*Future Internet* **2015**, *7*, 67-93; doi:10.3390/fi7010067

OPEN ACCESS

*future internet*

**ISSN 1999-5903**

www.mdpi.com/journal/futureinternet

*Article*

# Social Networking Privacy—Who's Stalking You? †

**Diane Gan ‡,* and Lily R. Jenkins ‡**

CSAFE Centre, Department of Computing & Information Systems, Faculty of Architecture, Computing and Humanities, University of Greenwich, London SE10 9LS, UK;
E-Mail: lilyrjenkins@gmail.com

† First presented at CFET 2014—7th International Conference on Cybercrime, Forensics, Education and Training, Christ Church Canterbury, UK, 10–11 July 2014, ISBN:97801909067158.

‡ These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: d.gan@gre.ac.uk; Tel.: +44-208-331-9708; Fax: +44-208-331-8665.

Academic Editor: Steven Furnell

*Received: 25 August 2014 / Accepted: 13 March 2015 / Published: 23 March 2015*

**Abstract:** This research investigates the privacy issues that exist on social networking sites. It is reasonable to assume that many Twitter users are unaware of the dangers of uploading a tweet to their timeline which can be seen by anyone. Enabling geo-location tagging on tweets can result in personal information leakage, which the user did not intend to be public and which can seriously affect that user's privacy and anonymity online. This research demonstrates that key information can easily be retrieved using the starting point of a single tweet with geo-location turned on. A series of experiments have been undertaken to determine how much information can be obtained about a particular individual using only social networking sites and freely available mining tools. The information gathered enabled the target subjects to be identified on other social networking sites such as Foursquare, Instagram, LinkedIn, Facebook and Google+, where more personal information was leaked. The tools used are discussed, the results of the experiments are presented and the privacy implications are examined.

**Keywords:** Twitter; Tweets; StreamdIn; Twitonomy; Creepy; geo-location; stalking; identity theft; data leakage; Facebook; Foursquare

## 1. Introduction

The social networking site Twitter first appeared on the Internet in March 2006 to provide users with a platform that they could use to communicate and connect with people around the world in as little as 140 characters. This is not limited to just friends and family, but also people you want to share with and connect to. Twitter encourages users to express themselves freely through the use of tweeting and hash tagging to "share what you see, feel and experience as it happens" [1]. This approach has enabled Twitter to become a central hub of real time news with breaking stories often before the media have picked them up, as well as finding out what your friends are doing. The London Riots were an example of this [2]. Twitter was used by the Police as a tool to track the rioters to try to prevent riot hotspots and other crimes that were occurring and also to keep the public informed [3].

Social media, whether it be for leisure or business has become a part of our everyday lives with Twitter and Facebook leading the way, with 645,750,000 active registered Twitter users on 11th July 2014 [4]. People tweet on their way to work and when they get home without realizing that they are also giving away information about where they are, where they live and when they are not at home to anyone with access to the Internet.

Twitter added the geo-location function to user profiles in 2009 [5] enabling followers to know exactly where an individual was tweeting from. It was only a matter of time before privacy issues regarding this information started to emerge. With geo-location now an integral part of Twitter, it is still unclear as to whether users know how to use this feature or how to protect themselves. There has been very little research to establish how many Twitter users do not realize that features such as their location and their identity are turned on by default. As most teenagers today have at least one profile on a number of social networking sites, such as Facebook or Twitter, they inadvertently reveal a lot of personal information about themselves that anyone can see and use. They do not appreciate the risks and so would not be aware that they are also exposing their private information [6]. From anecdotal observation it would appear that many users have the same picture (avatar) of themselves and the same username or "handle" on all of their social media sites. This makes it even easier to identify them and follow them through cyber space to harvest more personal information about them [7]. It could be said that anyone using Twitter would not necessarily realize that they could be tracked to their home simply because they uploaded a tweet or a picture to a social networking site. Furthermore, users are actively encouraged to enter a lot of personal information into social networking sites, so that this information can be shared with others [8]. This includes date and place of birth, what school they went to, name address and gender, which is all useful information for attackers. The dangers of this are identified in [9] whereby attackers could gain access to someone's Twitter account to spread malicious code or they could steal enough personal information about a user to assume their identity. This could even include gaining access to their credit card, bank details and passwords. The user may also unintentionally reveal sensitive personal information within the contents of their tweets [9,10]. Figure 1 demonstrates this, as it shows a picture of someone's credit card posted on Twitter. Anyone viewing the original post can see all the card details including the 16 digit number across the front, the person's name, the expiry date and in the case of this picture the three digit code on the back of the card is given away in the tweet. This is one of 151 million results returned from a search on Twitter for credit card pictures.


