Semantic Massage Addressing based on Social Cloud Actor's Interests

Reem M. Bahgat *r.bahgat@fci-cu.edu.eg*
Faculty of Computers and Information/
Department of Computer Science
Cairo University
Cairo, Egypt

Akram I. Salah *Akramsalah.21@gmail.com*
Faculty of Computers and Information/
Department of Computer Science
Cairo University
Cairo, Egypt

Hamada I. Abdul Wakeil *hamada.mabrouk@computer.miniauniv.edu.eg*
Faculty of Computers and Information/
Department of Computer Science
El_Minia University
El_Minia, Egypt

Abstract

Wireless communication with Mobile Terminals has become popular tools for collecting and sending information and data. With mobile communication comes the Short Message Service (SMS) technology which is an ideal way to stay connected with anyone, anywhere anytime to help maintain business relationships with customers. Sending individual SMS messages to long list of mobile numbers can be very time consuming, and face problems of wireless communications such as variable and asymmetric bandwidth, geographical mobility and high usage costs and face the rigidity of lists. This paper proposes a technique that assures sending the message to semantically specified group of recipients. A recipient group is automatically identified based on personal information (interests, work place, publications, social relationships, etc.) and behavior based on a populated ontology created by integrating the publicly available FOAF (Friend-of-a-Friend) documents. We demonstrate that our simple technique can first, ensure extracting groups effectively according to the descriptive attributes and second send SMS effectively and can help combat unintentional spam and preserve the privacy of mobile numbers and even individual identities. The technique provides fast, effective, and dynamic solution to save time in constructing lists and sending group messages which can be applied both on personal level or in business.

Keywords: Wireless Communication, Short Message Service, FOAF, J2ME

1. INTRODUCTION

Despite sophisticated collaboration environments being around for a long time already, SMS is still the main means for distributed collaboration. People still use it for maintaining to-do lists, tracking, organization, etc. The major reasons for this may be grounded in the ease of use, the negligible learning effort to be able to use it, the universal availability, and that literally everyone has a phone device uses it and can be contacted and share information and Multimedia via messages.

Telephone numbers, like email addresses, are opaque identifiers. They're often hard to remember, and, worse still, they change from time to time. Mobile numbers are a means to send SMS, but sending individual personalized SMS messages to long predefined list of customers

can be very time consuming and problem of predefined lists is its rigidity and that it requires constant maintenance. The frequency of change adds overhead to the process, wireless communications such as variable and asymmetric bandwidth, and geographical mobility and high usage costs. It would be more effective to find a way to dynamically and automatically create and update those lists.

This paper provides a solution to this SMS sending problems. Our approach is to send SMS to groups of people, either their mobile numbers are saved in or not in your mobile device, by matching a particular set of attributes, e.g., all people who are interested in marketing in an organization, or all people in a specific work place who are interested in maintenance and knows someone. In other words creating the list semantically, meaning the lists are to be built based on customer characteristics or customer behaviors. So from this point of view we need some kind of technology to ensure that the group of people is defined.

The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. Two important technologies for developing the Semantic Web are already in place: First, eXtensible Markup Language (XML). XML lets everyone create their own tags that annotate Web pages or sections of text on a page or any resource. Second, the Resource Description Framework (RDF) [1]. RDF used to express meaning of any resource and encodes it in sets of triples, each triple being rather like the subject, verb and object of an elementary sentence in which Subject and object are each identified by a Universal Resource Identifier (URI). The Resource Description Framework (RDF) technology such as the Friend-of-a-Friend (FOAF) [2] ontology that predicates a person's express properties such as name, email address, phone number, group memberships, employer, gender, birthday, interests, projects, and acquaintances, and also using the Meaning-of-a-Tag (MOAT) [3] Ontology represents the meaning of user resources.

Our method can be summarized as follows:

- We introduce a technique that utilizes Semantic Web metadata as well as social network techniques to discover a semantically specified group (community or group). Expansion of people extraction is out of the paper scope.
- We propose a technique to send SMS to semantically specified group of recipients.
- We conduct extensive experiments to demonstrate that our technique is effective.

