Detection of Suspecious URLs Using Real Time System on Social Networks

S. UmaMaheswari
Research Scholar
SCSVMV University
Kanchipuram, India
umarunn@gmail.com

S.K.Srivatsa
Senior Professor
St.Joseph College of engg
Chennai, India
profsks@rediffmail.com

ABSTRACT-Twitter is one of the famous social networking and information sharing service which allows users to connect with worldwide users. When twitter users want to share a URL with friends via tweets, they usually use URL shortening services to reduce the URL length because tweets can contain only a restricted number of characters. Malicious users often try to find a way to attack it. The most common forms of web attacks including spam, scam, phishing, and malware distribution attacks, have also appeared on twitter using URLs. A number of suspicious URL detection schemes have also been introduced. They use static or dynamic crawlers, and they may be executed in virtual machine honey pots such as Capture-HPC and Honey monkey to investigate newly observed URLs. These schemes are ineffective against feature fabrications or consume much time and resources. We propose an effective suspicious URL detection system for twitter. Our system investigates correlations of URL redirect chains extracted from several tweets. Because attackers have limited resources and usually reuse them, their URL redirect chains frequently share the same URLs. We develop methods to discover correlated URL redirect chains using the frequently shared URLs and to determine their suspiciousness.

Index Terms-Twitter, Social networking, URL, Malicious, Suspecious, crawlers,Correlations and Redirect chains.

1. INTRODUCTION

Social networking sites are the websites used to communicate and for expressing their interests with others in online. It gives ease of access to new trends /topics and faster communication over longer distances. Some of Internet users use SNS for meeting new friends. Some users use it to find old friend and relatives. SNS provide users with lots of benefits like sharing various level of information, media sharing (photo, video, fetch) and many other things. To the highest degree SNS also allows you to make your group based on your interest. It is an easy way to find friends and is similar to "one to many " or "many to many" relationships [24]. It resembles the aspects such as simple, User friendly setup and personalization. In general, social networking websites are easily accessible and globally available. Hence the social networking sites played an important role in the society and in the world.

1.1 What is social networking?

Social Networking is a grouping of individuals into particular groups . A social network can be defined as a network of interactions or relationships, where the nodes consist of actors and the edges consist of the relationships or the interactions between these actors [25].We have so many types of SNS. They posses salient features for different types of purposes. They are photo-sharing or video-sharing capabilities, built-in blogging and instant messaging technology etc. Some sites are designed with specific ethic, religious, political, or other identity driven categories in mind [26]..Nowadays we have more benefits by using social networking sites such as constant flows of information from updates and real time communication, forming communities of interest, publishing and sharing the contents, colloborating with others and providing added context and value for the knowledge. The current researches have mostly done on the concept of privacy concerns in SNS. The growth of usable tools for protecting personal data in social media is becoming prominent problem that has caught much attention recently.[24].

1.2 Twitter

Twitter is a famous social networking and information sharing service[9] that allows users to exchange messages of fewer than 140-character, also known as tweets, with their friends. When Twitter users want to share a URL with friends via tweets, they usually use URL shortening services [11] to reduce the URL length because tweets can contain only a restricted number of characters. Owing to the popularity of twitter, malicious users often try to find a way to attack it. The most common forms of web attacks including spam, scam, phishing and malware distribution attacks have also appeared on twitter. Because tweets are short in length, attackers use shortened malicious URLs that redirect twitter users to external attack servers [12], [13], [6], [14]. To cope with malicious tweets, several Twitter spam detection schemes have been proposed. These schemes can be classified into account feature-based [7], [5], [15], [4], relation feature-based [16], [3] and message feature-based [17] schemes. A number of suspicious URL detection schemes [1] have also been introduced. They use static or dynamic crawlers, and they may be executed in virtual machine honey pots such as Honey Monkey [21] and Wepawet [22] to investigate newly observed URLs. These schemes classify URLs according to several features including lexical features of URLs, DNS information, URL redirections and the HTML content of the landing pages.

Nevertheless, malicious servers can bypass an investigation by selectively providing benign pages to crawlers.

