AN EFFICIENT MECHANISM FOR BLOCKING THE UNWANTED MESSAGE IN SOCIAL NETWORKS

[1]S.SIDDIQ, [2]M.S.KOKILA.
[1]Research Scholar, Department of Computer Science, Kongu Arts and Science College, Erode.
[2]Assistant Professor, Department of Computer Science, Kongu Arts and Science College, Erode.

Abstract: Online Social Networks (OSNs) are today one of the most popular interactive medium to share, communicate, and distribute a significant amount of human life information. In OSNs, information filtering can also be used for a different, more responsive, function. This is owing to the fact that in OSNs there is the possibility of posting or commenting other posts on particular public/private regions, called in general walls. Information filtering can therefore be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. OSNs provide very little support to prevent unwanted messages on user walls. For instance, Facebook permits users to state who is allowed to insert messages in their walls. Though, no content-based partialities are preserved and therefore it is not possible to prevent undesired communications, for instance political or offensive ones, no matter of the user who posts them. To propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN userwalls.

1. INTRODUCTION

Information and communication technology plays a significant role in today's networked society. There is a need to develop more secured mechanisms for different communication technologies, particularly online social networks. Online social networks provide very little support to prevent unwanted messages on user walls. With the lack of classification or filtering tools, the user receives all messages posted by the users he follows. In most cases, the user receives a noisy stream of updates [1]. In my work, an information Filtering system is introduced. The system focuses on one kind of feeds lists which are a manually selected group of users on online social network. List feeds tend to be focused on specific topics; however it is still noisy due to irrelevant messages.

In online social networks, information filtering can also be used for a different, more sensitive, purpose. This is due to the fact that in online social networks there is the possibility of posting or commenting other posts on particular public/private areas, called in general walls. Information filtering can therefore be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. The aim of the present work is therefore to propose and experimentally evaluate an automated system, called filtered wall, which able to filter unwanted messages in online social network user walls. We exploit Machine Learning text categorization techniques to automatically assign with each short text message a set of categories based on its content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminate features.

2. LITERATURE SURVEY

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy n company strength. Once these things r satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system.

2.1 Content-Based Filtering

Content based filtering selects an item based on user interest. It uses items previously preferred by the user and then suggests the best matched item. Each user acts independently in content based system. This kind of system chooses item depending on relation between item content and user recommendations against collaborative system that selects item based on relation between people with similar preferences[14]. The content based system creates a content based profile of a user based on rated items of a user. Items features are weighted based on features preferred by the user and recommendations are given by the system accordingly.

In content based filtering, the main issue is whether the system is able to learn from user's actions related to a particular content source and use them for other content types. Text classification is similar to content based filtering as documents processed in such type of system are mostly textual. In online social network user's social profile has to be taken into account and this makes content based filtering system difficult to apply in online social network domain as a standalone system.

2.2 Limited Content Analysis

Content-based techniques are limited by the features that are explicitly associated with the objects that these systems recommend. Therefore, in order to have a sufficient set of features, the content must either be in a form that can be parsed automatically by a computer (e.g., text), or the features should be assigned to items manually. While information retrieval techniques work well in extracting features from text documents, some other domains have an inherent problem with automatic feature extraction. For example, automatic feature extraction methods are much harder to apply to the multimedia data, e.g., graphical images, audio and video streams. Moreover, it is often not practical to assign attributes by hand due to limitations of resources [9]. Another problem with limited content analysis is that, if two different items are represented by the same set of features, they are indistinguishable. Therefore, since text-based documents are usually represented by their most important keywords, content-based systems cannot distinguish between a well-written article and a badly written one, if they happen to use the same terms. When the system can only recommend items that score highly against a user's profile, the user is limited to being recommended items similar to those already rated. For example, a person with no experience with Greek cuisine would never receive a recommendation for even the greatest Greek restaurant in town. This problem, which has also been studied in other domains, is often addressed by introducing some randomness. For example, the use of genetic algorithms has been proposed as a possible solution in the context of information filtering [8].

2.3 Policy-Based Personalization Of Osn Contents

Policy based personalization is applicable in many different contexts. It adapts a service in specific context according user defined policies. In online social networking sites user oriented policies can define how communication between two parties or more can be handled. The policy based personalization system in focuses on Facebook [13][14]. It assigns a category to each tweet and shows only those tweet to the user which are of his interest. In this scenario, policy based personalization represent the ability of the user to filter wall messages according to filtering criteria suggested by him.

