International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 11 – No. 11, March 2017 – www.ijais.org

A Mapping Study to Investigate Spam Detection on Social Networks

Balogun Abiodun Kamoru
Department of Software and Information Systems,
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia

Azmi Jaafar
Department of Software and Information Systems,
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia

Marzanah A. Binti Jabar
Department of Software and Information Systems,
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia

Masrah Azrifah Azmi Murad
Department of Software and Information Systems,
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia

ABSTRACT
Social networks such as Facebook, Twitter and SinaWeibo have become increasingly important for reaching millions of user globally. Consequently, spammers are increasing using such networks for propagating spam. Existing research on filtering techniques such as collaborative filters and behavioral analysis filters are able to significantly reduce spam. In recent years, online social networks have become the most important medium of communication among individual and organization to interact. Unfortunately, driven by the desire to communicate, fraudster or spammers have produced deceptive spam or unsolicited commercial email(UCE). The fraudsters' or spammer activities mislead potential users and victims reshaping their individual life and general communication on social network platform.

The aim of this study is to understand, classify and analyze existing research in spam detection on social networks, focusing on approaches and elements that are used to evaluate the general framework of spam detection and its architectural framework from the users perspective, service provider and security analyst 's point of view. This paper presents a systematic mapping study of several spam detection techniques and approaches on social networks that were proposed to measure to evaluate the general framework of spam detection on social networks. We found 17 proposals that could be applied to evaluate spam detection on social networks, while 14 proposals could be applied to evaluate the users, service providers and practitioners. Various elements of spam detection on social networks that were measured are reviewed and discussed. Only a few of the proposed spam detection on social networks are soundly defined. The quality assessment of the primary studies detected many limitations and suggested guidelines for possibilities for improving and increasing the acceptance of spam detection on social networks. However, it remains a challenge to characterize and evaluate a spam detection and framework on social networks quantitatively. For this fact, much effort must be made to achieve a better spam detection approach in the future that will be devoid of problem anomaly detection, fault detection, malware detection and intrusion detection

General Terms
Spam detection, Security, Mapping study,Spam detection metrics.

Keywords
Social Networks; Spam techniques; Spam Approaches;Spam Strategies.

1. INTRODUCTION
Spam detection on social networks mainly focuses on anomaly detection, fault detection, malware detection and intrusion detection. If a considerable effort is not made to find a technological solution to the menace of spam. The internet email and social email is in danger as an important medium of communication[1].

Social spam is low-quality information on social networks that is similar to email spam in that it is unsolicited bulk messages that users do not ask for or specifically subscribe to. Such spam, is a nuisance to people and hinders them from consuming information that is pertinent to them or that they are looking for[2]

Spam detection on social networking has been a major problem globally. The current state of spam is worsening and more rigorous effort are required to stop them in an effective manner. 75.9% of email messages are spam, while social networks are the most vulnerable attacks [3]. Presently , spammers are trying a new approach to gain access through facebook, Twiiter and Sina Weibo through numerous events on the social networks.

In the literature, most previous work on social spam has focused on spam prevention on a single social network e.g Facebook, Twitter and Sinaweibo[4][5][6].

Social spam is a relatively new research area and the literature is still sparse[7]. A large number of classifiers have been used in spam detection but choosing the right classifier and the most efficient combination of them is still problem. Previous work by [8] , proposes a Bayesian framework, which is theoretical efficient and practically reasonable method of combination, when investigating the integration of text and image classifiers.

[9], there are limited studies on spam detection. Problem of effective, efficiency and accuracy in spam detection on social networks and email generally, they try to provide survey and algorithms method to solve the problem pose by the threat. In 2015, it was estimated that approximately one seventh of English web pages were spam [10], one consultancy estimated that Russian Spammers earned roughly US$2-3 million per year.

With the recent survey, it shows that social spam is about 355% [11],there are many problem of spam detection and spam filtering are ineffective with lots of content and behavior feature. Millions of users and waste invaluable resources and have been burden to email system[12]. Twitter is still growing with 25 million active users while Facebook is about 130 million active users daily[72], while Sina Weibo has about 500 million users[32]. Annual report published by the crime complain Centre shows that there is high rate of spam on email and social media [28]

The work described in this paper not only extends and updates the previous reviews [2][9][18] provides goal of supporting and directing future research. Our review differs from previous that represent the literature in the spam detection on social networks with quality evaluations with respect to the following elements:

- **Different goal**. The main aim of this review is to understand, classify and analyze the existing spam detection on social networks for measuring the quality of spam detection on social networks and its architectural framework, to direct and support future research, while other reviews [9] [2][7], [74] and [33] aim mainly at provide an overview of quality measure and evaluations. Certainly a difference in goals leads to a different focus.
- **Different scope and review perspective.** Spam detection on social networks involve not only the defining the novel approaches uses, algorithm methods used, statistic method use or classification for quality attributes but also the extent to which they are empirically validated. In this paper, our review is focused on spam detection on social networks. The reviews in [3]. [74] [7] covers a wider scope for spammers on social networks and victims. [69] talk about spam filtering to address different web services on social networks. [16][22] proposed and implement text classification using wikipaedia based co-clustering classification algorithm.
- **Systematic mapping review and more comprehensive approach.** We based our review on a systematic mapping review, which led to the identification of 36 studies. The review in [9] is based on only 3 articles, and that in [7] is based on only 4 articles. The review in [2] and [74] is based on 9 articles.[18] it is difficult to determine how many primary studies contributed to their study. None of the previous reviews present a systematic mapping review. Compared to a traditional literature review , a systematic review has advantages: a well-defined methodology that reduces bias and wider context that allow general conclusion [74]
- **Classification of studies**. We classify the identified with respect to the scope and spam detection review [9], the study context [23]. studies done in [44], [49] and [15] have used dataset of fake reviews and future research on improve the accuracy of detection systems. [2] proposed a framework to help users to decide whether a review is spam. It gives 5 criteria for review: rating consistency, questions in review, all capital letters review, comparative sentences, link spamming. [35] claimed that their method performed well with a high level of accuracy (for some criteria, more than 98%). In terms of detecting intelligent spam reviews, which are very common in opinion sharing websites, many aspect were not considered in the study. The systematic mapping review method has allowed us to identify the relationship between the researchers and the practitioners, to assess the current state of spam detection o social networks in the context of spam detection system and to identify areas that need improvement by outlining the limitation of current research. We believe that the results that are obtained from this mapping study are important for the community of researchers who want to know the gaps in the literature and who want to understand topics that have been researched. This review will also be useful for practitioners as an indication of maturity in the selection of spam detection and to remain up-to-date with the state-of-the-art. In addition, new and enhanced spam detection framework can be proposed on the research that already been performed in this area of research.

This paper is organized as follows: Section 2 discusses Spam detection Concept and Framework. Section 3 describes the methodology. Section 4 provides in more details the results of our research questions. Section 5 discusses and analyzes the results. Section 6 concludes the paper and identifies future trends.

1. Spam detection concept and framework from the perspective of defining Spam detection on social networks

Several definition of spam detection are given [9], [33],[7],[74]; each of definitions states different characteristics for the framework of spam detection on spam detection.

The social-spam detection framework can be split into three main components. Figure 1 shows an overview of the system and we provide a brief explanation for each part here:

1) Mapping and Assembly: Mapping techniques are used to convert a social network specific object into a framework defined standard model for the object e.g Profile, model, message model or webpage model. If associated objects can be fetched based on this object, it is assembled here;

2) Pre-filtering: Fast-path techniques e.g blacklists, hashing, and similarity matching are used to check incoming objects against known spam objects;

3) Classification : supervised machine learning techniques are used to classify the incoming object and associated objects. [39] Proposed the use of Bayesian technique to combine the classification results into spam or non spam.

As we mention earlier, the perspective by which the spam detection framework can be analyzed and classified based on the previous literature reviews. With the rise of social networks as an important medium of communication, spammers have increasingly targeted social networks with spam [12], In most social networks, spammers can send spam to other users in a number of ways, such as messages, friend requests, wall posts, tweets, weibo tag and profiles. In most cases spammers can also include links to a website where the user will take another

