ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/318298908

A Systematic Literature Review on SMS Spam Detection Techniques

Article *in* International Journal of Information Technology and Computer Science · July 2017

DOI: 10.5815/ijitcs.2017.07.05

CITATION
1

READS
93

2 authors, including:

Lutfun Nahar Lota
East West University (Bangladesh)
1 PUBLICATION **1** CITATION

SEE PROFILE

All content following this page was uploaded by Lutfun Nahar Lota on 01 November 2018.

The user has requested enhancement of the downloaded file.

I.J. Information Technology and Computer Science, **2017, 7, 42-50**
Published Online July 2017 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2017.07.05

A Systematic Literature Review on SMS Spam Detection Techniques

Lutfun Nahar Lota
Institute of Information Technology, University of Dhaka, Dhaka, 1000, Bangladesh
E-mail: bit0416@iit.du.ac.bd

B M Mainul Hossain
Institute of Information Technology, University of Dhaka, Dhaka, 1000, Bangladesh
E-mail: raju@du.ac.bd

***Abstract*—**Spam SMSes are unsolicited messages to users, which are disturbing and sometimes harmful. There are a lot of survey papers available on email spam detection techniques. But, SMS spam detection is comparatively a new area and systematic literature review on this area is insufficient. In this paper, we perform a systematic literature review on SMS spam detection techniques. For that purpose, we consider the available published research works from 2006 to 2016. We choose 17 papers for our study and reviewed their used techniques, approaches and algorithms, their advantages and disadvantages, evaluation measures, discussion on datasets and finally result comparison of the studies. Although, the SMS spam detection techniques are more challenging than email spam detection techniques because of the regional contents, use of abbreviated words, unfortunately none of the existing research addresses these challenges. There is a huge scope of future research in this area and this survey can act as a reference point for the future direction of research.

***Index Terms*—**SMS Spam Filtering, SMS Spam Detection, Systematic Literature Review, Machine Learning.

I. Introduction

Short Message Service (SMS) is the most frequently and widely used communication medium. The term "SMS" is used for both the user activity and all types of short text messaging in many parts of the world. It has become a medium of advertisement and promotion of products, banking updates, agricultural information, flight updates and internet offers. SMS is also employed in direct marketing known as SMS marketing. Sometimes SMS marketing is a matter of disturbance to users. These kinds of SMSs are called spam SMS. Spam is one or more unsolicited messages, which is unwanted to the users, sent or posted as part of a larger collection of messages, all having substantially identical content. The purposes of SMS spam are advertisement and marketing of various products, sending political issues, spreading inappropriate adult content and internet offers. That is why spam SMS flooding has become a serious problem all over the world. SMS spamming gained popularity over other spamming approaches like email and twitter, due to the increasing popularity of SMS communication. However, opening rates of SMS are higher than 90% and opened within 15 minutes of receipt whereas opening rate in email is only 20-25% within 24 hours of receipt [28]. Thus, a proper SMS spam detection technique has significant necessity. There are several researches on email, twitter, web and social tagging spam detection techniques. However, a very few researches have been conducted on SMS spam detection. Spam SMS detection is more challenging than email spam detection because of the restricted length of SMS, use of regional content and shortcut words and SMS contains less header information than an email.

We cannot use techniques of email spam detection as-is in SMS spam detection. Proper SMS spam detection technique is needed to be identified. This is an open and comparatively new research field. There is a huge scope of research work in this field. A Systematic Literature Review (SLR) is necessary for starting any kind of research in any research field. There is no SLR on this topic. For this reason we intended to write a SLR on the field of spam SMS detection. The purpose of this study is to review the current status of SMS spam detection, finding the approaches and techniques of SMS spam detection, their advantages and disadvantages, their performance and performance measurement process using available resources to conduct a systematic literature review within time period 2006-2016. Through this research we can summarize all the researches on SMS spam detection field. This will establish a baseline for the future research. Researchers will get an overview on this research area at a glance.

