t)$ is the bandwidth in use at node $j$ at time $t$ and $C_j$ is the transmission capacity (bandwidth) of node $j$. The DE for the bandwidth congestion price can be approximately evaluated through

$$\mu_j^B(t + \Delta) \approx \mu_j^B(t) \left(1 - \kappa \Delta \left(1 - \frac{c_j(t)}{C_j} \right) \right)$$

for some suitable value of $\Delta$ seconds, with a similar expression for $\mu_j^p(t + \Delta)$. The congestion prices are thus constant when the resource is fully utilized, and increase/decrease when the resource is over-/under-utilized. The congestion prices are adjusted every $\Delta$ seconds: the smaller the value of $\Delta$, the quicker the resources at each node become utilized as fully as possible.

2.2 The route prices

Let $\epsilon_{x}^w$ denote the energy used per unit flow per unit time (power used per unit flow) when receiving. Let $\epsilon_{x}^{ij} = \max(10^{-2}, 10^{-4}|z_i - z_j|^2)$ denote the energy used per unit flow per unit time (power used per unit flow) when transmitting from node $i$ to node $j$ where $z_j$ denotes the $(x,y)$ location of node $j$. Note that $\epsilon_{x}^{ij}$ is a non-zero function in the vicinity of the transmitter. If node $i$ cannot reach node $j$ then $\epsilon_{x}^{ij} = \infty$.

Let $f_r(i)$ denote the node that node $i$ forwards traffic to on route $r$. The price $\mu_j^r(t)$ expressed in credits/Mbit that node $j$ charges for forwarding a unit of flow to the next node $k = f_r(j)$ along route $r$ is

$$\mu_j^r(t) = \begin{cases} 
\epsilon_{x}^{ij} \mu_j^p(t) + \mu_j^B(t) \\
(\epsilon_{x}^{s} + \epsilon_{x}^{ik}) \mu_j^p(t) + 2\mu_j^B(t) \\
\epsilon_{x}^{ij} \mu_j^p(t) + \mu_j^B(t) 
\end{cases}$$

if $j$ is a source, transit or destination node on route $r$ respectively.

2.3 The credit balance

Let $\mathcal{R}_s^r(t), \mathcal{R}_t^r(t)$ and $\mathcal{R}_d^r(t)$ denote the set of routes that originate, transit and terminate at node $s$ at time $t$ respectively. Each node $s$ maintains a credit balance $b_s(t)$ with an initial value equal to the target balance $\phi$. The credit balance of node $s$ is adjusted as follows.

First, the node spends credits at rate $\sum_{r \in \mathcal{R}_s^r(t)} y_s(t)\mu_s^r(t)$ for the congestion costs (bandwidth, power) incurred in transmitting its own traffic through the source, destination and transit nodes on its outbound routes, where $\mu_s^r(t) = \sum_{j \in r} \mu_j^r(t)$ is the price that node $o(r)$ charges for forwarding a unit of flow end-to-end along route $r$ and $y_s(t)$ is the flow sent along route $r$ at time $t$.

Second, the node receives credits at rate $\sum_{r \in \mathcal{R}_t^r(t)} y_t(t)\mu_t^r(t)$ for the congestion costs (bandwidth, power) incurred in acting as a source node, transit node or a receiver node.

Third, that part of the credit balance that differs from the target balance $\phi$ is discounted using a factor $\beta$. Thus if $\beta = 0.01$, then over one time unit, the under-provisioned nodes that possess a credit balance of less than $\phi$ receive 1% of the credits that they lack and the over-provisioned nodes that possess a credit balance larger than $\phi$ surrender 1% of that part of their balance that exceeds $\phi$. The credit balance $b_s(t)$ satisfies the DE

$$\frac{db_s(t)}{dt} = -\beta(b_s(t) - \phi) + \Omega_s(t)$$

where the reimbursement

$$\Omega_s(t) = \sum_{r \in \mathcal{R}_s^r(t)} y_s(t)\mu_s^r(t) - \sum_{r \in \mathcal{R}_t^r(t)} y_t(t)\mu_t^r(t).$$

This DE can be approximately evaluated through

$$b_s(t + \Delta) \approx b_s(t) - \beta \Delta(b_s(t) - \phi) + \Delta \Omega_s(t).$$

A node at the edge of the network will attract little transit traffic and earn few credits for transporting transit calls. The flow rates for traffic originating from such nodes will be low. A node at the centre of the network will attract transit traffic and earn credits for transporting transit calls. The flow rates for traffic originating from such nodes will be larger.