**Figure 1.** Credit Card Picture and Tweet Posted on Twitter.

Being aware of how much social media sites such as Twitter impact on our daily lives, many users do not realize that businesses also use social media to reach out to clients to promote their brands and even to their own employees. What is posted onto social media sites could be included in how potential employers check a person's personal life prior to employing them. A classic example of this was the newly appointed Kent Police Commissioner's Youth Advisor Paris Brown who was forced to withdraw when her historical twitter content was made public [11]. Education of school children as to the dangers of social media in general and of Twitter in particular is essential and this work is being presented in local schools to warn children of the dangers of twitting.

The aim of this work was to determine how much information was leaked by three test subjects, who used Twitter normally and had geo-location turned on, which would enable us to track them through cyber space using only social networks. A preliminary study was undertaken to determine feasibility and the results of this are presented. Using publicly available data from their social networking sites we show that it is possible to harvest private information. We then identified three freely available applications for mining Twitter data from the large range of available online tools. The starting point for the investigations for each subject was a picture uploaded to their Twitter account which had geo-location information embedded within it. Having identified personal information and patterns of behavior the three test subjects were then tracked to other social networks, which not only verified the information already gathered using Twitter, but led to more information being leaked. These three test subjects are indicative of typical Twitter user behaviors and they were selected as suitable candidates on this basis.

The rest of this paper is laid out as follows. The privacy issues and related work are presented. An overview of the tools used is given in Section 2, the experiments are discussed and the results are presented in Section 3. Section 4 introduces a set of guidelines for Twitter users to help them to protect their personal information and the conclusion is given in Section 5.

### *1.1. The Privacy Debate*

In 2011 a number of high profile Twitter accounts were taken over by hackers and used to spread false information. This resulted in calls for Twitter to improve their security by making HTTPS standard for all users [12]. In response, Twitter announced that it was the user's responsibility to secure their own passwords and to manually select the HTTPS setting themselves. In May 2014, Twitter announced that it was improving the security of the password reset function. By providing a username, email or phone number, instructions can be sent out to the user to reset their password [13]. They also added a login history so that they could identify false login attempts, for example if the user appears to be on the other side of the world trying to log in. This is still not as secure as many would like [13]. Just how much online privacy protection should be provided by the sites themselves? It is the personal responsibility of the user to monitor what information is uploaded and shown. The ease of use goal of all social networking sites can overshadow their goals for privacy and security for the users [14]. There is a risk that users are unaware of this, as many prefer a social networking site that is easy to use and it is this naivety that enables personal information leakage to occur.

Twitter has an extensive privacy policy that a new user must agree to. At the time that this work was completed the policy clearly stated that all user profiles and subsequent tweets were by default public. Within the policy Twitter details how your information is used through their services like applications, websites and third parties. The use of many services enable Twitter to "collect, transfer, manipulate, store and disclose" [15] the information provided by their users. Any information provided can be used by Twitter or third party services to help identify users and it is the privacy settings that allows users to select how and if their information is used. Users have the option to include location information with their tweets which will be stored to provide "features" for services. Having location information in tweets allows the services of Twitter to customize and improve by providing the user with a "more relevant content like local trends, stories, ads and suggestions for people to follow" [15]. Hidden away in the location support page are details about how to use this setting effectively to get the best results. It also raises a caution to all users to be careful about the information they share and to keep locations such as home addresses private by not tagging them using the location setting [7]. There is a risk that the majority of users do not read this.

