ABSTRACT
Choosing appropriate information dissemination strategies is crucial in mobile ad hoc networks (MANET) due to the frequent topology changes. Flooding-based approaches like diffusion have a strong similarity with epidemic spreading of diseases. Applying epidemiological models to information diffusion allows the evaluation of such strategies depending on the MANET characteristics, e.g. the node density. In order to choose appropriate strategies at run time, the model should be easily evaluated.

In this paper, an epidemic model is developed for a simple information diffusion algorithm based on simulation results. We analytically investigate the impact of node density on information diffusion. The analytical model allows the evaluation at runtime, even on devices with restricted resources, and thus enables mobile nodes to dynamically adapt their diffusion strategies depending on the local node density.

Categories and Subject Descriptors
I.6.5 [Simulation and Modeling]: Model Development – modeling methodologies.

General Terms
Algorithms, Performance, Reliability.

Keywords
MANET, analytical modeling, adaptive information diffusion.

1. INTRODUCTION
Mobile ad hoc networks (MANETs) are constituted by mobile devices equipped with short range radio transmission. Communication is possible between devices within each other’s radio range. The mobility leads to frequent topology changes in such networks, which harden classical networking tasks, such as routing or information distribution.

In flooding-based approaches nodes forward a received message or information to all their neighbors. Subsequently, all nodes within the network should receive the message or information. Even though flooding might expose some unnecessary message overhead it should provide a robust basic strategy for information dissemination in networks with an unknown or changing topology. However, the characteristics of MANETs even prohibit that a flooding process reaches every node. If the density of nodes is too high the radio transmission will block out messages if too many nodes are repeating their incoming messages as it is done in flooding. This problem is referred to as broadcast storms [10]. Here flooding shows a worse performance than selecting a smaller number of nodes to propagate the message to, which is called selective flooding. The other extreme would be a MANET with a low node density, where nodes can not communicate since they are likely to be out of each other’s radio range. Hence, a flooding phase is likely to stop before all nodes are reached. Repeating the message transmission over a period of time, which is called hyperflooding [11], can help to cope with such network partitions. As a result, different situations require different propagation strategies. There is no single strategy working in all above mentioned situations.

Since the node density strongly influences the performance of flooding-based information dissemination, the selection of a dissemination strategy should be based on the density of nodes. The dynamics in MANETs, however, lead to a constantly changing density. This density cannot only change over time but also in space, where clusters of nodes in one area of the MANET can lead to a sparse populated area elsewhere. As a result, even incorporating the density of the MANET alone does provide a suitable criteria for the selection of the flooding strategy.

An adaptive approach based on the local density around a node is desirable in order to cope with the varying MANET characteristics. This leads to the problem of determining the quality of a dissemination strategy depending on the node density.
This quality can be expressed in spreading ratio, energy consumption, spatial coverage and the level of meeting the application requirements. In this paper we investigate the spreading ratio depending on node density and formulate the application requirement as reaching a certain ratio of nodes in a given time interval.

To enable adaptation at runtime mobile nodes should carry simulation data or use simple heuristics [17] to switch between the different dissemination strategies depending on the current density. The first approach is not suitable for resource limited devices, the second approach can not switch smoothly enough. In our approach we develop an analytical model for each dissemination strategy, which analytically expresses the spreading ratio depending on node density. Monitoring its local density at runtime, a node can use the provided analytical expressions to compute the spreading ratio for different available dissemination strategies. Finally the node selects the most suited dissemination strategy to the application requirements.

The remainder of this paper is organized as follows. Section 2 presents an overview of our approach. Section 3 defines our system model and our diffusion algorithm. A model for simulations as well as simulation results are presented in Section 4. Section 5 describes our analytical model for our dissemination strategy. We analytically express the single model parameter in dependency on node density, discuss the results and show how to use them for adaptation. Section 6 describes the related work. Finally Section 7 concludes the paper and discusses future work.

2. OVERVIEW OF OUR APPROACH

There is a basic similarity between the dissemination of information among mobile devices carried by users and the transmission of infectious disease between the individuals themselves [4]. Both are processes in which given a contact something is communicated. When a disease like influenza spreads quickly and infects many individuals it is called an epidemic. Our simulations of information dissemination in MANETs (Section 4) show an epidemic like behavior as well.

