Improving dynamic skills searching in virtual social communities using agent's network self-organization

A. Mercier and
M. Occello and
J.P. Jamont

LCIS Lab - Grenoble University
50 rue Barthelemy de Laffemas
BP 54 - 26902 Valence Cedex 9

annabelle.mercier@iut-valence.fr *, michel.occello@iut-valence.fr, jean-paul.Jamont@iut-valence.fr

Abstract: Our paper will focus on the use of a multi-agent approach for the introduction of a member in a social network. As for Multi-Agents Systems, a global view of a social network is not available and furthermore undesirable. To join a network, each member must build a local network representation according to its needs. Our contribution focuses on the detection of thematic areas to assist the introduction of a new user in the social network. To do so, we must define and identify the parts of the system (agents, groups of agents) that represent the interesting nodes for this user. Our approach is based on a decentralized interrogation of the network and on organizational structures detection.

Keywords: social networks; skill searching; multi-agent systems; self-organization.

1 Introduction

The communications on the Internet can be tracked to represent the social activity of people. A communication can be viewed from a person to another, from a person to a set of people and vice versa. For example, communications can be email, downloads on peer-to-peer network, publications or navigations on the web. Being a member of a social network which is the new popular web application, becomes the new use of Internet. Building his virtual social network on the web becomes more easy than in real life. For example, it is less time-consuming to post a message to a set of contacts instead of phoning each of them during 10 minutes to give a news. Moreover, it is easy to find again forgotten contacts, to put them in the "server-memory" instead of searching their postal address, sending a mail and waiting for a response. This kind of application (like Facebook) is building a private web where people expose their private life but it is also a big data set for analyse social behaviour.

Moreover, social network applications should be useful in the case of enterprise to give a map of the organization. For example, an employee can use the application to find where to put a student for a training period. With the growth of social applications, there are new data to analyse social behaviour with traditional methods described in the following but there is a lack of tools and services for these social applications. Our goal is not to understand social network from a sociological point of view but is to provide a tool to find information and member's skills in a social network. Our approach is placed on the user field : how can an user be introduced in the network ? How can he join a person or a group of the same interest ? How can he have a complementary knowledge ? How can he diffuse an information ?

Based on a multi-agent method, our method should be adapted to the large scale network like the virtual social network on the web. A social network can be viewed like a set of agents in interaction. Generally the methods in Social Network Analysis (SNA) are static : after keeping the data, or after modeling a graph of social network, clustering algorithms are used to extract for example social behaviours and communities. The means to study dynamics is to keep the depiction of social network at several times $_t$, $_{t+1}$ and to deduct some behaviours and properties of the graph. We don't provide a tool for analysing the graph properties but we try to provide a tool to assist an user to exploit his social network which manipulate the graph of the social network on line. We have a local view of the network and exploit it instead of studying the whole network. Moreover, the global view of the graph is not possible because in this kind of application there is a large-scale network where a centralized algorithm can not give result because of the large volume of data. In the literature, a social network is represented by a graph with users as nodes and links to represent the relationships Lebhar (2005). With this representation, a social network should be viewed like a set of agent in interaction, an agent represent a node and the links with the neighbourhood represents the relationship. By adopting a multi-agent method (MAS) approach, we take a decentralized view of the problem and provide a tool adapted to large scale network. In general, multi-agent method in SNA are used to model and simulate social network, our approach is quite different because we take profit of MAS to build a tool to assist the user of social network plateform. The main goal is to offer more services in such a platform.

The rest of the paper is organized as follow. Section 2 exposes our approach and shows the related work in social network analysis. Section 3 is devoted to present our self-organization multi-agents model in the context of social network and the results obtained by a simulation. Future work and the conclusion are developed in section 4.