The paper is structured as follows: Section 2 explains the preliminaries for the paper as follows: Subsection 2.1 gives an overview of how the recipients are extracted dynamically based on personal characteristics or behaviors. The Semantic Extraction Method is provided in Subsection 2.2. The Proposed Technique used for Sending Semantic SMS is introduced in Subsection 2.3. Subsection 2.4 shows our experimental results. Finally, conclusion and future work are presented in Subsection 2.5.

2. PRELIMINARIES

2.1 People Extraction Method

Recently, people use social networks to share the interests and contents, such as bookmarks, web blogs, questions/answers, photographs, music, and videos. Building user communities of the same interests, finding the domain experts in different subjects, identifying hot social topics, and recommending personal relevant contents is a fundamental problem of social networks.

Sending SMS to a specific group or community of people requires a method for collecting this group. There have been a plenty of methods and schemes aiming to find users with common interests. Zhao et al. [4] proposed a framework that focused on community extraction based on the strength on ties between members of a community and its ties to the outside world. Li et al. [5] proposed a social interest discovery approach based on user-generated tags. Mika et al. [6]

provided a method for extraction, aggregation and visualization of online social networks based on semantic technologies. Matsuo et al. [7] proposed a social network extraction system called *POLYPHONET*, which employs several advanced techniques to extract relations of persons, detect groups of persons, and obtain keywords for a person. Yan et al. [8] proposed an approach for community discovery based on the contents of social actors' personal interests and their social relationships.

Some of the above methods doesn't use the semantic technologies to form the communities, others used the semantic technologies to form the communities based on users interests but doesn't take in account the meaning of the users interests.

2.2 Semantic Extraction Group (SEGROUP)

We introduce a novel method that employs the semantic web technologies for reasoning with personal information extracted from the Friend-of-a-Friend (FOAF) documents that predicates a person's express properties such as name, email address, phone number, interests, projects, and acquaintances, as shown in FIGURE 1:

1. FOAF and Tag dataset: this dataset is collected by spidering the Semantic Web.
2. People discovery: based on the people's interests and other personal information.

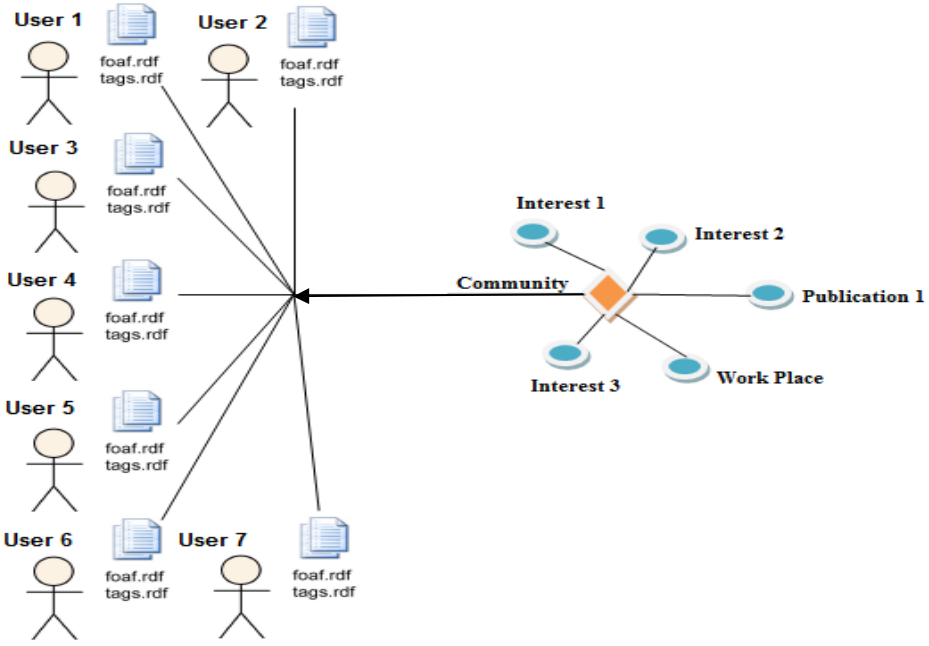

FIGURE 1: People Extraction Process

- **FOAF and Tag Dataset**

FOAF is a machine-readable ontology describing people, their activities, their relationships and objects. FOAF allows groups of people to describe social networks without a centralized database. FOAF is an extension to Resource Description Framework (RDF) and is defined using Web Ontology Language (OWL) [9].