There have been some proposals exploiting classification mechanisms for personalizing access in OSNs [8] [12]. For instance, in a classification method has been proposed to categorize short text messages in order to avoid overwhelming users of micro blogging services by raw data. The user can then view only certain types of tweets based on his/her interests. In contrast, Golbeck and Kuter propose an application, called Film Trust that exploits OSN trust relationships and provenance information to personalize access to the website[9][10]. However, such systems do not provide a filtering policy layer by which the user can exploit the result of the classification process to decide how and to which extent filtering out unwanted information [5].

2.4 Social Influence

The social influence of another element determining people's use of Facebook is social influence carried out a study among Singaporean respondents whose average age was 25. The researchers hypothesized that the user's likeability to join and use Facebook is directly related to the following: the number of friends using the social network, the belief that Facebook has the most active users in Singapore and globally, and finally the belief that Facebook is the most used networking site among an individual's peers[12][13]. In addition, predicted that the use of the website increases with the size of a person's social network and with the usefulness of the functions and applications on the website. The findings confirmed these hypotheses and the researchers concluded that peer effect has indeed a role in an individual's choice to use Facebook[15][16]. They grounded their research on a theoretical model developed by Kelman, which presents three modes of social influence. The first is compliance and is the process by which an individual starts using a specific technology under the influence of family members or friends. The second mode of social influence is internalization representing an individual's decision based on similarity of values with certain communities or groups of interest[10][11]. The third is identification and sees the individual's recognition emotional and evaluative of his place in a certain group of community.

3. PROBLEM FORMULATION

3.1 Overview

In the social media, like Facebook and Flicker have changed the way people interact with each other during past decades. Because people is human, their interactions online always show social activities, forming different groups according to their preferences, hobbies, education levels. Therefore, getting group information or categories are critical important for those advertisers.

3.2 Problem Definition

Online social network service that has not been provided so far. Indeed, today online social networks provide very little support to prevent unwanted messages on user walls. For example, Face book allows users to state who is allowed to insert messages in their walls. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather it requires to design ad-hoc classification strategies. This is because wall messages are Constituted by short text for which traditional classification Methods have serious limitations since short texts do not Provide sufficient word occurrences. The OSN contain short text is characterized by 1) Shortness in the text length 2) Sparsity in the terms presented, which results in the difficulty in managing and analyzing them based on the bag of-words representation only. Short texts can be found everywhere, such as search snippets, product reviews etc. Short text classification is to classify the short texts and assign the short text a label from predefined taxonomy. It takes web search snippets as a representative of short text. These search snippets are collected during web search transaction by using various phrases of different domains as issued queries. Specifically, a classifier can be built for assigning web search snippets labels such as Business, Computer etc. Based on this classifier, search results can be organized effectively, and as a result web users can be navigated to the needed information. One possible solution to handling sparsity of short text is to expand short texts by appending some words to the text based on the semantic relatedness between words. Such methods examine individual word only, without considering the context of the word, i.e., word co-occurrence within the short text, which usually is very meaningful for short text classification. To address this issue, a reasonable solution is to map the short text itself to topic space obtained from external corpus.

3.3 Proposed System

The aim of the present work is therefore to propose and experimentally evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from online social network user walls. We exploit Machine Learning (ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and discriminate features. The solutions investigated in this paper are an extension of those adopted in a previous work by us from whom we inherit the learning model and the elicitation procedure for generating pre-classified data.The original set of features, derived from endogenous properties of short texts, is enlarged here including exogenous knowledge related to the context from which the messages originate. As far as the learning model is concerned, we confirm in the current paper the use of neural learning which is today recognized as one of the most efficient solutions in text classification. In particular, we base the overall short text classification strategy on Radial Basis Function Networks (RBFN) for their proven capabilities in acting as soft classifiers, in managing noisy data and intrinsically vague classes. Moreover, the speed 2 in performing the learning phase creates the premise for an adequate use in online social network domains, as well as facilitates the experimental evaluation tasks.

4. SYSTEM METHODOLOGY

Our goal is to design an online message filtering system that is deployed at the OSN service provider side. Once deployed, it inspects every message before rendering the message to the intended recipients and makes immediate decision on whether or not the message under inspection should be dropped. This research achieved the above mentioned objectives with some algorithms and techniques.

Short Text Classifier Algorithm

Machine Learning Techniques

Filtered Wall Architecture

4.1 Short Text Classifier

Text representation using endogenous knowledge has a good general applicability, though in operational settings it is appropriate to use also exogenous knowledge. We introduce contextual features (CF) modelling information that characterize the environment where the user is posting. These features play important role in deterministically understanding the semantics of the messages [12]. According to Vector Space Model (VSM) for text representation, a text document dj is represented as a vector of binary or real weights dj = w1j, . . . , w│T│j, where T indicates the set of terms that occur at least once in at least one document of the collection Tr, and wkj € [0; 1] denotes how much term tk contributes to the semantics of document dj. In the BoW representation, terms are identified with words. For non-binary weighting, the weight wkj of term tk in document dj is computed according to the standard term frequency - inverse document frequency (tf-idf) weighting function, defined as (1) Where # (tk, dj) indicates the number of times tk occurs in dj, and #Tr (tk) indicates the document frequency of term tk, i.e., the number of documents in Tr in which tk occurs. Dp features are heuristically calculated; their definition stems from intuitive considerations, domain specific criteria and in some cases required trial and error procedures.

Correct words: it represents the amount of terms tk € T ∩ K, where tk is a term of the considered document dj and K is a set of known words for the domain language. This value is normalized by #(│ │ =1 k, dj).

Bad words: they are determined similarly to the correct words feature, where the set K is a collection of "dirty words" for the domain language.

Capital words: it represents the amount of words mostly written with capital letters, calculated as the percentage of words within the message, having more than half of the characters in capital case. For example, the value of this feature for the document "To be OR Not to BE" is 0:5 since the words "OR" "Not" and "BE" are considered as capitalized ("To" is not uppercase since the number of capital characters should be strictly greater than the characters count).

Punctuations characters: it is computed as the percentage of the punctuation characters over the total number of characters in the message. For example, the value of the feature for the document

"Hello!!! How're u doing?" is 5/24.

Exclamation marks: it is computed as the percentage of exclamation marks over the total number of punctuation characters in the message. Referring to the aforesaid document, the value is 3/5.

Question marks: it is computed as the percentage of question marks over the total number of punctuations characters in the message. Referring to the aforesaid document, the value is 1/5.

4.2 Machine Learning Techniques

A Machine learning approach learns from training data and creates classifiers for the classification of new dataset. The main task of text classification is to assign a predefined category with each text. Text classification is accomplished on the basis of endogenous collection of data. The machine learning, based classifier learns how to classify the categories of incoming data on the basis of features extracted from the set of training data. Below are the key methods which are commonly used for text classification.

Neural network classifiers
Support vector machines
Naive Bayes classifier
Decision tree

4.2.1. Naive Bayes Classifier

Naive Bayes classifier is a probabilistic classifier based on Bayes theorem with independence assumption [4]. Given a class variable, it assumes the presence or absence of specific feature is unrelated to the presence or absence of any other feature. For instance a fruit is considered to be cherry if it is red, round and small in size. Bayes classifier considers each of these features independently to the probability that the fruit is cherry regardless of the presence or absence of any other feature. The main advantage of this classifier is that it requires a small amount of training data to estimate the parameters required for classification.

4.2.2. Neural Network Classifier

Neural network classifiers consist of neurons arranged in layers converting an input vector into output. The most commonly used neural network is multilayer feed forward network in which a unit feeds its output to all the units of the next layer but there is no feedback to the previous layer. Radial basis function network is an artificial neural network which uses radial basis function as an activation function. The output of this network is a linear combination of radial basis functions of the inputs and neuron parameters. It is robust to outliers [5] and therefore more suitable in this context.

4.2.3 Support Vector Machines

The support vector machine classifiers analyze data and recognize pattern in it. They are based on supervised learning model and are able to perform nonlinear classification in addition to linear classification. The support vector machine classifier is suitable for large amount of unlabeled data and small amount of labelled data. The high dimensional input space, irrelevant features, sparse document vectors and linearly separable text classification makes support vector machine classifier suitable for text categorization.

4.2.4. Decision Trees

Decision trees classifiers are used for a hierarchical decomposition of the data space. It determines the predicate or a condition depending on attribute value. Class labels in the leaf data are used for classification. In order to reduce the over fitting data pruning is required in decision tree. This classifier requires iterative training procedure and is oversensitive to training data.