2 Our Approach

2.1 Existing network structure representations

Social networks are a kind of large interacting networks where local interplays are more well-known than global behaviours of the network. In the literature, there are a lot of methods to model and analyse big networks. Models' aim is to build mathematical objects reflecting real networks like social networks, electrical distribution networks, etc. First of all, each model adopts a graph representation with nodes and link between nodes. For Lebhar (2005), these models reproduce three main real networks properties: (i) small diameter, (ii) distribution degree follows a power low, and (iii) node clustering.

The small diameter represents the largest of the shortest path between two nodes of the network, it was observed that the diameter is about the logarithm of the number of nodes. Small world model, a most popular model in social network, is based on this idea Watts and Strogatz (1998).

Random graph of Erdos and Renyi (1960) is the first one for modeling real networks. It is used for social network analysis and it follows the second property.

Clustering coefficient which represents high local node density is introduced also in the model of Watts and Strogatz (1998).

Models are useful for networks analysis and simulation. Amblard (2002) classifies the models in three categories. The theoretical methods try to build graphs with properties of real network. First, theoretical models see the network as an exogenous entity without making differences between links types. There are cellular automata Wolfram (1986), random graph Erdos and Renyi (1960), small word Watts and Strogatz (1998), and scale-free networks Barbarasi et al. (2000). The latter takes into account the observation about the power-low distribution node's degree, the graph is built with an iterative method. A more recent is Kleinberg's model which reflects real network properties and shows an approach for navigation in large scale interacting networks Kleinberg (2000). Then, the sociological approach adopts statistical models, the focus is on the relationship between network members. There are Holland and Leinhardt (1981) model, metric model Banks and Carley (1994), triad completion model Heider (1958), and variance degree Blau (1967). Finally, there are methods based on agent theory Yolum and Singh (2003); Barton and Allan (2007). The agent can keep the members behaviour's, these methods are used for agent-based social simulation Falcone and Castelfranchi (2001). Other methods based on agent theory focuses on recommandation and reputation in social networks Sabater and Sierra (2002); Pujol et al. (2002). To summarise, most of these methods are used to simulate social networks by doing link analysis. Most of them don't take into account the member changes, the analysis is static. In our approach, proposed below, the focus is not on network analysis and members behaviour but on functionalities offered to social network plateform users. Face to the scale of the network, our approach is based on multi-agent theory and is decentralized. Positioned on a node which represents a member, we propose a tool for an user to transmit a message or search a skill with his local information and environment. In other words, we add a layer to system which deals with social network. This layer forgives new functions or services adapted to the user.

2.2 From a global point of view to a member point of view

An user of the social network system is represented by a node. With his neighbourhood - that is to say nearest nodes -, such a member can have access to services, functions and knowledge provided by the neighbours. We do not compute a path with the information of the whole graph but we are going to use the agent node's local view. For example, by resolve a query, the tool allows to find the border of the community which can answer the information need. So the tool is able to discover a community with a little semantic. Another tool characteristic is to take into account the member's changes. For example, a member can break connections to other, add knowledge or services. The new services bringing by the nodes can be accessed in real time, in this way, the tool should follows dynamic evolution of the agent community. If a service appears on a node, it can be accessed if it is queried. We haven't to get a photography of the whole graph but with the local information we have a way to access at the service. We don't make a static and centralized analysis of the whole graph but we know that we are able to get the information in a decentralized way from a node. This approach is adapted to a large scale graph. Moreover, it provides a decentralized tool to search knowledge based on the member's relashionship and the member's knowledge which represents the semantic of the nodes.

For the moment, the focus is only on finding the service or a way to the reach the service. The service relevance defined in information retrieval or the service recommandation level like in multi-agent system should be aborded in perspective.

The network structure emerges in function of the knowledge so the semantic on the node is exploited.