Existing mathematical models that describe epidemic processes can be useful for us as they do for medical researchers. Medical researchers use epidemic models both to describe the spread of disease within a population and to take preventive or treating measures. We use epidemic models both to describe and to adapt the information dissemination in MANETs.

In general a model depends on a few parameters whose values should be determined from observations. Table 1 gives the parameters of a simple epidemic model, the so-called SI-Model. We give their correspondent meaning for information diffusion (more details in Section 5).

In Section 5 we show that this model is suitable for our diffusion strategy and that the entire progress of \(I(t)\) can be described by an expression that only depends on the infection rate \(\alpha\), the population size \(N\) and time \(t\). To enable adaptation to the node density, we aim to express analytically the infection rate of the MANET depending on node density. We proceed herby as follows: We simulate the performance of our diffusion algorithm, i.e. the spreading ratio, for a given density and fit the obtained curve to the analytical one and get for the chosen density the correspondent infection rate. Repeating this procedure for different densities and interpolating the points (density, infection rate) we derive an analytical expression for infection rate depending on the node density.

### Table 1. Analytical model parameters in epidemiology versus diffusion

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Epidemiology</th>
<th>Information diffusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>Number of individuals</td>
<td>Number of mobile nodes</td>
</tr>
<tr>
<td>(S(t))</td>
<td>Number of susceptible individuals at time (t)</td>
<td>Number of nodes at time (t) that are interested in the information</td>
</tr>
<tr>
<td>(I(t))</td>
<td>Number of infectious individuals at time (t)</td>
<td>Number of nodes at time (t) that carry the information</td>
</tr>
<tr>
<td>(x)</td>
<td>Contacts per unit of time and per individual</td>
<td>Contacts per unit of time and per node</td>
</tr>
<tr>
<td>(\beta)</td>
<td>the probability of transmission in a contact between an infective and a susceptible</td>
<td>the probability of transmission in a contact between an infective and a susceptible</td>
</tr>
<tr>
<td>(a = \frac{\beta x}{N})</td>
<td>Infection rate</td>
<td>Information infection rate</td>
</tr>
<tr>
<td>(\alpha = a \times I(t))</td>
<td>Force of infection: The probability per unit of time for a susceptible to become infective</td>
<td>Force of diffusion: The probability per unit of time for a node to receive the information</td>
</tr>
</tbody>
</table>

3. SYSTEM MODEL

We are interested in studying large-scale MANETs composed of vehicles or a large number of mobile portable devices. A network considered within the scope of this work is a set of nodes in an environment, which all can communicate with each other either directly or through multi-hops within a time interval. A population describes the set of nodes that is part of any network (of interest) at any time. A population can be defined by fixing a geographic area and considering the set of nodes sojourning there. We tolerate network partitioning, so at some times the considered node population can consist of several networks. The population is closed given that the number of participating nodes is fixed and known and that no new nodes join and no existing nodes leave.

The mobile nodes forming the population are mobile devices such as notebooks and PDAs. We assume that they move in the considered area according to a random waypoint model [2]. Initially, each node chooses a destination uniformly at random in the considered area and a speed, and moves there with the chosen speed. Then the node pauses for a random period of time, before repeating the same process. We assume that every node has enough memory to save the information that will be diffused. In this paper we do not consider the different energy-saving operating modes and assume that every node always is in active mode and do not become out of energy.

In the population described above we aim to diffuse one static information entity. An information entity is a structure of data that is considered to be atomic for the diffusion, i.e. it will not be split during diffusion.
The most important factors which affect the characteristics of information diffusion (for example spreading speed, i.e. the time to spread the information entity to all nodes) are population size, communication parameters (transmission range and rate and discovery time, i.e. the time needed to establish a physical channel to a node located in the communication range), mobility parameters and the epidemic algorithm. We define the spreading ratio at time \( t \) as the ratio between the number of nodes that carry the information entity and the number of all nodes, i.e. population size. Let \( i(t) \) denotes the spreading ratio at time \( t \). \( i(t) \) varies between 0 and 1.