2. LITERATURE SURVEY

In detecting spammers on social networks, Gianluca Stringhini and Christopher Kruegel [23] have showed that spam on social networks is a problem. For study, he created a population of 900 honey-profiles on three major social networks and observed the traffic they received. Then he developed techniques to identify single spam bots as well as large-scale campaigns. He also showed how our techniques help to detect spam profiles even when they do not contact a honey profile. We believe that these techniques can help social networks to improve their security and detect malicious users.

In uncovering social spammers: social honeypots and machine learning Kyumin Lee and James Caverlee [28] has presented the design and real-world evaluation of a novel social honeypot-based approach to social spam detection. Our overall research goal is to investigate techniques and develop effective tools for automatically detecting and filtering spammers who target social systems. By focusing on two different communities, we have seen how the general principles of social honey pot deployment, robust spam profile generation and adaptive and ongoing spam detection can effectively harvest spam profiles and support the automatic generation of spam signatures for detecting new and unknown spam. Our empirical evaluation over MySpace and twitter has demonstrated the effectiveness and adaptability of the honey pot-based approach to social spam detection.

In beyond blacklists: learning to detect malicious websites from suspicious URLs .Justin Ma, Lawrence K. Saul [27]has proposed and described an approach for classifying URLs automatically as either malicious or benign based on supervised learning across both lexical and host-based features. We argue that this approach is complementary to both blacklisting which cannot predict the status of previously unseen URLs and systems based on evaluating site content and behavior which require visiting potentially dangerous sites. Further, we show that with appropriate classifiers it is feasible to automatically shift through comprehensive feature sets (i.e., without requiring domain expertise) and identify the most predictive features for classification. An open issue is how to scale our approach to handle millions of URLs whose features evolve over time. We address the issue in subsequent work by using online learning algorithms.

3. AN OVERVIEW OF EXISTING METHODOLOGIES

Many Twitter spam detection schemes have been introduced. Most have focused on how to collect a large number of spam and non spam accounts and extract the features that can effectively distinguish spam from non spam accounts[20]. In this paper[10], we propose a suspicious URL detection system for Twitter. Instead of investigating the landing pages of individual URLs in each tweet, which may not be successfully fetched, we considered correlated redirect chains of URLs included in a number of tweets. Because attackers' resources are limited and need to be reused, a portion of their redirect chains must be shared. We found a number of meaningful features of suspicious URLs derived from the correlated URL redirect chains and related tweet context information. We collected a large number of tweets from the Twitter public timeline and trained a statistical classifier with their features. The trained classifier has high accuracy and low false-positive and false-negative rates.

Account feature-based schemes use the distinguishing features of spam accounts such as the ratio of tweets containing URLs, the account creation date, and the number of followers and friends. The relation feature-based schemes rely on more robust features that malicious users cannot easily fabricate such as the distance and connectivity apparent in the twitter graph. Extracting these relation features from a twitter graph however, requires a significant amount of time and resources as a twitter graph is tremendous in size. The message feature-based scheme focused on the lexical features of messages. Twitter public timeline to detect accounts that post tweets with blacklisted URLs and yet others monitor twitter's official account for spam reporting at spam. But the existing system cannot catch suspicious URLs that repeat after long-time intervals and also the latency is bad.

4. PROPOSED METHODOLOGY

Our system consists of four components: data collection, feature extraction, training, and classification. The collection of tweets with URLs and crawling for URL redirections. To collect tweets with URLs and their context information from the twitter public timeline, this component uses twitter streaming APIs. Grouping of identical domains, finding entry point URLs, and extracting feature vectors. This component monitors the tweet queue to determine whether a sufficient number of tweets have been collected. Retrieval of account statuses and training of the classifier. Because we use an offline supervised learning algorithm, the feature vectors for training are relatively older than feature vectors for classification. The classification component executes our classifier using input feature vectors to classify suspicious URLs. When the classifier returns a number of malicious feature vectors, this component flags the corresponding URLs and their tweet information as suspicious. This suspicious will be delivered to security experts or more sophisticated dynamic analysis environments for an in-depth investigation. The proposed system can easily fabricate syntactical features of spam messages and Some simple modifications can also be applied to other services that can monitor a continuous URL stream.

5. SYSTEM ARCHITECTURE

A system architecture is the conceptual design that defines the structure and behavior of a system. An architecture

description is a formal description of a system, organized in a way that supports reasoning about the structural properties of the system.

