Future Generation Computer Systems 64 (2016) 98–107

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Spam filtering framework for multimodal mobile communication based on dendritic cell algorithm

El-Sayed M. El-Alfy [a],*, Ali A. AlHasan [b]

[a] King Fahd University of Petroleum and Minerals, College of Computer Sciences and Engineering, Dhahran 31261, Saudi Arabia
[b] Saudi Aramco, Dhahran 31311, Saudi Arabia

HIGHLIGHTS

- Proposed an intelligent framework for multimodal textual spam filtering for mobile devices.
- A novel hybrid machine learning approach and fusion with dendritic cell algorithm.
- Analyzed the discrimination of a rich set of content and style related features that can be easily extracted from received messages.
- Rigorously evaluated and benchmarked models on five email and SMS datasets using a variety of performance measures.
- Reduce complexity for feature extraction while preserving good performance.

ARTICLE INFO

Article history:
Received 1 October 2015
Received in revised form
15 February 2016
Accepted 24 February 2016
Available online 3 March 2016

Keywords:
Mobile spam filtering
Feature analysis
Hybrid machine learning
Information fusion
Immune system
Dendritic cell algorithm

ABSTRACT

With the continual growth of mobile devices, they become a universal portable platform for effective business and personal communication. They enable a plethora of textual communication modes including electronic mails, instant messaging, and short messaging services. A downside of such great technology is the alarming rate of spam messages that are not only annoying to end-users but raises security concerns as well. This paper presents an intelligent framework for filtering multimodal textual communication including emails and short messages. We explore a novel methodology for information fusion inspired by the human immune system and hybrid approaches of machines learning. We study a number of methods to extract and select more relevant features to reduce the complexity of the proposed model to suite mobile applications while preserving good performance. The proposed framework is intensively evaluated on a number of benchmark datasets with remarkable results achieved.

1. Introduction

In the past decade, applications of mobile technology have witnessed a booming interest from researchers and developers. The number of mobile users has drastically increased with almost seven billion cellular subscriptions worldwide at the current time [1]. Many mobile applications have emerged with support for new features in addition to normal phone communication. Mobile devices, such as smartphones, tablets and PDAs (Personal Digital Assistants), are now used for most daily activities such as web browsing, entertainment, mobile payment, accessing medical and personal records, banking and e-learning.

Mobile devices also enable a plethora of textual communication modes which provide a very convenient way for personal and business purposes. These modes include electronic mails and short messages in the form of IM (Instant Messaging) and SMS (Short Messaging Service). One of the main concerns of such a great technology is the alarming rates of spam messages. For instance, according to Symantec, one of the leading companies for cyber security products, the global rate of spam in email traffic in 2014 was 60% [2]. Spam may contain not only irritating content to end-users such as unwanted adult and violence material but it can also waste a lot of resources and become a vector for several security breaches.

Unlike desktop applications, effective security controls to protect mobile devices are not so mature. This can be attributed to the limited resources and processing power, and the lack of knowledge and awareness of many end-users regarding security controls. In contrary, mobile devices are likely to contain personal and confidential information such as credit card numbers, contact lists, emails, medical records and other sensitive documents. These reasons and more are making mobile devices more attractive to

* Corresponding author.
E-mail address: alfy@kfupm.edu.sa (E.-S.M. El-Alfy).

http://dx.doi.org/10.1016/j.future.2016.02.018
0167-739X/© 2016 Elsevier B.V. All rights reserved.

a landscape of cyber attacks including information leak. Spam or unsolicited messages can be the easiest way for target attacks. They can be used for phishing and as a carrier vehicle for other malware types such as worms, backdoors, and key loggers. The risk of spam could be operational or financial loss. Hackers can utilize the compromised mobiles to make calls to premium numbers without the end-users' permission, stealing contact data, or participating in fraudulent and botnet activities.

The future generation of mobile technologies will witness more emphasis on security-related issues. Besides reporting spam to service providers and carriers, spam filters should be deployed at the receiving end as well to directly block unwanted messages. Other mitigation techniques include white and black listings and challenge-response authentication. Content-based filtering has received considerable attention over the past years but the major focus was on spam emails. Relatively recent, some methods have been proposed for SMS spam [3–7]. However, the accuracy is still relatively low and further research is required to investigate new features and new lightweight ways of calculating and utilizing them.