By spidering the Swoogle Semantic Web [12] and collecting the information contained in FOAF files, we can build a large collection of data about people and their personal information. Ding et al. [10] gives a formal definition of a strict FOAF document D with the following four characteristic patterns:

1. D is a valid RDF document. This can be validated by a RDF parser.
2. D uses the FOAF namespace.

3. D contains an RDF graph pattern as shown in FIGURE 2. X and Z are two different instances of rdfs:Resource and Y is an instance of rdf:Property using FOAF namespace.
4. D defines only one instance of foaf:Person without referencing it as an object in any triples within D. D may additionally have some other instances of foaf:Person; however, each of them must be referenced as an object in at least one triple in D.

The above patterns, especially the fourth pattern, are quite strict and exclude many documents not dedicating to a person. Therefore, by removing the fourth pattern, Ding et al. [10] defined a general FOAF document as long as it contains of foaf:Person.

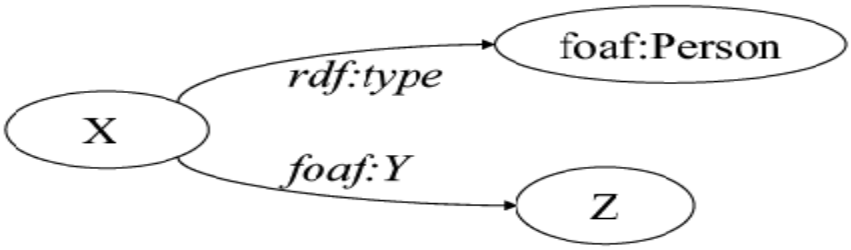

FIGURE 2: FOAF document pattern (image from [10])

- **People Discovery**

Shared Personal Information such as interests and other characteristics are a common method to discover people. But there is a problem faces the methods that depend on personal information, it is hard to extract the representative information from the FOAF properties because it is represented by URLs. One of the most important phenomena of Web 2.0 is ***tagging***, that let users play an important role in the process of creating content. Tagging goes a step further by letting them control the way they organize it. Tags generated by users represent their apprehension of the content of the websites. They are more expressive and accurate than the features extracted by machine [5]. But there are other problems faces the methods that depends on tags are the ambiguity and heterogeneity of tags and their lack of machine-understandable. Meaning can be a problem for information retrieval, especially when people use tags that can have different meanings depending on the context.

A concept of the Meaning Of A Tag (MOAT) [3] Ontology that provides a Semantic Web framework to publish semantically-annotated content from free-tagging using URIs of Semantic Web resources such as URIs from DBpedia [14] or any knowledge base. Those URIs are used to identify everything (pages, people, documents, books, interests, etc.) in a unique and non-ambiguous way. An example of the MOAT Ontology is shown in FIGURE 3.

```
<moat:Tag rdf:about="http://tags.moat-project.org/tag/paris">
<moat:name><![CDATA[paris]]></moat:name>
  <moat:hasMeaning>
    <moat:Meaning>
      <moat:meaningURI rdf:resource="http://dbpedia.org/resource/Paris_%28mythology%29"/>
      <foaf:maker rdf:resource="http://example.org/user/foaf/1"/>
    </moat:Meaning>
  </moat:hasMeaning>
</moat:Tag>
```

FIGURE 3: An example of MOAT Ontology (image from http://www.patrickgmj.net/blog/alexandre-passants-moat-meaning-of-a-tag-project)

We can use the Friend Of A Friend (FOAF) along with the Meaning Of A Tag (MOAT) Ontologies to represent the user information (foaf.rdf) and the user tags (tags.rdf), respectively and use the relationship between them on foaf:maker property in MOAT ontology to obtain related people who tagged a resource with the same tag name and to get related people who tagged a resource with the same meaning as shown in FIGURE 4 and 5 respectively. All further relevant information can also be easily retrieved in this simple way.

```
select DISTINCT ?maker ?tag_meaning ?tag_name where
       {
        ?x foaf:maker ?maker.
        ?y moat:name ?tag_name.
        ?z moat:meaningURI ?tag_meaning.
        filter(?tag_name = \""+your_tag_name+"\")
       }
```

