Information Diffusion in a Single-Hop Mobile Peer-to-Peer Network

Jani Kurhinen and Jarkko Vuori
University of Jyväskylä
Department of Mathematical Information Technology
B.O.Box 35 (Agora)
FIN-40014 University of Jyväskylä
FINLAND
[kurhinen,mimic]@mit.jyu.fi

Abstract

In this paper a simple and novel way of modeling the process of information diffusion in single-hop mobile peer-to-peer networks is presented. In the model observation concentrates on areas where mobile nodes visit rather than mobility of each mobile node at a unit level. This gives more realistic results compared to the traditional pure random movement based models. Since the required technology to study this is not yet available, the results and conclusions are based on computer simulations.

1. Introduction

Peer-to-peer (P2P) networks have attracted attention lately. The ability to harness different resources scattered in the Internet is probably the biggest appeal of the P2P technology. Taking advantage of P2P starts from a primary approach, in which all the tasks and responsibilities are shared between the peers. This means that there is no single control entity responsible for services, hence the name peer-to-peer network. Due to a growing number of providers, a greater amount and diversity of information can be attained from various P2P networks.

In addition to peer-to-peer networks, another technology has also lately attracted attention: mobile ad hoc networks (MANET). These networks were developed by a DARPA project in the early 1970's [6], and have later spread to civilian use as well. A mobile ad hoc network is a network with no predetermined structure. It consists of mobile wireless network nodes that are able to communicate with each other through intermediate nodes [2].

In this paper we study a new approach of using a combination of P2P and MANET technologies to enable information diffusion within a user population. The combination is referred as Mobile P2P, MP2P.

2. Information Sharing in a MP2P environment

Both peer-to-peer networks and ad hoc networks share the same common problem: delivering messages in a rapidly changing network topology is a challenging task. When examining this task more profoundly, we can find two opposite types of message delivering methods. In the first one a network node communicates proactively being an active party in transmitting or inquiring data. The second one is reactive communication, where communication link is established on-demand when required. Both of the message delivery methods, proactive and reactive, are applicable in the context of mobile peer-to-peer networks, but however, proactive transmission will cause substantial overhead compared to payload traffic if the network structure is dynamic. The reason for that comes simply from the fact that mobility of network nodes changes continuously the topology of the network, and thus, routes between network nodes must be updated frequently, and this new information must be delivered to all of the nodes whether or not they benefit from it.

Both of these cases share a common feature: the message is transmitted via other network nodes. This is, indeed, the only way to deliver a message from node to node with no direct data link between them. In MP2P, however, the network topology changes as mobile nodes move. Due to these changes, the nodes may end up near another node who possesses desired information or to whom they want to deliver information. This whole system can be simplified by allowing only direct single-hop links so that information is transmitted from the source node to the target without intermediate nodes, and. In this case the network must assume that the movement of the mobile nodes will deliver messages to the querying nodes.

In a scope of this paper we study data exchange in which only single-hop communication is used. By not allowing the usage of intermediate nodes between the two ends of the communication link, the complex routing problem is avoided. With this limitation the mobile node has two possible ways to get a required piece of information: either it finds its way into a location, where it can request the information reactively, or it can receive that information from a node who is serving it proactively. However, as the simulation results will show, this limitation does not prevent the information diffusion, but provides a solid base for applications that benefit from the information diffusion among MP2P nodes.

3. Information diffusion Models

The following part presents existing models of mobile nodes and information diffusion. However, it will be pointed out that these are not very well suited for depicting short-range communications in mobile peer-to-peer networks. Therefore, a new revised model will be presented in section 3.2.

3.1. Research with an epidemic diffusion model

In order to study a dynamic system, such as the MP2P network, often the first thing to do is to select a mathematical node mobility model best suited for the particular case. Yu and Li [7] list three most popular models to be random waypoint, random walk and random direction models.