2.3 An adaptive network structure representation

The network topology influences the efficiency of the member's requests Gaston and Des Jardins (2005). Several approaches try to model real networks with mathematical objects ; this type of graph are used for static link analysis. Contrary, our approach is devoted to follow the member's changes dynamically so the network topology has to be adaptive. Moreover, this topology should be built with local node informations in the context of large scale network. The MWAC model Jamont and Occello (2007) constructs a self-organized network which can be used in the field of social network, section 2.4.2 describes the modeling of social network with the MWAC model. Without global information of the social network, the structure network emerges from the nodes informations. It is not a thematic communities discovery but a network topology discovery. Then with this topology, routing queries and searching thematic communities should be efficient dynamically. In the opposite, several approaches are devoted to find communities by using a set of fixed data. For example, Diehl et al. (2007) is based on an email corpus analysis to find the communities' borders. Mobasher et al. (2000) do the discovery with log analysis. These kind of methods are not adapted to follow a member activity in real time in a large scale network. For scaling issues, a decentralized view is more adapted. The social network is represented with a set of nodes in interaction. A member in the social network is represented by a node which brings informations or skills, internal behaviour and links to others members (nodes). Step by step, a path to others nodes can be build to access needed services or skills.

2.4 A model to organise dynamically a social network

2.4.1 Agentifying the social network

An agent is a software entity evolving in an environment that it can perceive and in which it acts. It is endowed with autonomous behaviours and has objectives. Autonomy is the main concept in the agent issue: it is the ability of agents to control their actions and their internal states. The autonomy of agents implies no centralized control Wooldridge (1999).

A multiagent system is a set of agents situated in a common environment, which interact and attempt to reach a set of goals. Through these interactions a global behaviour, more intelligent than the sum of the local intelligence of multiagent system components, can emerge.

We consider that each member of the social network is an agent.

Each member holds services or knowledge. In a first approach, each service will be described by a set of keywords. Each agent will have a list of services :

$$service\_list = < [keyword]^*, service\_name >$$

The objective for a node is to look for other nodes able to achieve services or to give the information.

A privacy ratio p between the number of services or information the agent accepts to share and
$p = (nb\ public\ services \div total\ nb\ services)$

We define a perception field for each agent as all the agents it can perceive i.e. all the agents it is friend with. We assume that a social member is at least friend with the agent which has invited him in the network or with the agents of the interest group it joined at its introduction into the network.

Each agent owns knowledge about itself and about the others, as for example the list of friends.

We can define the sociability score S as proportional to the number of related agents:

$$S = K \times number\ of\ friends$$

For each node, we can define a workload whose objective is to express the activity of a member i.e. the computational resources it consumes (if artificial) or the free time it can have (if human).

The activity can be calculated from the current number of services an agent supply to its neighbours (through services relations). For human assistant agents we can measure this workload through the number of connections to the network on a given time period.

The higher a node activity is, the less time remains to it to contact friends.

We can take into account this availability to calculate the K value.

2.4.2 Adapting the MWAC model

In opposition to the case of a programmed functionality, the designer does not need to consider all the possibilities the system has to react to each situation.

The objective is to reduce the number of requests induced by the most simple topologies currently used in social networks (as simple graphs). We will use an organization structure based on the roles of the agents in requests transmissions. Our organizational basic structures are constituted by (see fig 1):

- one and only one *group representative agent* (r) managing the communication in its group,
- some *connection agents* (c) which know the different representative agents and can belong to several groups,
- some *simple members* (M) which are active in the communication process only for their own tasks (They do not ensure information relay).

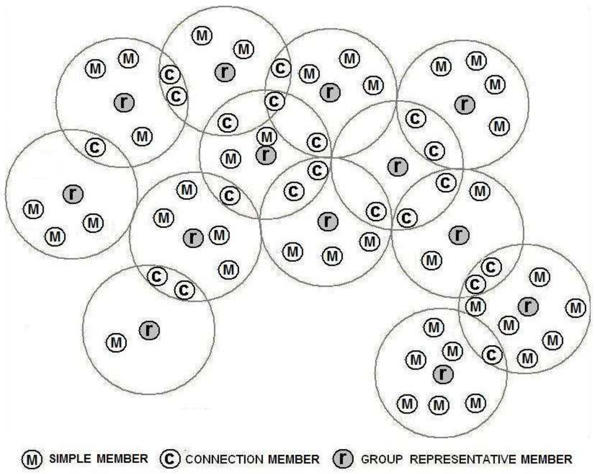

Figure 1 Our organizational structure

With this type of organizational structure, the message path between the source, a simple member (a) and the receiver a simple member (b) is $(a, r),^* [(r, c), (c, r)], (r, b))$. If the source is a representative agent the first term does not exist. If the receiver is a representative agent the last term does not exist.