FIGURE 4: Querying MOAT for persons tagged a resource with the same tag name

```
select DISTINCT ?maker ?tag_meaning where
      {
       ?x foaf:maker ?maker.
       ?z moat:meaningURI ?tag_meaning.
       FILTER regex(str(?tag_meaning),"+"\""+your_tag_meaning_URI+"\",\"i\")
      }
```

FIGURE 5: Querying MOAT for persons tagged a resource with the same meaning

2.3 The Proposed Technique for Sending Semantic SMS

The proposed technique has two principal processes.

1. Extracting the group(s) process **(SEGROUP)** that combines the content network and social network, and boosts Semantic Web technologies in current Web. Therefore we use the Friend-of-a-Friend (FOAF) ontology that is a first attempt at a formal, machine process-able representation of user profiles and friendship networks. FOAF profiles are created and controlled by the individual user and shared in a distributed fashion. Also the Meaning-of-a-Tag (MOAT) ontology that provides a Semantic Web framework to publish semantically-annotated content from free-tagging using URIs of Semantic Web resources in a unique and non-ambiguous way. Those two Semantic Web Projects FOAF (Friend of a Friend) and MOAT (Meaning of a Tag) can be combined to enable data portability between social media sites to allow us to create and extract the group(s) - the mailing list(s) - that define the SMS recipients.
2. Sending SMS to the selected group process. Sending SMS is no matter how it done, it can be from the mobile device directly or from joining a web site to send the messages. For this we build a server that deal with mobile program for sending.

The proposed technique uses a GUI based on J2ME [13] to set the criteria to collect the group(s) of recipient for the SMS. These criteria will send to a server build in Java. First, the server collects FOAF documents and its associated MOAT documents from the Semantic Web. Second, the server use the criteria send by the sender mobile to query the FOAF and MOAT documents for persons satisfy these criteria as shown in FIGURE 5.

The steps for the first process of the proposed technique can be described as follows:

Step 1: Collect the FOAF and associated Tags files from the Web.
Step 2: For each criterion (i.e., tag_name), query the Tags files to select similar persons who tagged their resources with the same criterion, and also select tagMeaningURI for this criterion.
- **2.1** If found then store the people's foaf.rdf URL in an array (i.e. person array) (with no duplication).
- **2.2** Use the tagMeaningURI to query tags files to select persons who tags have the same meaningURI, but not equal in tag names (i.e., the criterion).
  - **2.2.1** Do **2.1**.
- **2.3** If no person found then this criterion do not match any person.

In step 2, because the resource may have different tag name according to the person created it, so if a person found who tag a resource with the querying tag name (criterion) we select the meaningURI and use this URI to query for persons who have a tag with this meaning.

The steps for the second process of the proposed technique can be described as follows:

Step 1: Determine the criteria, the message, and the sender mobile number to select group of persons.
Step 2: For each person in person array the server query the corresponding FOAF file to select the person phone number or email address and send their names to the sender mobile or a group of the names plus the number of people if large.
Step 3: The sender mobile receives the person names whose satisfy the querying criteria, and have option to select some of them.
Step 4: The server receives the selected person names and begin to send SMS to them.

In step2, users could email large numbers of people and because each of the data ownership and authentication issues on the internet and inaccurate information. We therefore need safeguards to ensure that user errors won't result in a mass SMS spamming. Feedback to the user regarding his or her query definition could take the form of a list of people to which the SMS will be sent, assuming the number of people is sufficiently small that such a display is feasible. Otherwise, displaying the number of people to which the SMS will be sent along with a sample of those people could be useful. Such feedback would let the user catch errors in the query definition and give confidence that the SMS won't accidentally go to a large number of people for whom it's not meant.

2.4 Experimental Results

This section presents the results of evaluating the effective of the proposed technique that is based on J2ME. We consider the number of recipients as a criterion to evaluate the performance of the proposed method to generate community. The main purpose of the proposed technique is to semantically collect a specified group of recipients. We achieved this by selecting the recipients by using the proposed method for selection **SEGROUP** method. As FIGURE 6 shows the number of members in the community generated from a small tested FOAF and Tag dataset (1000 FOAF and Tag documents created manually). In clustering, the more we add factors that describe the group the more the number of members decreases.