2.4 The willingness-to-pay

Each node $s$ determines its resource usage according to its willingness-to-pay $w_s(t)$ at time $t$ for the congestion costs incurred in sending its traffic which is given by $w_s(t) = \alpha_s b_s(t)$ where $\alpha_s > 0$.

2.5 Flow allocation

Each source node $s$ can originate one call at a time to a random destination $d$. Consider a call in service on the least
that is at least equal to the time between successive price
binding contract which remains in force for a period of time
node to process a unit of flow for a particular call is thus a
are equal to the payments received. The price quoted by a
essarily match the credits that the transit/destination nodes
prices. In this case the credits paid by a source will not nec-
instanta-
neously known at all nodes. If the pricing information is
nodes charge for forwarding a unit of flow are instanta-
flight rate
expressions to the DE’s which determine the congestion prices
power and bandwidth restrictions are enforced indirectly since a constraint violation will cause the prices of power and of bandwidth to increase which in turn will lower the flow rate $y_t(t)$.

3. A SIGNALLING PROTOCOL

The fluid-level simulation models [2, 3] that were used to explore the model presented in Section 2 update the prices and flows sufficiently frequently so that the bandwidth and power constraints are violated infrequently, and by small amounts. However, in reality the prices and flows cannot be updated arbitrarily frequently. Furthermore, the information needed to update the flow on a route is to be found at the nodes along the route. This information must be sent to the originating node of the route, and the transfer of this information takes time and resources.

3.1 The credit balance

The computation of the reimbursement $\Omega(t)$ in the credit balance given by Eqn. (3) assumes that the prices that the nodes charge for forwarding a unit of flow are instantaneously known at all nodes. If the pricing information is not instantaneously available, then the nodes will base their cost calculations on different, slightly outdated congestion prices. In this case the credits paid by a source will not necessarily match the credits that the transit/destination nodes add to their own balances for relaying/receiving the call.

Payments should be conducted so that the payments made are equal to the payments received. The price quoted by a node to process a unit of flow for a particular call is thus a binding contract which remains in force for a period of time that is at least equal to the time between successive price updates for the call.

The discount term $\beta(\delta y_t(t) = \phi)$ in Eqn. (3) models a transfer of credit from nodes with a credit surplus to nodes with a credit deficit. This credit redistribution is decentralized and scalable. No mechanism is required to transfer the redistributed credit. A node that has a surplus of credits will, over a period of time $\Delta$, destroy a fraction $\beta$ of that surplus. Likewise a node that has an under-supply of credits will, over a period of time $\Delta$, create credits corresponding to a $\beta$ fraction of its deficit.

This paper does not address the issue of trust. Note that a node-specific discount rate $\beta$ will violate the model's property that the total credit balance is equal to the number of nodes in the network. If $\beta \neq \beta$ at one or more nodes then the total credit balance might not be equal to the number of nodes. For example, an increase in the total credit balance can be regarded as inflation: the increased amount of credits lowers the value of the credit.

3.2 The signalling protocol

The protocol uses two types of signal: an updateFlow signal and an updateCost signal. With reference to Fig. 1, a call $q$ connected on route $r = (j_0, j_1, \ldots, j_k)$ issues an initial updateFlow signal at node $j_0$ at time $t$ (say $t = 0$). The flow $y_t(t)$ is set to zero. The signal packet propagates from node to node along the route $r$: the resource prices and the resources in use are updated at each node along the route. The signal processing delay is $\epsilon$ seconds at each node. Note that in Fig. 1 the intervals $\Delta$ and $\epsilon$ are not drawn according to scale.

When the updateFlow signal arrives at node $j_n$ at time $t + n\epsilon$, an updateCost signal is issued at node $j_n$ on the reverse route $\bar{r} = (j_n, j_{n-1}, \ldots, j_0)$. The signal packet propagates from node to node along the route $\bar{r}$: the costs of processing a unit of flow at each node $i \in \bar{r}$ are computed and stored in the signal packet.

When the updateCost signal arrives at node $j_0$ at time $t + 2n\epsilon$ the credit balance at node $j_0$ is updated and an updateFlow signal is issued on the route $r$. This signal updates the flow $y_t(t + 2n\epsilon)$. The signal packet propagates from node to node along the route $r$: the resource prices and the resources in use, the credit balance at each transit node along the route and at the destination node $j_n$, are updated.

Additional updateCost, updateFlow signal pairs are sent at fixed intervals of time $\Delta$. A final updateFlow signal is sent when the call ends.