In this paper we consider a diffusion algorithm that follows the protocols for information dissemination in sensor networks (Sensor Protocol for Information via Negotiation: SPIN-1 [14]). When a mobile node discovers other mobile nodes, it advertises a summary of its information entities. The listening nodes then request the information entities which they are interested in. Finally the advertising node sends the requested data.

4. SIMULATION MODEL

We use our own simulator, which is written in Java and implements the following model. The Media Access Control (MAC) layer is as an abstraction of the IEEE 802.11b MAC layer. We assume a constant transmission rate \( r \). The communication range is also constant and denoted by \( R \). The discovery time \( t_d \) is chosen uniformly at random between two fixed values. The implementation of the MAC layer considers collisions caused by mutual transmissions.

The area of interest is a 1000m x 1000m field. Let \( N \) be the number of nodes in this area and \( d \) its node density (measured in \( 1/\text{km}^2 \)). We assume a closed population so \( N \) is constant during the simulation. For random waypoint movement we use speeds between 3 and 70 Km/h and a pause period between 0 and 100 s. Table 2 summarizes the simulation parameters.

At the beginning of the simulation just one node carries a single static information. We assume that every node is interested in this information. We ran 20 passes for the same simulation scenario and we considered the average. A simulation pass is run until a spreading ratio of 95% is reached. Thus we can avoid that a few isolated nodes unnecessarily enlarge the simulation time.

Table 2. Simulation parameters

<table>
<thead>
<tr>
<th>Area</th>
<th>1000m x 1000m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>( N ) (variable)</td>
</tr>
<tr>
<td>Communication range</td>
<td>( R = 75m )</td>
</tr>
<tr>
<td>Communication rate</td>
<td>( r = 2048 \text{ Kbit/s} )</td>
</tr>
<tr>
<td>Discovery time</td>
<td>Uniform between 2 and 3 s</td>
</tr>
<tr>
<td>Data size</td>
<td>510 bytes</td>
</tr>
<tr>
<td>Movement pattern</td>
<td>Random waypoint</td>
</tr>
<tr>
<td>Velocity</td>
<td>Uniform between 3 and 70 km/h</td>
</tr>
<tr>
<td>Pause</td>
<td>Uniform between 0 and 100 s</td>
</tr>
</tbody>
</table>

We simulated the spreading of the single constant information entity through populations with different node densities: \( d \) between 50 and \( 5000/\text{Km}^2 \). The result for \( d = 150 \), 180 and 220/ \( \text{Km}^2 \) is shown in Figure 1.

5. ANALYTICAL MODEL FOR OUR SIMPLE DIFFUSION ALGORITHM

In this Section we develop our analytical model and express its single parameter in dependency of node density. Then we discuss our results and show their relevance.

5.1 Analytical Model

According to the diffusion algorithm and to the assumption that every node is interested in the single static information a node follows the following two-states compartmental model: It either carries the information or not and once infected by the information a node remains infectious. Let \( S(t) \) denote the number of susceptible nodes, and \( I(t) \) the number of infective nodes. So we consider the two-states mathematical model shown in Figure 2. Each letter in a rectangle refers to a compartment in which a node can reside.

![Figure 2: Compartment diagram for the SI-Model](image)

Hereby \( \alpha \) is the diffusion force in the MANET. This parameter indicates the strongness of the diffusion process and has the dimension \( \text{time}^{1/2} \).

We concentrate on the behavior of large scale populations so we use a deterministic compartmental epidemic model. For small populations a stochastic model should be used [1]. To develop the solution, we need to write the mass balance equations for each compartment:

\[
\begin{align*}
\frac{dS(t)}{dt} &= -\alpha * S(t) \\
\frac{dI(t)}{dt} &= \alpha * S(t)
\end{align*}
\]

