Detecting the Offensive Post on Real Time Status

Manju Singh

M.tech scholar, Rewa institute of technology, Rewa M.P.

Abstract Social Networking sites are acquirement most recognition due to the perfect and well- organized way of contact between two parties. However, during communication between two parties it may perhaps feasible that several insulting words are to be used with the intention of may be deliberately or involuntarily is referred as Flaming during communication in the computer mediation. This project mainly made the centre of attention on the Flame detector model, which will incarcerate the status of the user with the help of the web services from social networking sites and store it in to the local host. In this project the working of the flame detector model had been described by taking the real time status of any user and detecting the flame and also generating the graph according to the intensity of the flame.

.Keywords: computer mediated communication, web service, flaming, flame detector, offensive message.

I. INTRODUCTION

The social networking sites also help the public in seeking the old friends, relatives, batch mates of schools as well as college.[26][28][29] With the help of social networking sites the people started to communicate with the people whom they do not know and make them friends[15][16]. Thus, Social Networking Sites have made the world too much shorter indeed with the help of the internet. Every Text communication when two strangers or even two familiar persons are doing gossips they cannot see each other. The text pattern of the chat depends upon the user behaviour that can be further classified in to two categories. One is offensive or aggressive behaviour and another is defensive or sophisticated behavior[3],[5],[6],[8]. Some research studies projected that these modes of communication invited more conflict through manipulative and disrespectful tendencies of users that escalates anger and incites tensions between factions (Friedman & Currall, 2003; Harrison & Falvey, 2002; Landry, 2000; Markus, 1994; Moore, Kurtzberg, Thompson, & Morris,1999; O'Sullivan & Flanagin, 2003)[2][4][7][10]. Flaming is as explained earlier is some kind of detection of abusive words present in the message, wall post, or email. [1] [12]. For example when messages are exchanged by the two parties for wishing each other on various occasions like good night, good morning, good noon, happy birthday, love feeling the user generally uses smiley or heart shape while communication that is also good examples of flaming but those can be considered in very affirmative way rather than abusive one[13] [18]. Today most of the websites related to the social networking like face book, Google+, Orkut etc., working hard for finding flaming words being used by various users over the world during text communication. This paper introduces the high focus on the offensive nature of the user when it posts or sends some abusive words to the wall or email of another user who is known or unknown by intentionally or unintentionally. In addition to this, the paper contains the flame detector model that will detect the flame and make aware of the abusive words used in their profile[9],[11],[14],[19].

II. LITERATURE SURVEY

Now-a-days Computer-mediated communication (CMC) is analyses as one of the most significant stage for social interaction, thereby eliminating the physical interaction constrictions. Apart, there are credits endorsed to CMC, while the other there is also a further surface in new mediated communication such as e- mail, online chat conversation and communication message in the appearance of antagonistic and aggressive hostile communicative manners, termed as offensive message or "flaming". [1],[4],[13].The utilize of CMC in the association or group or organization has raised severe doubts that whether it is valuable in good way and detrimental in worst way for the organizations due to offensive message in the chat conversation as flaming, which has turn out to be outstanding part of the social interaction. It is a wide-ranging faith that the offensive message as flaming may consists of hostile and insistent aggressive activities shown through the source or platform as CMC. Though, Lea et al. confront this credence.[2],[10],[12],[7]. Even so, scholars enclose jointly pushed onward to undertake to compute levels of flaming offensive behavior as well as to endow with justification and give the possible remedies[22],[23],[29].

III. . CONCEPT OF FLAMING

The research proverb a necessitate to classify offensive or flames on web especially social networking sites as it exposed some mold of replies based on the kind of flame or offensive posted by consumers. All statuses specified to the survey subjects can be classified largely into one of these groups. Several examples are taken from the survey[20],[24],[25].

Flaming inclination is distinguished to be highest when users purposely use offensive, incendiary and hostile message or communication against another client or users. This is foremost on diverse forms of CMC i.e. computer-mediated communication but is less perceived on web in social networking sites.[24] Users have a preference to

maintain their altercation private and not make known them to all their friends on their list to gratuitously. Users are additional cognizant of their actions on social networking sites. Yet, there are undersized groups who take such steps and utilize spot like status messages, comments, etc for doing offensive flaming. Such flaming outlines are quite notice on status messages but are additional predominant on discussions on groups or community venues. Further, other users with direct flames reciprocate such flames .

