This is the published version (version of record) of:


Available from Deakin Research Online:  
http://hdl.handle.net/10536/DRO/DU:30025776

Reproduced with the kind permission of the copyright owner. Best possible copy available.

**Copyright**: 2008 The People's University of Bangladesh
Intelligent User Interface for
-Classifier Email Categorization

Md Rafiquil Islam*
Wanlei Zhou**

tent user interface (UI) for filtering spam based on classification
e options to the user and the system to select the proper classification
user emails. Much work has been done to filter spam from legitimate
hm and substantial performance has been achieved with some amount
the case of spam detection FP problem is unacceptable sometimes. The
de multifarious approach for selecting the algorithms for categorizing
st of output emails, called "grey list (GL)", to user for making decision
s. It also provides appropriate feedback to the input for choosing
as well as parameters. It has been shown that the performance of the
with existing system, in order to reduce FP problems and accuracy.

hm, Spam, UI, FP, GL.

gest world wide problems facing the Internet today. The Internet is a
very day life and the email has become a powerful tool intended to idea
as for users' commercial and social lives. Due to the increasing volume
Internet Service Providers (ISPs) are facing lot of problems. The cost to
d email, and employee productivity has become a tremendous problem
ices. However, it is amazing that despite the increasing development
ies, the number of spam messages continues to increase rapidly.
le to control the problem in a variety of ways. Identification and spam
system allows end-users to regain a useful means of communication.
have been centred on the more sophisticated classifier-related issues.
spam classification is an important research issue. The success of
xt categorization has led researchers to explore learning algorithms in

problem, users and organizations analyse the tools with available to
spam in its environment. Tools with a flexible UI will provide an
well as organization. UI is a part of an application that the user sees and
thing the user sees and interacts with is part of the interface. The UI is
and architecture that make the system work. The interface includes the
, metaphors, online help, documentation and training. An intelligent
work or perform a task in the way that makes the most sense to them. It
In this paper an intelligence user interface has been presented for email categorization which will give the flexibility for interaction with user as well as the system. The user interface has been designed based on well known classification algorithms that will protect email services from infiltration by spam. The overall model of our proposed system will give better flexibility to user to filter spam and provide a list of output emails, called "grey list (GL)", that will remind user to make decision about the status of these emails for further classification. This process will reduce the FP problems which are considered one of the research challenges for spam filtering.

The organization of this paper is as follows: Section 2 will describe the overview of the techniques of email categorization and section 3 will describe the design and analysis of UI. Section 4 presents some screen shots of UI. Finally, the paper ends with conclusion and references in section 5 and 6 respectively.

2. Email Categorization Techniques

This section describes a brief overview of spam filtering techniques.

2.1 Rule Based Technique

The rule based filtering techniques is a set of rules to classify e-mail as spam or legitimate e-mail and it can be applied at either the MUA (Mail User Agent) level or the MTA (Mail Transfer Agent) level.

E-mail clients contain an element at the MUA level for categorising e-mail based on a set of rules determined by the user. These rules can be constructed to examine an e-mail message's header, body, for keywords or phrases given by the end-user. A common use of such rules is to categorise newly arrived e-mail into a specific folder called spam. While this technique does work well, it does have a serious problem. The rule set needs constant updating and refinement because most spammers use obfuscation techniques. Some common obfuscation used is misspelling words.

Filtering at the MTA level can achieve some economies of scale but it also triggers some problems. Since by nature, spam is sent in bulk, blocking the sender can dramatically reduce the number of spam needed to be stored and delivered. Some of the techniques described for MUA rule based filtering can be applied at the MTA level [3,4,7].

White-list (WL)

WL is an MUA level rule-based filtering technique, where a WL is a register containing a collection of contacts from which e-mail messages can be accepted. If an e-mail arrives but does not come from one of the contacts in the WL, then it is treated as spam and placed in the spam folder. While this technique is effective for some users, it has also drawbacks. Any email sent by a stranger will simply be incorrectly classified as FP. However there is a scheme that incorporates a challenge response mechanism to allow users to be added to a user's WL.

