# Neighborhood Matchmaker Method: A Decentralized Optimization Algorithm for Personal Human Network

Masahiro HAMASAKI[1] and Hideaki TAKEDA[12]

[1] The Graduate University for Advanced Studies,
Shonan Village, Hayama, Kanagawa 240-0193, Japan,
[2] National Institute of Informatics,
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan,

**Abstract.** In this paper, we propose an algorithm called *Neighborhood Matchmaker Method* to optimize personal human networks. Personal human network is useful for various utilization of information like information gathering, but it is usually formed locally and often independently. In order to adapt various needs for information utilization, it is necessary to extend and optimize it. Using the neighborhood matchmaker method, we can increase a new friend who is expected to share interests via all own neighborhoods on the personal human network. Iteration of matchmaking is used to optimize personal human networks. We simulate the neighborhood matchmaker method with the practical data and the random data and compare the results by our method with those by the central server model. The neighborhood matchmaker method can reach almost the same results obtained by the sever model with each type of data.

## 1 Introduction

Information exchanging among people is one of powerful and practical ways to solve information flood because people can act intelligent agents for each other to collect, filter and associate necessary information. The power stems from personal human network. If we need variable information to exchange, we must have a good human network.

Personal human network is useful for various utilization of information like information gathering, but it is usually formed locally and often indepently. In order to adapt various needs for information utilization, it is necessary to extend and optimize it. In this paper, we propose a network optimization method called *"Neighborhood Matchmaker Method"*. It can optimize networks distributedly from the arbitrarily given networks.

## 2 Related Work

There are some systems to capture and utilize personal human network in computer. Kautz et al. [1] emphasized importance of human relations for WWW and showed done primary work for finding human relations, i.e., their system called ReferralWeb can find people by analyzing bibliography database. Sumi et al. [2] supported people to meet persons who have same interests and share information using mobile computers and web applications. Kamei et al. supported to form communities using visualization relationship among participants[3].

In these systems, they assume a group as a target either explicitly or implicitly. The first problem is how to form such groups, especially how we can find people as members of groups. We call it "meet problem". The second problem is how to find suitable people in groups for the specific topics and persons. We call this problem "select problem". The bigger group is the more likely to contain valuable persons to exchange information. However, we have to make more efforts with these systems in order to select such persons from a lot of candidates in the group. It is difficult for us to organize and manage such the large group.

Therefore information exchanging systems should support methods that realize the above two requirements i.e., to meet and select new partners.

## 3 Neighborhood Matchmaker Method

As we mention in the previous chapter, if we need better relationship for information exchanging, we must meet and select partners more and more. It is a big burden for us, because we should meet all the candidates before we select them in advance. Since we do not know new friends before meeting them, we have no ways to select them. How can we solve this problem in our daily life? The practical way is introduction of new friends by the current friends. It is realistic and efficient because the person who knows both can judge whether this combination is suitable or not. Friends work as matchmaker for new friends. We formalize this "friends as matchmaker" as an algorithm to extend and optimize networks.

The key feature of this approach is no need for central servers. The benefits of this approach are threefolds. The first is to keep spread of information minimally. Information on a person is transferred to only persons connected to her/him directly. It is desirable to keep personal information secure. The second is distributed computation. Computation to figure out better relationship is done by each node, i.e., computers used by participants work for it. It is appropriate for a personal human network because we do not have to care the size of network. The third is gradual computation. The network will be converged gradually so that we can obtain the optimal network to some extent even if we stop the computation anytime.

## 4 Formalization

In this chapter, we introduce a model that can optimize networks by formalizing the method in our real life. We call that method *"Neighborhood Matchmaker Method (NMM)"* hereafter. Before explaining *NMM*, we define the network model for this problem. At first we define a person as a node, and a connection for information exchanging between people as a path. Here we assume that we can measure a degree of the connection between two nodes (hereinafter referred to as "connection value"). Then, we can define that making a good environment for information exchanging is optimizing this network. In *NMM*, the network is optimized by matchmaking of neighbor nodes.

We need the following two conditions to apply *NMM*.

- All nodes can possibly connect to each other
- All nodes can calculate relationship between nodes connected to them

A summary these conditions, all nodes can act as matchmakers for their connected nodes to improve the connection network. The behavior of a node as a matchmaker is as follows.

1. Each node calculates connection values between its neighbor nodes. (We call this node "matchmaker")
2. If it finds pairs of nodes which have good enough connection values by computation of connection value, it recommends them i.e., it tells each element of the recommendation pair that the pair is a good candidate for connection.
3. The node that receives recommendation decides whether it accepts or not.

We can optimize personal human network by iteration of this behavior. Figure 1 shows these behaviors. In the next chapter, we test this method with simulations.