The total freedom of movement, however, is very problematic if the model tries to represent a real world: in fact, the pattern of mobility have most of the time some limitations based on, for example, physical or cultural obstacles.

In addition to modeling mobility patterns, another part of the research problem is information diffusion among mobile nodes, which has been studied by, for example, Arai et al [1] and Khelil et al [3], and in both of these some free random mobility model is used. Arai et al also have included stationary information sources into their study. These source points were fixed locations where the spreading information were originated. Similar extension were also used in a study by Papadopouli et al. [4]. Although Papadopuli's study was not about information diffusion in a MP2P network, they also presented the idea of the need for stationary components to complement mobility models. In other words, one must take in account all the elements that might affect the result whether they are mobile or not.

When studying information diffusion in a system, in addition to the model for mobility of the network nodes, there is also requirement for modeling the information exchange process, i.e. the way a piece of information is transferred between nodes. One possible, and often used, model is an epidemic diffusion model, which, for example, above mentioned Khelil et al [3] have used. Although, the epidemic model is widely used in various areas of research, its main problem is its tight relation with random mobility models: it is assumed that an object which carries the 'infection' will meet other objects in a random pattern. When using this kind of model, the number of infected users as a function of time will have general S-shaped form, like the one from Khelil et al. This is based on the fact that with only random mobility it requires 'critical mass on infected objects' for the information to spread fast. Therefore at the beginning of the diffusion process, the information does not spread fast. As opposite to this traditional model, we claim that information can be expected to spread exponentially immediately when it is released into area where there are high density of possible new carriers of the information.

3.2. The Exchange Pipe model

The dynamic nature of MP2P network also leads into a situation where the node density is not constant, but varies during time [3]. Because of this variation, there in the simulation space will form denser and sparser subareas.

As was earlier pointed out, the single biggest problem in the current models is their tendency to always focus on modeling the movement patterns of mobile nodes. However, a MP2P network with a very limited communication range can not be expected to work, unless an adequate number of nodes is located in a limited area at a given moment. When this assumption is combined with the notion voiced in [3] that mobile nodes are typically packed on certain areas, it can be justified to change the focus of study from separate mobile nodes to defined observation areas.

Figure 1 compares the traditional model based on node movement with an observation area based model designed for this study. The traditional model on the left-hand side shows the direction and speed of each individual mobile node, depicted by the direction and length of the arrow. The right-hand side focuses exactly on these relevant areas where information can be exchanged due to high node density. The outlined areas are node thickenings, and the arrows represent nodes entering and leaving the dense area. Unlike the current node-centered models, this one does not contain information of the state of individual nodes, but only the stream of nodes through the node thickenings.

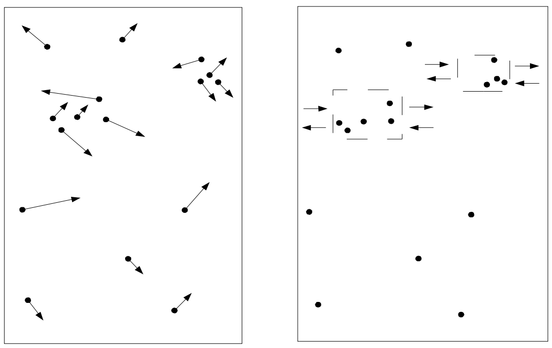

Figure 1. "Traditional" way of modeling and the exchange pipe model.

When planning the simulation model, the inspiration came from thinking of a busy promenade where the users of mobile devices are packed tight, whereas outside the promenade the density of them is substantially lower. People enter and leave the promenade from its ends and encounter a host of other users while on the promenade. The promenade was generalized into an information exchange pipe, hence the name. This exchange pipe model can be illustrated by detaching an outlined area from figure 1 and simplifying it slightly. In order to ease processing of the model, it can be assumed that the exchange pipe is one-dimensional, and it is also defined that there can be two of more streams of nodes going through the exchange pipe in any direction (see figure 2). The information exchange takes place between the nodes moving in the node streams.