Most social networking sites, including Twitter, use the opt out approach for many of their features which means that once someone is signed up, it is their responsibility to check the settings if they wish to protect their privacy, as a number are on by default. Using this approach it would be expected that these sites would use the opt-in approach and allow users to willingly and knowingly turn them on, so that they can acknowledge they might encounter a lack of privacy as a result [14]. Many people would assume this would apply to their online profiles as well but only half of social media users utilize the security settings and make their profile private [16]. Unfortunately, most of the information is made visible to the public by default which opens the door to identity theft and cyber stalking. On the other hand, users making an update of their whereabouts or posting any sensitive personal data provide opportunities for criminals. For example, when someone posts "out for dinner", "home sweet home", "at gram's", "away for the summer", "will be on the next flight to Paris", or "just got admitted to St. Paul's hospital" they have given away significant information to criminals such as burglars and stalkers.

The message about the dangers of social networking profiles that are viewable by anyone is not wide spread as 66% of Facebook users had no knowledge regarding privacy settings and 47% of teenagers had public Facebook profiles [17]. According to a survey conducted by Welter (2013), the starting point for as many as 81% of Internet related crimes is a social networking site [18]. Crime Wire have reported that 39% of social network users have been victims of profile hacking. Criminals are aware of the information leakage potential surrounding social network sites, with 78% of burglars having used Facebook, Twitter, Foursquare and Google Street-view to select their victims [18]. As many as 54% of burglars were alerted to empty homes because people posted their whereabouts and their statuses on these sites. 50% of child sex offenders admitted obtaining information about the victim from their social networking profile. Surprisingly, 38% of Facebook users were identified as being under the age of 13 in 2010 [17], in spite of Facebook's policy which requires that to be eligible to create an account all uses be over 13 years of age [15]. 85% of parents surveyed with children aged 13 to 17 said that their child used social networking sites. Statistically it has been identified that teenagers are more likely than adults to include their real age (50%), their photo (62%), their home town (41%), the name of their school (45%), videos of themselves (14%), their phone number (14%) and their exact location (9%) [17].

In 2014 the setting "Protect my tweets" is still off by default, so that anyone searching the Internet can identify an individual's tweets. There are also a number of other settings which are still on by default, such as Photo Tagging and **"**Let others find me by my email address" and the author of this article [19] could not think of any reason why someone would want to have this feature enabled. However, the Tweet location setting has now been turned to "off" by default. Neagu (2014) proposed that users need to defend their Twitter privacy by ensuring that they use strong passwords and by not posting information which discloses a current location, so there are clearly still issues with Twitter's security policy [20].

### *1.2. Geo-Location Tagging*

Geo-location tagging in social media over the past five years has revolutionized how Twitter users share information with their followers including their exact location using latitude and longitude coordinates which tells anyone exactly where the tweeter is located. Geo-location information can be gathered in many different ways, with photographs providing the most data. In almost all newer cameras geo-location data is collected automatically regarding where a photo was taken and then stored within the metadata EXIF (Exchangeable Image File Format) records of the image [21].

All smartphones have geo-tagging functionality built into them. Just like the social media site Twitter, smartphones have to specifically be given permission by the user to access their current GPS location [22]. With over three quarters of users using their mobile devices to access Twitter and many other applications, the geo-location settings on smartphones are easily turned on and forgotten about which gives rise to a lot of information leakage and it is this automated service that gives the most cause for concern regarding the disclosure of private information [23]. Once uploaded to Twitter anyone can pin point exactly where that photograph was taken. It is the automated service that seems to be highlighting the most concerns for the exposure of private information, even though it is driven by the end users, who may be unaware that this can result in unexpected disclosures [23]. Although geo-location tagging has many privacy issues it is still marketed "as a useful way to stay connected to others in real time" [24].

