International Journal of Engineering Research & Management Technology ISSN: 2348-4039
March-2016 Volume 3, Issue-2
Email: editor@ijermt.org www.ijermt.org

# An Approach for SMS Spam Detection

**Parshuram Raut, Sharad Kumar Singh, Abhishek Kumar**
Students
IIMT College of Engineering
Gr. Noida

**ABSTRACT:**
In the today's world the use of mobile increases tremendously. And hence the companies start to use of SMS for their advertisement. At the beginning the companies are send their promotional messages through SMS gateways. But due to increasing number of promotional messages the companies start the service DND, the DND service restrict only the SMS send through SMS gateways and hence the companies start to send their promotional advertisement messages through spammer's mobile phones. The approach presented in this paper detects these messages sent through spammers mobile and restrict it.

**KEYWORDS**: SMS, spam filtering, SMS spam, mobile spam

**INTRODUCTION:**
In the recent years, we have witnessed a dramatic increment in the volume of spam email. Other related forms of spam are increasingly revealing as a problem of importance, specially the spam on Instant Messaging services (the so called SPIM), and Short Message Service (SMS) or mobile spam. Like email spam, the SMS spam problem can be approached with legal, economic or technical measures. Mobile phone spam is a form of spam directed at the text messaging or other communications services of mobile phones. It is described as mobile spamming, SMS spam, text spam, m-spam or mspam. As the popularity of mobile phones surged in the early 2000s, frequent users of text messaging began to see an increase in the number of unsolicited (and generally unwanted) commercial advertisements being sent to their telephones through text messaging. This can be particularly annoying for the recipient because, unlike in email, some recipients may be charged a fee for every message received, including spam.
Mobile spam may be a relevant drawback in Far East countries since year 2001. In Korea, the quantity of mobile spam was already larger than the quantity of email spam at the tip of 2003. Spam messages square measure accustomed advertise chemical analysis services, premium rate numbers, or merchandising medication and software system. many countries have taken legal and technical measures to regulate the SMS spam drawback. Japan government filed 2 acts in 2002 that outlined and punished email and mobile abuse These laws, the hassle of self-regulation from Mobile Network operators, and a few technical limitations, have helped to cut back the quantity (but to not quit) mobile spam. All in all, consultants contemplate that mobile spam will solely get controlled through the mixture of technical and legal measures. SMS spam has been thought to be a minor drawback in Western countries, principally as a result of the price of causation spam messages is far larger than that of causation email spam. However in Europe, SMS electronic messaging is that the fashion: nearly all folks over fifteen years recent own a mobile, and a median user sends concerning ten SMS each day that produces SMS messages an ideal target for abuse. Moreover, botnets of zombie PCs square measure getting used to emulate real users once causation SMS messages through free SMS electronic messaging services at e.g. Russia. So, SMS spam price is decreasing. In different words, mobile spam will pay. In fact, over eightieth of users admit to possess received mobile spam

a spread of technical measures against spam email are already projected. Most of them will effectively be transferred to the matter of mobile spam.

Mobile phone spam is mostly less pervasive than email spam, wherever in 2010 around ninetieth of email is spam. Quantity of mobile spam varies wide from region to region. In North America, mobile spam has steady inflated from 2008 through 2013, however remains below one hundred and twenty fifth as of December 2013. In elements of Asia up to half-hour of messages were spams in 2013.The lesser and geographically uneven prevalence of transportable spam is due to geographic variation of prevalence of mobile vs non-mobile electronic communications, the upper price (to spammers) of and technological barriers to causation mobile messages in some areas, and to enforcement in others. Today, most transportable spam is distributed from mobile devices that have postpaid unlimited electronic messaging rate plans. whereas the speed plans yield unlimited electronic messaging, really the comparatively slow causation rate (on the order of magnitude of 1/s) limits the quantity of messages that will be sent before associate degree abusing mobile is clean up here.