Two quantities were defined to describe the attributes of the pipe: node stream flow (Φ) and node density (ρ). Node stream flow is defined as the number of nodes (*n*) passing a certain point during a given period of time as follows

$$\Phi = \frac{n}{t}. \qquad (1)$$

Density, on the other hand, is defined as the number of nodes (*m*) located at the same time in a certain space. The exchange pipe was earlier defined as a one-dimensional object, due to which the certain space can be replaced with a one-dimensional representative length (*x*). Let us define node density in a node stream as follows:

$$\rho = \frac{m}{t}. \qquad (2)$$

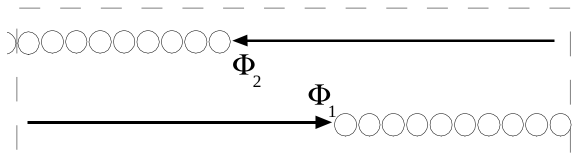

Figure 2. The information pipe with two node streams.

As mentioned in section 3.1, pure mobile environment sometimes needs fixed nodes to give more realistic results. The information that will be spread need to come to the system from somewhere, and in this model the information source is defined as such a fixed unit. In other words, while some people are walking on the promenade, others are queuing at an ice cream stand outside the promenade and will receive information about the current prices. The ice cream stand operates as a source of information and spreads its prices to the people who visit it. As time goes by, all the potential customers will eventually have visited the stand, and therefore received the information about the prices, making the information penetration 100%. However, if standing at the ice cream queue is the only way to receive information, it will take a long time to reach that penetration level.

As the customers re-enter the promenade after finishing their ice-creams, they become potential information distributors. They now have information, in which other pedestrians are also interested, and the ice-cream buyers are now able to spread this information to other pedestrials. MP2P networks thus enable information diffusion not only directly from the information source but also from one user to another, which then accelerates the diffusion process.

The exchange pipe model mimics the previous example of the diffusion of ice cream prices. The promenade refers to the dense area (the exchange pipe), in which the information diffusion is possible with the MP2P technology, and the stand is the original source of the information. Pedestrians at the promenade are mobile nodes, and depending of the current MP2P technology penetration, some number of the mobile nodes are able to communicate with each other.

The model can also be used to illustrate city centers, highways as well as any other place with a similar high contrast between crowded and sparsely populated spaces. By combining several pipe modules, an entire road network, for example, can be modeled. Different parameters (node stream flow and density) can be assigned to different parts of these networks in order to represent certain types of roads etc. The ends of these pipe modules are then representing crossroads. Outside these 'city centers' or 'highway networks' the users are so scattered that they are insignificant in respect of information diffusion.

3.3. The effect of number of active users

The innovation diffusion process in a social system has been found to follow a certain pattern. The members of a social system using the innovation draw an S-shaped curve when observed as a function of time [5]. The width of the curve varies depending on the speed of the diffusion process, but the basic shape always remains the same.

The theory of the diffusion of innovations was created by Rogers [5] at the beginning of 60's. According to his theory, people can be categorized based on their tendency to adopt innovations. There are five groups which adopt it in the following order: innovators, early adopters, early majority, late majority and laggards.

Only 2,5 % of a the members of a social system are considered innovators, followed by 13,5 % of early adopters. The so called average users, which are divided into two groups, the early (34 %) and late majority (34 %), constitute 68 % of the total population. After the early majority has adopted an innovation, its penetration level in the population has reached 50 %, and after the late majority has adopted it, the only remaining group is the laggards, with a 16 % share of the members of the social system.

By using the above mentioned numbers of Rogers' theory, we divided our simulations into five different phases: in each phase a new category of user have adopted an MP2P communication device. With this addition we can monitor the behavior of our simulated MP2P environment with a different level of participating users.