5. SYSTEM IMPLEMENTATION

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

5.1. Filtering Rules

In defining the language for FRs specification, we consider three main issues that, in our opinion, should affect a message filtering decision. First of all, in online social networks like in everyday life, the same message may have different meanings and relevance based on who writes it. As a consequence, FRs should allow users to state constraints on message creators. Creators on which a FR applies can be selected on the basis of several different criteria; one of the most relevant is by imposing conditions on their profile's attributes. In such a way it is, for instance, possible to define rules applying only to young creators or to creators with a given religious/political view. Given the social network scenario, creators may also be identified by exploiting information on their social graph. This implies to state conditions on type, depth and trust values of the relationship(s) creators should be involved in order to apply them the specified rules. All these options are formalized by the notion of creator specification, defined as follows.

5.2 Filtering Rules

User can state what contents should be blocked or displayed on filtered wall by means of Filtering rules. Filtering rules are specified on the basis of user profile as well as user social relationship.FR is dependent on following factors

Author |
Creator Specification |
Content Specification |

An author is a person who defines the rules. Creator Spec denotes the set of online social network user and Content Spec is a Boolean expression defined on content. Action denotes the action to be performed by the system on the messages matching content Spec and created by users identified by creator Spec.

5.3 Online Setup Assistant For Frs Thresholds

As mentioned in the previous section, we address the problem of setting thresholds to filter rules, by conceiving and implementing within FW, an Online Setup Assistant (OSA) procedure. OSA presents the user with a set of messages selected from the dataset discussed. For each message, the user tells the system the decision to accept or reject the message. The collection and processing of user decisions on an adequate set of messages distributed over all the classes allows computing customized thresholds representing the user attitude in accepting or rejecting certain contents. Such messages are selected according to the following process. A certain amount of non-neutral messages taken from a fraction of the dataset and not belonging to the training/test sets, are classified by the ML in order to have, for each message, the second level class membership values.

5.4 Blacklists

A further component of our system is a BL mechanism to avoid messages from undesired creators, independent from their contents. BLs is directly managed by the system, which should be able to determine who are the users to be inserted in the BL and decide when user's retention in the BL is finished. To enhance flexibility, such information is given to the system through a set of rules, hereafter called BL rules. Such rules are not defined by the SNM, therefore they are not meant as general high level directives to be applied to the whole community. Rather, we decide to let the users themselves, i.e., the wall's owners to specify BL rules regulating who has to be banned from their walls and for how long. Therefore, a user might be banned from a wall, by, at the same time, being able to post in other walls.

The first is related to the principle that if within a given time interval a user has been inserted into a BL for several times, say greater than a given threshold, he/she might deserve to stay in the BL for another while, as his/her behaviour is not improved. This principle works for those users that have been already inserted in the considered BL at least one time. In contrast, to catch new bad behaviours, we use the Relative Frequency (RF) that let the system be able to detect those users whose messages continue to fail the FRs. The two measures can be computed either locally, that is, by considering only the messages and/or the BL of the user specifying the BL rule or globally, that is, by considering all online social network users walls and/or BLs.

6. FINDINGS

6.1stchas Significant Speed Benefits Without Substantial Performance On Simulated Data

The main aim of this research is to increase the OSN Wall Security of the users. The STC mechanism is the new one to secure the user wall.

The above graph shows that the User access in the Social networks, the each user activity is listed here.

The "x" axis shows the Social networks in the web. "Y "axis shows the number of usage in the day. The methods all are is applied in the system. The CBF(Content Based Filtering), CF(Collaborative Filtering) , PBP (Policy – Based Personalization) is the old methods. The methods all are provide little bit security for the user wall. But not provide the full security of the personal wall.

In that chart, OSN1, OSN2, OSN3 and OSN 4 are the different type of Online social Networks, Each colour mention the text filtering mechanism in the OSN. As shown in that chart, PBP (Policy –Based Personalization) achieves 32.1% improvement over CBF on the overall accuracy.

6.2 The Table Shows That The Comparison Method Of User Usage And Stc Mechanism

The above graph shows that the STC usage in the OSN networks. The STC is the new mechanism to secure the wall of the users. The existing system is not providing the high security. The X axis denotes the Mechanisms CBF(Content Based Filtering), CF(Collaborative Filtering), PBP (Policy – Based Personalization) , STC( ). The graph value is GUI (Graphical User Interface), SNA(Social Network Application), and SNM(Social Network Manager). The Y axis mentions the users. As shown in that chart, STC performs consistently better for all Filtering Mechanism. It may be used with STC to have a better accuracy with an additional time cost of initial training. STC achieves 40%, 60%, 85% improvements over PBP. I believe that Short Text Classification would resolve the user wall issue.

