INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

A NEW APPROACH TO FILTER SPAMS FROM ONLINE SOCIAL NETWORK USER WALLS

M. Bhavish Santhosh Kumar[1], G. V. R. Reddy[2]

[1]*Computer Networks Information Security, bhavishsanthoshkumar@gmail.com*
[2] *Computer Science Engineering, gvr_reddi@yahoo.co.in*
JNTU Hyderabad, India

Abstract

Nowadays, Internet has become the part of our daily lives with many reasons. Social networking sites are the main reason to make interactions and to maintain human relationships by using the internet. To maintain healthy relationships user need to get some support from the Online Social Networks, to avoid spam messages posted on their walls. But now, due to the lack of support user can't handle these messages directly to discard from their walls. This little drawback leads to spoil the entire network of the user. I propose a system to filter spam messages from user walls which can be achieved through content-based filtering.

Keywords: spam, online social networking, messages, user walls, user control, content based filtering

1. Introduction

Technology has been developed and the internet users have been increased quite rapidly in the past 10 years [3]. Most of the internet users are spending much time on watching videos and social networking [4]. Now a day's online social networks have become the most common need to everyone since of the features facilitated by the social networks.

The features facilitated by online social networks are -

a. Multimedia :
Multimedia applications like photo sharing, video sharing, audio sharing are the elegant features.
- Photos can be shared through Flickr, Google plus, Orkut etc
- Videos can be shared via YouTube, Dailymotion etc
- Audio sharing is done with Imeem, Doregama, Songs.pk etc

b. Entertainment :
The social networks provide entertainment by playing online games
- Virtual Worlds: Second Life
- World of Warcraft, Zapak many more sites provide Online Gaming

c. News/Opinion :
News and expert opinions also useful to the users.
- Digg, Reddit, RSSfeed are the Social news providers
- Yelp, epinions are for reviews. For technical, analytical and mobile reviews (GSMARENA) also provided.

d. Communication:

This is facilitated by blogs, social networking services, and events. By this way many people can communicate each other. Technical queries can be shared through technical blogs.

- Microblogs: Twitter, Pownce
- Events: Evite
- Social Networking Services: Google+, Facebook, LinkedIn, MySpace

Social networks are used to make new relationships and to maintain relationships more effectively. Earlier no one is ready to share their views but now even a small thing can be shared to their friends. The usage of users made these social networks as powerful such that the importance of support facilitated by the online social networks to the users also increased a lot.

Since the usage of online social networks is more and it should play a vital role in users context but they failed to provide direct control to users over the messages posted on their walls. A little support is provided to online social network users to control the messages posted on their user walls. In order to provide support, a filtering mechanism could help. The filtering can be done is many ways but content based filtering is more efficient. Since to provide support over messages, depending on the content of the messages will give good results.

The filtering can be done based on content of the messages as earlier mentioned but the filtering is explained in the next chapters.

2. Related Work:

The main goal is already mentioned in the introduction section, filtering can help to design a new algorithm which could improve the user support by preventing spams on their walls. The filtering could be either content based or policy based.

Content based filtering is used to work with the words or phrases present in each individual message posted by every user, is a spam or not. Content based filtering can be defined as a Machine Learning (ML Filtering) process which learns from user's interest and recommend familiar suggestions to the user. [4] User recommendations can be shown as per the rating given by the user for particular item. The user interest might be weighted than other preferences. But here comes the challenge how to learn user interest.[2]

A word-based filtering can be the basic in Content-based filtering. Set of repeated or suspicious words can be used to compare, in the prevention of spams. In general spammers knowingly misspell the words to avoid this filtering. So there is a necessity of updating these black-words in filter by the IT people. [2]

Rule-based filters are little effective than the word-based filters, rather than using a suspicious word, multiple terms can be considered. When a message is scanned the vulgar terms can be weighted high, whenever the score reaches maximum message can be treated as spam. Large scope is there to find spams using it[2].

3. Blacklist Maintenance:

Blacklist maintenance is nothing but maintaining list of words which are used to make users to avoid the spams from their walls. The general spams on online social networks can be collected and maintained as blacklist. The list can be maintained at database but it won't be in a single list, it would be in multiple lists or tables. This blacklist can be maintained by the administrator. The administrator has the privilege to add a blacklist word i.e. spam word.