II. Background and Related Work

SMS spam detection is comparatively a new research area than email, social tags, and twitter and web Spam detection. Some of the researches of Spam detection includes [1], [2], [3] etc. These researches are mostly conducted after 2011. There are several established email spam detection techniques. SMS spam detection technique has some challenges over email spam detection

Copyright © 2017 MECS

such as restricted message size, use of regional and shortcut words and limited header information. These challenges need to be solved. There is scope of research in this field and some research works have been conducted on it. There are different categories of SMS spam filtering such as white listing and black listing, content-based, non-content based, collaborative approaches and challenge-response technique [4], [5], [12], [29]. The techniques are used in client side, server side or in both client and server side [4]. Several Machine Learning Algorithms such as Naïve Bayes, Support Vector Machine (SVM), Logistic Regression, Decision Trees, K-Nearest Neighbor are used to classify between Spam and legitimate SMSes named as Ham. Discussion about the machine learning algorithms, process and techniques of spam filtering is discussed in the following subsections.

A. Machine learning Algorithm

Bayesian is a probabilistic approach that starts with a prior belief, observes some data and then updates that belief. The probability being spam and not spam of a word can be calculated with the frequency of that word in ham and spam messages using the Bayesian algorithm [30]. A prior probability also needs to be assumed in this algorithm which is a shortcoming of this approach.

Support Vector Machines are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. If a set of training example containing spam and legitimate SMS is given, then an SVM training algorithm builds a model that can assign new examples into spam and legitimate category. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on [31].

The binary logistic model is used to estimate the probability of a binary response based on one or more predictor (or independent) variables (features). Logistic regression can be used in SMS spam detection on the basis of different feature variables [32].

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance of event outcomes. A decision tree can be used to make decision that whether a new message is spam or ham [33].

The k-nearest neighbors algorithm (k-NN) is a nonparametric method used for classification and regression. The input consists of the k closest training examples in the feature space. The output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors [34].

Random Forests grows many classification trees. To classify a new SMS from an input vector, the algorithm puts the input vector down each of the trees in the forest. Each tree gives a classification, called "votes" for that class. The forest chooses the classification having the most votes [35].

B. Spam Filtering Process

A manually classified spam and ham messages are input or training set for a spam filtering algorithm. The algorithm consists of the following steps [12].

Preprocessing: Removing irrelevant contents like stop words are the part of data preprocessing.

Tokenization: Segmenting the message according to words, characters or symbols called tokens. There are different tokenization approaches such as word tokenization, sentence tokenization, word or character N-grams and orthogonal sparse bigrams.

Representation: Conversion to attribute value pairs.

Selection: Selecting important attribute values which have impact on classification rather than choosing all pairs of attribute value.

Training: Train the algorithm with the selected attribute values.

Testing: Test the newly arrived data with the training model.

C. Content Based Filtering

Most of the works on SMS Spam detection are content based [1], [3], [11], [12]. Content based filtering is based on the contents of SMS like spam words, unusual distribution of punctuations and message length. Yadav et al. [1] proposed a user centric approach that used content based filtering using Bayesian machine learning algorithm with user generated features like blacklisting and white listing, preferred keywords to filter unwanted SMSes and reduced the burden of notifications for a mobile user.

Narayan et al. [3] developed a two level stacked classifier to classify between spam and legitimate SMS. The first level of classifier records a subset of words whose individual probability is higher than a threshold. After that second level of classifier is invoked, this takes the chosen words form first level as input. They took different combinations of machine learning classification algorithms in two levels such as Bayesian and SVM, SVM and Bayesian, Bayesian and Bayesian, SVM and SVM.

Ishtiaq et al. [11] proposed a SMS spam classification algorithm using the combination of Naive Bayes classifier and Apriori algorithm. They integrated association rule mining using Apriori algorithm with Bayesian algorithm. Apriori retrieves the most frequent words occurred together then Bayesian calculates the probability of occurring a word independently and together with other words, in spam or ham messages.

Gomez et al. [12] analysed to what extent Bayesian filtering techniques used to block email spam, can be applied to the problem of detecting and stopping mobile

Copyright © 2017 MECS

spam. They pre-processed the messages with different tokenization approach, selected features and tested them with different machine learning algorithms, in terms of effectiveness. They demonstrated that Bayesian filtering techniques can be effectively transferred from email to SMS spam with appropriate feature extraction.

D. Non-Content based filtering

Many proposed techniques used non-content based filtering [2], [7]. Warade et al. [2] detected the spam messages by checking mutual relation between the sender and receiver and the content of the messages. If no mutual relation is found between sender and receiver and message contains spam contents, then the system tags the message as spam and sends it to spam box. If mutual relation and no spamming content exist then it directly sends to inbox of the receivers mobile. It solved the problem of balance deduction and wastage of SMS memory. But calculating only mutual relation is not a proper solution. Spam detection algorithm needs both classification algorithm and this kind of feature extraction from contents.