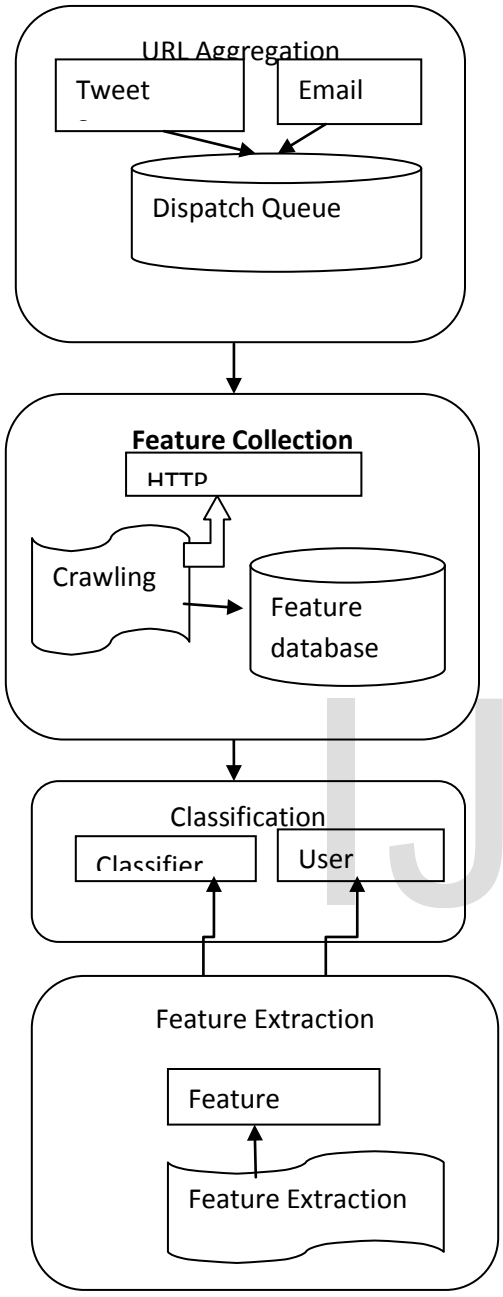

Fig.1. system architecture

The system components or building blocks and provides a plan from which products can be procured and systems developed, that will work together to implement the overall system.The proposed system architecture is depicted in Fig.1. The proposed system contains four stages like Suspicious URL Detection, Data Collection, Feature Extraction ,Training and Classification.

5.1 Data Collection

The data collection component has two subcomponents like Collection of tweets with URLs and Crawling for URL redirections. To collect tweets with URLs and their context information from the twitter public timeline, this component uses twitter streaming APIs. Whenever this component obtains a tweet with a URL, it executes a crawling thread that follows all redirections of the URL and looks up the corresponding IP addresses. The crawling thread appends these retrieved URL and IP chains to the tweet information and pushes it into a tweet queue. As we have seen, our crawler cannot reach malicious landing URLs when they use conditional redirections to evade crawlers. However, because our detection system does not rely on the features of landing URLs, it works independently of such crawler evasions. This is given in the following Fig.2.

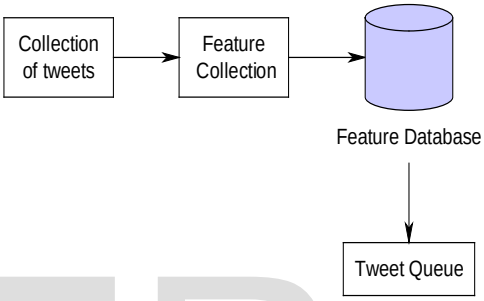

Fig.2. Data collection

5.2 Feature Extraction

The feature extraction component has three subcomponents such as Grouping of identical domains , Finding entry point URLs and Extracting feature vectors. This component monitors the tweet queue to determine whether a sufficient number of tweets have been collected. Specifically, our system uses a tweet window instead of individual tweets.

When more than w tweets are collected (w is 10,000 in the current implementation), it pops w tweets from the tweet queue. It is represented in Fig.3.

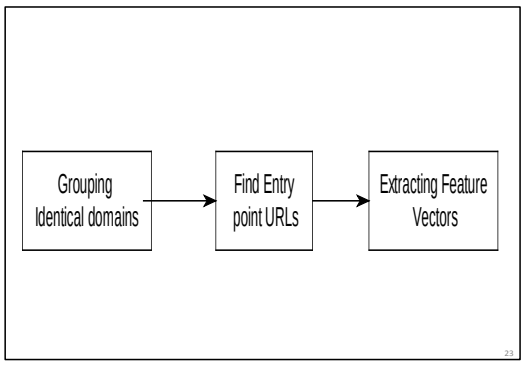

Fig.3. Feature Extraction

First, for all URLs in the w tweets, this component checks whether they share the same IP addresses. If several

URLs share at least one IP address, it replaces their domain names with a list of domains with which they are grouped. This grouping process enables the detection of suspicious URLs that use several domain names to bypass the blacklisting.

5.3.Training and Classification

The two subcomponents of training component are retrieval of account statuses and training of the classifier.Because we use an offline supervised learning algorithm, the feature vectors for training are relatively older than feature vectors for classification. To label the training vectors, we use the twitter account status; URLs from suspended accounts are considered malicious, whereas URLs from active accounts are considered benign.

We periodically update our classifier using labeled training vectors. The classification component executes our classifier using input feature vectors to classify suspicious URLs. When the classifier returns a number of malicious feature vectors, this component flags the corresponding URLs and their tweet information as suspicious. These URLs, detected as suspicious, will be delivered to security experts or more sophisticated dynamic analysis environments for an in-depth investigation. It is depicted in Fig.4.

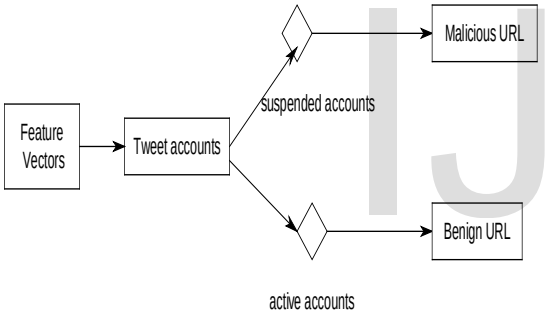

Fig.4. Training and Classification

6. DATA FLOW DIAGRAM

Data-flow diagrams (DFDs) were introduced and popularized for structured analysis and design. DFDs show the flow of data from external entities into the system, shows how the data moved from one process to another, as well as its logical storage. The data flow diagram is represented in Fig.5.

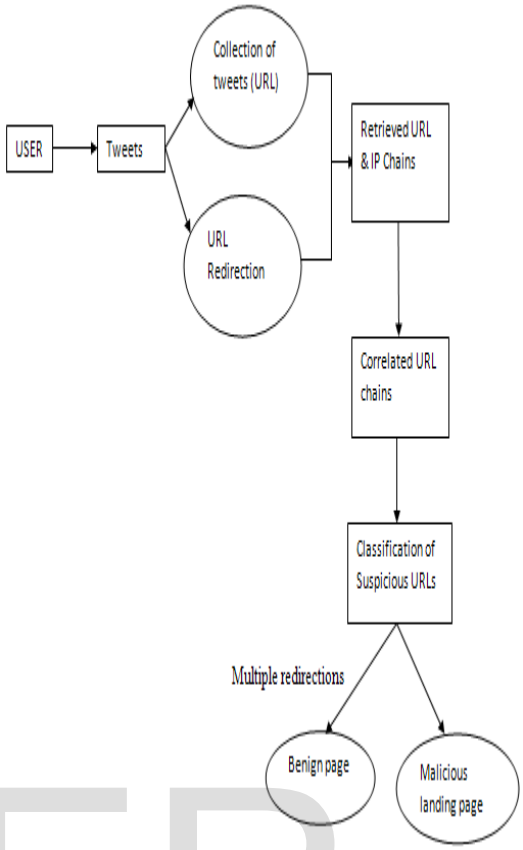

Fig.5. Data-flow diagram

7. EXPERIMENTAL SETUP

Some lightweight static detection systems focus on the lexical features of a URL such as its length, the number of dots, or each token it has [8], and also consider underlying DNS and WHOIS information [2], [19]. Therefore, we need dynamic detection systems [18], that use virtual machines and instrumented web browsers for in depth analysis of suspicious URLs.