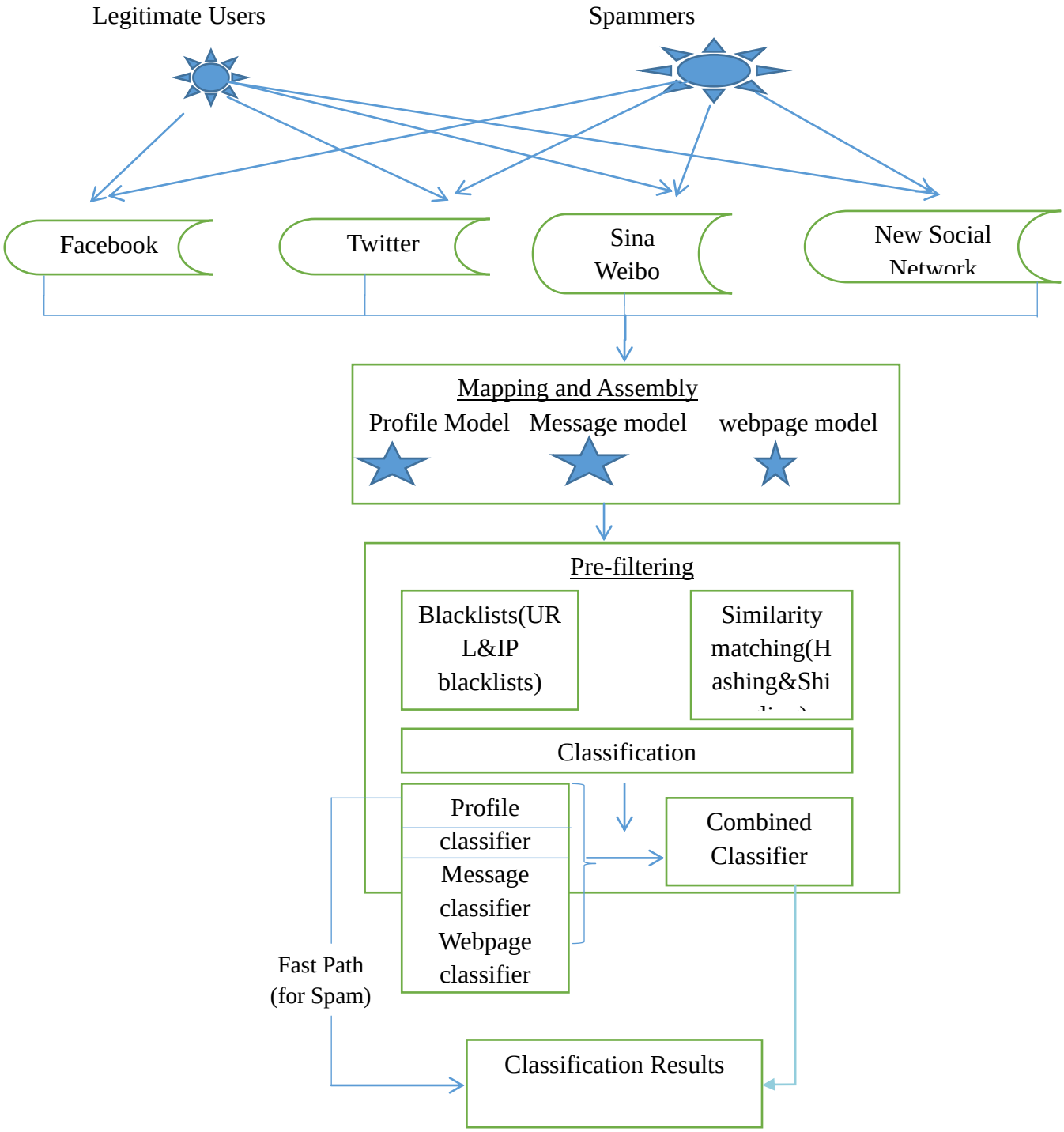

Figure 1: Architectural Overview of the spam detection framework (D.Irani, et al,2011)

Facebook, Twitter, Sina weibo, and other major social networks employ dozens of people fight on their network(Wang et al,2011). Most of these social networks use collaborative filtering (where users report objects that are spam and behavioral analysis (where logs of interactions are used to detect spamming patterns) to detect spam on their network. Such dynamic methods may be eventually able to detect social spam, but require a non-trivial amount of lag time to accumulate sufficient evidence.

Social networks will also employ classification based techniques which use labelled training data to find similar occurrence of spam on the social network. Due to the evolving nature of spam [8][17], these classification based technique need to be retrained and adapted to newer spam[68].

Although techniques to propagate spam may vary from one social network to another, due to specific of each social network, anecdotal evidence suggests that spam generally fall into the category of pharmaceutical, pornographic, phishing, stocks, and business promotion campaigns.

In this paper, we visualize spam detection concept and framework from the perspective of spammers and the victim of the spam in respect to service providers and stake holder on social network. Fig 1 provides a simplified architectural overview of the spam detection framework on social media platform. There is legitimate users and spammers which compose of Facebook, Twitter and Sina weibo and the new social networks. There are 3 component parts, mapping and assembly, pre-filtering and classification. In each component e.g mapping and assembly has profile model, message model and webpage model, pre-filtering e.g blacklists and similarity matching, classification e.g profile classifier, message classifier and web page classifier. Therefore, for a better understanding of the overview of the spam detection framework, we further refine the existing literature review on spam detection on social network. This strategy will provide a clear picture of the spam detection framework on social network.

2. SOCIAL-SPAM DETECTION FRAMEWORK

a) An overview of the framework is shown in figure 1 and we present three main parts in the following subsections.

a. Mapping and Assembly: to build a framework that is social network agnostic, we have to create a standard model for objects within the social network. We defined a model of an object as a schema containing the most common attributes of the object across social networks. Once a model is defined, we need to map incoming objects from social networking into objects of the model.

b. Models: our framework defines three models representing the most important objects in social networks, namely: profile model, message model and webpage model. We omit other models as they are not required to demonstrate the feasibility of the framework Wang et al(2011).

c. Classification: from previous scholars work, we identify they use classification and classifiers, one classifier for each model and use different types of supervised machine learning classifier, including standard algorithms such as naïve Bayes[19], support machine vector SVM [5] and LogitBoost [75]. After the classifier for each model involved return a decision, it is passed on to the combiner. There are four different combination strategies available for us to adapt in our framework: AND strategy, OR strategy, majority voting strategy and Bayesian strategy. AND strategy classifies an object as spam if all classifier, for each model, classifies it as spam. Bayesian strategy is a slightly modified version of a strategy from previous research on creating an anti-spam filter combination framework for text-and image email [8].

2.1.1. Data in spam detection

There are number of data and features pertaining to a review that can be used in techniques to detect if the review is spam. These data and features are categorized into three predominant types in [9].

i. Content of review: the text of a review is called the content of the review. The content of each review is the first thing to be considered in spam detection practice. Content of a review are significant in spam detection, the techniques based on them are not sufficiently comprehensive to detect all types of fake reviews.
ii. Meta-data of review: information about the review besides its actual content is called meta-data e.g,. the reviewer's identity, the geo-location of the reviewer's computer and its MAC and IP addresses. Through analyzing these types of data.
iii. Information about the product: information about a product is useful in spam detection such as, the product description. Furthermore, we can classify the data as public and site-private. Public data can be extracted from review websites. Private data refer to data that are not publicly available in the review websites.

2.1.2. Spammer detection techniques

Because the primary artifact in detecting a spam review is the review itself, several researchers have studied this problem by focusing on review, limited studies have been conducted in the area of detecting spammers.

A number of researchers assume that spammers usually allocate a specific time interval to post spam reviews, and uses this assumption to help detect spammers [44] ,[28]. spam attacks on social network are prominent on social platforms. Algorithm were mostly used in spam detection review: a Bayes change point detection algorithm to fit curves using time series, a template matching algorithm on the result of the previous algorithm to find burst patterns and sliding window to detect blocks in time series matched with a joint burst in all dimensions of the time series approaches.

2.1.3. Detection techniques for group spammers

Occasionally, spamming activities can be considered group spamming event; manufacturers may employ multiple spammers to do a job because of their ability to dominate all aspects, features and sentiments for a product or brand. A group of spammers could be formed [14]

2.1.4. Motivations of social spam

The first step toward analyzing and classifying spam detection on social network, the effective measure to detect and combat social spam is an understanding of the motivation behind. Based on our experience as well judging from past history of spam in other contexts, we argue that the most threatening motivation is financial gain. How can someone make money by abusing social network system? This question has not yet been thoroughly explored. The spammer probably make money when users visits Facebook, Twitter and Sina weibo, and therefore the spammer needs to attract the users to the site. Social spam is a cheap way to attract users. Others methods include email spam, search engine manipulation, and placing ads. The first is more expensive because there is

already an infrastructure in place against email spam: filters, black lists, and so on. Search manipulation is more expensive because search engines have a financial interest in preventing rank manipulation, and thus invest in spam detection algorithms. Finally, advertising has obvious monetary and disclosure costs. Social network are therefore a target of opportunity; an abuser can submit many spam annotations effectively, efficiently, cheaply and anonymously.

It is important at this point to briefly discuss the relationship between social spam and click fraud. Advertising networks and keyword tools are legitimate when used as intended. If a user tags with helpful keywords a legitimate site containing ads, this is not a case of spam. We consider social spam only those abusive uses of social network in which misleading tags are used, and fraudulent or malicious site is tagged or link.

2.1.5. Features of spam detection

The first issue to address spam detection on social network is class of objects should be seen as potential candidates for spam labelling. Spam can be injected to social networks at different levels. The traditional view is to classify pages or site as spam based on their content, that is, resources that users of the system perceive as non relevant or " Junk". The problem with this perspective is its subjectivity: what is spam to one person can be interesting to another. Secondly, we can focus on spam posts, i.e., on malicious associations between resources and tags. Finally, one can look at user accounts created with the goal of injecting foreign content into the system. Such accounts may or may not mix with legitimate content with spam, in order to mask spamming activity. Flagging users as spammers is the approach taken by some social networks regarding spam detection, such as BibSonomy. This approach is intuitive and easy from an administrator's point of view [33].

3. RESEARCH METHOD

3.1 Protocol development

This paper presents a systematic mapping study of spam detection framework and analysis based on guidelines that were proposed by [76], [28]. We started by reviewing of the existing systematic literature [10]; [12];[11];[71]; [69];[26]; we concentrated on developing a protocol for a systematic mapping study that has addressed questions that are related to the spam detection framework on 3 different social networks platform [77],[76], [2][8][9]. In the following sections, we will detail each process that we use.

3.2. Research questions and motivation

The following research questions have been addressed.

RQ 1. Are the performance measure for spam detection framework and Overview or are they based on individual or general social network platform or just 3 platform like Facebook ,Twitter and Sina weibo. It is important to ask, who is harmed by spammers who generate fake content? The advertiser gains, because real users click on ads and thus visit the advertiser's page, which is desired outcome [22]. [8] claimed that using anti filter combination framework for text and image email for incremental learning on social network for spam detection. To addressed this question, we investigated and classified existing research on spam detection to whether a performance measurement was performed on spam detection framework on social network; we describe the investigation in Section 5.2.