Example: "Hello, stupid! If you were not such an idiot, you would understand why you are wrong about everything. However, apparently you are so retarded that you cannot even spell right"

Indirect flaming is normally opting for to illustrate incongruity and hostility but posted in a language, which can only be understood by the participants concerned. Friends of the user who read such communication or post or messages would comprehend that there is some disagreement, but barely would be able to track accurate references or the users towards whom the offensive post or flame is projected. Such offensive flaming patterns can be seen on status messages that are made public to all friends. Such communicated messages or post are posted to show disagreement[21].

Example: "Actor X tweeted "My movie earned 100cr in a week". Actor Y tweeted "and still your movie was not worth watching.

When the references to places or situations or people, are evidently affirmed in any message posted by the user, it can be said as Straight flames. This style or strategy of offensive flaming is used alongside with direct or indirect flaming[18]. Thus the straight flaming is highly dangerous.[20]
Since straight flames are clearer, they have higher probability of drawing counter flames by consumers or users, which then augment the intensity of the subsequent flames by the users. When a customer uses statements that can have varied derivations aimed towards a particular person, at exact place or situation, thus it may be termed as Satirical Flames. Satirical flames are more complicated as the references made in these flames are hazy [27].Thus, reactions to such flames are normally enquiring of the details. Hot flames are typified as "incendiary messages" and "inflammatory remarks". Classic descriptions correspond to hot flaming as "rude or insulting" post or messages, "vicious attacks", "and nasty and often profane diatribe", "derisive commentary" [27]. In other words, hot flames can be recognized as post or messages that may add up to attacks such as name calling, swearing, insulting on other communicating party/parties. It is also described or characterized by the make use of rude behaviour (may be sexually oriented), offensive, aggressive and an angry tone. Use of words or intension or literature that may not be classify as offensive abusive or hostile but when we think the situation in which it is said, it means completely different and user receiving it feels insulted and humiliated[30].

IV. PROPOSED METHODOLOGY: DETECTING OFFENSIVE POST ON REAL TIME STATUS

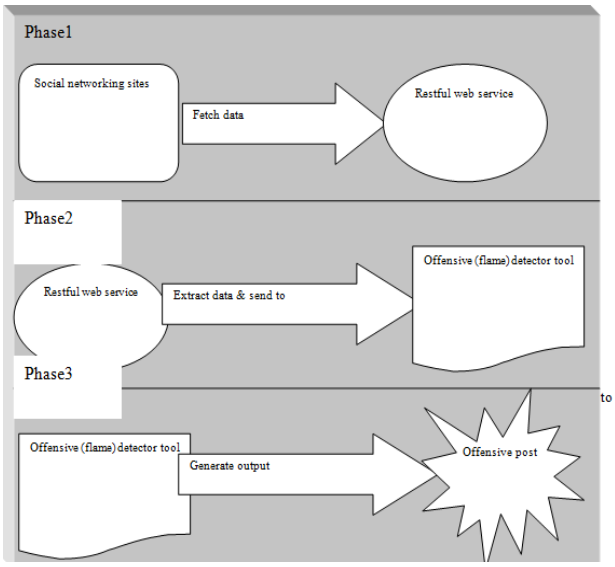

Fig. 4.1 Flame detector model

This model is very useful for the social networking sites and for the teenager group to know about the flaming, to know about the rules of communication, to aware about the intensity of the flaming and to avoid the flaming word or abusing word using in the social networking sites. Thus by the help of this model the rate of flaming in the social networking sites will easily decrease. The model use to have basically three phases i.e. status, web service and the flame detector. The communication in the social networking sites is done basically through the status i.e. wring comment. The entire user used to write the comment

on their profile to make aware to his friend's and the relatives that what is going in his life. The several users may also comment back to his status. Some of them write meaningful comments to him while others generally who are his non met friends will comment or write some flame words or abusive comment to him.

The non met friends are just ignorant of him that who he is, what will be reaction of him after seeing that type of comment they feels like. In the social networking sites the user should always use to avoid making the non- met friends. Thus this kind of friend use to write the comment whatever they feels like. In the social networking sites the user should always use to avoid making the non met friends. The status of the user will be grasping through the web service. The flame detector model used to have an application product of the social networking sites.

