Camouflages and Token Manipulations- The Changing Faces of the Nigerian Fraudulent 419 Spammers

LONGE, Olumide Babatope
Dept. of Computer Science,
University of Ibadan, Ibadan, NIGERIA
longeolumide@yahoo.com

CHIEMEKE, Stella Chinye.
Dept. of Computer Science,
University of Benin, Benin City, NIGERIA
schiemeke@yahoo.com

ONIFADE, Olufade .F. Williams
Laboratoire Lorrain de Recherche en Informatique et ses Application (LORIA)
LORIA- Campus Scientifique, B. P. 239, 54506 Vandoeuvre-Lès-Nancy, France
fadowilly@yahoo.com

LONGE, Folake .Adunni
The African Regional Centre for Information Science,
University of Ibadan, Ibadan, NIGERIA
adefolakelonge@yahoo.com

Abstract

The inefficiencies of current spam filters against fraudulent (419) mails is not unrelated to the use by spammers of good-word attacks, topic drifts, parasitic spamming, wrong categorization and recategorization of electronic mails by e-mail clients and of course the fuzzy factors of greed and gullibility on the part of the recipients who responds to fraudulent spam mail offers. In this paper, we establish that mail token manipulations remain, above any other tactics, the most potent tool used by Nigerian scammers to fool statistical spam filters. While hoping that the uncovering of these manipulative evidences will prove useful in future antispam research, our findings also sensitize spam filter developers on the need to inculcate within their antispam architecture robust modules that can deal with the identified camouflages.

Keywords: 419-Spammers, Bayesian, Outbound-filtering, SPAMAng, Filters, Nigeria.

1. Introduction

Some spammers are legitimate businesses, engaged in overly aggressive marketing efforts, because there are no formal limits on their actions. In spite of the challenges created by needing to work at an international level, there is a reasonable expectation that legal strictures, both laws and contracts, will constrain these businesses to a tolerable level [25]. In contrast, *rogue* spammers actively seek to avoid accountability, to subvert barriers to their traffic, and to acquire unwitting and unwilling participation of machines owned by others. Independent of the legal details, the best social model to use for analyzing this latter group is crime. Often the activities do not violate particular laws, but what is most important is that the style of a spammer's conduct is the same as that of a criminal.

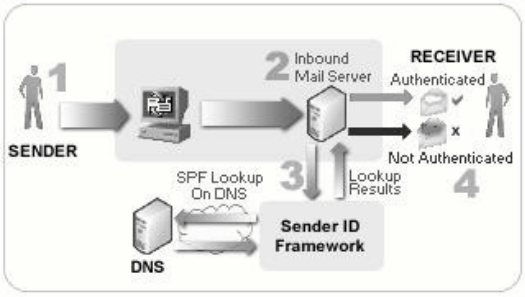

Figure 1: SPF Mechanism
(Source: Wong & Schlit [29])

Unfortunately, the technical and operational world of spamming has also developed in scale and sophistication. Spamming used to entail one sender and one sending machine. Its performance was limited by the capacity of that machine and the bandwidth of its Internet connection. Today, rogue spammers control vast armies of compromised systems, called *zombies*, as shown in Figure 1. Zombies are

owned by legitimate users who are unaware that their system has been compromised and is being used for spamming.

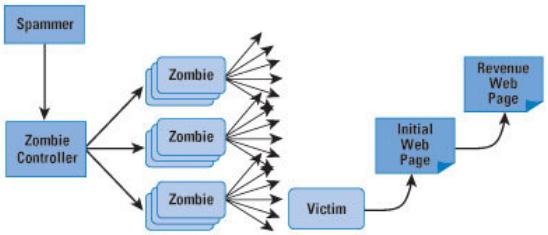

Figure 2: Rogue Spammer Control Network

The community of rogue spammers is remarkably well organized, and Nigerian 419 spammers fall into this category [24]. It has become an extensive, underground economy. Some participants specialize in developing methods for breaking through filters. Others take over machines and turn them into zombies. Others sell the use of a zombie collection for periods of spamming. The estimated number of zombie systems is in the many tens of millions. After spam delivery, recipients often "click" to a transaction Web page. Web hosting is provided at multiple levels, in order to obscure the server side of the process, further reducing accountability. Typically, spammers have the classic goal of selling products. However, they also can have political or religious motivations or even blatantly criminal intent, such as extortion. The ability to send very large number of messages to a specific destination gives spammers a tool that can be used to threaten an organization with a denial of service attack on their network [25].

2. Related Literature

Spam behavior is not simply a matter of one concept drifting to another in succession, but instead it is a superimposition of constant, periodic and episodic phenomena [17]. Researchers have shown that spam content changes over time. Swan and Allan [26] employed a 2-t test to discover "bursty" topics in daily news stories. Their test was designed to determine whether the appearance of a term on a given day was statistically significant. They concluded that that spam has complex time-varying behavior. Some terms recur intermittently, such as *adult, click,free, hot and removed*. Others are episodic, for instance the terms common in a "Nigerian 419 scam" such as *Nigeria, Lagos, assistance, beneficiary* and terms in a "pornstar video" such as *awesome, pornstars, jenna, lauren, nicki, orgy* also burst at some points. Terms such as *Christmas* burst late in the year and presumably reappears every year around the same time.