Qian Xu et al. [7] investigated ways to detect spam message senders based on non-content features that include temporal and graph-topology information but exclude contents because of user-privacy issues. They focused on the problem of identifying professional spammers based on the overall message sending patterns. Furthermore, they only concentrated on finding SMS spam on the server side, as the client-side detection is mostly content based.

E. Feature Engineering

The success of machine learning depends mostly on appropriate feature selection [6]. The feature can be both content based and non-content based. The ref. [2], [8] focused only on non-content based features like mutual relation of sender and receiver, user black-listing and white listing and user preferred keywords words. Whereas some researchers considered only content based features [6]. A proper spam detection algorithm needs both content and non-content based features. Non-content based features include static, temporal and network features [7]. Content based features are word frequencies [11] and keyword based features are presence of spam words and stylistic features are count of exclamation, count of alphanumeric word, average word length and many others [15].

III. Methodology

A systematic review collects and critically analyses multiple research studies or papers or journals and provides the summary of the existing literature on a specific research domain [9]. A review of existing studies is often quicker and cheaper than embarking on a new study. For conducting SLR, some steps need to be followed as mentioned in [9]. The steps include formulating research questions, finding and analysing researches that relate to the questions, answering to that questions and demonstrating a summarized result of literature survey. Details of these steps are discussed in the following sub-sections.

A. The Need for a Systematic Literature Survey

Email Spam detection is an established research field. Many researches and literature survey have been done on email spam detection as well as for twitter, web and social tag spam detection. There is insufficient systematic literature survey available on SMS spam detection because of its being comparatively new research area. Although SMS communication has started mostly in 2000, it gained its popularity in 2006 and even became more popular after the flourishment of android phones [19]. With the increase of the number of people using SMS as a communication medium, SMS spamming also gets more popularity to spammers. As a result, research on SMS spam detection had emerged with its necessity and researches on it have started mainly after 2007. Our goal with this SLR is collecting proper background knowledge on SMS spam detection field, gaining knowledge about the currently used algorithms for SMS spam detection, their advantages and disadvantages, identifying the evaluation measure for the spam detection algorithms, comparing the accuracy of the algorithms, identifying any gaps in current research in order to suggest areas for further investigation. The motivation for this work is to establish a basis for any research on SMS spam detection. Any kind of research starts on the basis of systematic literature review. This is the main rationale of this SLR.

B. Research Questions

Identifying research question is one of the important steps in a SLR. We have identified three research questions for this SLR. The questions and their motivation are presented in table 1.

Table 1. Research Question

RQ1. What are the current approaches of SMS spam detection?	To identify the algorithms used for SMS spam detection.
RQ2. What are the advantages and disadvantages of the algorithms?	To understand the convenience and drawbacks of the algorithms.
RQ3. What are the measurement policies of SMS Spam detection algorithms?	To identify existing measurement policies and metrics to evaluate the algorithms.

C. Searches for Studies

At first we searched with the term 'SMS spam detection' on Google Scholar. Then we identified keywords noted in the relevant papers. After that we identified alternative spelling and synonyms for search terms. Some examples of resulting search string are given below: “SMS Spam”, “SMS Spam Filtering”, “Machine Learning”, “Security and Protection”, “Text Analysis”, “Security in Mobile Communication”, “Short Message Service”, “Naive Bayesian Algorithm”, and “Anti-Spam Filtering”.

Copyright © 2017 MECS

D. Study Selection Procedure

To select relevant studies we primarily searched on google scholar. We have collected some papers from it. There are some other conferences and journals such as: IEEExplore, ACM, IJCSI, ITJ are found through google scholar tool. The list of journals and conferences from where we have found our relevant papers is presented table 2. We also performed manual google search. Selected paper contains many references; we also searched for the referenced papers and have taken some of them as our relevant paper. We used the google scholar's related articles and cited by feature for our searching procedure.

Table 2. Sources Searched

No.	Source	Abbreviation
1	IEEE	IEEE Xplore
2	ACM	Association for Computing Machinery
3	IJCSI	International Journal of Computer Science Issues
4	IJISS	International Journal of Information Security Science
5	ITJ	Information Technology Journal
6	IJRAT	International Journal of Research in Advent Technology
7	IJITCS	International Journal of Information Technology and Computer Science
7	CAE	International CAE Conference
8	Google Scholar	
9	Google	
10	Computers and Security	
11	Expert Systems With Applications	

E. Study Selection Criteria

There are inclusion and exclusion criteria for systematic literature review. SMS spam detection is a new research area and there are not much relevant studies in this field. That is why we chose most of the available articles.