The message sending saving comes owing to the fact that the sending is only directed to the representative agent of the groups and to some connection agents. To be more specific, a receiver path research with flooding techniques will cost, in the case of a traditional network, a number of emissions equals to the number of neighbours.

However, introducing in the network an organizational structure implies to build and to take care of the maintenance of routing tables. Generally, the adaptive features of these tables come from periodical exchanges between the different nodes.

In our approach we do not wish to use this technique to ensure the maintenance of the coherence. Indeed, our principle will be "if we do not need to communicate, it is useless to make efforts to ensure the coherence maintenance". However, we will thus use eavesdropping of surrounding agent communications. We extract from these message exchanges knowledge to update our beliefs about our neighbours. Moreover, our self-organization mechanism will integrate a management policy based on the node activity. The organizational structures will thus emerge.

The organizational basic structures (see fig 2) are based on three roles. A *group representative agent* managing the communication in their group. A *connection agent* knows different representative agents and they can belong to several groups. *Simple members* are active in the communication process only for their own tasks (They do not ensure information relay).

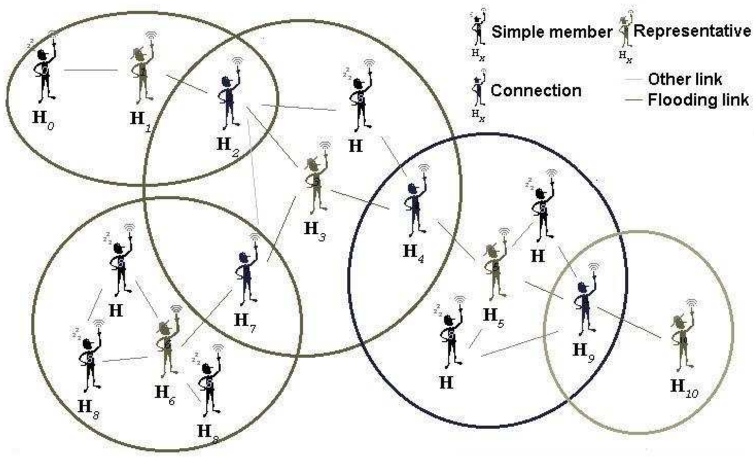

Figure 2 Group organization for communication management

A group is constituted by one *group representative agent*, some *connection agents* and possibly some *simple members*. The agents have a common scheme but their role allows them to adopt some behaviours with the messages of neighbour agents. In fact, all agents have the same introduction protocol, they work to detect and correct organization inconsistency since a message reception and want to obtain services from others.

However, on top of that, a representative agent must manage all messages sent to its group : when an agent satisfying the request is not in the sender's neighbourhood then the a route must be found by the sender's representative agents to pursue the search. It is thus more solicited than other agents. A connection agent must maintain relations between its neighbour representative agents. So, its role is to transmit its representative agent with the messages concerning recipients located in the group.

A representative agent is very solicited, from a communication point of view. It will spend on average more energy than the others to the relational activity : it will have so less time to satisfy service requests. The more members there are in a group, the greater the benefit in the reduction of transmitted volume induced by the flood is : the system thus reduces the number of messages. The ideal representative agent is the most social member (the one having the most important number of friends) and which is not overbooked, having the most important time available to solicit them (i.e. with the lowest workload). Representative have the highest centrality.

The attribution of role must integrate this information.
The organization is built according to an exchange of messages between agents. We use role allocation based self-organization mechanisms achieving the representative agent election. This function estimates the adequation between its desire to be the boss and its capacity to access to this position.