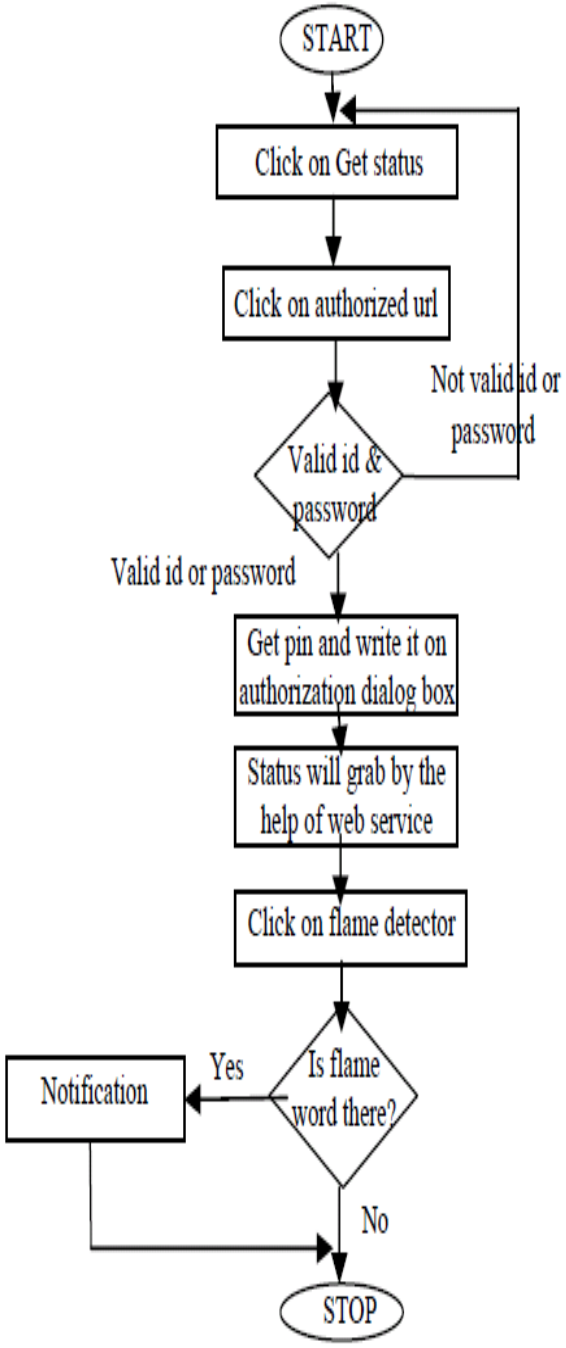

Fig. 4.2 Flow chart of the tool

The application product use to have an application product key and the authorization base url, consumer key and the consumer secret key following through the restful web service which make the user easily to access the part of social networking sites by having the valid user login and the valid user password. The authorization base url will get the valid pin number generated by the social networking sites for testing the authorization of the user. Thus after getting the pin number the tool will easily take out the status from the site for the further processing. The flame detector will now detect the flame from the status by the help of word to word matching algorithm. It will match all the word of the status from its database. After matching the word it will produce the warning message or the notification of the flaming word used in the status

V. EXPERIMENTAL RESULTS

Fig. 5.1 Main form of tool

The above figure 5.1 is the main form of the tool, which have the several buttons to perform the several operations, like online service, flame detector etc. User just needs a single click to perform t he required operation.

Step1: for the twitter login just go to the online service and click on t he twitter button as shown in figure 5 .2 below.

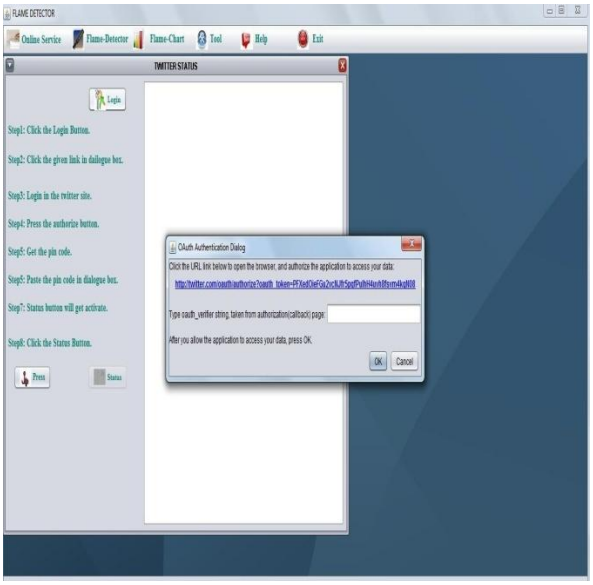

Fig. 5.2 Twitter status dialogue box form

After clicking on t he twitter t he tool will provide the twitter login form as shown in above figure 5.2. This form has the two button login and the status. The login button is for login into the twitter account by the help of the genuine login id and the password. The status button will grab out the status from t he twitter account after getting login in it