F. Data Extraction

Table 3 contains the extraction form used to gather extracted information from our study. This table demonstrates information about our chosen data such as chosen papers type, their published conferences, publication years, motivation and methodology of paper.

Table 3. Data Extraction Table

Data item	Value
Study identifier	S#
Paper type	Conference/ Journal
Name of e-library	e.g. ACM
Year of publication	2006- 2016
Name of journal	e.g: ITJ
Which RQ was answered	RQ1/ RQ1/RQ3
Outcomes of the paper	Summarized literature survey on SMS spam detection
Motivation of paper	Create a baseline for SMS spam detection
Method of paper	Techniques/ Approaches / Algorithms
Validation of paper	Analysis model

IV. Validation of the Study

Our SLR was conducted to investigate all the used approaches and techniques in SMS spam detection. The threats to the validity of our review are that there may be selection bias and lack of sufficient resources. We tried to reach all possible and relevant information resources. Some resources might not have been published directly. Another threat is some resources are not available for public use.

V. Result Analysis

At first we manually searched on google using the topic Spam Detection to gain an overview in spam detection field. It resulted in many email, twitter, web and SMS spam detection related papers. Then we customized our search using only SMS spam detection. It resulted in a few papers. Although there are SLR for other spam detection techniques but none of the search strings produces a SLR for SMS Spam detection. Through our study selection procedure we have chosen 17(S1-S17) papers published in different conferences and journals relating only to SMS spam detection. Among the 17 studies S1 and S11 are from same authors and S11 is an extension of S1. The ref. [20] is a journal which is an extension of the conference paper S10. S12 is an extension of [8]. As a result, in total we have studied 19 studies. Table 4 summarizes the reviewed papers Study ID with the reference no given in reference section, publication years, name of the conferences and journals where the papers published and the research questions they answered.

Copyright © 2017 MECS

Table 4. Summary of the Reviewed Literature

Study ID	Year	Conference/Journal	Answer Research Question
S1 [1]	2012	IEEE	RQ1, RQ2, RQ3
S2 [2]	2014	IJRAT	RQ1
S3 [3]	2013	ACM	RQ1,RQ2, RQ3
S4 [5]	2010	Computers and Security	RQ1,RQ2, RQ3
S5 [10]	2012	IJCSI	RQ1 RQ2, RQ3
S6 [11]	2014	IJMLC	RQ1 RQ2, RQ3
S7 [12]	2006	ACM	RQ1 RQ2, RQ3
S8 [7]	2012	IEEE	RQ1, RQ3
S9 [13]	2015	CAE	RQ1 RQ2, RQ3
S10 [14]	2011	ACM	RQ1, RQ3
S11 [15]	2011	ACM	RQ1, RQ3
S12 [16]	2007	ACM	RQ1, RQ3
S13 [17]	2008	ITJ	RQ1, RQ3
S14 [18]	2014	ASTL	RQ1
S15 [4]	2015	Information Security Journal	RQ1 RQ2, RQ3
S16[21]	2014	JBASR	RQ1 RQ2, RQ3
S17[22]	2013		RQ1, RQ3

