RESEARCH ARTICLE

Data Categorization for Detecting Posts in Social Network Using Sequential Tracking

M.PRIYADHARSHINI[1], S.VAIRACHILAI[2]

PG STUDENT, ASSISTANT PROFESSOR, NPR COLLEGE OF ENGINEERING AND
TECHNOLOGY, TAMIL NADU, INDIA
EMAIL: pshini1673@gmail.com, vairachilai2676@gmail.com

Abstract: Online Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. A system allowing OSN users to have a direct control on the messages posted on their walls. A flexible rule-based system, that allows users to customize the filtering criteria to be applied to their walls. Machine Learning-based soft classifier automatically labeling messages in support of content-based filtering. User can post neutral or acceptable content to their private wall. User can manually predict the content what they feel bad about the message.

Index Terms— Short text classification, Social Network, Filtered Wall, Black Lists, Content based Filtering, Topic Detection, Radial Basis Function Network(RBFN)

1. INTRODUCTION

A social network is a social structure made up of a set of social actors. it's becoming more popular in this growing world. The drawback in social network is privacy is less convenient. To avoid such drawbacks we propose Online Social Networks (OSNs) helps to give users ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Conventional approaches for topic detection have mainly been concerned with the frequencies of (textual) words. A term-frequency-based approach could suffer from the ambiguity caused by synonyms or homonyms. It may also require complicated preprocessing depending on the target language. We propose a probability model that can capture the normal mentioning behavior of a user, which consists of both the number of mentions per post and the frequency of users occurring in the mentions. This model is used to measure the anomaly of future user behavior. Using the proposed probability model, we can quantitatively measure the novelty or possible impact of a post reflected in the mentioning behavior of the user. We aggregate the anomaly

scores obtained in this way over. Hundreds of users and apply a recently proposed change point detection technique based on the sequentially discounting normalized maximum-likelihood coding. The main scope of this proposed system is to filter the unwanted messages that posted on the user's wall. The system contains a rule list in the user side which prevents the posting of unwanted messages on the user's wall. User can make the rules for content filtering User can add unwanted friends in black list and block them message too.

2. LITERATURE SURVEY

2.1. Competitive Agents for Information Filtering

Each filter is a rule with one or more conditions and a destination. Each filter rule can contain several conditions. The filters are applied in the order they are listed on the Mail Filter tab. Leave the message in the Inbox (no action).Move the message to another specified folder. Tag the message. Mark the message as read or as flagged. Discard the message.

2.2 Information Filtering: Overview of Issues, Research and Systems

BLs is directly managed by the system, which should be able to determine who the users to be inserted in the BL. are then decide when user's retention in the BL is finished. Information are given to the system through a set of rules, hereafter called BL rules. BL rules make the wall owner able to identify users to be blocked according to their profiles as well as their relationships in the OSN.BL mechanism to avoid messages from undesired creators, independent from their contents.

2.3. Clustering of Short Commercial Documents for the Web

Designing and evaluating various representation techniques in combination with a neural learning strategy to semantically categorize short texts. A hierarchical two-level strategy assuming that it is better to identify and eliminate "neutral" sentences, and non-neutral sentences. A hard classification in which short texts are labeled with crisp Neutral and Non-neutral labels. Such a list of grades is then used by the subsequent phases of the filtering process.

2.4. Social Network-based Trust in Prioritized Default Logic

When two individuals know each other, they can assess the trustworthiness of one another. Two people who are not directly connected do not have a foundation for knowing about trust. However, the paths connecting them in the network contain information that can be used to infer how much they may trust one another. Our algorithm looks at the trust values along paths connecting the source and sink to compute a recommendation to the source about how much to trust the sink.

2.5. Approximation by Radial Basis Function Network

The technique that allows realizing this operation is called vector quantization and the points that summarize the position of the nodes. The centroids are first randomly initialized in the space.
Multiplied by a factor k. We will explain in our application how to choose this parameter. When all other parameters are defined, these are determined by the solution of a system of linear equations. One of the disadvantages of the RBFN that we have presented is that they give an equal importance to all input variables.

3. CONCLUSION

A new approach to detect the emergence of topics in a social network stream. The basic idea of our approach is to focus on the social aspect of the posts reflected in the mentioning behavior of users instead of the textual contents. We have proposed a probability model that captures both the number of mentions per post and the frequency of mentioned. It is robust to rephrasing and it can be applied to the case where topics are concerned with information. The detection by the proposed link-anomaly based methods was earlier than the text-anomaly-based counterparts.

REFERENCES

[1] J. Allan et al., "Topic Detection and Tracking Pilot Study: Final Report," Proc. DARPA Broadcast News Transcription and Understanding Workshop, 1998.

[2] J. Kleinberg, "Bursty and Hierarchical Structure in Streams," Data Mining Knowledge Discovery, vol. 7, no. 4, pp. 373-397, 2003

[3] Y. Urabe, K. Yamanishi, R. Tomioka, and H. Iwai, "Real-Time Change-Point Detection Using Sequentially Discounting Normalized Maximum Likelihood Coding," Proc. 15th Pacific-Asia Conf. Advances in Knowledge Discovery and Data Mining (PAKDD' 11), 2011

[4] A. Krause, J. Leskovec, and C. Guestrin, "Data Association for Topic Intensity Tracking," Proc. 23rd Int'l Conf. Machine Learning (ICML' 06), pp. 497-504, 2006.

[5] C. Lesage, M. Cottrell eds ,"APPROXIMATION BY RADIAL BASIS FUNCTION NETWORKS" vol. 7, no. 4, pp. 203-214