A Survey on the Spam Issue in Twitter

Soufiane Maguerra*
Azedine Boulmakoul*
Lamia Karim**
Hassan Badir***

*LIM/IOS, FSTM, Hassan II University of Casablanca, Mohammedia, Morocco
{maguerra.soufiane,azedine.boulmakoul}@gmail.com,
**Higher School of Technology EST Berrechid, Hassan 1st University, Morocco
lkarim.lkarim@gmail.com
***National School of Applied Sciences Tangier, Abdelmalek Essaâdi University, Morocco
hbadir@gmail.com

Abstract. Social networks are being leveraged by cyber-criminals to cover a wider range of victims. In Twitter, spammers create several bots and behave in different patterns according to their desired aims. Particularly, spammers can spread malicious links leading to malware or phishing sites. Achievable by engaging in social bonds or responding to trending topics (hashtags). Spammers either spam in an individual manner, otherwise in coordinated communities with a clear insight. Decidedly, reinforcing cyber-security in Twitter is an indispensable fact. Several researchers have been studying the different aspects of spamming in Twitter. This paper includes a background over the information handled in Twitter, and a detailed survey over the papers dealing with the spam issue. The discussed papers have been published from 2010 to 2018. In contrast to other surveys, this paper is not limited to the detection of spammers but it also discusses the approaches to the detection of spam communities, compromized accounts, collective attention spam, and the extraction of cybercrime knowledge. Hence, this study can be considered as an essential step for the design of a unified spam detection framework.

1 Introduction

The misuse of social media lead to an exponential rise of cybercrime victims. Cyber-criminals leverage the privileges of social media to range over a wider field of victims. Consequently, malicious links are invading social walls. Crackers gather information about accounts to gain their trust; then, oblige them to pay ransoms for safety. Information concerning system vulnerabilities, hacking tools and techniques are been leaked to expand the cyber-criminals community. Thus, the primal necessity to secure social networks.

As a primal study, we focus on Twitter characterized by a total number of 1.3 billion accounts with 330 million active ones per month, 500 million tweets sent per day with 6000 sent every second, 100 million active users per day, and an estimate of 23 million bots (Sikandar G,