The updateFlow signal

Each call issues an updateFlow signal once when the call starts, once when the call ends and least once while the call is in progress, provided the call lasts longer than $2n\epsilon$.

The protocol assumes that the least cost routes are known. Consider a call $q$ that arrives to route $r = (j_0, j_1, \ldots, j_n)$. An updateFlow signal is scheduled at node $j_0$ and propagates node-by-node along the route. Consider an updateFlow signal for call $q$ at node $j \in r$ at time $t$. 

![Figure 1: The protocol timeline.](image-url)
If the signal starts or ends a call, then the route $r$ and the call $q$ are added to, or removed from, the appropriate route sets $R_j^r(t)$ where

$$ X = \begin{cases} S & j = j_0 \\ D & j = j_n \\ T & \text{otherwise.} \end{cases} $$

If the signal starts a call, then the signal is initialized to indicate that the node costs $\mu_r(t')$ and the route cost $\mu_r(t')$ stored in the signal are unavailable.

- The credit balance $b_j(t)$ is discounted by a factor $\beta$: $b_j(t) = b_j(t') - \beta \Delta (b_j(t') - \phi)$ where $\Delta = t - t'$.

- If $j \neq j_0$ then the cost $\mu_r(t')$ of processing a unit of flow from call $q$ at node $j$ on route $r$ is retrieved from the signal; the payment $y_r(t')\mu_r(t')\Delta$ to be made at node $j$ is computed and is added to the credit balance $b_j(t)$ at node $j$.

- If $j = j_0$ then the cost $\mu_r(t')$ of processing a unit of flow from call $q$ on route $r$ is retrieved from the signal; the flow $y_r(t)$ for call $q$ is updated according to $y_r(t) = \alpha b_j(t)/\mu_r(t')$. If the route cost $\mu_r(t')$ is unavailable, then the flow is set to zero.

- The congestion prices $\mu^B_j(t)$ and $\mu^R_j(t)$ are updated according to Eqn. (1).

- The resources in use $c_j(t)$ and $\gamma_j(t)$ are updated according to Eqs. (5) and (6).

- If $j \neq j_n$ then an updateFlow signal is scheduled after a short delay $\epsilon$ at the next node on route $r$.

- If $j = j_n$ and if this is the first updateFlow signal for call $q$, then an updateCost signal is immediately scheduled at the first node $j_n$ of the reverse route $\tilde{r} = (j_n, j_n-1, \ldots, j_0)$.

**The updateCost signal**

Each call issues at least one updateCost signal. Consider a call $q$ in progress on route $r = (j_0, j_1, \ldots, j_n)$. The first updateCost signal for call $q$ is scheduled at node $j_n$ when the first updateFlow signal for call $q$ arrives at node $j_n$. The updateCost signal propagates node-by-node along the reverse route $\tilde{r}$. Consider an updateCost signal for call $q$ at node $j \in \tilde{r}$ at time $t$.

- If the call has completed then the signal is dropped.

- If $j = j_n$, then the next updateCost signal is scheduled after a fixed delay $\Delta$ at node $j_n$.

- The cost $\mu_{r_j}(t)$ of processing a unit of flow for call $q$ at node $j$ along the forward route $r$ is computed from Eqn. (2) and stored in the signal.

- If $j \neq j_0$ an updateCost signal is scheduled after a short delay $\epsilon$ at the next node on $\tilde{r}$.

- If $j = j_0$: the signal contains the cost $\mu_{r_j}(t')$ of processing a unit of flow for call $q$ at each node $i \in r$; the payment to be made $y_r(t')\sum_{i=1}^m \mu_{r_j}(t')\Delta$ at the downstream nodes is computed and subtracted from the credit balance $b_{j_0}(t)$ at node $j_0$, and an updateFlow signal for call $q$ is scheduled at node $j_0$.

If $\Delta > 2\epsilon$ then the costs are fixed for a period of time sufficient to ensure that the credits paid by the source are equal to the credits transferred to the downstream nodes along the route. The signalling protocol ensures that the total credit balance $\sum_{t \in \mathcal{N}} b_j(t)$ is almost equal to the number of nodes in the network. In fact the total credit balance is slightly larger than the number of nodes since when a call terminates, some if not all of the downstream nodes on the route will have received their credits which have not yet been debited from the balance at the first node.