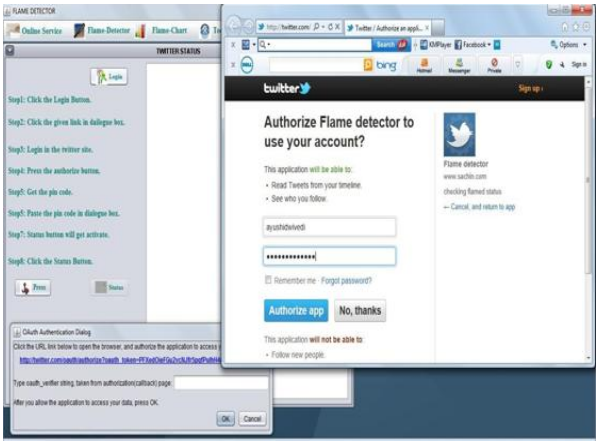

Fig. 5.3 Twitter status login form

Now as the user will click on the link, the website of the user will automatically release as shown in figure 5. 3. The user has to now enter the login id and the pas sword

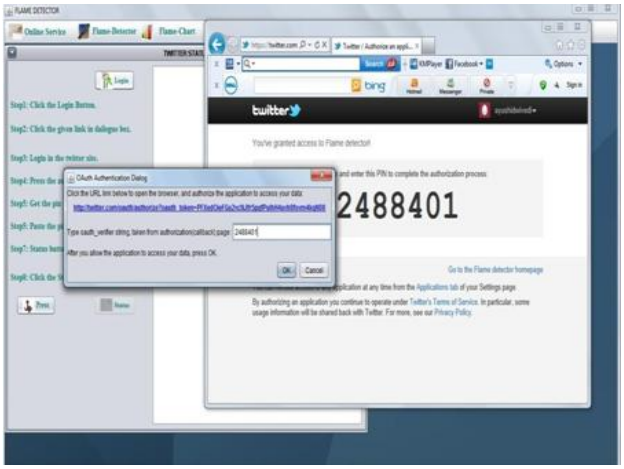

Fig. 5.4 Twitter status authentication pin form

The above figure 5. 4 shows the user has copied the pin number and paste in the dialogue box, for t he tight bound between the local host software and the twitter.

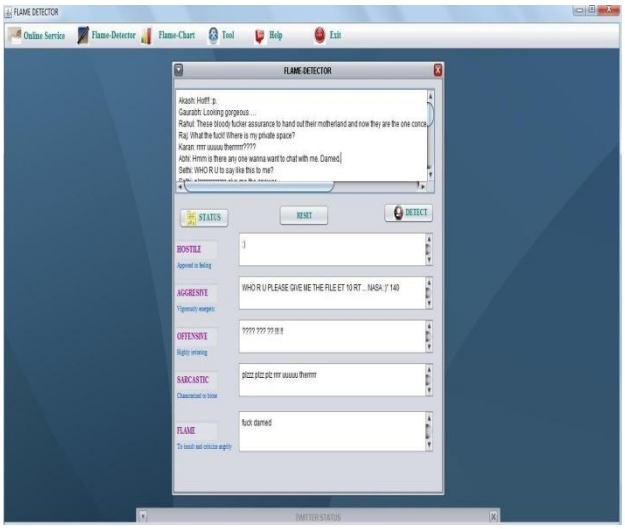

Fig. 5.5 Offensive Post Detection on real Time Status

Thus now the user will click on the status button of the flame detect or model , the status will be automatically copied to this form to detect the form . This form contains t h e two m ore butt on one to detect the fl am e and it will also categorize the all abusive words used in your profile. Thus it will also tell the user the intensity of the flam e that which type of flame is mostly used i n his profile. It also makes the u ser aware what are the flaming are what is the type of t he flaming. The intensity of the status that is grabbed output now by the user‘s profile is given in figure5 .5

VI. CONCLUSION

This research concludes that Social Networking Site users show a movement prototype of third order polynomial when they are faced with flames or hostile messages. Mass of the users favors non-hostile or mildly hostile replies to status messages and this become trend of fashion in day-to-day life. It was also noted that female users have lesser number of ‘Not-Met’ friends on their friend lists than male users. Thus with the help of this model the user will able to know the flame that what is it and how the intensity of the flame is get increased in these day to day life. They will also able to know the intensity of the flame and make their profile free from the flame that gives the pleasant look