Table 5. Summary of the Techniques Used by the Literature

Study ID	Techniques/ Algorithms/ Approaches	Description
S1 [1]	Content Based (Bayesian)	SMSAssassin: Android application uses content based filtering with user generated features to automatically filter spam SMSes resulting in different tabs.
S2 [2]	Mutual Relation	Based on the previous relation of sender and receiver.
S3 [3]	Two level stacked classifier	First level Records some words more than a threshold then sends them to the next level using Bayesian in both level and Bayesian in 1st level and SVM in second level.
S4 [5]	Hybrid Approach(Content Based and Challenge – Response)	Used upper and lower bound of threshold introducing an uncertain region for Bayesian filtering after that the messages which fall into uncertain region sent to the challenge – response technique which is user query based.
S5 [10]	Artificial Immune System	The phases of the algorithm are Building dataset, Message Matching and Affinity Calculation.
S6 [11]	Bayesian and Apriori Algorithm	Apriori retrieves the most frequent words occurred together then Bayesian calculates the probability of occurring a word independently and together with other words, in spam or ham messages.
S7 [12]	Bayesian	Message pre-processing and encoding, feature selection and then applying the classification algorithm.
S8 [7]	Non- Content Based	Non content-based features such as static, temporal and network features then classification with SVM and KNN.
S9 [13]	Bayesian with modified formula	Total number of spam SMSes are divided by the total occurrences of a word in Spam/Ham messages instead of the formula of occurrences of words divided by the total number of Spam and Ham messages. They also combined two formulas.
S10 [14]	Tokenization with various classifier	Two kinds of tokenization : separated by blanks and separated by special characters are used for classification in various machine learning algorithms.
S11 [15]	Bayesian and SVM	Tested the feasibility of applying both algorithms in mobile application domain.
S12 [16]	Content Based filtering	Machine learning algorithms with Lexical feature expansion such as words, orthogonal sparse word bigrams, character bigrams and trigrams.
S13 [17]	Feature Updating Protocol	At a regular interval on the basis of new arrival of SMSes Features will be updated using methods like document frequency, term frequency, information gain and mutual information.
S14 [18]	Virtual Ratio on Naïve Bayes, J-48 and logistic regression	VR is the relative ratio of average frequency of a keyword in spam and ham messages.
S15 [4]	Artificial Immune System	Consists of five modules: Innate mechanism, User feedback, Quarantine, Tokenizer, Immune Engine.
S16 [21]	Bayesian, Multilayer Perceptron Algorithm, Decision Tree	Selected four features and performed classification algorithms on them resulting in better performance in Bayesian
S17 [22]	Content based Filtering	Feature extraction and classification algorithms like Bayesian, SVM, K- Nearest neighbour, Random forest and Adaboost. There results concludes SVM outperforms other algorithms.

A. RQ1: What are the current approaches of SMS spam detection?

The used techniques, approaches and algorithms in spam detection and their short description with their study id is described in table 5. From the table we can see that, most of the approaches use content based filtering and for classification they used several machine learning algorithms mostly Bayesian and SVM. Study S1, S3, S6-S7, S9-S12, S14, S16-S17 used content based filtering. S4 is a hybrid approach, S8 is non content based, and S5 and S15 are based on artificial immune system. Most of the content based filtering used Bayesian as a classification algorithm.

B. RQ2: What are the advantages and disadvantages of the algorithms?

Table 6 demonstrates the result of RQ2. The advantages and drawbacks of the approaches are mentioned in the tables. From the table we can say that, content based filtering is more convenient than other non-content based and server side algorithms. Server side algorithms suffer from implementation complexity. Feature selection is also an important task for machine learning algorithms to work correctly. One important drawback is, some approaches do not use classification algorithm only focusing on user generated features. Classification algorithm is necessary for gaining better accuracy.

Table 6. Advantages and Disadvantages of Used Techniques

Study ID	Advantages	Disadvantages
S1 [1]	Combination of machine learning algorithms with user generated features	Users need to select features manually
S2 [2]		No classification algorithm is used
S3 [3]	Classification based on two algorithms	Threshold selection
S4 [5]	Combination of client and server side algorithms	Challenge-response technique suffers from server side traffic and user interaction problems
S5 [10]	Accurate as Naïve Bayesian with necessary feature extraction	Complex implementation
S6 [11]	Incorporating Apriori Algorithm	
S7 [12]	Used a weighting Mechanism to reduce false negatives	
S8 [7]		Suffers from implementation complexity
S9 [13]	Combination of two formulas gives better result in terms of false positives	
S10 [14]	Concludes SVM outperforms other algorithms and created a baseline for further comparison	
S11 [15]	Although SVM gives better results in Spam identification Bayesian is more feasible for mobile applications	Extensive feature engineering is needed for better accuracy
S12 [16]	Demonstrates the need of spam filtering in spite of having established email spam filtering	
S14 [18]	Lightweight and focuses on runtime	
S15 [4]		Server side, complex and suffers from updating issues
S16[21]		Implementation complexity

C. RQ3: What are the measurement policies of SMS Spam detection algorithms?

Accuracy of the SMS spam detection needs to be measured. In table 7, we have demonstrated the method or matrix to measure the algorithms for each study. Calculating accuracy from confusion matrix is one of the most commonly used measurement methods for classification algorithms. S3-S9, S11, S15, S17 used accuracy to measure their algorithm. Receiver operating characteristics (ROC) and Area under the curve (AUC) were also used to demonstrate algorithm accuracy. S7, S8, S11, S12 used ROC and AUC methods. True Positive rate, False Positive rate, F-measure, Precision, and Recall are also measurement methods for classification algorithms, which can be calculated from confusion matrix. Some of the studies also used these measures. S2 and S14 do not use any evaluation measure.

