Engineering Applications of Artificial Intelligence 39 (2015) 33–44

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

A combined negative selection algorithm–particle swarm optimization for an email spam detection system

Ismaila Idris[a], Ali Selamat[a,b,e,*], Ngoc Thanh Nguyen[c], Sigeru Omatu[d], Ondrej Krejcar[e], Kamil Kuca[e], Marek Penhaker[f]

[a] Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b] UTM-IRDA Digital Media Center of Excellence, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[c] Knowledge Management Systems Division, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
[d] Department of Electronics, Information and Communication Engineering, Faculty of Engineering, Osaka Institute of Technology, 5-16-1 Omiya, Asahiku, Osaka 535-8585, Japan
[e] University of Hradec Kralove, FIM, Center for Basic and Applied Research, Rokitanskeho, 62, Hradec Kralove, 500 03, Czech Republic
[f] Department of Cybernetics and Biomedical Engineering, Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 17. Listopadu 15, 708 33 Ostrava-Poruba, Czech Republic

ARTICLE INFO

Article history:
Received 9 February 2014
Received in revised form
29 October 2014
Accepted 1 November 2014
Available online 11 December 2014

Keywords:
Negative selection algorithm
Differential evolution
Particle swarm optimization
spam detectors

ABSTRACT

Email is a convenient means of communication throughout the entire world today. The increased popularity of email spam in both text and images requires a real-time protection mechanism for the media flow. The previous approach has been limited by the adaptive nature of unsolicited email spam. This research introduces an email detection system that is designed based on an improvement in the negative selection algorithm. Furthermore, particle swarm optimization (PSO) was implemented to improve the random detector generation in the negative selection algorithm (NSA). The algorithm generates detectors in the random detector generation phase of the negative selection algorithm. The combined NSA–PSO uses a local outlier factor (LOF) as the fitness function for the detector generation. The detector generation process is terminated when the expected spam coverage is reached. A distance measure and a threshold value are employed to enhance the distinctiveness between the non-spam and spam detectors after the detector generation. The implementation and evaluation of the models are analyzed. The results show that the accuracy of the proposed NSA–PSO model is better than the accuracy of the standard NSA model. The proposed model with the best accuracy is further used to differentiate between spam and non-spam in a network that is developed based on a client–server network for spam detection.

1. Introduction

Different techniques have been adopted to stop the threat of spam or to drastically reduce the amount of spam that attacks internet users across the world. An anti-spam law was enacted by legislating a penalty for spammers who distribute email spam (Bambauer, 2005). Additionally, two general approaches have been used in email spam detection: a knowledge engineering approach and a machine learning approach (Wamli et al., 2009). The knowledge engineering approach uses network information and internet protocol address techniques to determine whether a message is spam or non-spam; this approach is known as the origin-based filter. Sets of rules must be specified in the knowledge engineering approach to determine which email is to be categorized as spam or non-spam. Such rules could be created by the use of filters or by some other authority (Bambauer, 2005). An example of this process is a software company that provides specific rule-based spam filtering tools. However, the rules must be maintained continuously and must be updated, which is a waste of time and is inconvenient for most users (Thonnard et al., 2012). The machine learning approach is more efficient than the knowledge engineering approach (Guzella and Caminhas, 2009) and does not require specifying rules; instead, a set of pre-classified email messages is utilized. Specific algorithms are used to learn the classification rules from the email messages. Filtering techniques are the most commonly used methods; the system identifies whether a message is spam or non-spam based solely on the message content and some other characteristics of the message (Man and Mousoli, 2010). Despite the different approaches and techniques that have been adopted to fight the threat of email spam, the internet today still manifests an enormous amount of

[*] Corresponding author at: Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.
E-mail address: aselamat@utm.my (A. Selamat).

http://dx.doi.org/10.1016/j.engappai.2014.11.001
0952-1976/© 2014 Elsevier Ltd. All rights reserved.

spam (Zhang et al., 2004; Massey, 2003; Delany et al., 2012), and more attention is required with regard to how the threat can be drastically reduced if not totally eliminated. The battle against email spam is a very difficult battle; therefore, it makes sense to fight an adaptive email spam generator with an adaptive system. Most models emphasize applying and designing computational algorithms and techniques with the use of simplified models of different immunological processes (De Castro and Timmis, 2002; Dasgupta, 2006; Almeida and Yamakami, 2012; Sheikhan and Sharifi Rad, 2013). This paper proposes an improved solution to email spam detection by replacing the random detector generation in the negative selection algorithm (NSA) with particle swarm optimization (PSO). PSO is implemented with a local outlier factor as a fitness function to generate detectors in a negative selection algorithm.

The remainder of this paper is organized as follows. Section 2 discusses the related studies on negative selection algorithms. The proposed NSA–PSO and its constituent framework are presented in Section 3. An empirical study and dataset analysis are presented in Section 4. Section 5 presents the implementation, results and discussion. Finally, the conclusions and recommendations are presented in Section 6.

