Mat-2.177 Seminar on Case Studies in Operations Research

Modeling Mobile Peer-To-Peer Networks

Project plan

14/03/2007

Client: Nokia Research Center

Contact: Jukka Nurminen

Nokia Project Group

Emilia Etelä-Aho

Lauri Hyry

Heidi Kettunen

Eero Nevalainen

Table of contest

I. Introduction

This study is part of the course Mat-2.177 Seminar on Case Studies in Operations Research in spring 2007. The client of this project is Nokia Research Center which is a separate unit within Nokia.

The study is about modeling mobile peer-to-peer networks (MP2P). In the very beginning, the client did not provide us any specific problem to be solved. Instead, we were allowed to define the objectives of this study quite freely.

i. Project objectives

As we were allowed to set the objectives for this study by ourselves, we preferred to model cases that have practical applications. The main objective of this study is to model content propagation in MP2P networks. The main objective is divided in three smaller sub-objectives that all concentrate on modeling different kinds of behavior in MP2P networks. The sub-objectives are meant to be non-exclusionary: achieving the first sub-objective is prerequisite for achieving the others.

1. Modeling the propagation of new-year eves humorous text-messages in MP2P network

In this simplest sub-objective, we try to model the propagation of text-messages, when network users or nodes prefer all information sent to them in a same way. In this case, it is assumed that users also send the received text-messages further in a same way (all nodes receive and send all messages based on simple rules). This can be expanded later to include also user preferences and group memberships.

This is the first sub-objective to be simulated. Before trying to achieve the other two sub-objectives, the results of this model will be analyzed and possible model improvements made.

2. Modeling content propagation in MP2P network

In this second sub-objective, the network is assumed to consist of nodes that prefer some kind of information and nodes that do not have any preferences.

Modeling this case may be only a little step further from the objective described in the previous chapter. When modeling this case, we try to find answers to the following questions:

- How should advertisement be sent to the network in order to reach specific amount of network nodes?
- What is the relation of node properties and the overall network properties? E.g. how should we define the parameters of each node to achieve specific level of network reliability, scalability and bandwidth?

3. Modeling content propagation in MP2P network when network nodes have different preferences

In real life, a MP2P network may consist of older nodes that do not provide a way of displaying the node's preferences and of more modern nodes that prefer different kind of information based on the user's preferences.

When modeling this case, we try to find answers to the following questions:

- What is the relative amount of modern, preference-capability nodes network should have to receive certain efficiency increase in content propagation?
- How valuable the information of user preferences might be to the advertiser?

ii. Definitions

1. Nodes

In our model, the most important node property is going to be **membership**. The membership can be thought as belonging into a certain social group. The network of nodes may then consist of e.g. basketball fans, art lovers and human rights activists. The different groups prefer different kinds of information in a different way. E.g. nodes of basketball group may more actively send onwards and receive information about new basketball products, matches or news.

The use of memberships allows us to model the propagation of different kinds of content in network that consist of different groups. The membership raises many questions: how do nodes, which dynamically connect and disconnect the network, get to know each other? Do we need a specific network of networks that keeps information of nodes in it? How many nodes one node can get known to?

The memberships may also change and thus the users may move from one group to another in the course of time. The dynamic membership may not be easy target for simulation.

Some nodes in a network can also have more **credibility** than others. The information sent by a high-credibility node may be taken more seriously than the information sent by others. Different nodes may also be more **aware** and **active**. They may more actively either send or receive all kinds of information. There might also be limits on how much information one node can handle at same time. We describe the limitations in information handling as **buffering**.

2. Networks

Large, if not all, P2P networks include **routing**. This means that in addition to the direct information flow between nodes, some information may be routed from one node to another through intermediate nodes. We will not implement this feature into our model since it would increase the complexity of the model considerably. Therefore, we concentrate on modeling networks where nodes are able to communicate only with their neighboring nodes.

There may be some kind of information that will not receive others because its content will be thought as inappropriate. Let us call this property **filtering**. It is possible also to model the **reliability** of the network. The reliability analysis misses the objectives that consider the propagation of different kinds of information in the network with heterogeneous information preferring nodes. Thus, modeling this feature is also out of scope.

The **scalability** covers the subject of how the network can handle increase in network traffic and the joining and leaving of nodes in it. If the network nodes know only one other node, the network may handle greater amounts of traffic as all information received by a node is sent further only to one node. When one node leaves such a network, it may result in reducing the formerly bigger network to two isolated smaller networks.