### 3.3 The simulation model

We developed a simulator in Java using the DESMO-J simulation framework [6]. The simulation model is parameterized as follows: each node originates one call at a time; each call sets up a connection to a randomly selected destination node such that the call is connected on the route with the lowest (at the instant of the call initiation) costs, and the route is used for the duration of the call so that calls are not rerouted, which implies low signalling costs and no route flap; the call holding times and the call idle times are exponentially distributed with mean $0.5 \text{s}$; the delay $\Delta$ between successive updateCost signals is $0.1 \text{s}$ unless stated otherwise; the simulation lasts for $10,000 \text{s}$; the signal processing delay per node $\epsilon$ is $1 \text{ms}$. The values of the remaining simulation parameters are presented at the end of the paper.

### 4. NUMERICAL RESULTS

#### 4.1 A 10-node network, 1 node moving

In this experiment we use the 10-node network model [2] illustrated in Fig. 2. Node 1 moves from its initial position at the lower right of the $100 \times 100 \text{m}$ plane through the centre of the network and finishes at position $(40, 100)$ when some 8,000 of the 10,000 calls have completed.

The results of the experiment are shown in Fig. 3 which closely reproduce the corresponding figures in [2]. Node 1 initially attracts no transit calls when it is located at the edge of the network: the congestion prices, credit balance and throughput at node 1 are correspondingly low. As node 1 moves through the centre of the network it attracts transit calls: the congestion prices, credit balance and throughput at node 1 increase. Note that node 1 becomes an alternative to nodes 0 and 8 as a transit node; the power price and the credit balance at nodes 0 and 8 therefore de-
Finally, as node 1 moves away from the centre of the network it becomes less attractive as a transit node: the prices, credit balance and throughput at node 1 decline.

4.2 A 100-node network, 10 nodes moving

In this experiment we consider 100 nodes located at random on a $300 \times 300$ m plane. The nodes are partitioned at random into two groups. One group consists of 10 nodes which move according to a Gauss-Markov motion model [1]. The remaining 90 nodes are fixed (immobile). The purpose of the experiment is to show that a low signalling rate is sufficient to ensure that the throughputs remain approximately constant as the 10 nodes move extensively around the plane.

Table 1 shows how the inter- and intra-group traffics vary as the signalling delay $\Delta$ is varied over two orders of magnitude. At a high signalling rate of 100 $\text{updateFlow}$ signals per second ($\Delta = 0.01$ s) the model reproduces the results obtained from the original model where the prices are globally known, frequently updated and instantaneously available. The results obtained for a signalling rate of 10 $\text{updateFlow}$ signals per second ($\Delta = 0.1$ s) closely resemble the results from the continuous model. A signalling rate of 1 $\text{updateFlow}$ signal per second ($\Delta = 1$ s) represents an almost minimal signalling traffic of 4.3 signals per call. At this low signalling rate the traffic of the 90-node group shows high variability.

Fig. 4 shows that the group throughputs remain approximately constant as the 10 nodes move around the plane.

4.3 A Manhattan grid, 100 nodes moving

In this experiment we consider a $12 \times 12$ Manhattan grid on a $320 \times 320$ m plane. The nodes are partitioned into two groups. One group consists of 15 fixed nodes that are placed at the vertices of a hexagonal lattice. The vertices are separated by some 100 m (approximately twice the radio transmission range) which provides near optimal coverage of

Table 1: 100-node Network: Throughput

<table>
<thead>
<tr>
<th>$\Delta$ (secs)</th>
<th>10-node group</th>
<th>90-node group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>intra</td>
<td>inter</td>
</tr>
<tr>
<td>0.01</td>
<td>1.6 ± 6.3</td>
<td>13.4 ± 19.6</td>
</tr>
<tr>
<td>0.1</td>
<td>1.6 ± 6.1</td>
<td>13.5 ± 20.6</td>
</tr>
<tr>
<td>1</td>
<td>1.6 ± 5.9</td>
<td>13.5 ± 23.8</td>
</tr>
</tbody>
</table>
the plane if we assume that each node can transmit/receive omni-directionally over a disk of radius 56 m. The fixed nodes, which are not necessarily located next to street intersections or streets, represent a fixed infrastructure (FI). These nodes are logically fully meshed by broadband links: in addition, each fixed node is connected to the outside world via a router. The (single) router to the outside world does not have an air interface. The fixed nodes do not originate calls; they function as access points for the mobile nodes.