REFERENCES

[1] McKenna K, Bargh JA (2000). Plan 9 from cyberspace: the implications of the Internet for personality and social psychology. Personality and Social Psychology Review, 4, 57–75.
[2] Walther JB, Anderson JF, Park DW (1994), Interpersonal effects in computer-mediated interaction: A metaanalysis of social and antisocial communication. Communication Research, 460-487
[3] Norman A. Johnson, Randolph B. Cooper, Wynne W. Chin (2008). The effect of flaming on computer mediated negotiations. European Journal of Information Systems 17, 417-434.
[4] Friedman RA, Currall SC (2003). Conflict escalation: dispute exacerbating elements of e-mail communication conflict. Human Relations, 56 (11), 1325-1347.
[5] Harrison TM, Falvey L (2002). Democracy and new communication technologies. Communication Yearbook, 25, 1-33.
[6] Landry EM (2000). Scrolling around the new organization: The potential for conflict in the on-line environment. Negotiation Journal, 16 (2), 133-142
[7] Markus ML (1994). Finding a happy medium: Explaining the negative effects of electronic communication on social life at work. ACM Transactions on Information Systems, 12 (2), 119-149.
[8] Menninger K, Mayman M, Pruyser P (1963). The vital balance: the life process in mental health and illness. Viking, New York.
[9] Moore DA, Kurtzberg TR, Thompson LL, Morris MW (1999). Long and short routes to success in electronically

mediated negotiations: Group affiliations and good vibrations, Organizational Behavior and Human Decision Processes, 77 (1), 22-43

[10] Aiken M, Waller B (2000). Flaming among first-time group support system users. Information & Management, 37, 95-100.

[11] Alonzo M, Aiken M (2004). Flaming in electronic communication. Decision Support Systems, 36, 204- 213.

[12] Bernthal K (1995). Online transmission of inflammatory remarks. PC Novice, 6, 39-40.

[13] Boyd dm, Ellison, NB (2007). Social network sites: definition, history, and scholarship. Journal of Computer Mediated Communication, 13(1), article 11.

[14] Bruce A. Reinig, Roberto J. Mejias (2004). The Effects of National Culture and anonymity on flaming and criticalness in GSS-supported discussions.

[15] Chapman G (1995). Flamers. The New Republic, 13.

[16] Cherny L (1995), The MUD register: Conversational modes of action in a text-based virtual reality. Linguistics Department. Palo Alto, CA: Stanford University.

[17] Dvorak, J. C. (1994). The flaming of Madison Ave. Marketing Computers, 14, 22.

[18] ECAR Research Study 8, 2008. Social networking sites, Student and Information Technology.

[19] Ellison, NB, Steinfield C, Lampe C (2007). The benefits of Facebook ―friends:‖ Social capital and college students' use of online social network sites. Journal of Computer Mediated Communication, 12(4), article 1.

[20] Eveland JD, Bikson TK (1988). Work group structures and computer support: A field experiment,

[21] Transactions on Office Information Systems, 6, 354-37

[22] Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of de-individuation in a group. Journal of Abnormal and Social Psychology, 47, 382-389.

[23] Goodwin C, Heritage J (1990). Conversation analysis, Annual Review of Anthropology, 19, 283-307.

[24] Hinds P, Kiesler S (1995). Communication across boundaries: Work, structure, and use of communication technologies in a large organization. Organization Science, 6, 373-393.

[25] Horrigan J, Boase J, Rainie L, Wellman B (2006). The Strength of Internet Ties, Pew Internet & American Life Project report.

[26] Huber, GP (1990). A theory of the effects of advanced information technologies on organizational design, intelligence, and decision making. J. Fulk & C. W. Steinfield (Eds.) Organizations and communication technologies, Sage, 237-274.

[27] Joseph Kayany M. (1998).Contexts of Uninhibited Online Behavior: Flaming in Social Newsgroups on Usenet.

[28] Kiesler S, Siegel J, McGuire, T (1984). Social psychological aspects of computer-mediated communication,. American Psychologist, 39 (10), 1123-1134Name of Author(s), “Title of the research”, Citation Details, year.

[29] Name of Author(s), “Title of the research”, Citation Details, year.

[30] Name of Author(s), “Title of the research”, Citation Details, year.