D. Dataset Description

A training dataset is needed for any kind of machine learning classification algorithms. Results of the machine learning algorithms depend on the dataset. As a result spam detection algorithms can't run without a dataset. In table 8, we demonstrated different publicly available dataset used in different studies. Link of the dataset and some statistics such as total number of SMSes, number of Spam and Ham messages are shown in the table 8.

E. Performance Comparison

Most of the results of our studies demonstrated that Bayesian filtering is more suitable for spam detection. S3 showed that a two level stacked classifier using dataset referenced in [25] gives better accuracy of 99% with threshold 0.4 and 0.6 than the single classifiers. Hybrid approach of S4 demonstrates accuracy of 95%. S15 and S5 based on artificial immune system shows accuracy of 99% and 98% respectively. S6 gives 98%-100% on the

Copyright © 2017 MECS

dataset [26] but they did not consider all the data instead they choose small portions of the dataset and this accuracy is achievable only for some specific parts of the dataset. S9 shows 89% accuracy on some publicly not available Farsi SMS dataset with modified Bayesian formula. 97% accuracy is achieved by SVM with Spam Caught Rate 83.10% and Blocked Ham rate 0.18% on the dataset [26]. Whereas 98% accuracy is achieved by SVM on the same dataset [26] with Spam Caught Rate 92% and Blocked Ham Rate 0.31% in S17. This observation shows that results not only depends on classification algorithms and datasets but also on data preprocessing and feature selection process. S17 also demonstrates accuracy 98% on Bayesian with Spam caught rate 94% and Blocked Ham Rate 0.51%. S11 shows 97% ham accuracy and 72.5% spam accuracy on Bayesian and 93% ham accuracy and 86% spam accuracy on SVM on some publicly not available dataset. S16 showed 92 % correctly classified instances and 8% incorrectly classified instances on Bayesian which is better than Multilayer perceptron and Decision tree.

Table 7. Evaluation Measures of the Algorithms

Study ID	Evaluation Measure
S1 [1]	No evaluation measure only demonstrate their application
S2 [2]	
S3 [3]	Precision, Recall, F-measure and Accuracy
S4 [5]	Traffic amount, Accuracy
S5 [10]	Accuracy, False Positive Rate
S6 [11]	Accuracy
S7 [12]	ROCCH
S8 [7]	False Positive Rate, AUC
S9 [13]	Confusion matrix, Precision, Accuracy, F-measure
S10 [14]	Spam Caught/True Positive, Blocked ham/False positive, Accuracy and Matthews Correlation Coefficient (MCC)
S11 [15]	Ham, Spam identification Accuracy and Area Under the Curve (AUC)
S12 [16]	ROC, AUC
S15 [4]	Confusion Matrix, Accuracy, AUC
S16[21]	Correctly and Incorrectly classified Instances
S17[22]	Spam Caught(SC), Blocked Ham(BH), Accuracy (ACC)

Table 8. Dataset Description

Study ID	Available At	Total No. of Messages	Hams	Spams
S1 [1]	[24]	2000	1000	1000
S3 [3]	[25]	1450	730	721
S4 [5]			85.32%	14.75%
S6 [11]	[26]	5574	4827	747
S7 [12]	[27]			
S10 [14]	[26]	5574	4827	747
S11 [15]		4318	2195	2123
S12 [16]				
S14[18]	[26]/	5574	4827	747
S15 [4]		5240	2890	2350
S17 [22]	[26]	5574	4827	747