The fixed nodes are connected by wireless links to a group of 100 mobile nodes. The fixed nodes provide an interface from wireless to wired communication, and vice versa. The mobile nodes represent wireless devices carried by pedestrians or vehicles. The mobile nodes can communicate with each other as well as with the FI, provided the communicating nodes are within transmission range. Multi-hop routing using transit nodes ensures that destinations beyond transmission range can be reached. Two FI configurations are investigated namely a basic FI and an enhanced FI which are both presented in Table 2.

Initially, each mobile node is located at a random street intersection. The mobile nodes move as follows. The node selects a destination intersection at random different from its current position. In general, movement in both the E-W and N-S directions will be required to reach the destination. The node selects an initial movement direction at random: either E-W or N-S. Suppose without loss of generality that the destination is N-E of the initial location of the node and that the node initially moves N. The node selects a speed sampled from a uniform distribution $U(0.6, 1.4) \text{ m/s}$. The node moves N until it reaches the first intersection. The node is delayed for a period of time sampled from a uniform distribution $U(0, 400) \text{ s}$. When the delay expires, the node selects a new speed at random and moves N to the next intersection. The northwards journey continues until the node reaches the latitude of the destination intersection. The node now moves E, subject to random delays and random changes in speed until it reaches its destination. After another random delay the mobile node selects a new destination. The mobility model does not claim to generate realistic urban movement patterns. However, it does randomize the spatial distribution of the mobile nodes both with respect to the FI and with respect to themselves, since it prevents the mobile nodes from forming a cluster when traversing the same street, a formation that is well-suited for the intra-cluster relay of traffic.

The nodes communicate as follows. Mobile nodes originate one call at a time and may communicate with other mobile nodes or with the outside world: 25% of the calls

---

**Table 2: Manhattan Grid: Node Parameters**

<table>
<thead>
<tr>
<th></th>
<th>router to fixed node</th>
<th>mobile node</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>basic FI configuration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bandwidth $C_j$ (Mbps)</td>
<td>250.0</td>
<td>50.0</td>
</tr>
<tr>
<td>power $\Gamma_j$ (W)</td>
<td>0.0</td>
<td>5.0</td>
</tr>
<tr>
<td>target credit balance $\phi$</td>
<td>50.0</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>enhanced FI configuration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bandwidth $C_j$ (Mbps)</td>
<td>500.0</td>
<td>100.0</td>
</tr>
<tr>
<td>power $\Gamma_j$ (W)</td>
<td>0.0</td>
<td>10.0</td>
</tr>
<tr>
<td>target credit balance $\phi$</td>
<td>50.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

---

![Figure 5: Manhattan grid: mobile to mobile calls.](image-url)
that originate at mobile nodes are sent to the outside world. In this case the call is routed from the mobile source to the closest (in terms of the credits required to transport the call) FI access point. The inter-arrival times and the service times of outbound calls are sampled from exponential distributions with means 1.0 s and 0.5 s respectively. These call parameters could model web browsing activities.

A mobile node may receive many calls at a time. Calls from mobile nodes to mobile nodes select a mobile destination at random and are transmitted using the least cost route, which may or may not use the FI. Since spending fewer credits to send the call translates to spending less money (see below), this approach is the best choice from an economic point of view: since low prices will only occur where transmission resources (bandwidth and power) are available, the least cost route also is the route that is likely to accommodate the highest flow rate.

Monetary payments are conducted as follows. Both the fixed and the mobile nodes transfer credits amongst each other to compensate each other for relaying calls. No monetary transactions are required to compensate the nodes for these credit transfers. Recall that the target credit balance of a mobile node is $\phi$. Mobile nodes which have a credit balance less than $\phi$ and which are therefore under-supplied with credits will be charged a monetary amount when they create credits through the application of Eqn. (3). The money raised this way is used to compensate those mobile nodes which have a credit balance greater than $\phi$ and which are therefore over-supplied with credits: these nodes destroy credits through the application of Eqn. (3). Monetary transfer via the creation and destruction of credit also applies at the outside router. Since the fixed nodes do not originate calls, they can only destroy credit. Table 3 shows that the transfer of monetary units is almost zero-sum.