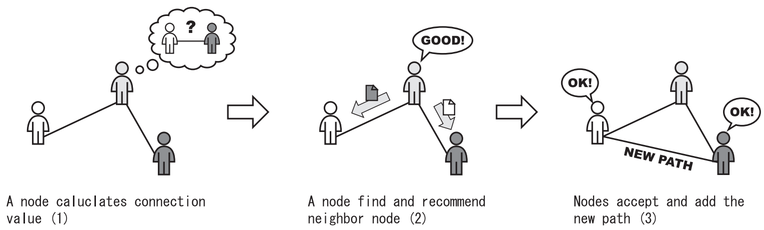


**Fig. 1.** behavior of nodes

## 5 Experiments

Since *NMM* just ensures local optimization, we should investigate the global behavior when applying this method. We test the method by simulation. We simulate optimization with *NMM* using the random data and the practical data.

### 5.1 The Procedure of the simulation

In the previous chapter, we introduce *NMM* as the three steps, but the third step, i.e., decision is free to choose any tactics for recommendation nodes. In the simulation, we choose a simple tactics. Each node wants to connect to other nodes that have better connection values i.e., if a new node is better in connection than the worst existing node, the former replaces the latter.

Figure 2 shows the flow chart of this simulation. At first, we create nodes each of which has some data to represent a person. In this experiment, the data is a 10-dimensional vector or WWW bookmark taken by users. We initially put paths between nodes randomly. We fix the number of paths during simulation. It means that addition of a path requires deletion of a path.

One node is selected randomly and exchanges paths in every turn. In this simulation, all nodes take the following tactics for exchanging paths. A node must add the best path recommended by matchmakers. If a node adds a path, it must remove the worst path instead. So that, the size of paths in the network is fixed. The adding path must be better than the worst path already had. If all nodes cannot get a new path using matchmakers, the network is converged. At that time, this simulation is concluded.

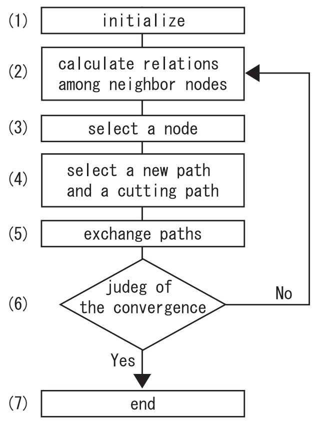


**Fig. 2.** Flow chart of the simulation

### 5.2 The Measurement

Since the purpose of the simulation is how our method achieves optimization of the netowrk, we should define what is the optimized network. We adopt a simple criterion. The best network for $n$ paths is the network that includes $n$

best paths in connection values [3]. A good news is that this network can be easily calculated by collecting and computing information for all nodes. Then we can compare this best network and networks generated by our method. Of course this computation requires a central server while our method can be performed distributedly.

We compare two networks in the following two ways. One is cover rate that is how much paths in the best network is found in the generated network. It means how much similar in structure two networks are. The other is reach rate that is comparison of the average of connection values between the best and generated networks. It indicates how much similar in effectiveness two networks are. These parameters are defined as the following formulas:

$$cover\ rate = \frac{\mid \{P_{current} \cap P_{best}\} \mid}{N}$$

$$reach\ rate = \frac{\sum_{l=1}^{N} f(p_l | p_l \in \{P_{current}\})}{\sum_{m=1}^{N} f(p_m | p_m \in \{P_{best}\})}$$

| | | |
|---|---|---|
| $p$ | : | a path |
| $N$ | : | the size of paths |
| $\{P\}$ | : | a set of paths |
| $\{P_{best}\}$ | : | the best set of paths |
| $\{P_{current}\}$ | : | the current set of paths |
| $f(p)$ | : | a value of path |

## 6 Simulation Results

There are two parameters to control experiments. One is the number of nodes and the other is the number of paths. In this experiment, we set the size of nodes from 10 to 100 and the size of paths from the 1 to 5 times the number of nodes. The simulation is performed 10 times for each set of parameters, and we use the average as the results.

The graphs in Figure 3 plot the average of cover-rate against turn. Figure 3-a shows the results when the size of paths is fixed as thrice and Figure 3-b shows the results when the size of nodes is fixed to 60.

In our formalization, we cannot know whether the network will converge. However, we can see that all graphs became horizontal. It implis that all networks were converged using matchmaking. And we can find the average of measurements and the turn of convergence are effective the size of nodes and paths.

We observed similar results on reach-rate. The difference from reach-rate is less dependent on size of paths and nodes.

[3] This criterion may not be "best" for individual nodes, because some nodes may not have any connections. We can adopt other criterion if needed.