VI. Discussion

In light of the above discussion, we can say that most of the research studies answered RQ1. They mostly used content based filtering with various machine learning algorithms. Eleven research studies used content based filtering, two studies used artificial immune system, one of them used hybrid approach, two of them focused on feature engineering and two of them focused on real world data set. Content based filtering suffers from challenges like short content, abbreviated words and user content safety. All of the studies tried to solve some challenges of SMS spam detection. For example, some studies solved real world data extraction process, some studies proposed hybrid approach to give better accuracy, some studies tried to overcome the challenges over email spam detection. None of the techniques solved the challenge of the use of regional content and shortcut words. These challenges lead to the future researchers to further investigation on the used approaches and techniques. Also most of the studies used Bayesian filtering for classification algorithm. Bayesian algorithm also suffers from traditional threshold selection problem, dataset dependency, assuming prior probability. Despite having those shortcomings, Bayesian is declared as the most suitable algorithm for spam filtering. Solving these problems of Bayesian also can be a research direction. This can result in better performance in Bayesian algorithm. SVM also gives better accuracy but suffers from implementation complexity. Other algorithms are less suitable for SMS Spam filtering.

VII. Conclusion

This paper presents the results of the systematic literature review on SMS spam detection techniques. We chose a total of 17 research papers on this field and reviewed their proposed techniques, advantages and disadvantages and challenges they addressed. We also examined their evaluation procedures. We demonstrated the publicly available dataset information which is a prior need for a spam filtering algorithm. We also discussed the background of this topic. In our systematic literature review, we have discussed the search and selection procedure, their publication years and the journals and conferences where those studies were published. Our results show the summary of the used techniques and advantages and disadvantages of the approaches. We have performed a performance comparison on the studied literature. In addition, we have found that none of the studies solve the challenges of use of regional contents and shortcut words. We have also discussed the problems of traditional machine learning algorithms. There is scope of further research in this filed and our systematic literature review can serve as a reference point for future researches.

References

[1] K. Yadav, S. K. Saha, P. Kumaraguru, and R. Kumra, "Take control of your smses: Designing an usable spam