**Fig.1-A snapshot of a received spam SMS on an Android phone**

## APPLICATION DESIGN:

We have designed and developed an initial version of SMS Assassin system for filtering spam SMS [8] which we deployed in real world for a month in 20 different participants' phone. Based on the usage patterns and user study participant feedback, we have made changes in the previous design of the application and here, we will present the new design of SMS Assassin application

### A. IMPLEMENTATION DETAIL:

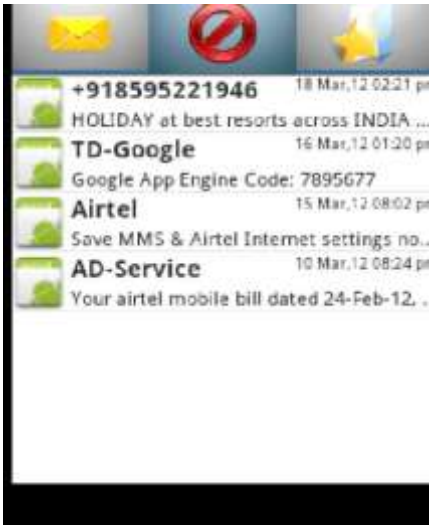
We have developed SMS Assassin application for Android and Symbian phones both. For Symbian, mobile application code was written in Qt and Simian C++ with around 5K lines of code and Android application was written in Java with nearly 7K lines of code. Both applications are developed as a standalone messaging Inbox application which can compose, read, delete and filter SMS according to user preferences. Applications have capability to read an incoming SMS and notify a mobile user based on her filtering preferences. We have developed our mobile application in such a way that it can replace the traditional Inbox of the phone with a SMS Inbox with spam filtering capability.

### B. BAYESIAN FILTERING:

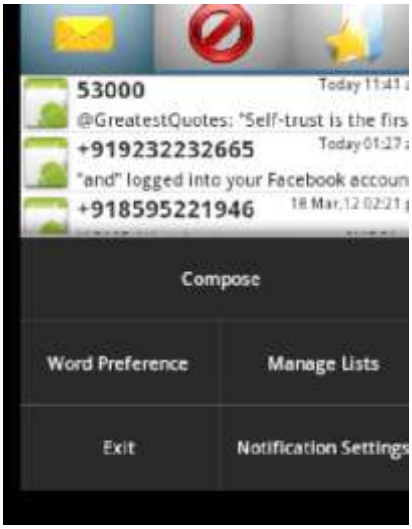
Bayesian filtering is an integral part of SMS Assassin application that is used to automatically classify SMS with user generated features. We will discuss classification of SMS based on user generated features in later

part of this section. We used Bayesian filtering in our application because it requires less computational resources and can be trained on the fly which suits for resource constrained mobile devices. Bayesian filtering is a content based supervised machine learning technique which determines whether a SMS is ham or spam based on the occurrence of keywords in the SMS. Bayesian filtering need a seed dataset containing both ham and spam SMS to train itself. 6 As part of training, Bayesian filtering computes count of each word's occurrence in both ham and spam SMS (training dataset) and store them.
Whenever a new SMS comes, applications extract all the words from it and compute a score with the help of training file. This score is then compared with a threshold parameter δ 7 to decide on a SMS is spam or ham.

**C. FEATURES:**
Our mobile application works like a messaging inbox with spam filtering capability. Hereby, we will describe some of the features of the application which is not usually found in messaging inbox of the phone, some of these features also acts as classification rules and populated by users which enables personalized filtering.
**1)** Different Tabs: The application provides three different tabs; one each for ham (Inbox) and spam SMS (Spam Box) respectively and third one is for user preferred SMS as shown in Figure 2a. Whenever a new SMS comes, the application automatically decides on corresponding tab for it based on its filtering mechanism. This kind of tabular interface and segregation
of SMS makes management and viewing of SMS every easy for the users on small screen mobile devices. All three different tabs provide flexibility to move SMS across different tabs to train the system in making future decisions.
**2)** Sender blacklisting and Sender White listing: Sender blacklisting feature is used to block a particular sender. It is a user driven feature where user blacklists.