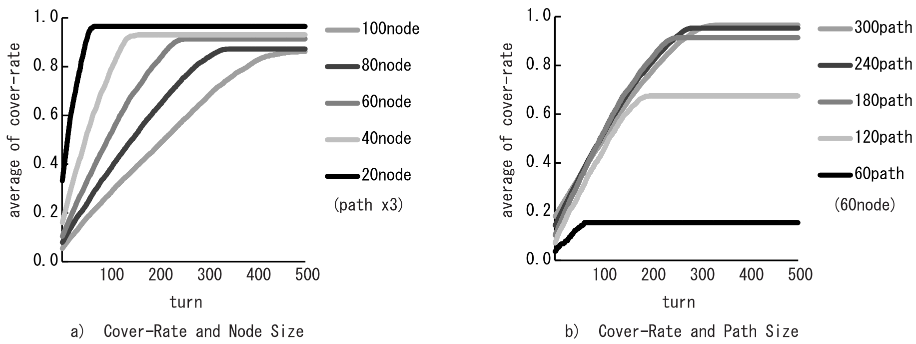


**Fig. 3.** Cover-Rate in the random data

We also examine the relevance between the size of networks and the turn of convergence. After iteration of simulation varying size of nodes and paths, we obtain the graph in Figure 4 plots the average of convergence turns against the size of nodes. This graph indicated that the turn of convergence increases linearly when the size of nodes increases. In this simulation, only a single node can exchange paths in a turn, so the times of exchanging per node do not became so large.

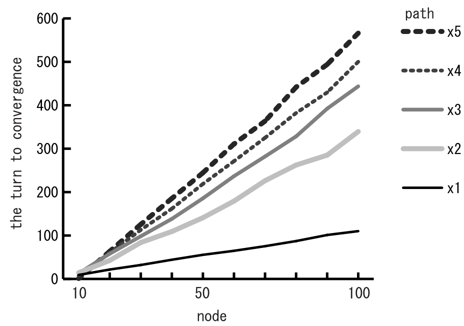


**Fig. 4.** Average of Convergence Turn

Let me estimate the complexity computation of the algorithm roughly. When the average of the number of neighborhood nodes is $r$, this algorithm calculates connection values $2r$ times in every turn. When the size of nodes is $N$ and the number of turns of convergence is $kN$ according to Figure 4, the calculation times to converge is $2rkN$ using *NMM*. In the centralized model the calculation times is $N^2$ because we have to calculate connection values among all nodes.

Since $r$ and $k$ are fix value, the order is $O(N)$ using *NMM*. It is less than $O(N^2)$ using the centralized model.

We also used the practical data generated by people. We use WWW bookmarks to measure connection values among people. Users always add a web page in which she/he is interested, and organize topics as folder in WWW bookmark. So it can be said that WWW bookmark represents the user profile. In this simulation, we need to calculate relationship between nodes. We use a parameter called *"category resemblance"* such as a value of relationship between nodes [4]. This parameter is based on resemblance of folder structure of WWW Bookmark. We examine the average of measurements and convergence turns. We found that there is the similar tendency with the random data. These results indicate that the network could be optimized in the practical data.

## 7 Conclusion

In this paper, we propose the way to obtain a new person who is a partner for exchanging information and proposed a method called *"Neighborhood Matchmaker Method (NMM)"*. Our method use collaborative and autonomous matchmaking and do not need any central servers. Nevertheless, by examining our experiment results, the optimal personal human network can be obtained. In this simulation we need the number of paths that is 2 to 3 times of the number of nodes and the number of turns that is 1.5 to 2 times the number of nodes in order to optimize the network sufficiently.

It is applicable to any size of community, because it calculates relationship among people without collecting all data at his server. It is possible to assist bigger groups that are more likely to contain valuable persons to exchange information. And it is less computational cost. Furthermore it is an easy and quick method because we can start up anytime and anywhere without registration to servers. We can assist to form dynamic and emergent communities that are typical in the Internet.

Now, we are developing the system using this proposed method. It is the system for sharing hyper-links and comments. In the real world, personal network changes dynamically through the exchanging information among people. A further direction of this study will be to experiment with this system and investigate effectiveness for it in real world.

## References


1. H. Kautz, B. Selman, M. Shah. ReferraWeb: Combining Social Networks and Collaborative Filtering. In *the Communications of the ACM*, Vol. 40, No. 3, 1997.
2. Y. Sumi, K. Mase. Collecting, Visualizing, and Exchanging Personal Interests and Experiences in Communities. *the 2001 International Conference on Web Intelligence (WI-01)*, 2001.
3. K. Kamei, et al. Community Organizer: Suppoting the Formation of Network Communities throuh Spatial Representation. In Proceedings of *the 2001 Symposium on Applications and the Internet (SAINTÓ1)*, 2001.
4. M. Hamasaki, H. Takeda. Experimental Results for a Method to Discover Human Relationship Based on WWW Bookmarks. In Proceedings of *the Knowledge-Based Intelligent Information & Engineering Systems (KESÓ1)*, 2001.