- Blacklist is the collection of spam words.
- In the database "a_table" to "z_table" tables are maintained and in "a_table" the words start with "a" can be inserted, in "b_table" the words start with "b" can be inserted. To do like this the adding spam word first character can be checked.

```
char c=spam.charAt(0);
insert into "+c+"_table values ('"+spam+"');
```

- "a_table" contains words start with character a. In the similar way every table contain the words.
  In this way the blacklist can be maintained.

4. Spam Filter Algorithm:

To design a spam filter algorithm the major thing used is text classification. When user posts any message that may be small message or big message but it needs to be split into words and then only the comparison can be done so easily.

- Step 1: User message can be split into words. Here the user message can be split into words using java-String split() method.

```
String[] m=msg.split(" ");
```

- Step 2: For every word first character can be checked

```
char c= m.charAt(0);
```

- Step 3: The words can be compared with spam words listed in the blacklist table starts with same character.

```
ResultSet rs = stmt.executeQuery("select * from "+c+"_table");

while(rs.next()) {

String em =rs.getString(1);

System.out.println(em );

if(m.equalsIgnoreCase(em)){} }
```

- Step 4: If any word matches with the black list table it can be replaced with "------" , otherwise it won't be replaced.

```
if(spm.check(m[i]))

{

 s+="---- ";

 }
```

- Step 5: After all words have been checked, the words can be merged and the message can be posted.

```
 s+=m[i]+" ";
```

5. Test results:

Coming to test results we tested this algorithm implementation. To test this here we just took *stupid* as spam word. The word was inserted into the database. The screenshot of user posting spam message is as follows.

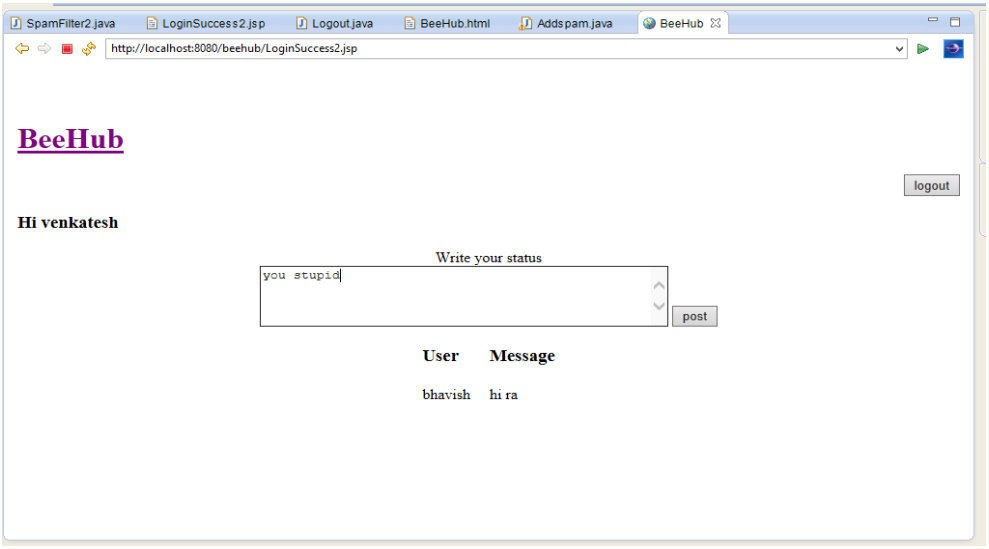

Figure: Posting a spam

After posting spam message the spam part can be replaced by "-----" and not spam part can be shown as normal no replacement is needed. The result is as follows-