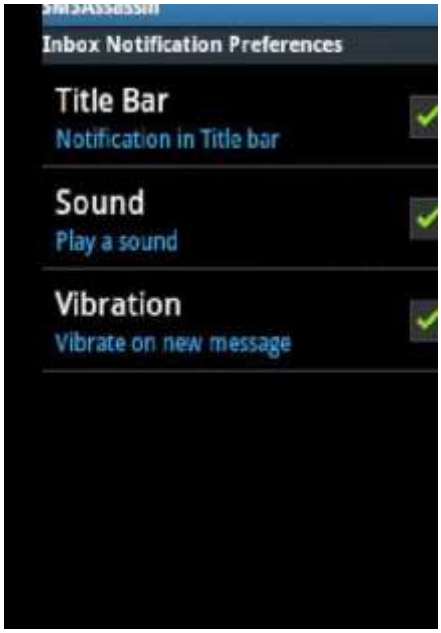



(a) Main Snapshot with three different tabs for different kind of SMS.




(b) Different features of the application.


**(c) Blacklisted senders by the User.**




**(d) Notification preferences for the Inbox tab**

**Figure 2: Snapshots of running SMS Assassin mobile application in Android based Samsung Galaxy Y phone**.

**3) REPORT AS HAM OR SPAM:**

If a SMS is spam and wrongly put into Inbox by the application then the user can report it as spam to tell application about wrong decision. The application gives an option to blacklist sender if a SMS is moved from Inbox to Spam Box. Similarly, if a ham SMS is put into Spam Box then user can report it as ham and subsequently move into Inbox.

**4)** User Preferences: In this feature, user can bind any preferred keyword (for instance, "Pizza") or phone number(s) to user preferred tab. All the SMS which are from the user preferred sender or contains the user preferred word will come to user preferred tab (rightmost tab in Figure 2a). This feature provides an alternative viewing and storing space to the user apart from Inbox which is flooded with different kind of SMS. In our previous user studies, most of participants demanded this kind of feature as searching of SMS becomes difficult as number of SMS in the messaging Inbox grows. We have seen some more specific use cases of this feature, which is described as follows:

**a)** Alice wants all the preferred promotional SMS to keep separate from the Inbox of the phone because she uses them often to get discounts. Whenever she receives a Pizza offer or a snapdeal offer, it should directly go to user preferred tab. In the application, she would just create a word or sender filter so that the incoming SMS comes into correct tab.

**b)** Bob has a group of friends, whenever he receives a SMS from any of them, he wants to keep it separate from the normal Inbox of the phone so that searching of a particular communication is easier and he can later refer to it quickly. In this case, he will bind phone number of all those friends to the user preferred tab of the application.

**5) CUSTOMIZED NOTIFICATIONS:**

A notification generated by unwanted SMS is very annoying to the users. Application provides a customized notification mechanism for each tab which can be set by user according to her own requirements, Figure 2d shows a snapshot of the application to customize notifications for all the SMS which will come to Inbox, it has similar Interface for other two tabs also. Typically, all mobile OSes provides three different types of

notifications; title bar (desktop notification), ring and vibration. Our application provides interface for managing these three different kind of notification for each tab. For instance, a user may set different notification mechanisms to different tabs; she sets no notification for spam SMS and sets high level (title bar + ring + vibration) to the SMS coming to user preferred tab or Inbox.

**6) AUTOMATIC FILTERING AND CROWD SOURCING:**

If a SMS does not get classified by sender blacklisting/ white listing or any of user generated rules then it is passed through a trained classifier based on Bayesian filtering to make a decision whether it is ham or spam. Bayesian filtering mechanism is already described in Section II-B. Since, Bayesian based filtering operates on keywords and spam SMS keywords keeps on changing according to current trends due to different festivals etc. The application uses crowd sourcing to keep itself updated with latest trends. It logs all the spam SMS of the mobile device, sends it to an aggregation server periodically which aggregates reported spam SMS from all the users using the system and generates a new training files every few days which can be synced by all the mobile devices for better spam filtering.