2. Related studies

The understanding of the artificial immune system (AIS) approach, which is based on the mammalian immune system, is vital for this study. A comprehensive artificial immune system survey has been provided in (Dasgupta et al., 2011). This paper discusses the history, recent developments and four major AIS algorithms. The main goal of the immune system is to distinguish between non-self and self elements, which is the basis of our implementation with the negative selection algorithm (NSA). This research will replace ‘self’ in the mammalian immune system with ‘non-spam’ in our system and ‘non-self’ in the mammalian immune system with ‘spam’ in our system. Most of the work on the negative selection algorithm (NSA) and particle swarm optimization (PSO) solves problems in anomaly detection and intrusion detection. No previous research implements PSO to generate detectors in a negative selection algorithm. The implementation of particle swarm optimization with the negative selection algorithm to maximize the coverage of the non-self space was proposed by Wang et al. (2009) to solve the problem of anomaly detection. In Gao et al. (2007), the focus is on non-overlapping detectors that have fixed sizes, to achieve maximal coverage of the non-self space; this approach is initiated after the generation of detectors by a negative selection algorithm. The artificial immune system (AIS) is a new mechanism that is implemented for the control of email spam. Pattern matching was used to represent detectors as regular expressions by (Oda and White, 2003a) in the analysis of messages. A weight is assigned to the detector; this weight is decremented or incremented when observing the expression in the spam message, and the classification of the message is based on the threshold sum of the weight of the matching detectors. This system is intended to be corrected by either increasing or decreasing all of the matching detector weights with 1000 detectors generated from a spam-assassin heuristic and a personal corpus. The results were acceptable based on the small number of detectors that was used. A comparison of two techniques to determine the message classification using a spam-assassin corpus with 100 detectors was proposed by (Oda and White, 2003b). This approach is similar to previous techniques, but the difference is the increment in the weight when there is recognition of patterns in spam messages. A random generation of detectors does not help in solving the problem of finding the best selected features; however, the feature weights are updated during the matching process. The weighting of the features complicates the performance of the matching process. More experiments were performed by Oda and White (2005) with the use of a spam-assassin corpus and a Bayesian combination of the detector weights. The messages are scored by the simple sum of the message matched by each non-spam in the detector space and also the use of Bayes scores. Words from the dictionary and patterns extracted from the set of messages are considered in the detector generation in addition to the commonly used filters, to be assured of the message classification. It was finally observed that the best results emerged when the heuristic was used and that it had a similar performance to the other two techniques. The immune system classifies correctly 90% of the messages. More specifically, it classifies 84% of the spam and 98% of the non-spam. The approach of scoring features or feature weighting during and after the matching process creates ambiguity in the selection of important features for spam detection due to its computational cost.

The work of Wamli et al. (2009) studies the possibility of using negative selection in email spam detection without prior information of the email spam. The negative selection algorithm is divided into four concurrent working modules with two repositories: the random detector generation module, the detector maturing module, the antigen matching module and the detector aging module, with a selves' repository and a detectors repository. The TREC07 corpus (Cormack and Lynam, 2007) was used in its implementation. After the initial 1/3 of the time during the learning period, the spam detection rate is over 80%, and it is over 70% most of the time. A new solution to solve the spam detection problem, which is inspired by the adaptive immune system model, is called the cross-regulation model and was presented in Abi-Haidar and Rocha (2008). This research shows the relevance of the cross-regulation model as a biologically inspired algorithm in the detection of spam. The Enron corpus was used in its implementation with the 70% spam experiment. The accuracy and *F*-measure are 83% and 79%, respectively.

The analysis of major work performed on negative selection algorithms with a combination of two different algorithms in a hybrid email spam model is contained in Sirisanyalak and Sornil (2007). An AIS-based module that extracts features was designed and further used for a logistic regression model; the set of detectors was initially generated using terms that were extracted from the training message and using data from matched detectors that were used in the regression model. The experiment uses spam-assassin. A genetic optimized AIS culled old lymphocytes (replacing the old lymphocytes with new ones) and also checked for new interests for users, using an approach that was similar to that presented in Hamdan and Abu (2011), to update intervals such as the number of received messages. An interval is updated with respect to time, user requests and other factors; many choices were used in selecting the update intervals, which was the aim of using the genetic algorithm. The experiment was implemented with a spam-assassin corpus that had 4147 non-spam messages and 1764 spam messages. The optimized spam detector with 600 generated detectors gives a false positive rate of 1.1% and a false negative rate of 3.7%, while spam detection with AIS and 600 generated detectors gives a false positive rate of 1.2% and a false negative rate of 4.9%. Other optimized algorithms are presented in Yildiz (2013), Mazhoud et al. (2013), Tenne (2012). A proposed anti-spam filter with an evolutionary algorithm is presented in Yevseyeva et al. (2013). The scores of the anti-spam filters are optimized to improve their accuracy. The optimization problem is considered in a single- and multi-objective problem formulation. Rough set theory, which is a mathematical approach for approximate reasoning, was proposed in Wenqing and Zili (2005) to group messages into three classes while targeting a low false positive rate. The selection of features into spam, non-spam or