RQ 2. Which elements of spam detection are being measured? How were these elements defined and validated? How these elements have affect each social platform like Facebook, Twitter and Sina weibo. We will discuss this question in Section 5.3 based on architectural overview of spam detection with respect to the number of issues, as follows:

- The danger of an ambiguous definition of elements can make it difficult for the spam detection to get metrics data reliably and could lead to an incorrect interpretation of the metric values.
- Whether the elements being measured are visible to spam detection developers.
- Whether the spam detection definition and formulation are validated.
- Limitation that restrict the practical use of performance measurement on spam detection [33]

RQ 3. Are the limitations of the current research? The aim of this question is to identify any gaps in the current research, to suggest areas for future research. We will discuss this question in Section 5.4 with respect to the limitations that were identified by this mapping study.

3.3. Search process

To determine how many primary studies relate to these research questions, we conducted an automated search to collect papers on spam detection on social network. The results obtained are shown in fig.2.

In step 1, based on our experience and the terms used in [9]. [74]. [77], and [8], we identify the following search strings:

1. Measure OR metric OR quality OR evaluation OR attribute,
2. Spam detection AND Framework,
3. Spam mechanism elements and Overview,
4. Social network AND spam detection review.

To make the search comprehensive and precise, an expert librarian was consulted. All of the possible combinations of these identified search strings were tested in the following databases: ACM Digital Library, IEEE Explore, Springer Link, Scopus, Sciencedirect, Elsevier, Microsoft academic research and Google scholar. These databases were selected because they are accessible to our library.

In step 2, a quick review of the title resulted in 455 papers that looked relevant to spam detection in general (email and social network). Step 2 was planned to ensure that any important articles are not missed. For ease of access during review. In step3, a more detailed review of the title, keywords and abstract using the exclusion and inclusion criteria defined in Section 3.4 was performed. Basically, only studies about the evaluation and metrics of spam detection on social network were selected.

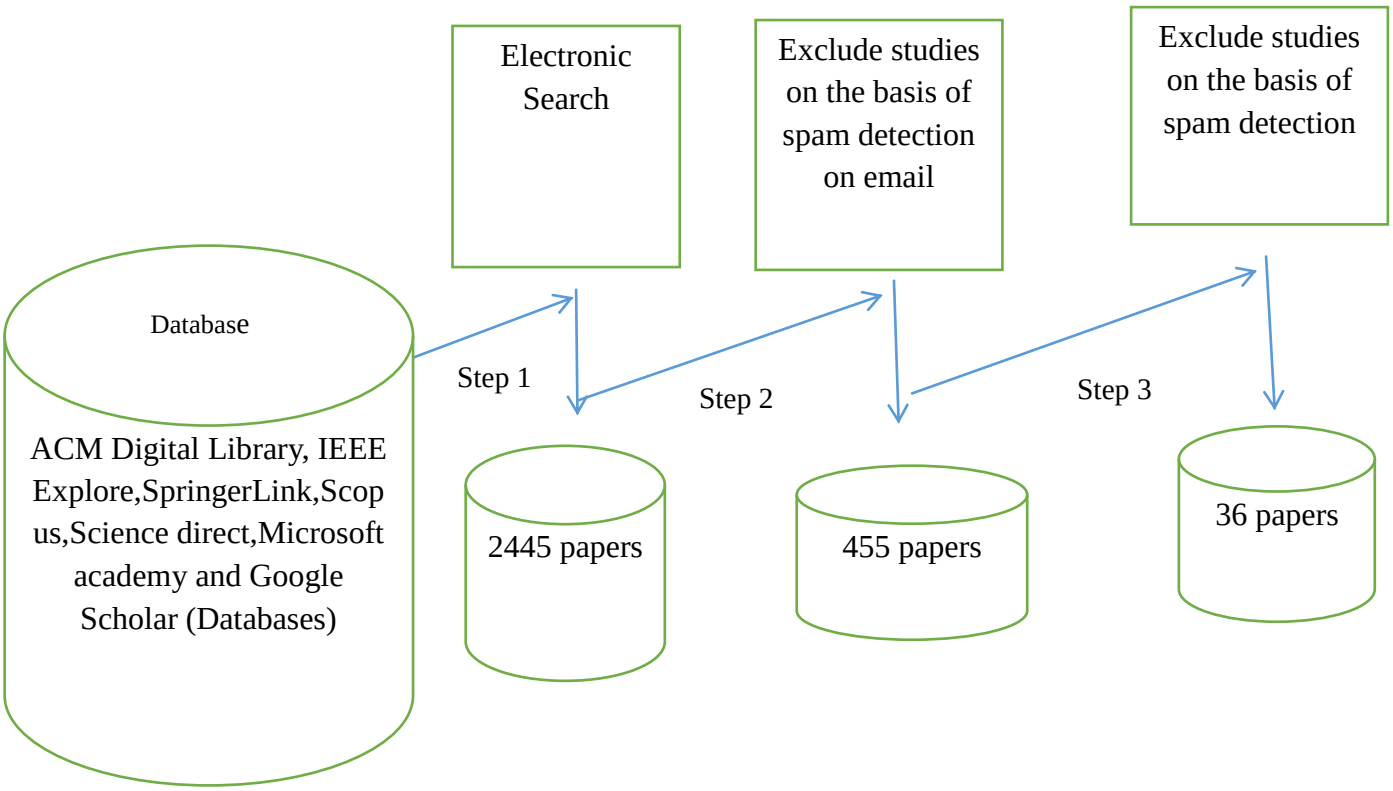

Fig.2. steps of the search strategy (Heydari.Atefeh, et al;2015)

Then , the reference lists containing the primary studies identified in the first step were searched manually. This step resulted in a list of 36 papers. A total of 31 of the 36 studies were primary studies, while five were secondary studies. Other researchers [23][67][50] and [44]used similar search approach.

3.4. Inclusion and exclusion criteria

With the respect to the research questions that are addressed in this paper, we excluded the following:

(a) In step 2. Irrelevant studies or papers that lie outside the field of spam detection on email.

(b) In step 3:

- The studies that are related to email and other platforms on spam detection.
- The studies on process spam detection.
- Duplicate publications of the same study in different journal, articles or publishers. This step is necessary because SCOPUS indexes IEEE,ACM and Elsevier publications.
- Implementation performance measure on spam detection one mail. dendritic cell algorithms for mobile phone spam filtering [39] [21]and spamming the internet of things: A possibility and probable solution [45]. As such, it cannot be claimed to be spam detection on social network.

In contrast, papers on the following topics were included:

- Both Spam detection and social network were included.
- Specifically, we focus on the spam detection reviews on social network specifically on Facebook, twitter and sina weibo that were proposed to evaluate the internal and external quality elements.
- Papers published before 2002 till date.

3.5. Quality Assessment Questions (QAQ) of primary studies

It is not essential to include an assessment of quality in mapping studies, as discussed in Heydari et al.(2015). However, in this study, the goal of the quality evaluation is to assess whether the proposed spam detection are meaningful, and the findings that were presented well would be of use practitioners. While the research questions (RQ) aim to characterize each spam detection on social network according to the basic principles of spam detection and the representation of performance measure of spam review generally, the QAQs are an attempt to provide a brief overview of the proposal and to measure the quality of the reporting of a study's concept, aims, context, data collection and analysis. Taken together, these QAQ could represent the concerns of the researchers and practitioners of the spam detection. Therefore, the importance of such QAQs is not only to improve the quality of on-going studies but also to encourage researchers to assess their proposal before submitting it for publication. To address our goal, we used the following questions:

QAQ 1. Did the authors justify the need for spam detection or state what the problem the spam detection on social network are intended to solve and provide a clear statement of the aims of the proposal?

QAQ 2. Did the authors appropriately present a research design to address the aims of spam detection with respect to the underlying framework for the elements in social network.

QAQ 3. Did the authors provide a specific hypothesis to be tested, state it clearly prior to defining the spam detection and discuss the theory from which it is derived? Without an underlying theory and a shallow hypothesis, we cannot understand the spam detection. Consequently, we use inconsistent approaches and obtain inconsistent results. A good example of a defining hypothesis from a theory can be found in [28].

QAQ 4. Did the authors provide a clear unambiguous definition of spam detection and explain how performance measure could be use to validate social network? The clear definition of the element, attributes and the metrics are more consistent and very important [8].

QAQ 5. Did the authors clearly identify who the performance measure on spam detection is? It is pertinent important to identify the spammer and the victims on social networks. The victim are the users of the Facebook, Twitter and Sina weibo. While the service provider has it own share of the exploit too.

QAQ 6. Did the authors specify the context in which spam detection would be used on social network like Facebook, Twitter and Sina weibo? For example did they specify the point in which the spam detection are more dangerous to the users, it is difficult to understand and apply performance measure if the context of a study is not fully defined. For example, for the context of social spam detection, see [33].

QAQ 7. Did the authors explain how the spam detection on social network could be gathered? For example, did they explain the appropriate data collection tool and how the spam detection values could be interpreted [77], provided a good discussion of many problems with data collection. Without a clear data collection template.

QAQ 8. Did the authors identify any pre-conditions that must be met, of constraints/limitations that are related to spam detection on social network or how the validity is assured ? For example are they appropriate only to identify the limitations on social networks and email while examining the classification and categorization of the spam detection. [2]; [29], and researcher may want to replicate the studies in different contexts.

Table 1: The answers scored criteria

The answer	ordinal scale of the answers
The answers are explicitly written in the primary study	Yes
The answers can be mostly inferred from the primary study	Mostly
The answers can be somewhat inferred from the primary study	Somewhat
The answers are undetectable in the primary study or unknown	No

Table 2: Summary of primary studies: an overview of the approaches/methodology that were followed to develop the spam detection metrics and framework.

Approach	Meaning of the approach
Best-effort approach	paper that attempts to measure and identify malicious Spam activities on Facebook, Twitter and Sina weibo which is complex in nature
Weka classifiers approach	paper that attempts to measure the various algorithm Model that are elements of spam detection to Precision, accuracy based on the fraction of users and spammers activities on social network
Spam detection techniques approach	paper that attempts the techniques for detections of Spam on social network that is based on specific Model and quality.
Statistical and unstatistical approach	paper that attempts to analyze discriminative Properties, identified features like algorithm Classification, limited to spam profile.
Text mining and corpus analysis approach	paper that attempts to discuss the evaluation with Empirical analysis of spam detection on social Network
Others	paper that attempts to discuss the requirement for spam detection metrics and framework. No actual Performance measurement that are proposed. Instead the paper measured through framework

The questions from QAQ1 to QAQ 8 were answered on an ordinal scale, as shown in Table 1.

3.6. Data extraction

The candidate studies were collected, and all of the data that is related to the research questions and the broader aims of this study were extracted. The information that was extracted from each primary study included the following:

1. Whether the proposal applies to users and spammer of spam detection on social network.
2. Spam detection context: to state the goal of the paper, the spam detection elements that are measured and how the authors defined the architectural framework.
3. Spam detection framework on social network.
4. Approach to the evaluation of spam detection on social network (see Table 3 and 4).
5. Spam detection keywords and acronyms.
6. Spam detection descriptions: which state and what the measurement approach is used, and how it is operationalized.
7. Spam detection assumptions and interpretation guidelines: to explain how the spam detection values could be interpreted, to curb the activities of the spam on social networks using spam detection techniques.

8. Target spam detection on social network and the proposed performance measure.

3.7. Data analysis

Based on [9] and [7],74], [18], we first classified the identified papers into evaluation papers and spam detection evaluation papers according to the information that was extracted from performance measure on spam detection definition context on social network.

We further extended the classification with respect to the users and the spammers of the social network using spam detection techniques to evaluate the performance measurement. It can be applied to measure the victim and the users of social medium to be aware of spam activities. We base our classification on the author's view and paper objectives.

The data extracted in section 3.6 are analyzed with respect to RQs and QAQs as stated in fig.3., the answer of the RQ1 is extracted from in " whether the proposal is apply to spam detection detail on social network".

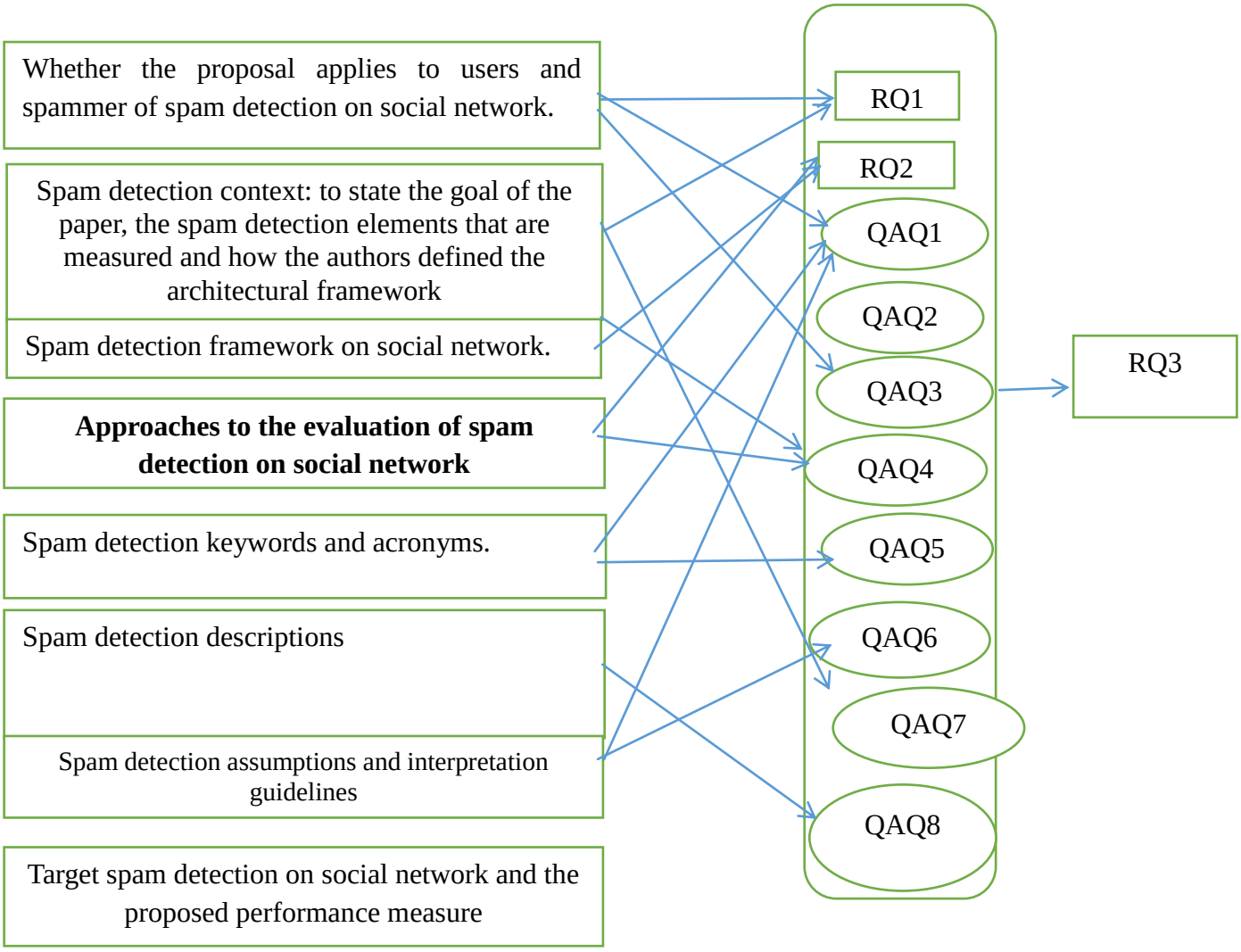

Fig.3. Mapping between research questions and quality assessment questions with data extraction

4. RESULTS

The results against each research question are presented in the subsequent sections.

4.1. Primary study background

The summary data were generated by categorizing the research studies, as shown in Table 2. Of the 36 papers that were identified to be research studies, 31 were primary studies, while five were secondary studies.

- 18% of the studies assume that an evaluation and quality of spam detection on social, they may be additional assumptions that are related to spam classification and text categorization.
- Another 18% of the studies assume that the spam detection can be enhanced by using algorithm model and un algorithm model, to identify the problem of spam detection on social network using learning approach mechanism [18], it has discussed some unresolved problem in spam detection techniques.
- 36% of the research papers were specification based spam detection on social network. These papers have

assumed the view of architectural framework of spam detection on three social platform(see fig.1).

- The rest of the papers have adopted a more sophisticated approach for introducing the proposed spam detection on social network. They argued that, previous researchers have exclusively work on spam detection email. We need to ensure that spam detection on social network were discuss and evaluate.

Table 3: Approaches to the evaluation of spam detection on social network.

Primary studies	context	performance measurement	level of validation
Gao, Hu&Wilson (2010)	They defined, informally, detecting characterizing social spam	URL-Obfuscation, redirection analysis	small experiment
Lam & Yeung (2007)	They informally identified the need for spam detection using Learning approach to detect spam On social network.	Gaussian similarity Score Scaling Mitigation scheme	Industrial Experiment
Markines et al. (2009)	Social bookmarking sites for spam detection using various machine Learning for algorithm classification	Validlinks, dataset	Industrial Experiment
A.Heydari et al. (2015)	Clues to detect spam review, to analyze the impact or feature used Extent literature to identify the most Effective tier of features	measurement rate, level of accuracy, review spam, spammer spammer group	Small Experiment
Goh& Singh(2015)	Random forest has proven to be Powerful classifier than data mining tool	Rate of accuracy, Measurement	Independently Validated by Goh&Singh
Ahmed& Abulaish (2013)	Analyze discriminative properties identified features like algorithms classification	Measures, features	Anecdotal
Alhassan&alfy (2015)	To analyze and evaluate several features sets, which can be extracted from social Network spam detection using mobile phone	Empirical, measures	Industrial Experiments
Zheng et al.(2015)	spam detection and spam filtering Are ineffective with lots of content.	empirical measure	Industrial Experiment
Miller et al.(2014)	To develop an anomaly detection system For identifying spammers	scale of measure, Nominal, ratio	independent validated
Schmid et al.(2015)	To propose data mining method To address email attribution On social network	Real data set	Industrial Experiments
Heyman et al.(2007)	Gap analysis approach to combat Spam on social network	measurement	Anecdotal
Gomes et al.(2005)	selection, which is calculated based On the quality attributes of spam Detection on social network.	Graph, measurement	Anecdotal

Table 4: Approaches to the evaluation of spam detection on social network

Primary studies	Context	Performance measurement	Level of Validation
Puttaswamy et al (2009)	Static and dynamic aspects for the privacy users of social network against Intersection attacks	Empirical, measurement	Industrial experiments
Schneider et al (2009)	Architecture complexity of Spam detection and understanding of social network from Network perspective.	Real data set	Small experiments
Singh et al (2009)	Structural complexity in privacy control on social network	Empirical Evidence	Anecdotal
Swamynathan et al (2008)	A function point like approach is named; through spam detection on e-commerce And social network	Data set	Industrial experiments
Webb et al (2008)	Complexity of spam detection on social network through honeypots	Real data set	Ancedotal
Xie et al(2008)	Analysis of dependency and complexity	Measurement	Anecdotal
Yardi et al (2010)	A link-list base technique to detect Spam on social network(twitter)	Real data	Industrial experiments
Zhou et al (2003)	based on concept of peer-to-peer systems, a set of metrics to measure Spam detection on social network	Empirical measurement	Industrial experiments
Cattuto et al (2007)	Analysis on bookmarking systems on social network for spam detection	Dataset	Small experiments
Witten et al (2005)	Coupling and mining of data on social network on spam detection	Real data set	Industrial experiments
Xu et al (2006)	A collaborative tagging and spam Detection on social network	Virtual data	Anecdotal
Lambiotte et al (2005)	Interaction complexity of spam detection on social network and collaborative Of junk email	Real data set	Industrial experiments
Markines et al (2009)	Evaluating the measures for emergent of spam activities on social network	Ordinal, ratio	Anecdotal
Bhaskar et al (2008)	A comparison of image spam using duplicate detection on social network	Ordinal. Empirical	Anecdotal
Wang et al (2007)	Based on concept of Image with near duplication detection on social network	Empirical	Industrial experiments
Androutsopoulous Et al (2000)	Analysis of using Bayesian to to filter spam on social network	Real data set	Industrial experiments
Gyongyi et al (2004)	Complexity of Trustrank in spam detection on web spam and social network	Empirical	Small experiments
Le et al (2004)	Analysis and evaluation of spam filtering	Empirical	Small experiments

Yan et al (2008)	Acosutic definition of spam detection	Real data set	Anecdotal
Wang et al(2006)	evaluation and measurement of spam Detection on Chinese spam (weibo)	Empirical	Small experiments
Sun et al (2005)	A novel classification of spam detection	Data set	Anecdotal

4.2. Quality assessment of primary studies

We assessed the primary studies for quality using QAQs that were addresses in Section 3.5. the quality assessment for each primary study is shown in Appendix A. The assessment was extracted in three steps. First , the first author selected the candidate studies and extracted all of the answers that were related to the quality assessment questions. We then randomly allocated 11 papers to each author of this study to assess independently. Second, all of the answers collected from each primary study were scrutinized and check properly by the author.

Fig.4. present an overview of the quality levels for each of the QAQs that are described in the previous section. This step is an attempt to measure how strong a case the original authors made when presenting their proposed spam detection. Our point is that it is possible to define and validate spam detection without clearly stating the addressed QAQs. In this chart, from left to right, we present each QAQ; from the front to the back, we present each of the analyzed rating scales; on the vertical axis, we have the quality level of each question. The overall low level of quality throughout the several ratings presented in our QAQs suggests that the spam detection described in these papers have a number of limitations. The most interesting part is that QAQ1 identifies 64% and 25% of the primary studies, giving a total of 89% (30 papers) scored "Yes" and "Mostly", respectively. These scores suggest that there is strong justification for the need for spam detection evaluation approaches. In contrast, the most disappointing aspect of the primary studies is the methodological weakness in their research process, which occurred in QAQ2 when 58% and 10% of the primary studies scored" somewhat" and "No" respectively. This result may have occurred because most of the primary studies are conference papers and would have had limitations on the number of pages.

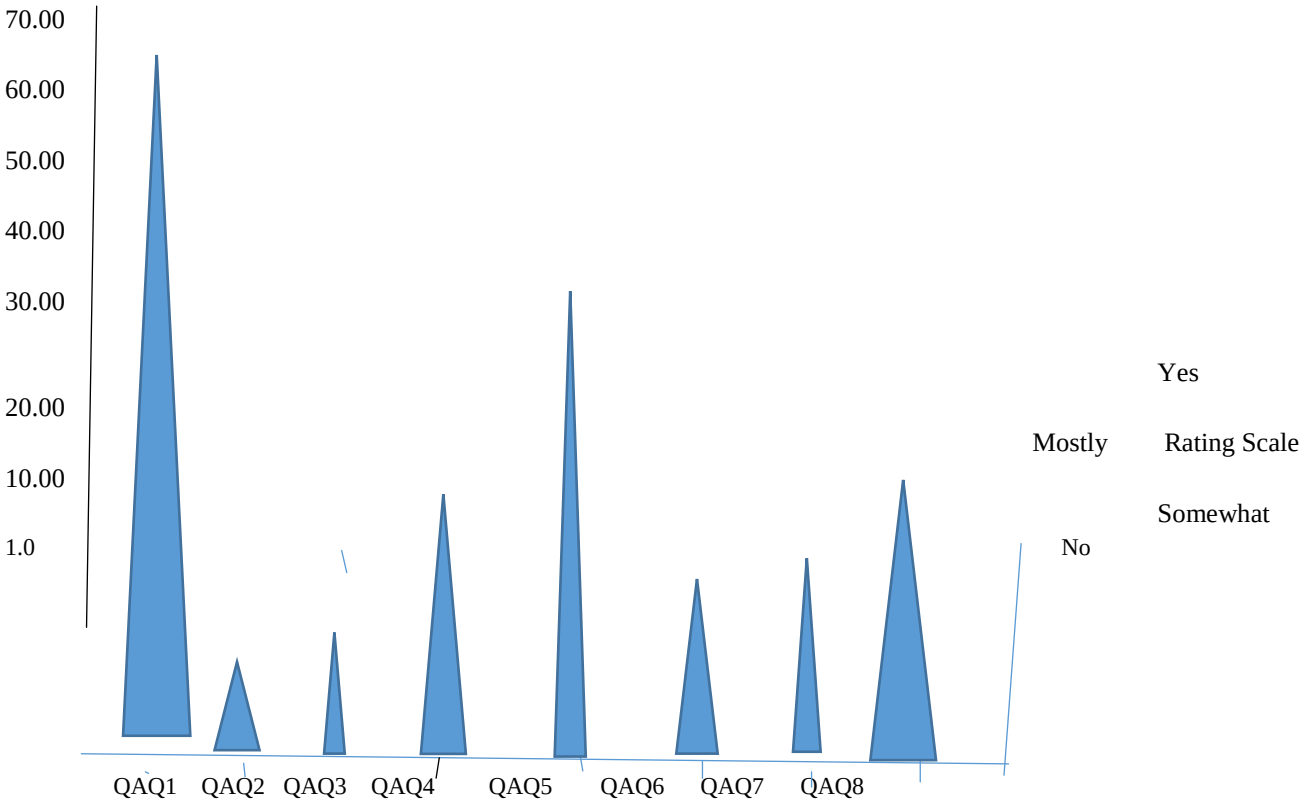

Fig.4. Overall quality assessment

4.3. Spam detection evaluation versus social network performance measure(RQ1)

In this sub-section, we provide detailed information on the set of primary studies that are included in this paper, as shown in Tables 3 and 4. For each paper, we identified the spam detection context, the framework of spam detection on social network, the validation of spam detection.

The spam detection context column summarizes the aim of paper with respect to the spam detection attributes being measured and how the authors treated the spam detection framework. We present a definition of the spam detection framework that is adopted in each paper, to avoid any confusion that may arise in their absence.

It is interesting to note that 28% of the primary studies explicitly adopted Markines's et al definition, while 33% implicitly adopted it by treating spam detection on social networks as main target that's why we ignored email spam detection.

The column (labeled "Spam detection framework") states the architectural overview of spam detection on social network, it will show exclusive details on how spam detection work on social network. The last column (labeled " level of validation") represents the extent to which the proposed spam detection have been validated. The level of validation is classified according to criteria presented in [74] as follow:

1. Anecdotal: example is provided to motivate the usefulness and applicability of the proposed spam detection framework.
2. Small experimental : an experimental is conducted to assess the proposed spam detection performance measure, but the sample of data does not allow generalization conclusion.
3. Industrial experimental: an experimental with a significant sample of real-world application is conducted.
4. Independently validated: an experiment made by third-party team confirm the conclusion made by the original authors.

To collect above information, we conducted a citation analysis using Google Scholar and Scopus to look up papers that would complement the validation of performance measure presented in an earlier paper. The results of the citation analysis were summarized in the column (labelled the "level of validation"). The fields that are marked with a dashed line are fields for which we did not found any one of the above criteria in the proposed spam detection on performance measure. It should be noted that several of the proposed spam detection metrics are from research still in progress.

The studies presented in Table 3 are mostly targeted evaluation of spam detection on social network. The obvious viewpoint for this situation would be that a victim of spam on social network and the spammer. Whereas the studies presented in Table 4 are mostly targeted at the evaluation of the spam detection metrics on social network.

In tables 5-7, for each spam detection performance measure, we identified a reference of the primary study to facilitate the discussion of the spam detection on social network. In the second column,

Table 5 Example of Spam detection on social network with reference to description and assumption

Reference	Metric name	Description	Assumption and guidelines
Anderson et al (2007)	Total number of hosting internet	This metric counts total number of internet hosting	Spam detection on social network need extra effort To be corrected
Gao et al (2010)	Average Number of detecting And characterize social spam Campaign	This metric is estimated by dividing the total number of detecting and characterize spam	It indicates that it need correction to errors.
Markines et al (2009)	Component of measurement and its density of metric	This metric is estimated by measuring its density	Spam detection need need to be detected.
Mika et al (2005)	Unified model and semantic	The ratio of actual number of unified model to the Available spam detection	A unified model and semantic need to be functional
Bergholz et al (2008)	Improved phishing detection using model based	The metric involve used model features	A comment that has phishing activities
Bilge et al (2009)	Automated identity	The metric that automated the attacks on social network	Automated functional
Bonneau et al (2009)	Link Critically metric	social graph on public listing	Critically required
Brown et al	Spam critique metric	actual metrics on spam	Critical metrics

(2008)			
Gross et al (2005)	Privacy issue on spam detection specifically for Facebook	Information retrieval	Mining techniques
Jones & Soltren (2005)	Threats to privacy	spam detection on social network	Spam detection
Benevenuto et al (2010)	Detecting spammer on Twitter, Measurement	Analysis of spammer and Victims.	Spam techniques
Cao et al (2012)	Aiding detection of fake account On social network	Comparison of fake and real account social network	Critical comparison

we presented the spam detection followed by brief description of the spam detection on social network. We also stated the values of the performance measure on spam detection (Markines et al,2009).

We categorize each spam detection metric according to classification criteria mentioned in Section 3.7, as follows:

- Example of spam detection performance measure that can be collected at the social network and email.
- Example of spam detection performance measure on social network.
- Example of spam detection performance measure on Facebook ,Twitter and Sina weibo.

Base on the table 7. Performance bottlenecks and the ability to get it resolved.

Table 6 Example of Spam detection on social network with reference to description and assumption

Ref.no	Metric name	Description	Assumption and guidelines
Cao et al (2012)	Social graph based	Sybil defenses with user Negative feedback	A component of graph Sybil defenses
Metsis et al (2006)	Spam filtering	filtering spam with Bayes	Naïve Bayes Classification
I.Rish (2001)	Empirical, Classifier	Empirical study of Bayes	Chosen empirical method
A.Seewald (2007)	Evaluation with data Analysis on content base	Evaluation and measurement of spam detection classification	Intelligence data analysis
Amitay et al (2003)	Structural evaluation of detecting spam activities	structural pattern	Detection techniques review

4.4. Which elements of spam detection are being measured? How were these elements defined?

The results of this research question were summarized in terms of measurement and detecting the social spam Gao et al. 2010[74] and the framework presented in D.Wang, et al.2011[77] as shown in Table 8. In the first column of Table 8, we answer the first part of question. Based on the analysis of the spam detection performance measure that are presented in Table 5-7 (Section 4.3), we then answer the second part of the research question and summarize it based on the elements that a spam detection attempts to measure Gao et al.,2010; D.wang et al., 2011[74][77].

5. DISCUSSION

In this section, we discuss the implications of the quality assessment of the primary studies and the results with respect to our research questions. Overall, the results in this paper are mostly similar to the Gao and Wang report [74], [77]. However, the study unit here is based on primary studies and spam detection performance measure, whereas the study by Gao and Wang is a primary study only.

5.1. Quality assessment of primary studies

Because the total possible quality score is 100% for each QAQ (i.e., the answer of each QA for the 36 papers is "Yes"), we have clearly identified a number of common problems with spam detection metrics that help to explain the current state of affairs. However, most of these problems are not specific to spam detection performance measure or metrics only. Indeed, it is common in much of information system research De Wang,et al,2011[77]

We think that the greatest deficiency in these primary studies is the absence of any serious consideration of QAQ3 and

QAQ8. These problems reduce the soundness of their conclusions. Perhaps the most serious problem is the QA3. Without underlying theory and a shallow hypothesis, we cannot understand the spam detection metric [22].

Most of the papers fail to grant the required quality score for QAQ7, which occurs when 29% and 29% of the primary studies have scored "No" and "somewhat", respectively. These quality scores suggest that approximately 18 papers might fail to discuss how to collect spam detection data and how the performance measurement values could be interpreted to guide practitioners to the needed information. This result is consistent with the view of [74][18]. The spam detection metrics discussed in these primary studies are, therefore, more likely to be unreliable than the performance measure that are discussed in other primary studies, according to the detection of review spam by [9] have clearly discussed the mapping and literature review survey from the specification of email and social network.

5.2. Are the performance measure for spam detection framework and Overview or are they based on individual or general social network platform? (RQ1)

Some authors [71]; [50]; Krause et al,2008 claimed that the best spam detection framework and analysis on social network are required functionality, under different composition contexts. The claim in those references is broad, in the sense that the best context of spam detection may not be the best candidate for composition in all composition scenarios. The overall idea is that knowing how good to detect the spam on social network. We mapped the primary studies that are based on the conference papers and journal proceeding, groupings is not unlike the set of spam detection performances measure in [29]. In this case, we need information from the spam detection framework..

Accordingly, a total of 17 studies out of 31 were proposed to measure the spam detection framework. This result consistent with the view of [77] and [74], in that we are more interested in the context of overall spam detection on social network rather than the context of the single platform of the social network but of larger platform of social network like Facebook, Twitter and Sina weibo.

Any modification to a past work on spam detection or an existing research scholars have been identified and highlighted, on the other hand, a total of 14 studies were proposed to evaluate spam detection on social network. The overall idea is that measuring the performance of spam detection would facilitate the concept, framework and general overview of spam detection on social platform especially Facebook, twitter and Sina weibo. We carry out exclusive investigation based on the previous and existing research on spam detection on social network [9], [60],[18].

Table 7 Example of Spam detection on social network with reference to description and assumption

Ref.no.	Metric name	Description	Assumption and guidelines
Jindal et al (2007)	Measurement of spam Review	Analyzing and detecting spam review analysis	identify the spam review
Mobasher et al (2006)	Spam attacks measure	identify the spam attack	collaborative filtering
Ntoulas et al (2006)	Content analysis	spam detection on web page and social network	Spam techniques

5.3. Which elements of spam detection are being measured? How were these elements defined and validated?(RQ2)

How these elements have affect each social platform like Facebook, Twitter and Sina weibo. As we discussed in Section 4.4, we may want to assess the spam on social network, how to detect the spam on the social network? And counting number of methods. For example, should we count property methods or should we count event methods? Without a clear definition of a performance measure, its application is likely to lead to different results. Moreover, we understand that there is an importance of the theoretical and empirical validation. Finally, without classification, text categorization and algorithm model level, any interpretation of the measurement is difficult. This scenario means that we cannot have more larger or smaller performance measure on social network using spam detection [59]

In the same way,, with respect to the discussion above, overall the definitions of existing spam detection metrics or performance measure and elements are ambiguous and unclear for most if not for all of the metrics. However, in fact, measures are not only instruments with strength but also have limits and constraints. In this context, the characterization and evaluation of spam detection on social network and in general and spam detection are not the easiest job. Consequently, before a measurement can be developed, a clear specification of what is being measured, why it is to be measured and how the metric value could be measured must be formulated, to provide real information from spam detection metrics rather than only numbers.

5.4. Are there limitations on the current research?

Although the set of spam detection performance measure presented in this paper are indeed useful for the characterization of the spam detection, the above analysis clearly provided a judgement regarding the following"

- The lack of a widely accepted performance measure of spam detection and quality attributes of the spam detection on social network. This lack may arise because most performance measure definitions were performed in an ad-hoc fashion, rather than meeting information requirements of spam detection framework upon which we classified and interpret

spam detection on social network especially Facebook, twitter and Sina weibo, the data collection and interpretation of spam detection metrics becomes subjective. In addition, most of these proposal have not achieved an industrial level validation.

- The poor quality of some papers identified in the quality evaluation section, which reduces the trustworthiness of the proposed spam detection metrics.
- The poor quality of some spam detection metrics definitions, which make it difficult for researchers or practitioners to ensure the correct collection of measurements that were initially intended by the spam detection experts. Overall, many spam detection performance measure have insufficiency either in their formulation, collection, validation or applications.
- The elements of spam detection metric definitions those are not visible to spam detection experts, including elements that are incompatible with the standard concepts of spam detection on social network especially Facebook, Twitter and Sina weibo..
- Elements that do not exhibit the attributes that the researchers claim to have been measured.
- The limitations or constraints that are required to map the specifications of target spam detection framework according to the definition of spam detection metrics.
- Insufficient information, such as non-stated hypotheses and inadequate context provided by the original studies, which may cause subjectivity in their replication or interpretation.
- As far as we know and understand, most of the proposals were just proposed theoretically, without any correlation with any external quality attribute. A few proposals were only tested by their authors, limiting knowledge sharing. For example, it is worth noticing that only one independent validation was performed by [74] [9]. This is mainly due to difficulties in experimental replication. A third party validation of spam detection metric is a fundamental and very much desirable for their proof of usefulness before common acceptance is sought. Thus, there is insufficient experimental validation.
- The poor experimental validation leads to lack of established performance measure threshold values, which obscure their value to practitioners.
- Most of the existing proposals to evaluate a general spam detection on social network and may not be single social network.

5.5. Limitation of the study

Our aim was to cover papers that were published between 2000 and 2015. However, regarding the search process, we may have overlooked certain papers because of the accessibility of their publisher sites and the limitations of our library. The limitations of this study are primary study selection bias, inaccuracy in data extraction, misclassification and quality assessment bias. To avoid a selection bias, a multistage process was used in the search strategy that involved many searches and three steps for the inclusion and exclusion criteria. To minimize the chance misclassification of the spam detection metrics and misinterpretation of the terms, the data collected from each primary study (Section 3.6) were also checked by the other authors independently. The procedure of having one extractor is not consistent with the standards [9]. With respect to the quality assessment criteria, there is a possibility that the extraction process may have caused some bias in the results. The other authors were to independently check the assessments.

6. CONCLUSIONS AND FUTURE WORK

To provide an overview of spam detection on social network and identify the right performance measure of the spam detection and spam detection framework, we have presented a systematic mapping study of existing spam detection on social network. We contribute to filling the gap on current approaches to spam detection metric and framework and in general concept. From victim perspective, spammers and to service provider of the social network especially Facebook, Twitter and Sina weibo, it is essential to spam detection review on social network, but is it beyond the scope of the review performed in this paper.

We think that the benefits of a spam detection metrics and framework cannot be achieved without performance measure for effectively evaluating spam detection. We found 20 proposals be applied to evaluate spam detection, while 19 proposals could be applied to evaluate the victim/users of the social network , spammers and the service provider. This task resulted in plethora of performance measure or metrics for spam detection framework and metrics; however, most of them may not be of much relevance to the spam detection on social network. We also investigated and presented various elements that are measured in this field. Even worse, inconsistency in component definitions can frequently be found among the many studies by measurement researchers. Two main factors could help to solve this conflict: first, we need agreement on which element of a component is to be measured. Our work provides a clear discussion in this respect, and it can serve as a starting point for further discussions (see Section 2). Second, we need to define, without any ambiguities, the elements of spam detection architectural overview that are to be measured. For example, what exactly is an element, framework or method?

We also contribute a good framework for systematic review comparison and quality assessment of spam detection metrics proposals by independent research teams. This framework can be further refined and adopted, to provide more details concerning the spam detection metrics definition that could mitigate many of the identified problems.

We do not claim that our review resolves all of the limitations and is agreed on by all parties, but rather that serves as a basis for further discussion from where spam detection measurement community can start paving the way to future agreements. From an academic point of view, we believe that this study can act as the starting point further primary studies as well as for more detailed secondary studies, which could lead to an empirically based body of knowledge. For practitioners, the results of the studies can be used as an indication of maturity for the current research.

We believe that several questions are raised by this investigation, and areas for future research presented. An interesting area for further research involves revising the existing spam detection performance measurement and architectural over view for better precision in measurement. Another interesting area is to develop a more sophisticated approach, such as combining more than one spam detection metrics based on logical conditions by which a subset of problems is detected, to characterize and evaluate spam detection with real information. To obtain an overview, see [9][39], [77]. We also note that there are no automated support tools that facilitate the collection and calculation of spam detection metrics or performance measurement. Last , but not least, the majority of the spam detection metrics discussed here were either insufficiently validated or not validated at all in their original proposal. Due to space constraints, we have left this concern for future work.

7. ACKNOWLEDGEMENTS

We Thank Associate Professor Dr.Marzanah A. binti Jabar for her ideas, comments, suggestions, and support as we prepared this paper. This work will not be possible without the help of Associate Professor Dr.Azmi Jaafar and Associate Professor Dr Masrah Azrifah Azmi Murad for her support during the cause of writing the paper. We also thank all those researchers whose works are referenced. Finally, we wish to thank the anonymous reviewers for their valuable comments.

Appendix A . Quality assessment of the primary studies

Primary studies	QAQ1	QAQ2	QAQ3	QAQ4	QAQ5	QAQ6	QAQ7	QAQ8
Gao et al (2010)	Somewhat	No	No	Somewhat	Somewhat	Somewhat	No	Somewhat
Heydari et al(2015)	Yes	Yes	Yes	Mostly	Mostly	Yes	Yes	Mostly
Wang et al (2011)	Somewhat	No	No	No	Somewhat	Mostly	No	No
Markines et al(2009)	Mostly		Somewhat	Mostly	Yes	Yes	Mostly	
Gao et al (2010)	Mostly	Somewhat	Somewhat	Somewhat	Somewhat	No	Somewhat	
Lam&Yeung (2007)	Yes	Mostly	No	Mostly	Yes	Mostly	Mostly	Mostly
Mika (2005)	Mostly	Somewhat	No	Somewhat	Mostly	Somewhat	Somewhat	Somewhat
Krause et al (2008)	Yes	Mostly	No	Somewhat	Yes	Yes	Mostly	Mostly
Cobb (2006)	Yes	No	Yes	No	Mostly	Somewhat	Yes	Somewhat
Lynam et al (2008)	Mostly	Somewhat	Yes	Yes	Somewhat	Mostly	No	Somewhat
Felt et al (2011)	Yes	Mostly	No	Somewhat	Mostly	Yes	Mostly	Yes	No
Jagatic et al (2007)	Yes	Somewhat	Yes	Somewhat	Mostly	No	Somewhat	Yes
Liu et al (2008)	Mostly	No	Yes	Somewhat	Somewhat	No	Yes	Mostly
Sureka (2011)	Yes	Mostly	Yes	No	Mostly	Somewhat	No	Somewhat

Primary studies	QAQ1	QAQ2	QAQ3	QAQ4	QAQ5	QAQ6	QAQ7	QAQ8
Stringhini et al (2010)	Yes	Somewhat	Mostly	No	Yes	Mostly	No	Yes
Hayati et al(2010a)	Yes	No	No	Somewhat	No	Yes	Mostly	Mostly
Hayati et al(2010b)	Mostly	Yes	Mostly	Yes	Yes	No	Somewhat	Yes
Shin et al (2011)	Yes	Somewhat	No	Somewhat	Mostly	Yes	Mostly	Mostly
Ramachandran et al (2011)	Mostly	Mostly	Yes	Somewhat	No	No	Yes	Somewhat

Sukanta et al (2012)	Yes	Mostly	No	Mostly	No	Somewhat	Yes	Yes
Dutta et al(2011)	Mostly	Yes	Somewhat	Somewhat	Yes	No	No	Yes
Debajyoti et al(2003)	Mostly	No	Yes	Yes	Somewhat	Mostly	No	
Kong et al (2005)	Mostly	Yes	Somewhat	No	Mostly	Yes	Yes	No
Li et al (2004)	Mostly	Yes	Yes	Somewhat	No	Yes	No	Somewhat
Pfleeger et al (2005)	Yes	Mostly	No	Yes	Mostly	No	Yes	No

8. REFERENCES

[1] Ismaila.I; Ali.S.,"Improved email spam detection model with negative selection algorithm and particles swarm optimization. In: Proceeding of Applied Soft Computing"; Applied Soft Computing 22 (2014) 11-27

[2] Irani.D.; S.Webb,; C.Pu, and K.Li ."Study of trend-stuffing on twitter through text classification. In Collaboration, Electronic messaging, Anti-Abuse and Spam Conference" (CEAS 2010),2010.

[3] Ahmed.F., Abulaish,M., ." A generic statistical approach for spam detection in Online Social Networks". Computer Communications 36 (2013) 1120-1129. Science direct (Elsevier).

[4] Gossier and Guadeloupe . "Social networks as an attack platform: Facebook case study". In Proceedings of the Eight International Conference on Networks.2009

[5] Irani.D; S.Webb, and C.Pu." Study of static classification of social spam profiles in myspace". In Proceedings of the International AAAI Conference on Weblogs and Social Media.2010

[6] Stein.T, Chen.E; and Mangla.K., ." Facebook immune system. In Proceeding of the forth ACM EuroSys Workshop on Social Network Systems" .2011

[7] Markines. B; Cattuto.C., Menczer., Benz.D., Hotho.A, and Stumme.G.," Evaluating similarity measures for emergent semantic of social tagging". In Proceeding 18th WWW Conference. 23-34pp.2009

[8] Byun.B;Lee.C;Webb.S;Irani.D; and Pu.C." An anti-spam filter combination framework for text-and-image emails through incremental learning" : In Proceedings of the Sixth Conference on Email and Anti-Spam (CEAS).2009.

[9] Heydari.A;Tavakoli.M.A;Salim.N;Heydari.Z.;".Detection of review spam: A survey". Expert Systems with Applications 42(2015) 3634-3442.2015.

[10] Goh.K.L.; Singh.A.K.; "Comprehensive Literature Review on Machine Learning Structures for web classification" In Proceeding 4th International Conference on Eco friendly computing and communication systems(ICECCS).Procedia Computer Science 7-(434-441) 2015.

[11] Zheng.X.; Zeng.Z.; Yu.Y.; Rong.C." Detecting Spammers on social networks". In Proceeding of the 26th Annual Computer Security Applications Conference, ACM.pp1-9.2010.

[12] Dinh.S;Azeb.T;Fortin.F;Mouheb.D."Spam campaign detection,Analysis and Investigation"Science direct.com. 2015.

[13] Li.Y; Wang.,T; X. Zhang,. A. Zhou,." Towards online review spam detection. In proceeding of the companion publication of the 23rd International conference on world wide web companion (pp341-342). International world wide web conference steering committee" .2014a

[14] Liu.B;"Sentiment Analysis and opinion mining, synthesis lectures on Human Language Technologies", 5(1),1-167.2012

[15] Ott.M., et al.;. "Finding deceptive opinion spam by any stretch of the imagination". In the 49th annual meeting for the computational linguistic. (pp.11).2011

[16] Xie.S; Wang.G; Lin.Y; & Yu.P; "Review spam detection via temporal pattern discovery". In Proceeding of the 18th ACM SIGKIDD. International Conference on Knowledge discovery and data minning". ACM.2012a

[17] Xie.S, et al,.. "Review spam detection via temporal pattern discovery". In Proceeding of the 18th ACM SIGKIDD. International Conference Companion on World Wide Web. ACM. 2012b

[18] Lam.H.Y; Yeung.DY.; " A learning approach to Spam Detection on Social networks". In Proceeding of 4th Conference on E-mail and Anti-spam, August,2-3,2007. CEAS 2007.

[19] Amitay.E.; Et al., "The connectivity sonar: detecting site functionality by structural patterns". Hypertext'03. 2003

[20] Castillo.C; Donato.D; Becchetti.L, .Boldi.P, Leonardo.S;, SantiniM;, Vigna.S; " A reference Collection for web spam", SIGIR forum'06.2006

[21] Gyongi.Z;Garcia-Molina.H; "Web spam Taxonomy. Technical report". Standford University. 2004

[22] Liu.B; JindalN;, "Analyzing and detecting review spam".ICDM 2007.

[23] Cao.Q, et al.,. "Aiding the detection of fake accounts in large scale social online services". In proceeding NDSI 2012.

[24] Zhang.H; "The optimality of naives bayes. 2004. In FLAIRS conference". Pp 562-567.2004

[25] Zhang.C.M & V. Paxson., 2011. "Detecting and analyzing automated activity on twitter". In PAM, pp 102-111.2011

[26] Wang.G, et al; " You are how you click: clickstream analysis for Sybil detection". In proceeding of the 22nd USENIX security symposium, pp241-256. 2013

[27] Seewald A.K;. " An Evaluation of Naïve Bayes Variants in content –based learning for spam filtering intelligent data analysis", 11(5): 497-524.2007

[28] Schneider.K.M; 2 "A comparison of event models for naïve bayes anti-spam email filtering". In EACL, pp 307-314. 2003

[29] Metsis.V; Androutsopoulos.I; G.Paliouras.; "Spam filtering with Naïve Bayes". In CEAS. 2006.