Our goal is to develop a suspicious URL detection system for twitter that is robust enough to protect against conditional redirections. Consider a simple example of conditional redirections, in which an attacker creates a long URL redirect chain using a public URL shortening service. When a user or a crawler visits the initial URL, he or she will be redirected to an entry point of the intermediate URLs that are associated with private redirection servers. Some of these redirection servers check whether the current visitor is a normal browser or a crawler. If the current visitor seems to be a normal browser, the servers redirect the visitor to a malicious landing page. If not, they will redirect the visitor to a benign landing page. Therefore, the attacker can selectively attack normal users while deceiving

investigators. For implementation Java is used in the system which posses Pentium Iv processor with 512 MB RAM. It is depicted in Fig.6.

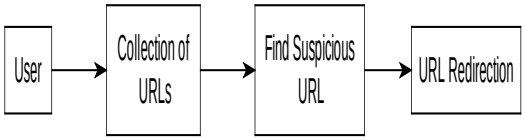

Fig.6. Suspicious URL detection

8. CONCLUSION

when protecting against conditional redirection, because it does not rely on the features of malicious landing pages that may not be reachable. Instead, it focuses on the correlations of multiple redirect chains that share the same redirection servers. We introduced new features on the basis of these correlations, implemented a near real-time classification system using these features, and evaluated the system's accuracy and performance. The evaluation results show that our system is highly accurate and can be deployed as a near real-time system to classify large samples of tweets from the twitter public timeline.

9. FUTURE ENHANCEMENT

For the spammers classified as legitimate users, we observed that most of the users use their account as a legitimate user and act as a spammer only for some of its tweet responses. Most of these users have tweet responses that are not spam, tricking the classifier in some attributes. In the future, we will extend our system to address dynamic and multiple redirections. We will also implement a distributed version of the proposed system to process all tweets from the twitter public timeline.

We propose an effective suspicious URL detection system for twitter. Instead of investigating the landing pages of individual URLs in each tweet, which may not be successfully fetched, we considered correlations of URL redirect chains extracted from a number of tweets. Conventional suspicious URL detection systems are ineffective in their protection against conditional redirection servers that distinguish investigators from normal browsers and redirect them to benign pages to cloak malicious landing pages. Unlike the conventional systems, this system is robust

REFERENCES

[1] D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: A Fast Filter for the Large-Scale Detection of Malicious Web Pages," Proc.20th Int'l World Wide Web Conf. (WWW), 2011.

[2] J. Ma, L.K. Saul, S. Savage, and G.M. Voelker, "Beyond Blacklists : Learning to Detect Malicious Web Sites from Suspicious URLs," Proc. 15th ACM SIGKDD Conf. Knowledge Discovery and Data Mining (KDD), 2009.

[3] C. Yang, R. Harkreader, and G. Gu, "Die Free or Live Hard?Empirical Evaluation and New Design for Fighting Evolving Twitter Spammers," Proc. 14th Int'l Symp. Recent Advances in Intrusion Detection (RAID), 2011.

[4] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida,"Detecting Spammers on Twitter," Proc. Seventh Collaboration,Electronic Messaging, Anti-Abuse and Spam Conf. (CEAS), 2010.

[5] K. Lee, J. Caverlee, and S. Webb, "Uncovering Social Spammers:Social Honeypots þ Machine Learning," Proc. 33rd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, 2010.

[6] S. Chhabra, A. Aggarwal, F. Benevenuto, and P. Kumaraguru,"Phi.sh/$oCiaL: the Phishing Landscape through Short URLs," Proc. Eighth Ann. Collaboration, Electronic Messaging, Anti-Abuse and Spam Conf. (CEAS), 2011.

[7] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting Spammers on Social Networks," Proc. 26th Ann. Computer Security Applications Conf. (ACSAC), 2010.

[8] D.K. McGrath and M. Gupta, "Behind Phishing: An Examination of Phisher Modi Operandi," Proc. First USENIX Workshop Large-Scale Exploits and Emergent Threats (LEET), 2008.

[9] H. Kwak, C. Lee, H. Park, and S. Moon, "What Is Twitter, a Social Network or a News Media?" Proc. 19th Int'l World Wide Web Conf.(WWW), 2010.