Our representative election function integrates some data on neighbours and activity levels. The decision algorithm is very simple, in case of conflict a mechanism of election is applied according to some criteria (activity level, number of neighbours...) taking the one having the high number of neighbours and the lowest workload.

```
IF neighbourNumber>0 THEN
  * One has neighbours
  IF neighnorRepresentativeNumber=0 THEN
    * None of our neighbours is representative: one decides to become one. This
    * case intervenes when one has just created the agent or when he is isolated.
    * One does not proceed to a vote because one makes the system unstable (the
    * member surely goes to carry on its path)
    myRole = REPRESENTATIVE;
  ELSE IF myRole = REPRESENTATIVE THEN
    * I am a representative agent too: I enter in conflict with the other applicants
    * to this role an election will take place and the agent with the best score will
    * remain in place.
    RepresentativeElectionProcedure()
  ELSE IF neighnorRepresentativeNumber=1 THEN
    * One of our neighbours is representative: one subjects oneself to its authority
    * and this even if the organization is less effective than otherwise. One privi-
    * leges, for the moment, stability to performance in the organization. One will
    * await the member leaves the network or wishes to leave its mandate.
    myRole = SIMPLEMEMBER;
  ELSE
    * There are, in our vicinity, several representatives: one becomes connection
    * agent for these representatives
    myRole = CONNECTION
  ENDIF
ELSE
  * One does not have a neighbour: one has no role any more
  myRole = NOTHING
ENDIF
```

This algorithm can be adjusted by other agents'suggestions such as an organization inconsistency.

In numerous approaches using organizational structure for routing the adaptivity depends on a configuration sending frequency which allows nodes to verify the data of theirs neighbours and their routing table. In our case, the organization is modified only when a problem occurs but we do not try to maintain it if we have no communication between agents. To realize this, all agents can use messages they send receive or relay from which they can extract some authorized information like the receiver, the sender, the type of message.

The messages acknowledgment is also used in this process of adaptation. When a message is not acknowledged, the representative agents concerned remove the road of their tables and a new path search procedure is started.

2.4.3 Experiments

In order to test and evaluate our approach, we have adapted the MASH simulator we presented in Jamont and Occello (2009) that we developed for the MWAC evaluation.

The simulation deals with 1200 agents following the distribution shown on figure 3.

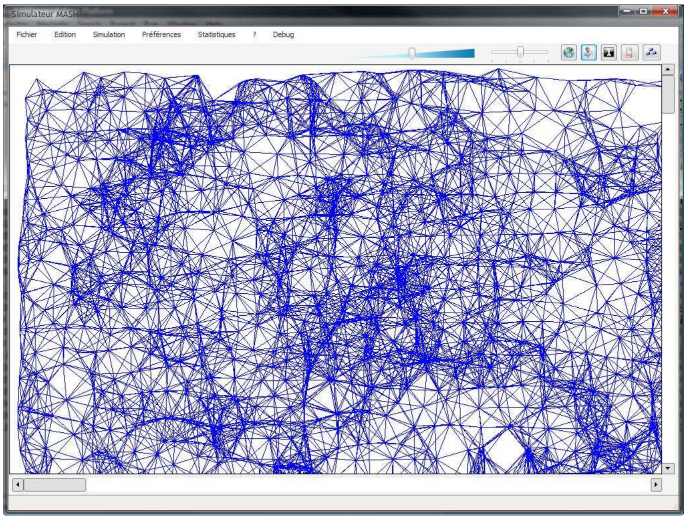

Figure 3 The social network structure

We considered that a neighbour (related by a link) is a friend. We spatializes the problem but we did not define any distance notion between the agents. The distance is always equal to 1, because the more important thing is the aquaintance.

On the figure 4 we can see the emergent topology. The remaining number of links is on this experiment a third of the initial total number. This number of link is the one really used for seeking as we are going to explain in the following section.