Many of Social Network Sites provides FOAF profiles for their users in different formats such as Facebook, Twitter, Flickr, Myspace, Last.fm. And to exporting FOAF data from different sites as follows:

1. Facebook: http://www.dcs.shef.ac.uk/~mrowe/foafgenerator.html
2. Twitter: http://semantictweet.com/
3. Flickr: http://apassant.net/blog/2007/12/18/rdf-export-flickr-profiles-foaf-and-sioc/
4. And many more (Drupal 7, WordPress plug-ins, etc.).

Because of these differences we chose the format presented by the FOAF project as a standard format to create our dataset. Most of the FOAF profiles generated by social sites may not contains all properties of FOAF and also doesn't provide the tag files for their users, so we created a small FOAF and corresponding Tag dataset manually to use more of the FOAF properties. But the proposed technique can be applied on the FOAF documents collected from different social sites.

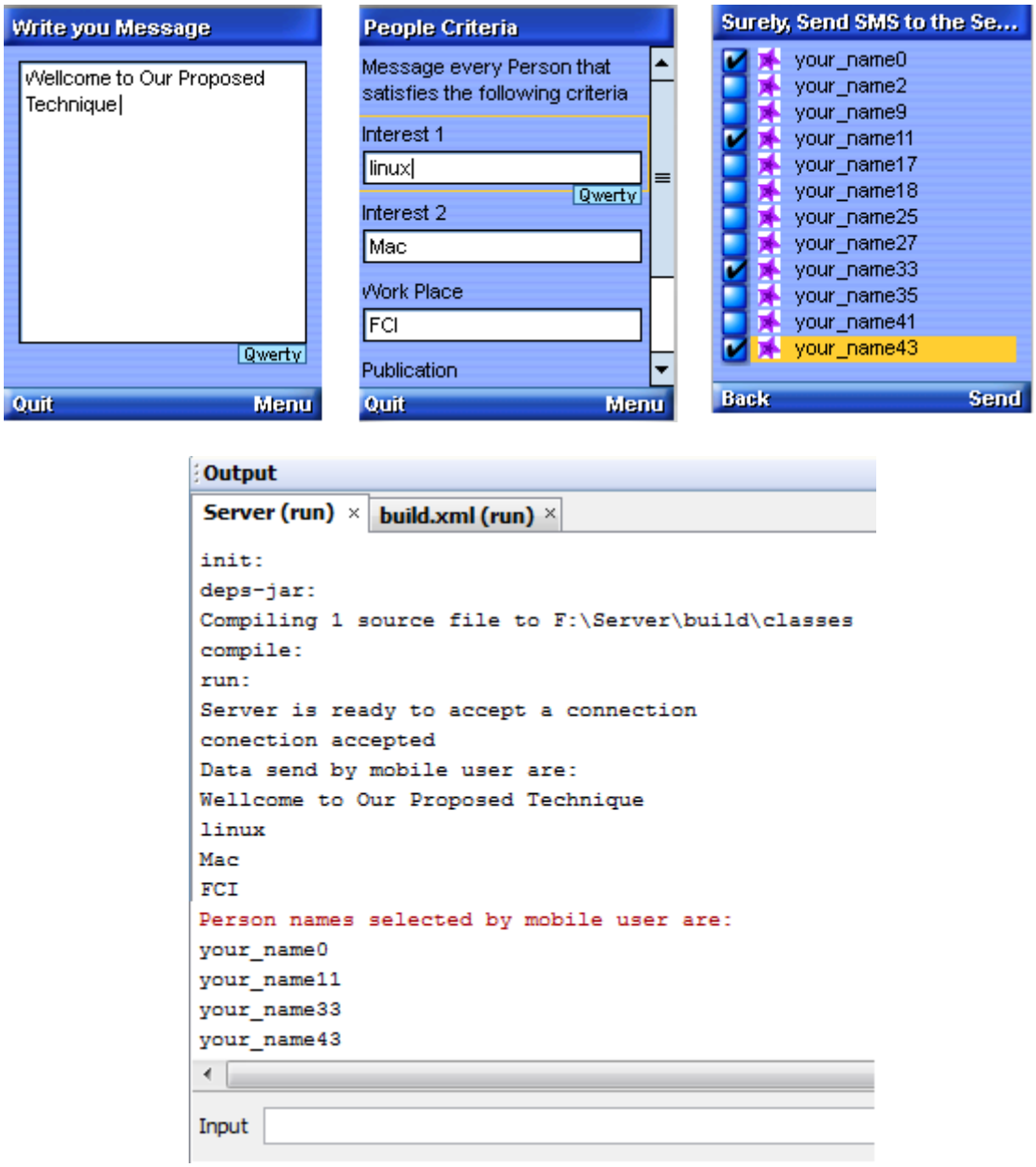

FIGURE 5: Simple example for the proposed technique

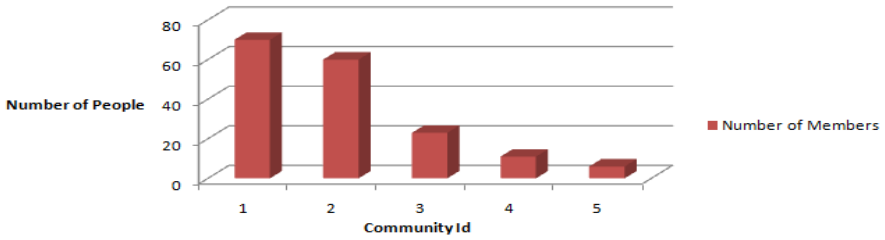

FIGURE 6: Community Size

Category	Type		System Type		User Info.					Community Types	
Method	User Centric	Object Centric	Centralized	Distributed	General Doc.	Network	Profiles	Registration Data	Others	Dynamic	Static
Community Extraction for Social Networks [4]	√		√			√					√
Tag Based [5]		√	√					√	√	√	
Collective Contexts [15]		√	√		√					√	
Polyphonet [7]	√		√		√						√
Topic and Role [16]		√	√		√				√	√	
Focused Social Extraction [17]	√	√	√		√			√	√		√
Social Actor's Interests [8]	√	√		√			√			√	
Flink [6]		√		√	√		√		√	√	
SEGroup	√	√		√			√			√	

TABLE 1: Comparing between methods

TABLE 1 summarizes the main features of some of the semantic extraction methods. The main focus is on the following aspects:

1. Method Type: Means according what the social network is established?
   - 1.1 User Centric: Means detecting social interests based on the social connections among users.
   - 1.2 Object Centric: Means detecting social interests based on the common objects fetched by users.