### *1.3. Related Literature*

Many applications that use Twitter's own APIs (Application Programming Interface) have come under scrutiny because of privacy issues. An example of this is the location application "Girls Around Me". This API was very controversial and particularly worrying. A user running the application was presented with a map showing the profile pictures of anyone (male or female) in the same area who had recently checked in with Foursquare and who had a publicly visible Facebook profile. Foursquare is an application which allows users to check in wherever they go so that their friends can find them. The user could then select to view all females or all males without their consent. The people identified by this application were unaware that they were being targeted. The application took data from Foursquare to plot the locations of the people and the photographs displayed were from their Facebook accounts. The pictures could then be used by anyone to gain access to that person's Facebook details [25]. This application has since been removed from app stores due to the complaints received. However, there is no evidence to suggest that this application was breaking Twitter's API terms of use.

There are a number of papers which explore the use of Twitter for teaching in very diverse fields, from engineering and medicine in higher education to secondary schools to enhance the student experience [26–31]. Pentina *et al.* compared Twitter users in America with their counterparts in the Ukraine. The focus here was comparing how following well-known brands in these two cultures affected trust [32]. Twitter has also become a reporting tool for journalists so that they can report breaking news, as demonstrated in two papers. The first paper discusses how national and international journalists used Twitter to share information and images during the four days of riots in the summer of 2011 [33]. The second paper describes how the journalist Andy Carvin used Twitter to post messages and links to images supplied by demonstrators during the 2011 Tunisian and Egyptian uprisings [34]. These are examples of how Twitter enables news to propagate around the world as it happens. Using Twitter to aid the Police in preventing and detecting crime is also a common theme and there is ongoing research into the use of Twitter to help to detect crime. Malleson *et al.* conducted a study in Leeds, England to investigate violent crime and to determine the main population at risk. Their study is based on "crowd-sourced" data to identify crime hot spots [35].

The approach taken by Gabielkov *et al.* (2014) was to study the structure of the Twitter social graph to investigate the way that tweets propagate. They used a very large sample retrieved from Twitter to identify different Twitter feeds [36]. Their paper demonstrated that it was possible to determine the location that a user was tweeting from when there was no available geo-location data. Using the language content of a person's tweets, the location of the tweeting device and the time zone, an accurate location can be identified and the challenges of this approach are discussed [37]. In our work we are using the geo-location data as the starting point for our investigation.

There are also a large number of papers on the use of Twitter for medical reporting and also for political campaigning in Australia but these are outside of the scope of this work.

The paper by Jeong and Coyle investigates the perceptions of young peoples' concerns about privacy on Twitter and Facebook. Their study identified that the greatest concern was that people in authority, parents, teachers *etc.* might gain access to their information, rather than complete strangers [38]. The strategy used by the "42 middle school students" [39] in another survey was to use false information on their profiles. This small group were aware of issues in online privacy but they had been selected as

being the most "digitally engaged students", which is not a representative sample of the whole cohort of students. They advocate teaching online privacy in schools in their conclusion [39].

Zhang *et al.* (2014) have undertaken a long term study of Twitter users' concerns using surveys and interviews with adult subjects. The results outlined the concerns that the users of Twitter and other social network sites had and how these had changed over the duration of this study. They also included the users perception of the impact of targeted advertising related to what the user was posting at that time. The subjects of this study were aware that their data and posts were being mined and used for directed advertising. The term "creepy" in the title of the paper [40] relates to the perceptions of the people interviewed to the targeted advertising using Twitter and not the application used in our work.

Mao *et al.* based their work on the analysis of leaks in Twitter on a very large sample obtained over a nine month period [41]. Their research was to identify privacy leaks related to three distinct categories of tweets, namely vacation plans, tweeting while being drunk and divulging medical conditions. They identified that people were tweeting on the assumption that this information was private as it was shared only with friends and family, which they classified as primary leaks. However the information leakage occurred during retweeting, which they classified as secondary leaks. They also identified that 76% of alerts related to vacations were easily available to burglars [41].