[30] Freeman.D.M; "Using Naïve Bayes to detect spammy names in social networks". In proceeding of AISEC '13, pages 307-314. 2013

[31] Xu.Z; et al;. "Towards semantic web: collaborative tag suggestions". In proceeding WWW'06 Collaborative web tagging workshop. 2006.

[32] .Mika.P; " Ontologies are: A unified model of social networks and semantics". In proceeding ISWC'05. Vol37299 of LNCS, pages 522-536.2005

[33] Markines.Benjamin ; "Efficient assembly of social semantics". In proceeding 19th ACM Conference on hypertext and hypermedia (HT), pp 149-156.

[34] Krause., et al., 2010. The anti-social tagger: detecting spam in social bookmarking systems. In Proc.4th Int'l workshop on adversarial information retrieval on the web (AIR web), pages 57-64.

[35] Kim.C, & Hwang. K.B; . "Naïve bayes classifier learning with feature selection for spam detection in social bookmarking". In proc. Europe Conference on Machine Learning and principles and Practice of Knowledge discovery in databases.(ECML/PKDD).2007

[36] Heyman.P; et al., 'Fighting spam on social web sites: A survey of approaches and future challenge"s. IEEE Internet Computing'11 (6): 36-45.2007

[37] Gkanogiannis.A;Kalamboukis.T; "A novel supervised learning algorithm and its use for spam detection in social bookmarking systems". In Proc.Europe .Conference On Machine Learning and Principles and practice of knowledge discovery in databases (ECML/PKDD).

[38] Chavalier.J;Gramme.P; "RANK for spam detection ECML-Discovery Challenge. In Proc. Europe Conference on Machine Learning and principles and practice of knowledge discovery in databases" (ECML/PKDD).2008

[39] Caverlee.J; et al.,. "Socialtrust: tamper-resilient trust establishment in online communities". In Proc.8th ACM/IEEE-CS Joint Conference on digital libraries(JCDL) pages 104-114.2008

[40] Benevenuto.F; et al; "Identifying video spammers in online social networks". In Proc. 4th Intl. Workshop on adversarial Information Retrieval on the web(AIR web),2008 pp 45-52.2008

[41] Bian.J; et al; "A few Bad votes too many?: towards robust ranking in social media". In Proc.4th International workshop on Adversarial Information Retrieval on the web (AIR web'08) pp 53-60.2008

[42] Gomes.L.H; et al., 2005. "Comparative graph theoretical characterization of networks of spam and legitimate email. In Proc. 2nd Conference on email and anti-spam. 2005. http://www.ceas.cc/papers-2005/131.pdf.

[43] Segal.J;, et al.;." Spamguru: An enterprise anti-spam filtering system". In 1st Conference on email and anti – spam CEAS 2004.

[44] Pfleeger. S.L ; G. Bloom., 2005. Canning Spam: Proposed solutions to unwanted email,security and privacy magazine, IEEE,3(2):40-47.2008

[45] Harris. E;. "The next step in the spam control war: Greylisting". Aug.2003. Retrieved: Aug.2006. 2003

[46] Bilge.L;, et al;. "All your contacts are belong to us: Automated identity theft attacks on social networks". In proceeding 18th International World wide web conference.2009

[47] Bonneau.J;, et al. "Eight friends are enough: social graph approximation via public listings". In proceeding of the 2nd ACM EuroSys workshop on social network systems, pages 13-8, ACM.2009

[48] Brown.G, et al .,2008.Social networks and context-aware spam. In proceeding of the ACM 2008 conference on computer supported co-operative work, pages 403-412, ACM NY,USA.

[49] Gross.R; Acquisiti.A. ; "Information Revelation and privacy in online social network (the facebook case)", in proceeding of 2005 ACM workshop on privacy in the electronic society, pages 71-80. 2005

[50] Jagatic.T;, et al. 2005. Facebook: Threats to privacy. Project MAC: MIT project on Mathematics and Computing.2005.

[51] Jones.G, Soltren.J; "Facebook: Threats to Privacy, Project MAC: MIT Projects on Mathematics and computing".2005

[52] Nazir.A; et al., 2008. "Unveiling Facebook: A measurement study of social network based applications" . In IMC'08: Proceeding of the ACM SIGCOMM conference on internet measurement, pages 43-56.NY USA.ACM.2008

[53] Hayati.P; et al; "Definition of spam 2.0: New spamming boom. In digital ecosystem and Technologies" (DEST),Dubai , UAE,2010. IEEE Computer society. 2010.

[54] Hayati .P. Potdar .V.;" Evaluation of web 2.0 Anti-spam Methods". In 7th Proceeding IEEE International Conference on Industrial Informatics, Cardiff Wales. 2009.

[55] Chu.Z,Gianvecchio.A;Hanning.S.S;Jajodia.S;"Who is Tweeting on Twitter Human, Bot or Cyborg"? In Annual Computer Security Application Conference .Austin Texas USA, December 6-10,2010, ACSAC'10, ACM.

[56] Markines.Benjamin;, et al. 2009. Social spam detection. In fifth International workshop on adversarial

Information Retrieval on the web. Madrid Spain, April 21,2009. AIRweb'09 ACM.

[57] Shin.Y; et al. " Prevalence and mitigation of forum spamming". In the 30th IEEE International Conference on Computer Communication. Shanghai, China, April 12-14,2011. IEEE INFOCOM 2011. IEEE Computer Society.Shanghai China.2011

[58] Thomason.A; "Blog spam: A Review, In conference on Email and Anti-spam" (Mountain View, California, August 2-3,2007, CEAS 2007.

[59] Hayati.P; Potdar.V, "Spammer and Hacker, Two Old friends". In 3rd IEEE International Conference on Digital Ecosystems and Technologies IEEE-DEST 2009, Istanbul, Turkey,2009.

[60] Sean. 2010. CPA lead Spam on Youtube. http://www.f-secure.com/weblog/archives/0002019.html

[61] Ridzuan .F; . et al., 2010. Key Parameter in Identifying cost of spam 2.0. In Proceeding of the 2010, 24th IEEE International Conference on Advanced Information Networking and Applications. IEEE Computer Society, pp 789-796.

[62] Liu.Y; et al., 2008. Identifying webspam with user behavior analysis. In fourth international workshop on adversarial information retrieval on the web. Beijing China, April 22,2008. Air Web'08. ACM.

[63] Sureka.A.; "Minning User Comment Activity for Detecting Forum Spammers in Youtube". In the 1st International Workshop on Usage Analysis and the web of data in the 20th International World wide web conference, Hyderabad, India March 28,2011.

[64] Stringhin.G; C. Kruegel.,Vigna.G;"Detecting spammers on social networks". In annual Computer Security Applications conference", Austin Texas, USA, ACSAC'10. ACM.2010

[65] Hayati.P.; et al. 2010. Web spambot detection based on web navigation behavior. In 24th IEEE International Conference on Advanced Information Networking and Application .AINA'2010, Perth, Australia.

[66] Hayati.P; et al. "Behaviour Based Web Spambot Detection by Utilising Action Time and Action Frequency". In the 2010 International Conference on Computational Science and Application. ICCSA'10, Japan,2010.

[67] Shin.Y; et al. 2011.The Nuts and Bolts of a forum Spam Automator. In the LEET'2011 Proceeding of the 4th UNISEX Conference on large scale Exploit and Emergent threats. Berkeley,CA, USA,2011.

[68] Bergholz.Andre; et al. 2010. New filtering approaches for phising email. Journal Computer Security; pp18(1):7-35.

[69] Bergholz.Andre.;. et al. 2008. Improved Phising detection using model-based features. In CEAS;2008.

[70] Anderson.D.S;Fleizach.C;Voelker.G.M;"Spamcatter:Characterizing internet scam hosting infrastructure" in Proceeding of 16th USENIX security symposium, SS'07;pp10:1-10:14. 2007

[71] Krebs.B; , 2012. Zeus Trojan Author in Spam kingpins. http://www.theage.com.au/it-pro/security-it/zeus-trojan-author-in-with-spam -kingpins-20122022-1tmqp.html

[72] Miller.Z; Dickson .B; Deitrick.W; Hu.W.; Wang. A.H; " Twitter Spammer Detection Using Data stream clustering ". Information Sciences 260.pp 693-695.Elsevier.2014.

[73] Grier.C; Thomas.K; Paxson.V,;.Zhang,S; "The Underground on 140 charactera or less. In 17th ACM Conference on Computer and Communications" security,Chicago,Illinois.USA,Oct 4-8,2010. CCS'10 ACM.2010.

[74] Gao.Hongyu; Hu. Jun; et al; "Detecting and Characterizing Social Spam Campaigns" IMC'10,Nov 1-3,2010. ACM 2010.

[75] Carreras .X; Marquez.L; "Boosting trees for anti-spam email filtering". Arxiv preprint, 2001.

[76] Pu.C; Webb.S; "Observed trends in spam construction techniques: a case study of spam evolution". In Proceedings of the Third Conference on Email and Anti-Spam (CEAS 2006), 2006.

[77] Wang.De; Irani. Danesh;Calton.Pu; " A Social -Spam Detection Framework" In Proceeding CEAS'11. Electronic Messaging, Anti-Abuse and Spam Conference.pp 46-54. 2011.

[78] Felt.Porter;Matthew.Finifter;Erika Chin; Steve Hanna; David Wagner. "A survey of mobile malware in the wild ". In Proceedings of the 1st ACM workshop on security and privacy in smartphones and mobile devices pages 3-14. 2011