4. Simulation

4.1. The model for diffusion of information

In order to study information diffusion process in a mobile peer-to-peer network a simulator was created. In the simulator there is four types of objects: 1) The mobile node is a unit, which seeks information. The nodes do or do not have the capability to share information with other nodes. This attribute represents the share how many of the users have adopted the required technology for MP2P as was stated in section 3.3. 2) Node streams, which consist of mobile nodes. Each stream has two attributes, flow and density, and the values of these attributes can be set differentially for each stream. The stream is used to move mobile nodes in the exchange pipe. 3) The exchange pipe is the place where mobile nodes can exchange their data if they have the capability to do that. The exchange pipe is the actual observation area, i.e. it is the place where mobile nodes are situated close enough to each other in order to exchange messages with the short-range MP2P communication devices. Each of the exchange pipes can have two or more node streams. 4) The information source is a place where the diffusing data is originated. Through the information source there are one or more node stream flows. From the information source all the visiting nodes are able to receive current information, and thus it is the only place to receive that for such a mobile nodes, which are not able to share data with other mobile nodes.

When the simulation starts mobile nodes are randomly selected to enter the node streams that moves them through the exchange pipes. The random selection mimics the random movement of the mobile nodes outside the observation area. But when a node enters to one of the node streams, its mobility pattern changes to follow the attributes of that particular node stream. When the nodes inside the exchange pipe encounters another node, they compare the data they are carrying, and if one have more or newer data it shares those with the another node. The encountering point-to-point connection, and the simulation environment is normalized to follow the communication range of the MP2P nodes: thus when the MP2P nodes share the same one-dimensional discrete coordinate, they are within a communication range.

The simulation was run using various parameter combinations. For the simulation graphs presented below the simulator was set to be extremely simple and to have one information source and one exchange pipe. The population was 100 000 from which 50 % was able to use MP2P technology. In the exchange pipe these were two node streams in which there was 20 nodes. The node stream flow in the information source was one half of the stream flow in the exchange pipe.

The simulation results were noticeably different from the ones that other researchers have attained with the epidemic information diffusion model with some pure random mobility pattern of the network nodes. While the old results share the common S-shaped form, this new model have an exponential growth behavior: the growth rate reaches its maximum right after the spreading information was released at the exchange pipe, i.e. when the first node, that is capable of sharing data with other nodes and is carrying the information, enters the exchange pipe, the information diffusion rate starts to grow explosively (see figure 3).

With the assumptions of the exchange pipe model, the exponentially growing number of information carriers model reality better than the S-shaped curve. The old models have required a critical mass of infected users in order to attain fast growth. However, when we are dealing with areas with a very high user density, like city centers and busy promenades, the critical mass of users is available right from the beginning.

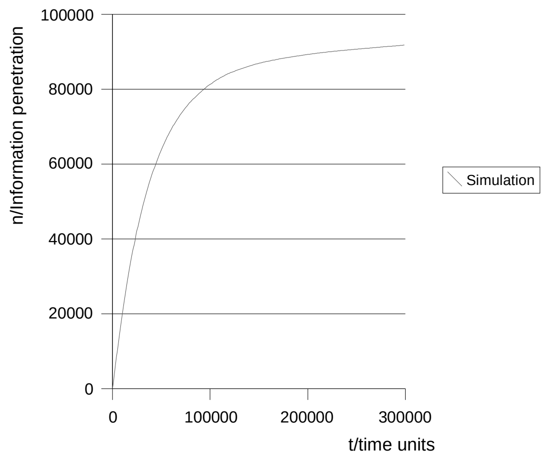

Figure 3. Explosive growth of the information diffusion simulation.