Most of the existing approaches for spam filtering focus on email spam and often treat the problem as a document categorization or genre classification problem where individual messages are preprocessed and represented by term weight vectors such as TF–IDF (which uses the product of Term Frequency (TF) and Inverse Document Frequency (IDF)) [8]. Then, statistical or machine-learning models are built using a training corpus to determine whether a particular message is spam or legitimate (ham or non-spam). However, spam spans a wide range of topics and hence it will be more effective to consider characteristics that are not only content related but also stylistic features [9]. Moreover, although both email spam and short message spam share a lot in common, there is little material in short messages for content-based filtering due to the limited message size, less contextual information, and use of idiosyncratic language with abbreviations, phonetic contractions, bad punctuation, and emotional symbols [5]. In this paper, we analyze several stylistic and content-related features and study their impact on three machine-learning algorithms used separately and as a majority-vote committee. Inspired by the danger theory and the immune-based systems, we propose a novel approach based on the Dendritic Cell Algorithm (DCA) for fusing the results of Naïve Bayes (NB) and Support Vector Machines (SVM). DCA is a relatively recent approach in machine learning inspired by the function of the biological immune system dendritic cells (DCs) [10]. Using three spam email datasets and two SMS datasets, we evaluate and compare the effectiveness of individual feature sets and their impact on the classification performance for their impact on the classification performance of the proposed model. Then, we combine the top two relevant feature sets and build a lightweight model.

The remainder of this paper is organized as follows. Section 2 briefly reviews related work. Section 3 describes the methodology and Section 4 presents the empirical analysis and results. Finally, Section 5 concludes the paper.

2. Related work

Machine-learning based solutions for mining network data has received considerably growing attention from the security community to strengthen the resilience of information systems against various types of malicious activities. Empirical evaluation of some machine-learning algorithms on benchmark corpora are presented in [11–16]. These algorithms belong to various categories including probabilistic, decision tree, support vector machines and lazy algorithms.

Carpinter and Hunt [17] have reviewed current and potential future tools for automated spam filtering including machine-learning and non-machine-learning approaches. Several machine-learning approaches for spam filtering have been also discussed in [18]. Content-based SMS spam filtering has been an active area of research [5]. In [19], Bozan et al. have presented an SMS spam filtering approach based on text classification and SVM, Bayesian and KNN classification methods. In [20], an email spam filtering has been proposed based on the group method of data handling networks. Support vector machines have been applied to filtering spam emails [21] and short message spam [22]. Another approach based on Bayesian classification for SMS spam has been proposed in [23]. New content-based features have been investigated to improve the performance of SMS spam detection [24]. Several variants of boosting trees have been studied for filtering spam emails in [25]. In [26], Chen et al. have proposed an SMS spam control based on trust evaluation through the analysis of spam detection behaviors and SMS traffic data.

Various forms of aggregations and hybrid solutions have been explored in the literature. For instance, a multi-layer pipeline for spam filtering has been studied in [27]. The stages consist of DNS blacklists and filters based on SYN packet features, traffic characteristics and message content. Another hybrid technique has been presented in [28] called symbiotic filtering consisting of distinct local filters from several users. This approach has been shown to be robust against both dictionary and focused contamination attacks. To address the difficulty to obtain negative training examples and the limitation of single class learning, Wei et al. have proposed a two-stage framework [29]. Ying et al. have proposed an ensemble approach to classify spam emails based on decision tree, support vector machine and back-propagation network [30]. In our earlier work [31], we have proposed a two-stage classifier using dendritic cell algorithm for filtering SMS messages. The proposed method was evaluated on two benchmark SMS datasets. In [32], a hybrid approach of content-based filtering and challenge-response has been discussed. The content based filter classifies a message as spam, harm or uncertain. If a message is classified as uncertain, it is checked further by sending a challenge to the message sender. The idea is that an automated spam generator is unlikely to reply with a correct response which is an indication that the message is spam. The simulation results have demonstrated that this approach can achieve high accuracy regardless of the algorithm used for content-based filtering.

3. Methodology

Fig. 1 shows a high-level outline of the core steps in the proposed spam filtering framework. Textual messages are received through various communication means including email, SMS and IM. Using a corpus of pre-classified messages, the training phase, as indicated by solid thick work flow path, constructs a classification model. During this phase messages are preprocessed and analyzed for extracting relevant features. Each message is represented with a vector. As indicated by the dashed thin lines, validation operation can be optionally performed while training a classifier. Once a classifier is constructed, it will be deployed to predict the class of newly received messages, as indicated by the solid thin lines. The details of preprocessing, feature extraction and training are provided in the following subsections.

3.1. *Corpus analysis and representation*

3.1.1. *Preprocessing*

The preprocessing phase includes the following steps. To enrich messages, we added two types of semantic information tagging: