Information Processing and Management 000 (2016) 1–21

Contents lists available at ScienceDirect

Information Processing and Management

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

# Recent developments in social spam detection and combating techniques: A survey

Manajit Chakraborty[a,*], Sukomal Pal[a,1], Rahul Pramanik[b], C. Ravindranth Chowdary[a]

[a] Department of Computer Science & Engineering, Indian Institute of Technology (BHU), Varanasi - 221 005, India
[b] Department of Computer Science & Engineering, Indian School of Mines, Dhanbad - 826 004, India

ARTICLE INFO

*Article history:*
Received 26 February 2014
Revised 17 April 2016
Accepted 26 April 2016
Available online xxx

*Keywords:*
Social spam
Social networks
Spam filtering

ABSTRACT

Spam in recent years has pervaded all forms of digital communication.The increase in user base for social platforms like Facebook, Twitter, YouTube, *etc*., has opened new avenues for spammers. The liberty to contribute content freely has encouraged the spammers to exploit the social platforms for their benefits. E-mail and web search engine being the early victims of spam have attracted serious attention from the information scientists for quite some time. A substantial amount of research has been directed to combat spam on these two platforms. Social networks being quite different in nature from the earlier two, have different kinds of spam and spam-fighting techniques from these domains seldom work. Moreover, due to the continuous and rapid evolution of social media, spam themselves evolve very fast posing a great challenge to the community. Despite being relatively new, there has been a number of attempts in the area of social spam in the recent past and a lot many are certain to come in near future. This paper surveys the recent developments in the area of social spam detection and mitigation, its theoretical models and applications along with their qualitative comparison. We present the state-of-the-art and attempt to provide challenges to be addressed, as the nature and content of spam are bound to get more complicated.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Spam or *Electronic Spam* in computer science jargon refers to the unwanted or unsolicited messages sent or received electronically by means of e-mail, instant messenger, blogs, newsgroups, social networks, web search, mobile phones *etc.* for the purpose of advertising, phishing, spreading malware, *etc*. (Wikipedia (2015)). As is evident from the definition of the term, spam is intended for malice and usually accounts for a viable but fraudulent source of income for some individuals or organizations. An individual involved in sending such spam messages is generally termed as a “spammer”.

Although initially targeted and limited to e-mail, spam has effectively invaded all electronic platforms across all media. Specifically, spam has proliferated (but is not restricted to) the following media:

* Corresponding author.
*E-mail addresses:* cmanajit.rs.cse14@iitbhu.ac.in, chakraborty.manajit@gmail.com (M. Chakraborty), spal.cse@iitbhu.ac.in (S. Pal), rahulpramanik@cse.ism.ac.in (R. Pramanik), rchowdary.cse@iitbhu.ac.in (C. Ravindranth Chowdary).
[1] the author relocated during the second revision of the paper from *b* to *a*.

http://dx.doi.org/10.1016/j.ipm.2016.04.009
0306-4573/© 2016 Elsevier Ltd. All rights reserved.

Please cite this article as: M. Chakraborty et al., Recent developments in social spam detection and combating techniques: A survey, Information Processing and Management (2016), http://dx.doi.org/10.1016/j.ipm.2016.04.009

- *e-mail spam*:- E-mail spam, also known junk mail or unsolicited commercial e-mail, is the custom of sending unwanted e-mail messages, frequently with commercial content, in bulk quantities to an indiscriminate set of recipients.
- *Spam in instant messaging*:- Instant messaging spam use Instant Messengers (IMs) for spam delivery. Although considerably subtle than its e-mail counterpart, it has the tendency to annoy users of IMs like Skype©[2] and Yahoo!® Messenger[3] with unsolicited messages from advertisers *etc*.
- *Spam in newsgroups and forums*:- Repetitive multiple posting on Usenet newsgroups are termed as Newsgroup spam. Similarly, creation and posting of advertisements on the Internet forums irrelevant to the topic of discussion are termed as Forum spam.
- *Spam in mobile phones*:- This form of spamming uses the Short Messaging Services (SMS) as its medium. Sometimes the customers are charged with premium services for being foxed into some sort of false subscription and scam.[4] Unsolicited text messages tend to interrupt inadvertently and irritate customers.
- *Search engine spam*:- Also known as *Spamdexing*, refers to the practice of manipulating the search engine ranking and relevancy algorithm to promote a particular website or web page.
- *Blog and wiki spam*:- Spam in blogs is also known as *splog* and refers to the comments that are extraneous to the topic of discussion. These comments are often embedded with URL links to some commercial sites. Some splogs are written like verbose ads for the websites they promote; others have no original content, featuring either nonsense or content stolen from legitimate websites (Rouse (2015)). Similar kinds of attacks are also observed in Wikis and other guestbooks which accept comments from general users.
- *Spam on video sites*:- Social media websites like YouTube are also infested with spam which usually involves comments and links to some pornographic or dating site or some unrelated videos. Sometimes these comments are automatically generated through bots.
- *Social networking spam*:- Even the latest forms of online communication like social networking sites are not spared from spam. Every now and then messages containing spam links are encountered by users. Social networking giants like Facebook, Twitter, Digg, Foursquare, *etc*., all suffer from one or the other kind of spamming.
- *Spam in online game messaging*:- Although what qualifies as spam on online video game sharing websites is quite subjective, instances of message flooding, request to join some particular group, violation of copyright terms and conditions *etc*., are sometimes referred to as spam.
- *Spam over internet telephony*:- Also known as *SPIT*, it uses Voice over Internet Telephony (VoIP) to spam. Usually, when the receiver fallibly receives a spam call, a pre-recorded message is played. This platform is a vulnerable target for spammers since VoIP is cheap and can be anonymized easily.

Spam is inevitable in almost all forms of online communication today and is known to hamper the productivity of the medium on which it appears. Various measures have been taken to improve the robustness of different electronic media against an array of spam attacks. These measures are better known as the anti-spamming techniques or spam combating techniques. While a lot of work has been done in the area of spam fighting, especially e-mail and web spam, the premise of spam combating in social networks and social media is still not mature enough. This is because social networks like the web graph is uncontrollable (Albert, Jeong, & Barabási (1999)) and ever evolving. Ease of content generation without any moderation and monitoring policy and rapid dissemination to a large set of users have made the online social media as a conducive breeding ground for spam. The marketers aggressively use this cheap and new mode of communication to advertise or sell their products. A survey carried out by a brand protector and compliance company revealed that the rate of social spam has increased by 355% within the period of January–July 2013.[5] Another report[6] also proves that proliferation of spam is rampant. With time, a lot more users of the Internet are expected to engage in communication through social networks. As spam hampers the productivity and performance of social websites causing erosion in the user base and thus associated financial loss, it is of utmost importance that social spam is contained within acceptable limits if not completely eliminated. All these factors are motivation behind our work on social spam and its combating.

Due to differences in the characteristic features of social websites from that of usual search engines or e-mails, spam fighting here is multifaceted and challenging. A number of innovative approaches have been employed to fight spam on social websites in addition to the already existing ones. A brief account of them can be found in Castillo and Davison (2011). This monograph discusses various aspects of adversarial web search having a complete chapter dedicated to user-generated content. It comprehensively accounts for all the social spam combating techniques till 2010. However, as a fast evolving field, social networks have substantially changed in the last five years matched by similar endeavors by spammers. Anti-spam strategies, thus, also need a major upheaval to mitigate them. We, therefore, started from the point where Castillo and Davison (2011) left. In this paper, we survey the emerging and evolving technologies and techniques that have been proposed in the recent years (2011 onwards) to combat spam in social network websites.

[2] http://www.skype.com/en/
[3] https://messenger.yahoo.com/
[4] http://www.androidheadlines.com/2016/03/state-mobile-spam-according-att.html
[5] http://info.nexgate.com/nexgate-social-media-spam-research-report
[6] http://www.baselinemag.com/security/malware-attacks-and-phishing-scams-increase.html

Please cite this article as: M. Chakraborty et al., Recent developments in social spam detection and combating techniques: A survey, Information Processing and Management (2016), http://dx.doi.org/10.1016/j.ipm.2016.04.009

Download English Version:

https://daneshyari.com/en/article/4966484

Download Persian Version:

https://daneshyari.com/article/4966484

Daneshyari.com