2. System Type: Means the user profile who control it?
   2.1 Centralized: Means the information is under the control of the database owner.
   2.2 Distributed: Means the information (profiles) are created and controlled by the individual user and shared in a distributed fashion.
3. User Information: Means the user data are collected from what?
   3.1. General Document: Means the user interests are gathered from publications, email archives, and co-authors papers.
   3.2. Network: Means the social communities are extracted from the user network.
   3.3. Profiles: Means the social interests are gathered from the user profiles or user home page.
   3.4. Registration Data: Means the interests are gathered from the user registration data in social networks.
4. Community Type: Specify the ability to update and expand the community or not.

The main ideas of Web 2.0 is to let users play an important role in the process of creating content, so comparing the **SEGROUP** method with all methods that described as centralized in the above table we have better resulting to the **SEGROUP** since it support for the Web 2.0 technologies that contribute to improve the performance of user selection.

The **SEGROUP** method like Social actor's interests [8] from TABLE 1, like the Social Actor's Interests, the **SEGROUP** use the user generated tags to represent the contents. But instead of creating the tag file set (each tag file has representative words corresponding to a foaf:interest), the **SEGROUP** takes into account the meaning of tags to overcome the tag problems, to easy to meaningfully search, and compare or merge similar collective tagging data on different sources through the using of the Meaning-Of-A-Tag (MOAT) ontology that provides a uniform structure and semantics of a set of tags and promotes their global sharing. Using the tag meaning instead of tag file set will minimize the time and effort of search to discover a community. Also lead to discover users who are closer to what we search for.

