Chapter 13
Recognizing User Portraits for Fraudulent Identification on Online Social Networks

Sudha Senthilkumar
School of Computer Science and Engineering, VIT University, Vellore, India

Satha Sivam S.
School of Information Technology and Engineering, VIT University, Vellore, India

Brindha K.
School of Information Technology and Engineering, VIT University, Vellore, India

ABSTRACT

Online social networks (OSNs) are increasingly influencing the way people communicate with each other. Well known sites such as Facebook, LinkedIn, Twitter, and Google+ have millions of users across the globe. With the wide popularity there are lot of security and privacy threats to the users of online social networks (OSN) such as breach of privacy, viral marketing, structural attacks, malware attacks, and profile cloning. Social networks have permitted people to have their own virtual identities which they use to interact with other online users. It is also completely possible and not uncommon for a user to have more than one online profile or even a completely different anonymous online identity. Entity resolution (ER) is the task of matching two different online profiles potentially from social networks. Solving ER has an identification of fake profiles. The solution compares profiles based on similar attributes. The system was tasked with matching two profiles that were in a pool of extremely similar profiles.

INTRODUCTION

People can now have their own virtual identities on social networks, which they can use to connect with other online users. Millions of people use social media sites like Facebook, Twitter, and Google+. Facebook, one of the most popular social networks, recently completed one of the largest initial public

DOI: 10.4018/978-1-6684-3991-3.ch013

offerings in Internet history. These social networks enable real-life users to construct online profiles based on the data they provide. The profiles are online personas that can exist independently of their real-life counterparts. Among these profiles, users communicate directly by posting and sharing content, expressing opinions about one another's posts, etc. The concept of a social network can be described as a graph composed of nodes and vertices, where friends are the nodes and friendships are the vertices. During the registration process, these profiles are created. As the average social network's registration process usually requires the user to type in their information manually, it is very easy and not uncommon to create a profile with inaccurate or fake information. The public information of the profiles from different social networks could be of interest to multiple parties in order to match and correlate data in order to identify a single entity with different profiles. Entity Resolution is the process of matching profiles to create a single entity that represents one real-world entity. We present an alternative form of comparing profiles that takes advantage of other information that is available without using the training phase, while providing a more detailed source of information about people, searching the web for people across social networks, helping businesses know their candidates better before hiring them, improving marketing strategies, detecting fake profiles, etc. The objective is to present an alternative form of comparing profiles that takes advantage of other information that is available without using the training phase. By comparing other types of information if it was publicly available, we went further than just comparing image-based features between profiles to solve ER. A string comparison technique was used for image-based features such as profile images and posted images.

Scope of Proposed System

The proposed system is the first to detect this fraud automatically. We aim to prevent romance scammers from creating fraudulent profiles or engaging with potential victims before they start a romantic relationship. Earlier research found that romantic scam victims score highly on idealized romantic beliefs scales. These beliefs are captured using a combination of structured, unstructured, and deep-learned features.

To solve fake profiles, we designed a method that uses two primary modules. Using a "data hiding" module integrated into the "Comparison distributor" component, we acquire datasets of profiles from social networks, compare the attributes, and discover similarities between them. Through the "Match Selector" component, the "Profile Matching" module works. Based on the previously determined similarities between profiles, the purpose of this second module is to identify potential matches.

LITERATURE SURVEY

Title 1: Design And Evaluation Of A Real-Time Url Spam Filtering Service
Authors: Kurt Thomas, Chris Griery, Justin Ma, Vern Paxsony, Dawn Song
Year: 2011
Description:

Scams, phishing, and malware have all increased as a result of social networks, URL shorteners, and other web services. Although extensive research has been conducted, email-based spam filtering techniques generally fail to protect other web services. URLs are crawled and analyzed in real-time while they are submitted to web services to determine whether they direct to spam. As a result of the diversity of web service spam, we evaluate the viability of Monarch and the fundamental challenges it faces. So-