Blacklist (BL)

BL contains lists of known spammers. Essentially when a user gets spam, the user adds the sender of the spam to the BL. The entire domain of the sender of the spam can be added to the BL. Newly arrived e-mails are checked, and if the sender is on the BL, the e-mail is automatically classified as spam. The major problem stems from the fact that spammers tend to forge header information in their spam. The sender information is generally forged, meaning that perhaps innocent people are added to a BL but more importantly the effect which the BL will have, is diminished dramatically.

A distributed blacklist is a network tool for anti-spam engines. Distributed blacklists maintain a collection of common spam messages on a central server. The filter is shared amongst the subscribers, so if one person identifies a message as spam then all others benefit. When a message arrives, it is compared to the database of known spam and deleted if a match is found. This method is low in FP, but false-negatives
duce error rates as legitimate e-mail would not be blocked even if the ISP
which it originated, is on a real-time block list. In addition, the presence
se the e-mail to be classified as spam.

as:

ich as SVM (Support Vector Machine), NB (Naïve Bayes) and Boosting
spam filtering. Each algorithm can be viewed as searching for the most
h space that contains all the classifiers it can learn. All machine learning
nstance representation. The instances are messages, each message is

where xi, ..., xm are the values of the attributes X1, ..., Xm,
odel in information retrieval [1,2]. In the simplest case, each attribute
'threshold'), of Boolean variables:

tributes, another two attribute vector representations is considered here.

ne frequency attributes are more informative than Boolean ones. With
of Xi in each message d is: xi = ti(d)/l(d), where ti(d) is the number of
represented by Xi, and l(d) is the length of d measured in token

'single tokens the n-grams of tokens with n /{1, 2, 3}/, that is sequences of
ve been examined. In that case, ti(d) is the number of occurrences in
sented by Xi, while l(d) remains the number of token occurrences in d

sification using machine learning algorithms can be categorized into two
N labelled training examples : {x1, y1}, ..., (xn, yn), x Rd where d is the
[7,8]. d training examples : {x1, y1}, ..., (xn, yn), x Rd where d is the

r is to propose a spam filtering model using an intelligence user interface,
and feedback to the user and the system for getting better accuracy and
y the user interface provides two things;

ulate a system and
duce the effects of the users' mani
affects the amount of effort the user must expend to provide input for the
tput of the system, and how much effort it takes to learn how to do this.
istic of the user interface, but is also associated with the functionalities of
ree to which the design of a particular user interface takes into account the
ology of the users, and makes the process of using the system effective.
ly, it describes how well a product can be used for its intended purpose by
effectiveness, and satisfaction, also taking into account the requirements
considerations and the user interface. It will be illustrated further.
3.1 Descriptions of the model

The proposed model for spam filtering using UI is shown in Figure (1). This user interface is linked with four main domains of the systems which include initial transformation of incoming email, feature extraction & selection, email data classification and the analysis the classification result. These four domains are interacting with UI through input and output sections. The details of the functions of UI are as follows:

**Input section**

Firstly, the model will collect individual user emails that are considered as both spam and legitimate. Then the email corpus is transformed or indexed using learning algorithms, which is considered as an initial transformation. The transformed incoming email is considered as an input of UI and the UI will collect mail corpus and send it to the classifier domain for categorization.

In the input section of UI, there are four user options named as "User-Opt-1, User-Opt-2, User-Opt-3, and User-Opt-4" as shown in the figure 2.
the preference to select the filtering approach, i.e., whether the system classifier classification approach. In the first case, the email will be classified by the algorithm and for the second case, the email will be classified using multiple algorithms as shown in the figure 2.

In this approach, the user will have further preference to select the type of classifier classification, which is mentioned as User-Opt-2, the user has to select one of the classifiers. There are number of classification algorithms, named as AI-N”. (Initially we have used three classification algorithms, i.e. N=3). If multi-classifier classification approach, the user also has the preference, User-Opt-3, to choose which combinations of classification algorithm(s) will be used. The combinations of classification algorithms, named as “ClassAl-1 2, ClassAl-1 3, N”, as indicated in figure 2.