2.5 Skills or Information searching improvement

2.5.1 Optimising requests using an extended-MWAC organized social network

We have seen that the topology has an impact on the performance tasks performed through the network Gaston and Des Jardins (2005).

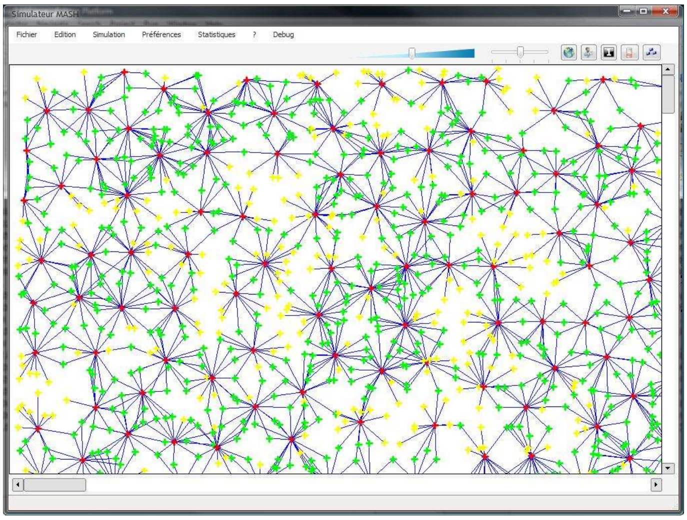

Figure 4 Simplified emergent network structure

No agent has a global view of the network and attempts to build one. Every agent in the network must be able to launch a skill, service or information search, and each request will use dynamically the network itself as a distributed database. We argue that such a dynamical decentralized approach is pertinent in a context where the network can be considered as infinite and the services location not stable.

Requests are launched on the network without any synchronisation with other members. The approach imposes to adopt an asynchronous messages exchange technique implying messages memorization.

$$msg ::=< id-msg, date, id-sender, id-receiver, status, < content >>$$

with : $status ::= request|reply$ and

$$\begin{aligned} content ::= \quad & [keyword]^*| \\ & id-req, id-req-sender, id-service-owner, service\_name \end{aligned}$$

In order to avoid loops in messages traffic each agent keeps a trace of the received messages. If a message has already been treated, it is forgotten.

To find nodes owning the right services or skills, each node must be able to achieve some of the following tasks : possessing the right service, knowing directly someone possessing the right service, being able to relay the request to another member pursuing the search (routing the request).

Moreover, agent own data depending on its role:

- A group representative agent has some information on the neighbouring group and other representative agents. They have a connection agent circular list for each neighbouring group.
- A group connection agent has some information on the neighbouring group representative.

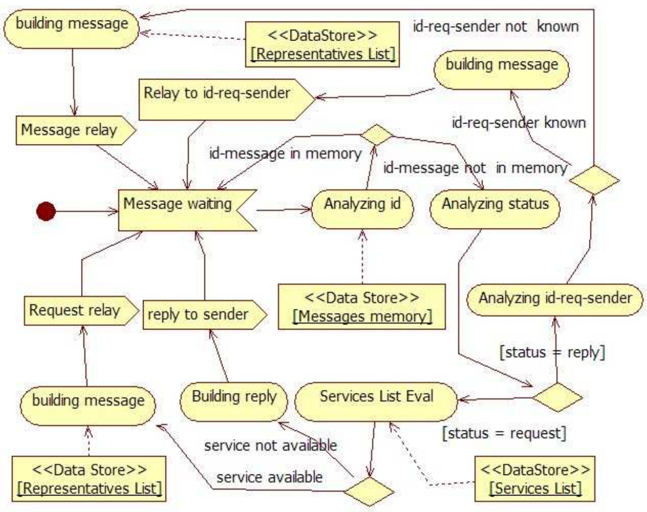

Figure 5 Message processing decision cycle in an agent

- A simple member, as any agent, has a neighbours'table with the identifiers of the members, their role and their group number.