sms filtering system," in 2012 IEEE 13th International Conference on Mobile Data Management. IEEE, 2012, pp. 352–355.
[2] S. J. Warade, P. A. Tijare, and S. N. Sawalkar, "An approach for sms spam detection."
[3] A. Narayan and P. Saxena, "The curse of 140 characters: evaluating the efficacy of sms spam detection on android," in Proceedings of the Third ACM workshop on Security and privacy in smartphones & mobile devices. ACM, 2013, pp. 33–42.
[4] A. S. Onashoga, O. O. Abayomi-Alli, A. S. Sodiya, and D. A. Ojo, "An adaptive and collaborative server side sms spam filtering scheme using artificial immune system," Information Security Journal: A Global Perspective, vol. 24, no. 4-6, pp. 133–145, 2015.
[5] J. W. Yoon, H. Kim, and J. H. Huh, "Hybrid spam filtering for mobile communication," computers & security, vol. 29, no. 4, pp. 446–459, 2010.
[6] S. J. Delany, M. Buckley, and D. Greene, "Sms spam filtering: methods and data," Expert Systems with Applications, vol. 39, no. 10, pp. 9899–9908, 2012.
[7] Q. Xu, E. W. Xiang, Q. Yang, J. Du, and J. Zhong, "Sms spam detection using noncontent features," IEEE Intelligent Systems, vol. 27, no. 6, pp. 44–51, 2012.
[8] G. V. Cormack, J. M. G. Hidalgo, and E. P. S ´anz, "Feature engineering for mobile (sms) spam filtering," in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2007, pp. 871–872.
[9] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," in Technical report, Ver. 2.3 EBSE Technical Report. EBSE, 2007.
[10] T. M. Mahmoud and A. M. Mahfouz, "Sms spam filtering technique based on artificial immune system," IJCSI International Journal of Computer Science Issues, vol. 9, no. 1, pp. 589–597, 2012.
[11] I. Ahmed, D. Guan, and T. C. Chung, "Sms classification based on na ̈ıve bayes classifier and apriori algorithm frequent itemset," International Journal of machine Learning and computing, vol. 4, no. 2, p. 183, 2014.
[12] J. M. G ´omez Hidalgo, G. C. Bringas, E. P. S ´anz, and F. C. Garc´ıa, "Content based sms spam filtering," in Proceedings of the 2006 ACM symposium on Document engineering. ACM, 2006, pp. 107–114.
[13] M. Poorshahsavari and O. Pourgalehdari, "Enhancing the rate of accuracy and precision in spam filtering in farsi sms."
[14] T. A. Almeida, J. M. G. Hidalgo, and A. Yamakami, "Contributions to the study of sms spam filtering: new collection and results," in Proceedings of the 11th ACM symposium on Document engineering. ACM, 2011, pp. 259–262.
[15] K. Yadav, P. Kumaraguru, A. Goyal, A. Gupta, and V. Naik, "Smsassassin: crowdsourcing driven mobile-based system for sms spam filtering," in Proceedings of the 12th Workshop on Mobile Computing Systems and Applications. ACM, 2011, pp. 1–6.
[16] G. V. Cormack, J. M. G ´omez Hidalgo, and E. P. S ´anz, "Spam filtering for short messages," in Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. ACM, 2007, pp. 313–320.
[17] Q. Sun, H. Qiao, and Z. Luo, "The feature updating algorithm for short message content filtering," Information Technology Journal, vol. 7, no. 5, pp. 790–795, 2008.
[18] S.-E. Kim, J.-T. Jo, and S. [18] S.-E. Kim, J.-T. Jo, and S.-H. Choi, "A spam message filtering method: focus on run time," 2014.
[19] A Brief History of Text Messaging, http://mashable.com/2012/09/21/text-messaging-history/#F4V9_15QGkqx. [Last Accessed: 05-11-2016]
[20] Almeida, Tiago, Jos é Mar á G ómez Hidalgo, and Tiago Pasqualini Silva. "Towards sms spam filtering: Results under a new dataset." (2013): 1-18.
[21] Mujtaba, G., and M. Yasin. "SMS spam detection using simple message content features." *J. Basic Appl. Sci. Res* 4 (2014): 275-279.
[22] Shirani-Mehr, Houshmand. "SMS spam detection using machine learning approach." (2013): 1-4.
[23] Ahmed, Ishtiaq, et al. "Semi-supervised learning using frequent itemset and ensemble learning for SMS classification." *Expert Systems with Applications*42.3 (2015): 1065-1073.
[24] http://precog.iiitd.edu.in/resources.html [Last Accessed: 05-11-2016]
[25] https://github.com/okkhoy/SpamSMSData. [Last Accessed: 05-11-2016]
[26] http://www.dt.fee.unicamp.br/~tiago/smsspamcollection/ [Last Accessed: 05-11-2016]
[27] http://www.esp.uem.es/jmgomez/smsspamcorpus/ [Last Accessed: 05-11-2016]
[28] https://www.cloudmark.com/en/s/resources/whitepapers/sms-spam-overview [Last Accessed: 05-11-2016]
[29] Iqbal, Muhammad, et al. "Study on the Effectiveness of Spam Detection Technologies." (2016).
[30] http://fastml.com/bayesian-machine-learning/ [Last Accessed: 05-11-2016]
[31] https://en.wikipedia.org/wiki/Support_vector_machine [Last Accessed: 05-11-2016]
[32] https://en.wikipedia.org/wiki/Logistic_regression [Last Accessed: 05-11-2016]
[33] https://en.wikipedia.org/wiki/Decision_tree [Last Accessed: 05-11-2016]
[34] https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm [Last Accessed: 05-11-2016]
[35] https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm [Last Accessed: 05-11-2016]

Authors' Profiles

Lutfun Nahar Lota is a graduate student at the Institute of Information Technology (IIT), University of Dhaka, Bangladesh. Currently, she is pursuing her Master of Science in Software Engineering (MSSE). She earned her Bachelor of Science in Software Engineering (BSSE) from the same institution. Her core areas of interest are Software Engineering, Security and Machine Learning.

Dr. B. M. Mainul Hossain is Assistant Professor at the Institute of Information Technology (IIT), University of Dhaka, Bangladesh. He received his Ph.D. degree in computer science from University of Illinois at Chicago, USA. Before that, he earned his Bachelor of Science and Master degrees from the department of Computer Science & Engineering, University of Dhaka, Bangladesh. He has the experiences of working both in industry and academia. He worked as a Software Engineer in Microsoft Corporation

(Redmond, USA) & Accenture Technology Lab (Chicago & California). His core areas of interest are Software Engineering, Security, Data Mining and Machine Learning.

How to cite this paper: Lutfun Nahar Lota, B M Mainul Hossain,"A Systematic Literature Review on SMS Spam Detection Techniques", International Journal of Information Technology and Computer Science(IJITCS), Vol.9, No.7, pp.42-50, 2017. DOI: 10.5815/ijitcs.2017.07.05

Copyright © 2017 MECS

View publication stats