Chapter 14
Deception Detection on the Internet

Xiaoling Chen
Stevens Institute of Technology, USA

Rohan D.W. Perera
Stevens Institute of Technology, USA

Ziqian (Cecilia) Dong
Stevens Institute of Technology, USA

Rajarathnam Chandramouli
Stevens Institute of Technology, USA

Koduvayur P. Subbalakshmi
Stevens Institute of Technology, USA

ABSTRACT

This chapter provides an overview of techniques and tools to detect deception on the Internet. A classification of state-of-the-art hypothesis testing and data mining based deception detection methods are presented. A psycho-linguistics based statistical model for deception detection is also described in detail. Passive and active methods for detecting deception at the application and network layer are discussed. Analysis of the pros and cons of the existing methods is presented. Finally, the inter-play between psychology, linguistics, statistical modeling, network layer information and Internet forensics is discussed along with open research challenges.

INTRODUCTION

The Internet is evolving into a medium that is beyond just web search. Social networking, chat rooms, blogs, e-commerce, etc. are some of the next generation applications that are gaining prominence. A darker side of this growth that has an immense negative impact on the society at large is the overt or

DOI: 10.4018/978-1-60566-836-9.ch014

covert support for deception related hostile intent. Deception is defined as the manipulation of a message to cause a false impression or conclusion (Burgoon & Buller, 1994).

Hostile intent and hostile attack have some differences. Hostile intent (e.g., email **phishing**) is typically passive or subtle and therefore challenging to measure and detect. However, hostile attack (e.g., denial of service attack) leaves signatures that can be easily measured. Note that intent is typically considered a psychological state of mind. *How does this deceptive state of mind manifest itself on the Internet?* Is it possible to create a statistically based psychological Internet profile for a person? To address these questions, ideas and tools from cognitive psychology, linguistics, statistical signal processing, digital forensics and network monitoring are required.

Deception based hostile intent on the Internet manifests itself in several forms including:

- promoting hostile ideologies—promoting false propaganda and psychological warfare;
- exploitation—deception with predatory intent on social networking web sites and Internet chat rooms;
- email **phishing**—a user is falsely asked to change the password or personal details in a fake web site, etc.

Clearly, the negative impact of these hostile activities has immense psychological, economical, emotional, and even physical implications. Therefore, quick and reliable detection or prediction of hostile intent on the Internet is of paramount importance.

To prevent e-commerce scams, some organizations have offered guides to users, such as eBay's spoof email tutorial, and Federal trade commission's **phishing** prevention guide. Although these guides offer sufficient information for users to detect **phishing** attempts, they are often ignored by the web surfers. In many email **phishing** scams, in order to get the user's personal information such as name, address, phone number, password, Social Security number etc., the email is usually directed to a deceptive web site that has been established only to collect a user's personal information, which may be used for identity theft.

Due to the billions of dollars lost due to **phishing**, anti-phishing technologies have drawn much attention. Carnegie Mellon University (CMU) researchers have developed an anti-phishing game that helps to raise the awareness of the Internet **phishing** among web surfers (Anti-Phishing Phil, 2008). Most e-commerce companies also encourage customers to report scams or **phishing** emails. This is a simple method to alleviate scams and **phishing** to a certain level. However, it is important to develop algorithms and software tools to detect deception based Internet schemes and phising attempts. Many anti-phishing tools are being developed by different companies and universities, such as Google, Microsoft, McAfee, etc. The first attempts to solve this problem are anti-phishing browser toolbars, for example, Spoofguard and Netcraft toolbars (Fette, Sadeh, & Tomasic, 2007). However, study shows that even the best anti-phishing toolbars can detect only 85% of fraudulent web sites. This performance is known to be far from being an acceptable level of security (Anti-Phishing Guide, 2008). Most of the existing tools are built based on the network properties, like the layout of website files or the email headers. For instance, Microsoft has integrated Sender ID techniques into all of its email products and services, which detects and blocks almost 25 million deceptive email messages every day (Anti Phishing technologies, 2008). Microsoft **Phishing** Filter in the browser is also used to help determine the legitimacy of a web site. Also, a PILFER (**phishing** identification by learning on features of email received) algorithm was