[10] S. Lee and J. Kim, "WarningBird: Detecting Suspicious URLs in Twitter Stream," Proc. 19th Network and Distributed System Security Symp. (NDSS), 2012.

[11] D. Antoniades, I. Polakis, G. Kontaxis, E. Athanasopoulos, S.Ioannidis, E.P. Markatos, and T. Karagiannis, "we.b: The Web of Short URLs," Proc. 20th Int'l World Wide Web Conf. (WWW), 2011.

[12] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Who Is Tweeting on Twitter: Human, Bot, or Cyborg?" Proc. 26th Ann.Computer Security Applications Conf. (ACSAC), 2010.

[13] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@spam: The Underground on 140 Characters or Less," Proc. 17th ACM Conf. Computer and Comm. Security (CCS), 2010.

[14] F. Klien and M. Strohmaier, "Short Links under Attack: Geographical Analysis of Spam in a URL Shortener Network,"Proc. 23rd ACM Conf. Hypertext and Social Media (HT), 2012.

[15] A. Wang, "Don't Follow Me: Spam Detecting in Twitter," Proc.Int'l Conf. Security and Cryptography (SECRYPT), 2010.

[16] J. Song, S. Lee, and J. Kim, "Spam Filtering in Twitter Using Sender-Receiver Relationship," proc. 14th Int'l Symp. Recent Advances in Intrusion Detection (RAID), 2011.

[17] H. Gao, Y. Chen, K. Lee, D. Palsetia, and A. Choudhary, "Towards Online Spam Filtering in Social Networks," Proc. 19th Network and Distributed System Security Symp. (NDSS), 2012.IEEE Transactions on dependable and secure computing, Vol 10, No. 3, May/June 2013.

[18] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and Evaluation of a Real-Time URL Spam Filtering Service," Proc.IEEE Symp. Security and Privacy (S&P), 2011.

[19] J. Ma, L.K. Saul, S. Savage, and G.M. Voelker, "Identifying Suspicious URLs: An Application of Large-Scale Online Learning,"Proc. 26th Int'l Conf. Machine Learning (ICML), 2009.

[20] Sangho Lee and Jong Kim," WarningBird: A Near Real-Time Detection System for Suspicious URLs in Twitter Stream", IEEE Transactions on dependable and secure computing, Vol 10, No. 3, May/June 2013.

[21] Y.-M. Wang, D. Beck, X. Jiang, R. Roussev, C. Verbowski, S. Chen,and S. King, "AutomatedWeb Patrol with Strider HoneyMonkeys: Finding Web Sites That Exploit Browser Vulnerabilities," Proc. 13th network and Distributed System Security Symp. (NDSS), 2006.

[22] M. Cova, C. Kruegel, and G. Vigna, "Detection and Analysis of Drive-by-Download Attacks and Malicious JavaScript Code," Proc. 19th Int'l World Wide Web Conf. (WWW), 2010.

[23] Gianluca Stringhini ,Christopher Kruegel and Giovanni Vigna,"Detecting spammers on social networks",Proceedings of the 26th Annual Computer Security Applications Conference,Pages 1-9,ACM-2010.

[24] Abdulrahman M.Al-Senaidy,Tauseef Ahmad, Mohd Mudasir Shafi,Privacy and Security Concerns in SNS: A Saudi Arabian Users Point of View,International Journal of Computer Applications (0975 – 8887),Volume 49– No.14, July 2012.

[25] Aggarwal CC. (Ed.) Social network analytics data, Springer, 2011.

[26] Rashmi A.Zilpelwar, Rajneeshkaur K.Bedi, Vijay M.Wadhai,An Overview of Privacy and Security in SNS, International Journal of P2P Network Trends and Technology- Volume2Issue1- 2012

[27] Justin Ma ,Lawrence K. Saul ,Stefan Savage and Geoffrey M. Voelker," Identifying Suspicious URLs: An Application of Large-Scale Online Learning", Appearing in Proceedings of the 26th International Conference on Machine Learning, Montreal, Canada, 2009.

[28] Zhiyuan Cheng, James Caverlee, Kyumin Lee and Daniel Z. Sui," Exploring Millions of Footprints in Location Sharing Services", Association for the Advancement of Artificial Intelligence,2011.