2.5 Conclusion and Future Work

In this paper a new technique for sending SMS for semantically specified group of recipients is introduced. The proposed technique based on Web 2.0 Social Network application to extract the group of recipients using the FOAF documents and using MOAT ontology to represent the meaning for the user contents. The advantage of this technique First, the technique doesn't require the user to save long list numbers in her/his mobile device. In addition, the technique doesn't requires the user to update the list individuals by subscribe or unsubscribe them. The experimental results show that the technique can send SMS effective and can help combat unintentional spam and preserve the privacy of mobile numbers and even individual identities. Future work is required to add Security, solving the issues of ownership and authentication on the internet and using the mobile Agent to overcome the wireless network problems.

3. REFERENCES

[1] S-F. Luis and F-G. Norberto. "*The Semantic Web: Fundamentals and A Brief State-of-the-Art*". The European Journal for the Information Professional Vol. VI, No. 6, December 2005.

[2] "*The FOAF Project*". Internet: http://xmlns.com/foaf/0.1/, [2011, March].

[3] A. Passant, P. Laublet. "*Meaning Of A Tag: A Collaborative Approach to Bridge the Gap Between Tagging and Linked Data*". LDOW2008, April 22, 2008.

[4] Y. Zhao, E. Levina, and J. Zhu. "*Community extraction for social networks*". arXiv: 1-15,May 18, 2010.

[5] X. Li, L. Guo, Y. E. Zhao. "*Tag-based Social Interest Discovery*". Presented at the 2nd Int. World Wide Web Conference Committee (IW3C2). Beijing, China, 2008.

[6] P. Mika. "*Flink: Semantic Web Technology for the Extraction and Analysis of Social Networks*". Elsevier Science, May 14, 2005.

[7] Y. Matsuo, J. Mori, M. Hamasaki. "*POLYPHONET: An Advanced Social Network Extraction System from the Web*". Presented at the 2nd Int. World Wide Web Conference Committee (IW3C2). Edinburgh, Scotland, 2006.

[8] F. Yan, J. Jiang, Y. Lu, Q. Luo, M. Zhang. "*Community Discovery Based on Social Actors' Interests and Social Relationships*". In Proc. 4th Int. Conf. on Semantics, Knowledge and Grid, 2008, pp. 79-86.

[9] M. K. Smith, C. Welty, and D. McGuinness. "*Owl web ontology language guide*". W3C recommendation, [On-line]. Available: http://www.w3.org/TR/owl-guide [May 13, 2011].

[10] L. Ding, L. Zhou, T. W. Finin, and A. Joshi. "*How the semantic web is being used: An analysis of foaf documents*". HICSS, 2005.

[11] A. Seaborne and E. Prud'hommeaux. "*Sparql query language for rdf*". W3C recommendation, [On-line]. Available: http://www.w3c.org/TR/rdf-sparql-query/ [Jan. 1, 2011].

[12] L. Ding, T. Finin, A. Joshi, R. Pan, R. Cost, Y. Peng, P. Reddivari, V. Doshi, and J. Sachs. "*Swoogle: A search and Metadata engine for the semantic web*". Presented at the Int. Conference on Information and Knowledge Management. Washington, DC, USA, 2004.

[13] J. Keogh. "*J2ME: The Complete Reference*", Brandon A. Nordin, 2004.

[14] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. "*DBpedia: A Nucleus for a Web of Open Data*". Presented at the 6th Int. Semantic Web Conf. November 2007.

[15] J. Mori, T. Tsujishita, Y. Matsuo, and M. Ishizuka. "*Extracting Relations in Social Networks from the Web using Similarity between Collective Contexts*". ISWC 2006, 487-500, 2006.

[16] A. McCallum, A. Corrada-Emmanuel, and X. Wang. "*Topic and Role Discovery in Social Networks*".

[17] M. Hamasaki, Y. Matsuo, K. Ishida, Y. Nakamura, T. Nishimura, and H. Takeda. "*Community Focused Social Network Extraction*". Springer Verlag, Berlin Heidelberg, 2006.