Table 3: Manhattan Grid: Transfer of Monetary Units

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Mobile Nodes</th>
<th>Outside Router</th>
<th>Fixed Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>fixed</td>
<td>paid by</td>
<td>paid to</td>
<td></td>
</tr>
<tr>
<td>basic</td>
<td>-7.19</td>
<td>-2.11</td>
<td>9.29</td>
</tr>
<tr>
<td>enhanced</td>
<td>-6.77</td>
<td>-4.19</td>
<td>10.90</td>
</tr>
</tbody>
</table>

Fig. 5 shows the average of 15 independent replications of a simulation of the Manhattan grid model. The 95% confidence intervals are too small to be shown. Figs. 5(a) and 5(b) present histograms of the number of mobile to mobile calls carried as a function of the planar distance from the origin to the destination of the call. These figures show that (1) some 150, 000 mobile to mobile calls connect nodes that are separated by a distance between 0.2 and 0.3 times the network diameter (425 m): the routes that connect these calls are on average 3.6 hops long; (2) longer routes make more use of the FI; and (3) if the bandwidth and power resources at the fixed nodes are increased, then the mobile to mobile calls make more use of the FI since it is now cheaper to use.

Figs. 5(c) and 5(d) present histograms of the number of calls carried as a function of the length of their routes as
measured in terms of hop count. These figures show that (1) only mobile to mobile calls are connected on routes of length 1; (2) most of the calls connected on routes of length 2 are from mobiles to the outside world: one wireless hop to the nearest access point followed by one hop from the access point to the router to the outside world; (3) by construction, mobile to mobile calls connected on routes of length 1 or 2 hops do not make use of the FI; most calls connected on routes of length 3 or longer are mobile to mobile calls and, as the routes become longer, make increasing use of the FI; and (4) if the bandwidth and power resources at the fixed nodes are increased, then the mobile to mobile calls make more use of the FI since it is now cheaper to use.

Table 4 shows that the enhanced FI allows the mobile to mobile calls to make more use of the FI which is cheaper, since the FI resources are now more abundant. Note that the average distance moved by a mobile node is some 5.4 km which corresponds to an average speed of some 0.5 m/s.

Fig. 6 shows that the total credit balance is almost equal to 50 × 15 + 5 + 100 = 225 and that the throughputs remain approximately constant as the 100 mobile nodes move extensively around the plane.

5. CONCLUSIONS

This paper presents a model of an ad hoc network where the nodes are given incentives to collaborate. The model is based on a framework developed in [2]. Each node has a credit balance: credits are spent when a node sends traffic and credits are earned when a node acts as transit or destination node. The problem of flow allocation is indirectly solved by adjusting the bandwidth and power prices such that the flow rates are increased where possible and decreased when necessary, due to limitations on the bandwidth and power available at the nodes.

Direct application of [2] to model a mobile ad hoc network is constrained since the model assumes (1) that the prices and the amount of resources in use are updated very frequently, and (2) that knowledge about the prices and the utilization of bandwidth and power at each node are available globally and instantaneously. The original model therefore yields network performance results that represent an upper bound to what can practically be achieved in a mobile ad hoc network.

We introduce a model of a simple distributed signalling protocol. The prices and flows are now updated less frequently, and solely by using locally available information. An appropriate choice of signalling rate yields a performance that is only slightly inferior to the upper bound.

The distributed signalling protocol is applied to larger network models where the mobile nodes move according to different motion models. We demonstrate that the least-cost routing approach of the protocol can be used to efficiently route calls that may or may not use a fixed infrastructure. Long-distance mobile to mobile calls (three or more hops) are most likely to use the fixed infrastructure.

Each node has an incentive to make its resources available to the community in return for using the resources at other nodes. However, nodes that do not attract traffic, so that their credit balances remain lower than their target balances, create credits in order to send data while well-provisioned nodes that spend less credits than they earn have to destroy a portion of their credits. To reward a well-provisioned node for taking part in this system, we require monetary payment for the creation of credit and monetary compensation for the destruction of credit, which provides a tool for monetarily evaluating node positions.

6. ACKNOWLEDGMENTS

This work was supported by the South African National Research Foundation, Nokia-Siemens Networks and Telkom SA Limited.

Constants

\[ \begin{align*} 
\alpha_j & \quad 0.3 \text{ s}^{-1} \quad \text{willingness-to-pay factor at node } j \\
\beta & \quad 0.01 \text{ s}^{-1} \quad \text{credit discount factor} \\
\kappa & \quad 0.05 \text{ s}^{-1} \quad \text{bandwidth and power congestion price factor} \\
C_j & \quad 10.0 \text{ Mbps} \quad \text{bandwidth available at node } j \\
\Gamma_j & \quad 0.5 \text{ W} \quad \text{power available at node } j \\
\delta & \quad 56 \text{ m} \quad \text{connectivity threshold} \\
e^{x} & \quad 10^{-3} \quad \text{power consumed per unit flow when receiving.} 
\end{align*} \]

7. REFERENCES