When observing the simulation results more specifically we noticed that he curve of the simulation results resembled function

$$n(t) = P - Pe^{-kt}, \qquad (3)$$

but it was not as smooth, and the parameters *P* and *k* could not be fitted to suite that function. However, from the output of the simulator one can separate three different factors: 1) The nodes that do not have technology for MP2P, and therefore are able to receive information only from the information source. 2) The nodes that have received their data from another node inside an exchange pipe. 3) The nodes which have the required technology, but received the information from the information source. The nodes who form the third factor act as a carrier and therefore they are the ones who start the explosive information diffusion process at the exchange pipe. By drawing the curves of these three factors separately we can see that each factor draws a curve that is form of (3) and it is possible to fit the parameters *P* and *k* to follow the simulation results (see figure 4). The curves with s-label are simulation results and the curves with n-label are the ones calculated numerically with (3). The number in the label represents the above mentioned factor group.

It can thus be argued that information diffusion in a short range mobile peer-to-peer network can be modeled with function *N(t)* as a sum of (3) with three different parameters:

$$N(t) = \sum_{i=1}^{3} n_i(t) = \sum_{i=1}^{3} \left(P_i - P_i e^{-k_i t}\right) \qquad (4)$$

in which P_i is the number of nodes in the group *i* and k_i are constants which values depends on ϕ and ρ.

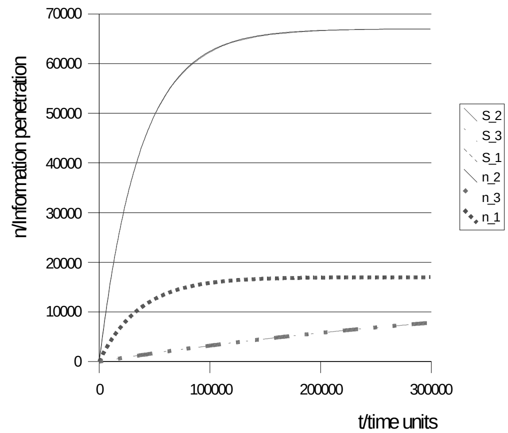

Figure 4. The factors of the simulation and N(t).

4.2. Innovation diffusion levels

In the previous section there was introduced a function which can be used for modeling information diffusion in MP2P. It was stated that the result depends on two parameters: the node stream flow and the node density. However, it was not described how these variables are selected. To be able to determine these parameters more accurate the innovation diffusion model presented in the section 3.3 can be used.

As it was stated earlier, the penetration level of a new innovation will spread following an S-shaped curve. By monitoring the diffusion of information with MP2P using different assumptions for the level on innovation penetration, it is thus possible to predict information diffusion behavior. In the figure 5, there is collected information diffusion curves of MP2P technology penetration of 0 %, 50 % and 100 %.

Although figure 5 already shows substantial difference between the usage levels, by taking the first derivate of time from (4), we can see more precisely the effect of the different number of MP2P users. This derivate can be written as follows:

$$\frac{dN(t)}{dt} = \sum_{i=1}^{3} P_i k_i e^{-k_i t} \qquad (5)$$

By using (5) with technology penetration levels 0 %, 50 %, and 100 % Figure 6 was created. It illustrates clearly the change in the speed of the information diffusion from the technology penetration level of 0 % to the level of 50 % and finally to the level of 100%.

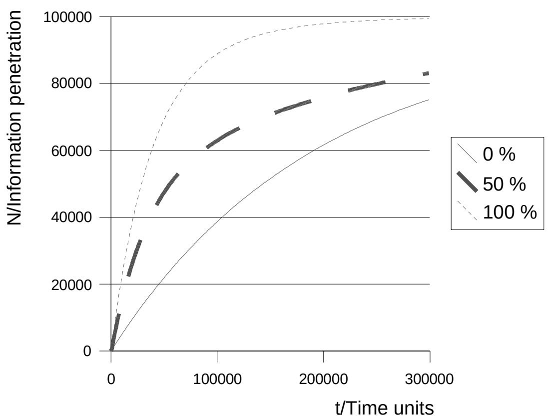

Figure 5. The effect of increased number of MP2P users.