6.3 Comparison Of The Filtering Concepts And Stc Concepts

The above graph is mentioning the X axis the User usage values and Y axis is the Filtering Concepts applied in the text mining of the OSN in previous years. The STC is the most powerful mechanism to secure the

OSN users from the unwanted posts. The SNM, SNA, GUI all are successfully run in only STC.

The existing mechanism of PBP, CF, CBF all are only support some of data. It's not securing the user wall.

Anyone can post any of content in the user wall. But in our Short text classification algorithm secure the user wall. No one post the unwanted message to the user wall. The STC store the unwanted words data in the Table with Classification. It's Split the data in Violence, Sex and some other category. If the post message data is match the table means its not allowed to post the content.

Social Network Manager, Social Network Application, Graphical User Interface all are together support the STC mechanism. So it's powerful mechanism to secure the user wall.

7. CONCLUSION

The present batch learning strategy, based on the preliminary collection of the entire set of labeled data from experts, allowed an accurate experimental evaluation but needs to be evolved to include new operational requirements. The development of a GUI and a set of related tools to make easier BL and FR specification is also a direction we plan to investigate, since usability is a key requirement for such kind of applications. In particular, we aim at investigating a tool able to automatically recommend trust values for those contacts user does not personally known. To believe that such a tool should suggest trust value based on users actions, behaviors and reputation in online social network, which might imply to enhance osn with audit mechanisms. However, the design of these audit-based tools is complicated by several issues, like the implications an audit system might have on user's privacy and/or the limitations on what it is possible to audit in current online social networks.

8. FUTURE ENHANCEMENT

In particular, future plans contemplate a deeper investigation on two interdependent tasks. The first concerns the extraction and/or selection of contextual features that have been shown to have a high discriminative power. The second task involves the learning phase. Since the underlying domain is dynamically changing, the collection of pre-classified data may not be representative in the longer term. As future work, we intend to exploit similar techniques to infer BL rules and FRs. Additionally, we plan to study strategies and techniques Limiting the inferences that a user can do on the enforced filtering rules with the aim of bypassing the filtering system, such as for instance randomly notifying a message that should instead be blocked or detecting modifications to profile attributes that have been made for the only purpose of defeating the filtering system.

REFERENCES

[1] A. Adomavicius, G.and Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extens ions," IEEE Transaction on Knowledge and Data Engineering,vol. 17, no. 6, pp. 734–749, 2005.

[2] N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: Two sides of the same coin?" Communications of the ACM, vol. 35, no. 12, pp. 29–38, 1992.

[3] P. Bonatti and D. Olmedilla, "Driving and monitoring provisional trust negotiation with metapolicies," in In 6th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY 2005). IEEE Computer Society, 2005, pp. 14–23

[4]]M. Chau and H. Chen, "A machine learning approach to web page filtering using content and structure analysis," Decision Support Systems, vol. 44, no. 2, pp. 482–494, 2008.

[5] P. J. Denning, "Electronic junk," Communications of the ACM, vol. 25, no. 3, pp. 163–165, 1982.

[6] P. W. Foltz and S. T. Dumais, "Personalized information delivery: An analysis of information filtering methods," Communications of the ACM, vol. 35, no. 12, pp. 51–60, 1992.

[7] J. Golbeck, "Combining provenance with trust in social networks for semantic web content filtering," in Provenance and Annotation of Data, ser. Lecture Notes in Computer Science, L. Moreau and I. Foster, Eds. Springer Berlin / Heidelberg, 2006, vol. 4145, pp. 101–108.

[8] P. J. Hayes, P. M. Andersen, I. B. Nirenburg, and L. M. Schmandt, "Tcs: a shell for content-based text categorization," in Proceedings of 6th IEEE Conference on Artificial Intelligence Applications (CAIA-90). IEEE Computer Society Press, Los Alamitos, US, 1990, pp. 320–326.

[9] P. S. Jacobs and L. F. Rau, "Scisor: Extracting information from online news," Communications of the ACM, vol. 33, no. 11, pp. 88–97, 1990.

[10] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," in Proceedings of the European Conference on Machine Learning. Springer, 1998, pp. 137– 142.