Spammers typically use purpose-built applications to distribute their spam [27]. Greylisting tries to deter spam by rejecting email from unfamiliar IP addresses, by replying with a soft fail (i.e. 4xx). It is built on the premise that the so-called 'spamware' does little or no error recovery, and will not retry to send the message. Any correct client should retry; however, some do not (either due to a bug or policy), so there is the potential to lose legitimate email. Also, legitimate email can be unnecessarily delayed; however, this is mitigated by source IP addresses being automatically whitelisted after they have successfully retried once. An analysis performed by Levine [28] over a 7-week period (covering 715,000 delivery attempts), 20% of attempts were greylisted; of those, only 16% retried. Careful system design can minimise the potential for lost legitimate email; certainly greylisting is an effective technique for rejecting spam generated by poorly implemented spamware.

SMTP path analysis [21] learns the reputation of IP addresses and email domains by examining the paths used to transmit known legitimate and spam emails. It uses the 'received' line that the SMTP protocol requires that each SMTP relay adds to the top of each email processed, which details its identity, the processing timestamp and the source of the message. Despite the fact that these headers can easily be spoofed, when operating in combination with a Bayesian filter, overall accuracy is approximately doubled.

Tom [17] opined that it will be difficult to estimate how much we can expect spam as a concept to drift over time, in part because no metric of concept drift has been adopted by the antispam community. Cunningham et al. [18] construct a case-based reasoning classifier that can track concept drift. They propose that the classifier both adds new cases and removes old cases from the system collection, allowing the system to adapt to the drift of characteristics in both spam and legitimate mails. An initial evaluation of their classifier suggests that it outperforms naive Bayesian classification. This is unsurprising given that naive Bayesian filters attempt to learn a "unified spam concept" that will identify all spam emails; spam email differs

significantly depending on the product or service on offer. In [19] the Teiresias pattern discovery algorithm was applied to email classification. Given a large collection of spam emails, the algorithm identifies patterns that appear more than twice in the corpus. Negative training occurs by running the pattern identification algorithm over legitimate email; patterns common to both corpora are removed from the spam vocabulary. Successful classification relies on training the system based on a comprehensive and representative collection of spam and legitimate emails. Experimental results are based on a training corpus of 88,000 pieces of spam and legitimate emails. Spam precision was reported at 96.56%, with a false positive rate of 0.066%.

2.1 Techniques Adopted By the Nigerian Spammers

In order to reach a large volume of users, Spammers require an equally large number of e-mail addresses. These are usually collected in three different ways. By using mail harvesters to scavenge for e-mail addresses listed on web sites and message boards (particularly USENET groups), by performing a dictionary attack (pairing randomly generated usernames with known domain names to 'guess' a correct address) or by purchasing address lists from other individuals or organizations. Once they have addresses, Spammers can use programs known as bulk mailers to automate the sending of Spam. These programs can send huge volumes of e-mail messages in a small amount of time. Some bulk mailing programs use open-relays (e-mail servers that allow unauthorized users to send e-mail) to send messages, effectively hiding the true address of the Spammers. Bulk mailers can also fabricate the *from* address in e-mail message headers to further hide the identity of the Spammer [31].

Another technique spammers utilize to send e-mails is with the use of *zombie network*s, also know as *bot network*s. Zombie is the term given to a computer that has been infected by a virus, worm, or Trojan Horse, which allows remote entities to take control and use it for their own (usually illegal) purposes. A large amount of these computers, usually called a *network* or *army* can be co-opted to send spam e-mails, requiring little of the spammer's own computing power and network bandwidth. This technique is also popular as it protects the identity of the spammer. What follows is a short summary of some of the major techniques and tactics adopted by the Nigerian 419 scammer.

2.2 Current Filters Against 419 Mails

Dealing with the 419 Spam problem at the destination will not always yield the desired results as Spammers have evolved different measures to beat Spam filters at these nodes.. This is particularly so with 419 mails. Available benchmarks though promising, still reveals some degrees of inefficiency when compared with the rate of false positives (Process Software, 2007). SPAMAng, an outbound filtering to combat Nigerian Fraudulent 419 mails from the source was proposed in [24] Other major challenges are discussed below:

(a) The fact that the design of efficient filtering systems depends on spamicity measures means that accuracy must be a goal when designing antispam systems. Accurate spamicity measures can only be obtained if the filter designers have a balanced corpus of Spam and real mails (ham) from the domain at which the filtering efforts are targeted. Unfortunately, most Spam designers have easy access to the Nigerian 419 mail corpus but are disadvantaged at having a corpus of ham mails originating from the source at which the Spam mails have emanated. The consequences of these are the occurrence of false positives, false negatives and lack of efficiency in the filtering systems. The contribution of Spam mails in training the text classification engine is as important as that of ham mails.

(b) Generic filtering, where a single filtering package is designed to filter different types of Spam becomes disadvantageous in a situation where there is serious deviations in the test categories and manipulations by Spammers in specific domain. This is a major challenge of using existing filters to deal with the Nigerian 419 mails. Rule-based approaches are as deficient as non-domain specific approaches in filtering junk mails that have diversified text categories. This is due to the fact that logical rule sets usually make rigid binary decisions as to whether to classify a given message as junk [30].