2.5.2 Experiments

After having built an initial topology using the MWAC mechanisms detailed previously, the aim is to use the network to search services, skills or information.

Using the simulator and the population (1200 members) build previously, here are the experimental conditions and some results.

We affected forty services to each agent. At each iteration three skills or services search are launched on agents randomly chosen (25 to 100 agents by iteration). We considered hundred iterations.

We have for objective to minimize the number of message circulating between the members for a service search.

We choose to compare the efficiency of the approach to a Time To Live (TTL) flooding classically used on Internet Routing. What we call flooding TTL is a flood which seeks jump by jump. If you can not find what you are looking at a jump, we go to 2 jumps and so on. Normally everything is less than five jumps. We express the benefit by calculating the ratio between the total number of messages relayed to identify a service (and not all the members that have the service) using extended-MWAC and TTL flooding solutions.

The figure 6 shows interesting results since MWAC uses up to hundred times less messages. This result is issued from a first phase of experimentation including only one simulation with only one topology. Studying an average on a great number of initial topology could give a more precise evaluation. We can notice too that we did not take into account in the performance evaluation the duration needed to create the initial organization.

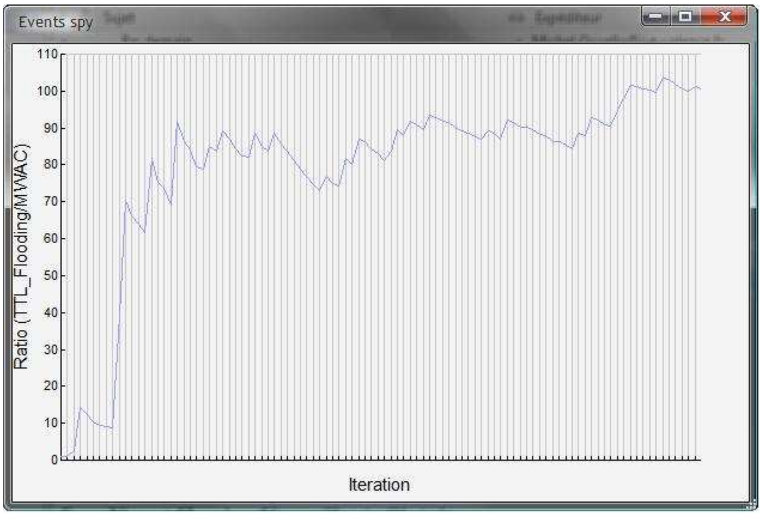

Figure 6 Skills searching experimentation results

3 Conclusion

This paper attempted to propose a model as a base for an innovative social software for services or skills seeking for social network tools.
The main idea was to make an on-line analysis of the network to find interesting nodes, using non explicit groups on-demand emergence to manage the search process.

The advantages of this approach using an adaptation of the MWAC model are :

- the local vision of the network,
- the independence of nodes (allowing a physical decentralization),
- the absence of path (definitively defined) to take into account user preferences and available members skills which can evolve rapidly.

We presented a first attempt for a model and a first evaluation. We can improve experimental studies but current results are already promising. Minimizing the number of visited nodes aims to improve the search activity more efficient and thus to enhance the quality of service.

To improve the model, an issue would be the transposition of this MAS model to a semantic level. In a first phase, the model can be improved by

- introducing semantic information in the choice of receiver,
- completing the friend definition by limiting it to a given subjects by envisaging valuation of the relation defined according to dedicated themes or interests for example,
- adding a memory of requested keywords with the recipient having declared their interest,

- introducing recursion in the multi-agent model to improve efficiency and representation.

With the introduction of semantic in the model, we will be able to focus on the service relevance and not only on searching the service. Like in information retrieval, a dynamic score, based on semantic information should be calculated for each agent's service and should be used to choose the way to the best service or skill in the social network.

This work could finally be the base for further applications. Considering the social networks under an artificial angle, this approach can be used for web services discovery. By exploiting in an adapted way the requests, and using the recursion properties it can be extended toward a tool usefull to cartography and to visualize large social networks.