Figure 1: Simulation results

Table 2. Simulation parameters
The problem is that $\alpha$ is not a constant, but depends on the number of susceptibles and infectives and the probability of transmitting the information in a contact. Suppose each susceptible makes $x$ contacts per unit of time that are of the information transmitting type (both nodes are within each other’s communication range). Then the susceptibles in total make $xS(t)$ contacts per unit of time. Hence a random waypoint movement model is used, we assume that the contacts are at random with members of the total population: $N=S(t)+I(t)$ (the epidemiologists speak about homogeneous mixing model). Then only the fraction $I(t)/N$ of the contacts are with infectious individuals. Let $\beta$ be the probability of information transmission in a contact between an infective and a susceptible. Then the rate of susceptibles that become infected is

$$\beta(xS(t)) \frac{I(t)}{N}.$$

Thus the diffusion force is

$$\alpha = \frac{\beta x}{N} I(t).$$

We substitute

$$a = \frac{\beta x}{N}$$

and call it infection rate. Thus

$$\frac{dl}{dt} = a * S(t) * I(t)$$

With $S(t)=N-I(t)$ we get:

$$\frac{dl}{dt} = a * I(t) *[N - I(t)] = a * N * I(t) - a * I^2(t).$$

This first order ordinary differential equation has the following general solution:

$$I(t) = \frac{N}{1 + C * N * e^{-aNt}}.$$

Where $C$ is a constant of integration that depends only on the initial conditions. $C$ is computed as follows. At the beginning ($t=0$) we assume that we have just one node that carries the information. So $I(t)$ should fulfil $I(0)=1$.

$$I(0)=1 \Rightarrow \frac{N}{1 + C * N} = 1 \Rightarrow C = \frac{N-1}{N}.$$

Thus the final solution of (2) is:

$$I(t) = \frac{N}{1 + (N-1) * e^{-aNt}}.$$

According to the definition of spreading ratio (see system model):

$$i(t) = \frac{I(t)}{N} = \frac{1}{1 + (N-1) * e^{-aNt}}$$

(3)

For $a = 0.04$ 1/s and $N=100$ nodes, $i(t)$ is plotted in Figure 3.

Figure 3: Visualisation of (3) with $a=0.04$ 1/s and $N=100$ nodes

The spreading ratio gained from the simulation and shown in Figure 1 looks very similar to the spreading ratio gained from the analytical model in Figure 3. This is a confirmation of the usefulness of existing epidemic models to describe the spreading of information in MANETs.

In this simple mathematical model the infection rate $a$ describes the progress of the information dissemination in the MANET. Information dissemination depends on node density, communication model, mobility model, information model and the epidemic algorithm. So $a$ depends on node density, the communication model parameters (transmission rate and range and discovery time), the movement model parameters (minimum and maximum velocity and the maximum pause time), information model parameters (information size) and the epidemic algorithm. Equation (1) shows that $a$ depends on population size ($N$), mean number of contacts per node per unit of time ($x$) and probability of information transmission given an adequate contact ($\beta$). We can easily conclude that $x$ is the most relevant feature of our mobility model and $\beta$ is the most relevant feature of communication model, information model and diffusion algorithm that impacts the information diffusion.

The challenge is now to analytically express the parameter $a$ in dependency on the MANET parameters and the diffusion algorithm. Through our hierarchical modelling approach this task can be reduced to the determination of $x$ from the random waypoint movement model and $\beta$ from the underlying communication and information models and the diffusion algorithm. In this paper we study the dependency of the infection rate on the population size or node density. So we fix the communication, movement and information model parameters and vary just the node density. We use the same values as fixed in the simulation model.

5.2 Determination of the Infection Rate

In the following we present a simulation-based empirical approach to determine the infection rate $a$. Many simulations are run to get enough points in the space $a = a(d)$. Using interpolation we then get an analytical expression for $a(d)$.

We used the least squares method, a procedure for finding the best fitting curve to a given set of points, to fit the simulation results to the formula (3). Thus we determined for a given...
MANET with node density $d$, the correspondent infection rate $a_i$ (see Figure 4). We used mathematica [18] to process this fitting procedure.

5.3 Results

We varied the node density and computed each time the correspondent infection rate. The results that we got are shown in Figure 5.

5.4 Interpretation

It is clear from Figure 6 that there exists an optimum node density $d_{opt}$ that maximizes the infection rate and therefore also the diffusion speed. We observe a maximum infection rate by $d = d_{opt} = 620 \frac{1}{km^2}$. With increasing node density the infection rate $a$ first increases, reaches a peak and then decreases. We interpret the observed fact as follows: If $d$ is too small, the network is very sparse and partitioned. Only through the movement of infective nodes between partitions the diffusion can go on. If $d$ is too large, then the interference becomes dominant and the algorithm diffuses more slowly. Concurrent wireless transmissions in an ad-hoc network namely limit its throughput capacity, because they create mutual interference (collisions).