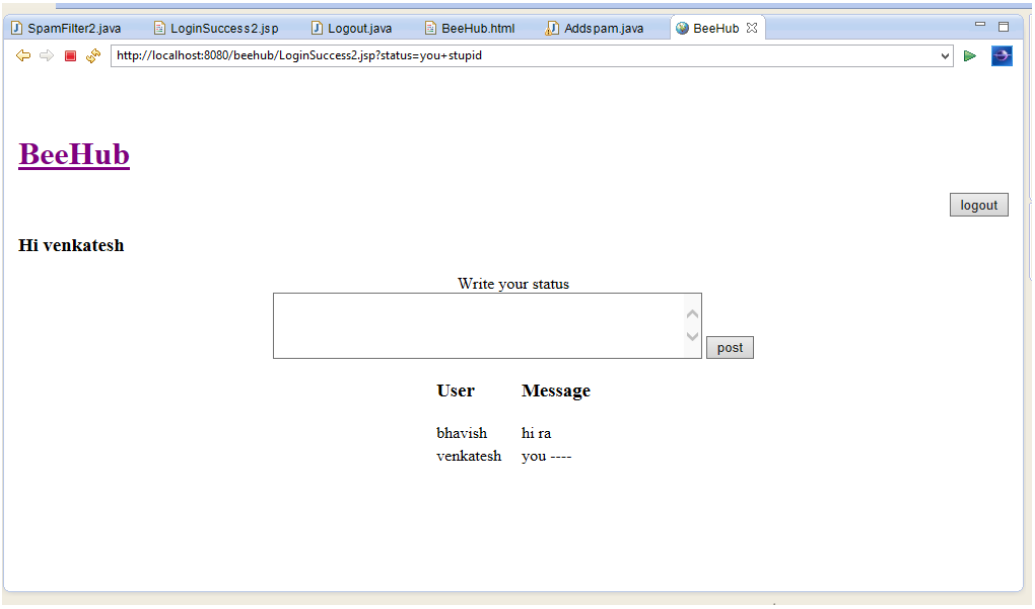

Figure: Spam filtered

6. Conclusion and future scope:

The flexibility of a system can be enhanced through refining rules and blacklist management. In this context, the underlying domain is dynamic and the collection of pre-classified data provided for training purpose may not be valid for a longer time. A preliminary work in this direction has been done in the context of trust values used for social network access control purposes.

However, we would like to remark that, the system proposed in this paper represents just the core set of functions needed to provide a sophisticated tool for social network message refining. Even if we have complemented our system with an online assistant to set refining rule thresholds, the development of a complete system easily usable by average social network users is a wide topic which is out of the scope of the current paper. As such, the face book applications to be meant as a proof-of-concepts of the system core functions, rather than a fully developed system.

Finally I conclude that with the content based filtering, spam filter has been designed and by that way user can avoid spam messages posted on their user walls. Future scope is to design it for messages posted in other languages, image spam and video spam filtering.

Acknowledgement:

This work is sustained and done under the esteemed guidelines of Dr. G. Venkat Rami Reddy garu, who is a lecturer and head of the department of Software Engineering at School of IT, Jawaharlal Nehru Technological University, Hyderabad.

REFERENCES:

[1] "A System to Filter Unwanted Messages from OSN User Walls" Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, Moreno Carullo Department of Computer Science and Communication University of Insubria 21100 Varese, Italy.
[2] "A Survey on Spam Filtering for Online Social Networks", IJCSMC, Vol. 3, Issue. 4, April 2014, pg.1130 – 1135.
[3] Internet Live Stats (elaboration of data by International Telecommunication Union (ITU) and United Nations Population Division) go to http://www.internetlivestats.com/internet-users/

[4] http://www.chinainternetwatch.com/3378/china-internet-users-online-time/

[5] http://www.process.com/psc/fileadmin/user_upload/whitepapers/pmas/intro_bayesian_filtering.pdf

[6] "A Framework to Refine Awful Tweets from e-Networks"Ratna Kailashnadh Singamsetty et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (1) , 2014, 740-

A Brief Author Biography

M Bhavish Santhosh Kumar **–** is pursuing his post graduation from School of Information Technology, JNTU Hyderabad. He did his B.Tech from ADAM's engineering college. His research area interests include information security, computer networks.

G V R Reddy **–** is presently Associate Professor in Computer Science and Engineering at school of Information Technology. He is more than 11 years of experience in Teaching, and Software Development. His areas of interests are: image Processing, Computer Networks, and Analysis of Algorithms, Data mining, Operating Systems and Web technologies.