Feature extraction (FE) is another domain of our proposed UI. It is a way to extract the original features of incoming email. It reduces the number of features from the original data. The kernel function plays an important role in some feature extraction. But appropriate kernel selection is an important part for performance. Sometimes kernels are chosen according to the characteristics of the user, which is mentioned as User-Opt-4, to choose appropriate
Linear: \( k(x_i, x_j) = x_i^T x_j \)

Polynomial: \( k(x_i, x_j) = \gamma x_i^T x_j + r \^d, \gamma > 0 \)

RBF: \( k(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2), \gamma > 0 \)

Sigmoid: \( k(x_i, x_j) = \tanh (\gamma x_i^T x_j + r) \)

Here, \( \gamma, r, \) and \( d \) are kernel parameters.

Before starting the classification of email corpus, the user needs to learn the classifier algorithm(s) using training data sets. The training data can be spam data or legitimate data or both. In the main window of the UI, there is a menu bar named "Data Classification" and under this menu bar there is a sub-menu as train classification algorithm(s). Based on the information of the training data set the test data will be classified accordingly.

Output Section

The output of classifier comes into the UI through analyse domain according to the user selection. The main activity of the output process is shown in the following figure 3

![Flow diagram of output sections.](image)

As we have discussed in the previous section that the user can select single or multi-classifier classification approach. In the case of single classifier algorithm selection, the system will send the individual output of the classifier to the output section as a spam or legitimate email. This is a simplest
f multi-classifier classification selection, the classifier will send the analyser will differentiate the filtering output into two categories;

from different classifier, i.e. \(yi(l)\), from the classes of outputs \(yi(l)\in\text{Out-Al-I} \wedge 2 \ldots \wedge N\) is very effective because all the classifier has its will send it directly to the spam or legitimate email database. This either TP (True positive) or TN (True negative).

f outputs \(yi(l)\in\{l\}\); i.e. common legitimate outputs from different outputs \(yi(l)\in\{l\}\); i.e. common spam outputs from different classifier from different classifiers i.e. not common predictions comes from the re called GL outputs.

e classes of outputs \(yi(l/s)\in\{-1,1\}\); i.e. mixed output from different \(\text{Out-Al-1} \wedge 3, \ldots \text{Out-Al-1} \wedge 2 \ldots \text{N-1}\).

f the combined classifiers can be represents as follows:

\[ yi(l) \]

\[ \text{to different database named as GL, for getting feedback from the user.} \]

\[ \text{id make decision whether it is spam or legitimate. After user feedback it nate database. The classifier will also consider the feature of these GL sput section after collecting the features from the mixed ranked outputs ion accuracy further.} \]

valuation of the classification effectiveness after analysing all types of The evaluation will be done by considering the total cost in terms of complexity as well as the accuracy of the output in terms of reducing FP

tion model using UI (figure 1), emphasis has been given on the basic in different aspects of anti-spam filtering, especially the learning-based us, because many machine learning techniques for spam filtering can a some amount of FP tradeoffs which are generally expensive in real y through different classification algorithms and found that sometimes producing the final result with same email corpora. Keeping this mind, model using multiple classifiers (the basic architecture shown in figure ter through multiple classifiers. The proposed system will also provide lly from the analysis of the GL emails. This technique enhances the ur experiment shows that using these techniques a substantial amount of in order to implement this model a intelligent UI is primary requirement
Secondly, GL analysis of this UI, the GL is the list of the emails which are not TP or TN. The term GL is related to BL and WL and considered as the middle of them, i.e. not sure about WL or BL. Because the analyser cannot make final conclusion whether it is spam or legitimate. Actually, it is the output of the Out-Al-1\textdagger, Out-Al-1\textdaggerdbl, and Out-Al-2\textdaggerdbl. These GL are sent to the specific folder to make decision about the status of these email by getting user feedback where the UI makes an important interface between user and the system. After getting user response it would be considered either WL or BL and the information of this feedback will be added to the input FS & FE domain for further accuracy.

4. Screen shot of UI

In this section the main two screen shots of our UI have been presented one for individual classification and other for combined classification.
I for spam filtering model based on classification algorithm has been given to this model based on different aspects of learning based anti-spam systems. This UI will give better flexibility to the user in order to selectation algorithms for categorizing spam from legitimate emails. The paper to find a framework for an intelligent solution to spam filtering.


...am filtering using ML Algorithms" Proceedings of the WWW/Internet