This analytic expression is valid for densities not in proximity of 0 (the smallest population size we considered was 50).

- $d \geq 620 \frac{1}{km^2}$:
  
  Ansatz: $a = a_0 + \frac{a_1}{d} + a_2 * d^2$

  Result:
  
  $$a = 6.4 * 10^{-4} + \frac{25.5}{d} - \frac{9081.6}{d^2}$$

  where $a$ in $s^{-1}$ and $d$ in $km^{-2}$.

results of work [15,16,5] confirm our results. If we assume uniform distribution of nodes, the optimal mean number of neighbors in our MANET is $n_{opt} = \pi R^2 d_{opt} = 10.95$. Relying on the works [13,16] we can expect that infection rate decreases but not toward zero. The interpolation in (5) expects that infection rate for infinite density is $a_{\infty} = 6.4 \times 10^{-4} \gamma'$. 

5.5 Relevance of the Results
The results presented here can be generalized for any algorithm that shows the same compartmental SI-Model and that is run by nodes that follow similar communication and movement patterns.

In [6] the authors studied empirically a simple epidemic algorithm, i.e. flooding, in stationary large-scale multihop networks. Their experimental results show an epidemic-like information propagation and second that epidemic models are appropriate to model information diffusion processes in ad hoc networks.

Analytical models for information diffusion allow succinct descriptions of information dissemination in MANETs using very few parameters with analytic expressions. Assuming a node can perceive its environment (e.g. local node density and population size), it can process the infection rate of a given epidemic algorithm, using the analytical expression (4) and (5). Running different diffusion algorithms, for which the node has a SI-Model, the node is able to predict the progress of information spread over time for these algorithms (i.e. using (3)). The node will be able to switch between these algorithms to adapt information dissemination to the needs of the application and to the current environmental situation in the MANET, e.g. the local node density.

6. RELATED WORK
In [12] a simple stochastic epidemic model is used to analytically determine the delay until data spreads to all mobile devices. This model is a pure birth process which is a simple continuous-time Markov process. The authors also assumed two compartments S and I for a node, but they only considered small populations (5 nodes) for which stochastic models are recommended. We are interested in large scale networks for which deterministic compartmental epidemic models are more suitable. In our work we investigated the entire progress of dissemination over time and analytically expressed the model parameter in dependency on node density.

In [13] a diffusion-controlled model is used for data dissemination between mobile queriers and static servers, i.e. the static trapping model. The stationary information servers are modelled as traps (sinks) and the mobile nodes as particles that performs diffusive motion in the space and are absorbed on the traps when they step on them. The authors also fitted the simulation results to the analytical results to determine the model parameter (lattice-dependent constant) but they do not express this parameter analytically in dependency from important population parameters. The static trapping model is not suitable for our system model because we do not consider fixed nodes and we assume the cooperation between mobile nodes. But we are looking for exploring other diffusion-controlled models to model information propagation in MANETs.

Mathematical epidemiology has been also used outside of the biological field to model how ideas propagate [7] and to develop algorithms for maintaining replicated databases [3]. IBM Research applied these techniques to model the spreading of computer viruses in the Internet and to derive anti-virus strategies [8].

7. CONCLUSION AND FUTURE WORK
In this paper we applied a two-compartments analytical model, i.e. the so-called SI-Model, for a simple diffusion algorithm. We showed that information dissemination using this simple epidemic algorithm and assuming a random waypoint movement model can be well modelled by that deterministic compartmental epidemic model and can be described by its parameter: Infection rate. Finally we derived an analytical expression for the infection rate in dependency on node density, which is a major indicator for the suitability of information dissemination strategies.

In future work we will look for investigating the impact of other underlying model parameters on the infection rate. As movement model we used the simple random waypoint model, which allowed us to assume a homogenous mixing of nodes. In more realistic scenarios, like vehicles moving on a road map, this assumption can fail. That is why we look for exploring data dissemination on different network topologies like small-world graph. In more realistic scenarios population will not be close. For open populations we have to introduce birth and death rates to our compartment diagram.

8. REFERENCES