Acknowledgements

A part of this research is supported by the Web Intelligence Project of research cluster ISLE of *Region Rhone-Alpes*, France.

References

Amblard, F. (2002). Which ties to choose ? a survey of social networks models for agent-based social simulations. In *International Confrence On Artificial Intelligence, Simulation and Planning in High Autonomous Systems*, pages 253–258.

Banks, D. L. and Carley, K. M. (1994). Metric inference for social networks. *Journal of Classification*, 11:121–149.

Barbarasi, A.-L., Albert, R., and Jeong, H. (2000). Scale-free characteristics of random networks : the topology of the world wide web. *Physica A*, 281:69–77.

Barton, L. and Allan, V. H. (2007). Methods for coalition formation in adaptation-based social networks. In Klusch, M., Hindriks, K. V., Papazoglou, M. P., and Sterling, L., editors, *Cooperative Information Agents XI, 11th International Workshop, CIA 2007, Delft, The Netherlands, September 19-21, 2007, Proceedings*, volume 4676 of *Lecture Notes in Computer Science*, pages 285–297. Springer.

Blau, P. M. (1967). *Exchange and Power in Social Life.* Wiley, New York.

Diehl, C. P., Namata, G., and Getoor, L. (2007). Relationship identification for social network discovery. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, July 22-26, Vancouver, British Columbia, Canada*, pages 546–552.

Erdos, P. and Renyi, A. (1960). On the evolution of random graph. *Publications of the Mathematics Institute of Hungarian Academy of Science*, pages 17–61.

Falcone, R. and Castelfranchi, C. (2001). *Social trust: a cognitive approach*, pages 55–90. Kluwer Academic Publishers, Norwell, MA, USA.

Gaston, M. E. and Des Jardins, M. (2005). Agent-organized networks for dynamic team formation. In *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 230–237. ACM.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. Wiley, New York.

Holland, P. W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, 76:33–65.

Jamont, J.-P. and Occello, M. (2007). Designing embedded collective systems: The DIAMOND multiagent method. In *IEEE International Conference on Tools with Artificial Intelligence - ICTAI 07*, pages 91–94, Patras Greece. IEEE Computer Society.

Jamont, J.-P. and Occello, M. (2009). A multiagent tool to simulate hybrid real/virtual embeded agent societies. In *Proceedings of IEEE International Conference on Intelligent Agent Technology - IEEE/IAT'2009*, Milano, Italy.

Kleinberg, J. (2000). The small-world phenomenon : An algorithmic perspective. In *32nd ACM Symposium on Theory of Computing (STOC)*, pages 163–170.

Lebhar, E. (2005). *Algorithmes de routage et modeles aleatoires pour les graphes petits mondes*. PhD thesis, Ecole Normale Superieure de Lyon.

Mobasher, B., Cooley, R., and Srivastava, J. (2000). Automatic personalization based on web usage mining. *Communications of the ACM*, pages 142–151.

Pujol, J. M., Sangüesa, R., and Delgado, J. (2002). Extracting reputation in multi-agent systems by means of social network topology. In *The First International Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2002, July 15-19, 2002, Bologna, Italy*, pages 467–474. ACM.

Sabater, J. and Sierra, C. (2002). Reputation and social network analysis in multi-agent systems. In *The First International Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2002, July 15-19, 2002, Bologna, Italy*, pages 475–482. ACM.

Watts, D. and Strogatz, S. (1998). Collective dynamics of small world networks. *Nature*, pages 440–442.

Wolfram, S. (1986). *Theory and Applications of Cellular Automata*. World Scientist.

Wooldridge, M.-J. (1999). Intelligent agents. In Weiss, G., editor, *Multiagent systems: A modern approach to Distributed Artificial Intelligence*. MIT Press.

Yolum, P. and Singh, M. P. (2003). Personalized emergent communities in referral networks. In *Proceedings of the IJCAI Workshop on Intelligent Techniques for Web Personalization*.