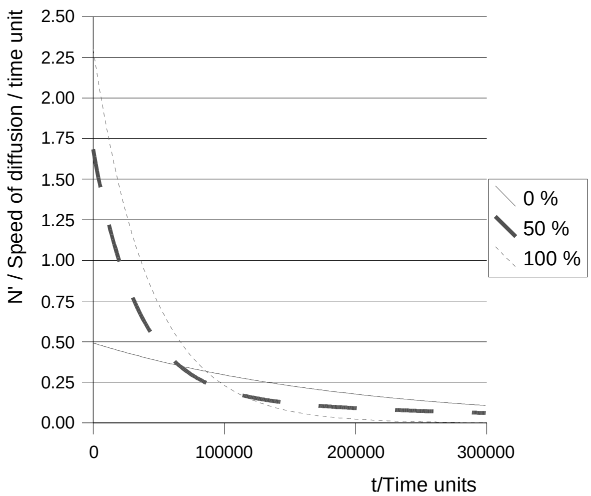

Figure 6. The speed of information diffusion with different number of MP2P users.

As figure 6 implies the MP2P system has the biggest effect on the information diffusion right after the information is released to the user. The result can thus be claimed to be “reasonable” for situations where people are packed tight in a relatively small area, and thus enables explosively growth. The figure also shows that even 50 % technology penetration level provides massive increase in the speed of information diffusion.

5. Conclusion

The research began by examining previous studies of information diffusion in a peer-to-peer environment. The models used in them were, however, considered not so well suitable for MP2P, since they were based on random mobility patterns, not suited to model real world mobility. In addition to this, monitoring mobile nodes located far from another offers no added value whatsoever. A simulation model was created for the purpose of observing the thickenings of mobile nodes in which information is able to transmit from node to node.

Contrary to the previous results, the diffusion curves of the new model were not S-shaped, but the most intense growth took place at the very beginning of the diffusion process. At this point, when a piece of information was released in a thickening of mobile nodes, the information diffusion was almost explosive. It seems that our model is better suited dense urban areas mobile penetration is high. In the low populated areas it is still required first to attain a critical mass of “infected” users, and there the old models with random movement patterns are still valid, but when observing such areas where a population density is high, the exchange pipe model provides a new way to describe the information diffusion process.

The current implementation of the exchange pipe model has its limitations. The model, for example, leaves the sparse area out from the observation. However, we can assume that random contacts between MP2P nodes will affect the information diffusion at some level. The next logical step would thus be combining these two different models. In this kind of combination, the interest rests on this question: does the fast diffusion in the exchange pipe change the old results where in random movement based models it took some time to gain the critical mass needed to accelerate the diffusion process.

6. References

[1] Arai, T., Yoshida, E., Ota, J. 1993. Information Diffusion by Local Communication of Multiple Mobile Robots. Proc of International Conference on Systems, Man and Cybernetics. 'Systems Engineering in the Service of Humans'.

[2] G. Ding, and B. Bhargava, "Peer-to-peer file-sharing over mobile ad hoc networks", The Second IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004, Proceedings, pp. 104-108.

[3] Khelil, A., Becker, C., Tian, J., Rothermel, K. 2002. An Epidemic Model for Information Diffusion in MANETs. Proc. of the 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems.

[4] Papadopouli, M., Schulzrinne, H. 2001. Sensor networks and energy management: Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices. Proc. of the 2nd ACM international symposium on Mobile ad hoc networking & computing.

[5] Rogers, E., M. 1983. Diffusion of Innovations, 3rd Ed. The Free Press, New York.

[6] R. Schollmeier, I, Gruber, and F. Niethammer, "Protocol for peer-to-peer networking in mobile environments", The 12th International Conference on Computer Communications and Networks, 2003, ICCCN 2003, Proceedings, 20-22 Oct. 2003, pp. 121-127.

[7] Yu, D., Hui Li. 2003. Proceedings of International Conference on Communication Technology, ICCT 2003, Volume: 2.