II. About group work

We have three different roles in our group:

Research group will look through the literature and look for material that could be used in our modeling process. This is really important job because we are not willing neither to build a model nor get results that already are available. It is also important to keep an eye on the *development group* so that they will not get lost when playing with their simulation and modeling toys. When new interesting ideas and objectives that suit the modeling process come up, it is important that they will be brought up as early as possible. The research group is formed by Emilia Etelä-Aho and Heidi Kettunen.

Development group will build up the simulation model based on the objectives set previously. Good communication with the research group is necessary. It should be kept in mind that the results, the development group is going to bring, should answer the questions set in objectives. On the other hand, the results that cannot be explained by the group should not be completely terminated without good reason. The development group has one full-time member Eero Nevalainen and a part-time member Lauri Hyry.

There is also a leader of the group. Main task of the leader is to coordinate the team so that everyone always has the newest information available and a rough idea what the others are currently doing. The group leader will also keep the client informed about the proceedings. Lauri Hyry will work as a group leader.

III. Schedule and tasks

In this chapter, we present a rough schedule for the project. The project has been divided in three parts. It is important to keep the client aware of all the important decisions and proceedings encountered during every part.

Table 1 Tasks of phase 1

Phase 1: startup
Group meeting: getting to know each other and choosing a leader First meeting with client Brainstorming Literature reviews: what has already been done? Preferably we would like to do something new Locking the direction: simple model - what do we want get out from it? Biggest challenge: choosing right direction
Phase ends - project plan is published

Table 2 Tasks of phase 2

Phase 2: hands-on
Literature reviews: can we achieve necessary goals with our model. Do we want to extend our model? Is the core of our model sufficient? Bringing up new ideas and questions Keeping in good touch what is going to come out from our model; what do we need more? Start analyzing first results
Biggest challenge: keeping feet on ground while scanning the horizon
Choosing relevant tools First simulation model Deepening and expanding our model First results from model
Biggest challenge: keeping the expandability and achieving critical mass
Phase ends - midterm report is published

During the second phase, we have separated tasks for both the research and development groups. Also, the first simulation model will be built. Some time must be spent for choosing the right tools: it is important to choose tools that allow the building of MP2P network of hundreds of nodes. Failure in choosing a relevant tool would be harmful as moving to a different tool later was really time-consuming. The model built and its results must be communicated to the whole group before proceeding towards the second and third objectives.

The properties of nodes and network must be chosen on some level from the beginning. It should be possible to expand the model later so that we can move towards the second and third objectives or choose other objectives if better ones come up (client opinion is important). It is important that the research group will constantly bring up new ideas and analyze current results to make sure the direction we are moving to is right.

It is, of course, not possible to keep the whole group aware of all the details but the different units should have good understanding on the *overall picture*. The overall picture should be composed and updated by the group leader. The second phase ends to the review of mid-term report.

Table 3 Tasks of phase 3

Phase 3: results
Analyzing documenting modeling results Explaining the result Deepening our understanding on the subject Writing the report
Biggest challenge: bringing up real results with statistical significance
Extending the model Documenting the model
Biggest challenge: verifying the model
Phase ends - final report is published

The most important task of the third phase is to analyze the results of the simulation model and write the final report. The focus of the final report will be on presenting the results of the more complex simulation models and cases.

IV. Risks

The research area, in general, is quite complex. Therefore, there are many reasons that may cause the project to fail. In the following sections, we address some of the most important ones.

i. Time

When dealing with complex issues, like MP2P modeling, it is always easy to forget important issues or lose the overall picture. This can happen easily if one concentrates on the subject only now and then – short periods of time. Thus, there should be a good overview on the work kept by the leader.

ii. Subject

As we were asked to define the objectives of this project quite freely, it is possible that we are trying to model too complex networks. We have tried to minimize this risk by analyzing the set objectives carefully. However, it is possible that we have missed some important aspects which may later complicate the modeling considerably.

iii. Tools

The complexity of the objectives, we are trying to model, brings lot of demands to the tools used. In order to achieve statistical significance, the networks should consist of at least several hundreds of nodes, and there should be a way to generate these networks at a short notice. It is not sure, if there are right tools available for this task.