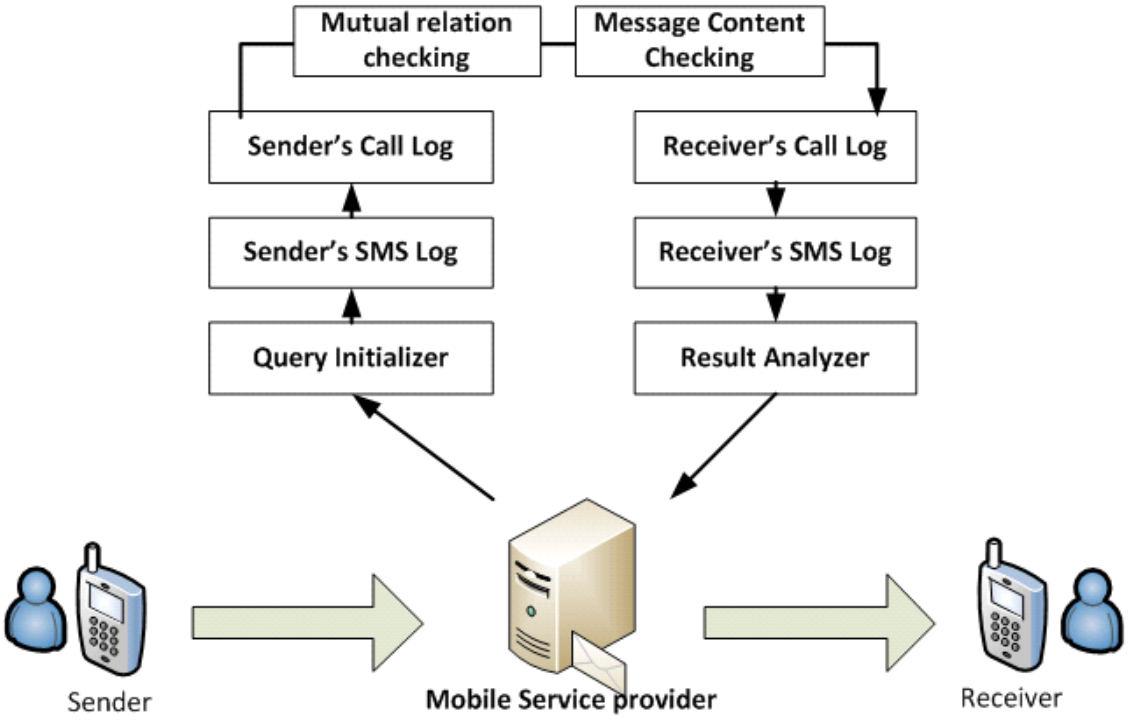



**Fig. System Architecture**

**CONCLUSION:**

Mobile phone spam is a form of spam directed at the text messaging or other communications services of mobile phones. By using SMS service provider level SMS spam detection system the system will first look up in SMS and call log data base and check a direct or the mutual relation between sender and receiver if system found no relation and if the message content are found spamming then it will treat message as a spam message and forward message with spam tag or directly reject it. By using this system, the problems occur due to spam messages like balance deduction, wastage of SMS memory etc. is get solved.

## REFERENCES:

1. F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida. Detecting spammers on twitter. CEAS 2010.
2. G. V. Cormack, J. M. G´omez Hidalgo, and E. P. S´anz. Spam filtering for short messages. In CIKM '07, pages 313–320. ACM, 2007.
3. Chao Gao and Jiming Liu, "Modeling and Restraining Mobile Virus Propagation", IEEE Transactions On Mobile Computing, Vol. 12, No. 3, March 2013.
4. Zi Chu, S. Gianvecchio, Haining Wang, and Sushil Jajodia, "Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg?", IEEE Transactions On Dependable And Secure Computing, Vol. 9, No. 6, November/December 2012.
5. K. Yadav, A. Malhotra, P. Kumaraguru, R. Khurana, and D. Singh. Take control over your smses: A real-world evaluation of a mobile-based spam sms filtering system. In Manuscript submitted in APCHI'12, 2012.
6. D.-N. Sohn, J.-T. Lee, and H.-C. Rim. The contribution of stylistic information to content-based mobile spam filtering. In ACL-IJCNLP Short papers, pages 321–324.
7. D.-N.Sohn,J.-T.Lee,K.-S.Han,andH.-C.Rim, "Content based mobile spam classification using stylistically motivated features," Pattern Recognition Letters, vol. 33, pp. 364–369, 2012.
8. S. J. Delany, M. Buckley, and D. Greene, "Sms spam filtering: methods and data," IEEE Expert Systems with Applications, in press.
9. K. Uysal, S. Gunal, S. Ergin, and E. SoraGunal, "Detection of sms spam messages on mobile phones," Proc.ofIEEE20thSignal Processing and Communications Applications Conference, 2012.