Mearns *et al.* (2014) developed a framework to retrieve tweets which enabled researchers to investigate trends and topics using time stamps, geo-location data and real-time streaming. This is looking at information flows and patterns between Twitter users rather than at individual users [42]. Another paper presented a framework for determining a user's location to within 10 Km by using the textual content of their tweets [43]. Jin *et al.* (2013) undertook a survey into user behavior on social networks such as Facebook, Twitter, Google+, LinkedIn, and Foursquare [44]. They were using social graphs to identify relationships within the Twitter data based on friendships, followings, interaction between users and latent interaction, which is described as someone browsing a profile. They also included an investigation into traffic activity on these sites and the use of mobile devices for social networking sites. Their coverage of malicious behavior relates to spam and to attacks at the network level rather than individual misbehaving users, as in our work.

Retweeting is an important aspect for Twitter users, which is used to spread information amongst users. There is a body of research which has investigated the extent to which users retweet. Shi *et al.* [45] use a dataset harvested from Twitter which, on analysis, showed that retweeting displays a "chain effect" and that retweeting is extremely popular because it is easy to use within Twitter applications.

Analysis of retweeting behaviour and the factors that influence how retweets propagate, such as the user themselves, the content of the tweet and the timing factor determine whether a tweet is retweeted is presented [46]. They concluded that each user in their dataset retweeted in a seven day period, and that the majority do so at a low frequency, with the average number of retweets being 197. The 3% of users who retweeted more than 1,000 times in the same period were classified as retweet-aholics [46].

Tweet and retweet behaviors were gathered in a micro blogging web site located in China, which has similar functionality to Twitter. They collected a data sample of 1.7 million users with 4 billion flowing relationships. 300,000 micro blogs were extracted which consist of the original "tweet" and all of the associated "retweets". They determined that each "tweet" had been "retweeted" up to 80 times. They then modeled the "retweet" behavior, which demonstrated the same behavior is found on similar micro blogging sites as on more famous sites such as Twitter [47].

The speed and degree by which information is disseminated through the "Twittersphere" is modeled in a number of papers. Jiang *et al.* [48] used a dataset which comprised 1.7 million Twitter users. They identified 23.7 million retweets that originated from these users and categorized them according to their popularity. They defined four levels of popularity which ranged from "unpopular" to "highly popular" based on the number and the speed of diffusion of retweets within the data sample. Lee *et al.* [49] have developed two models to identify the content of peoples' tweets to identify how likely they would be to retweet and also the likelihood as to whether they would propagate information via retweets when asked to do so by a stranger. Pervin *et al.* [50] have modeled "retweetability" using their "Information Diffusion Impact" model which examines the roles of the Twitter users involved in retweeting. These are the "information starter", "amplifier", and "transmitter". They have validated the results from their model with a dataset from the retweeting that occurred directly after the Boston bombings and the Japanese earthquake. Lia *et al.* [51] have developed a framework to measure the way that dynamic information propagates using retweets. They compare a large scale dataset from Twitter with their experimental results to evaluate their framework. Modeling retweeting behaviors is presented in this paper [52] to predict how many times a tweet is retweeted. Liu *et al.* use a two-phase model to predict how many times a tweet will be retweeted. They have also used a dataset from Twitter to test the accuracy of their framework when compared to real data.

The results of several retweet behavior models which have been used to evaluate Twitter data are presented by Macskassy *et al.* [53]. Their dataset consisted of over 768,000 tweets which originated from 30,000 Twitter users and they concluded that there are a lot of factors that influence retweeting behavior and this complexity cannot be duplicated with a single model.

Current research into social networks is mainly focused on investigating the mining of information from Twitter to identify trends and themes, often using social graphs. Retweeting is part of normal behavior amongst Twitter users and follows a number of models, as the studies discussed above have shown. Some tweets are massively retweeted while others are only retweeted a few times. Our work differs from the approach taken by the majority of researchers in that we demonstrate the danger to an individual user of leaking even the smallest piece of information from social networks which could expose them to being stalked or even having their home burgled. The test subjects in this work were completely unaware of their exposure and they are all adults over 21 years of age. We believe our approach to this research area is novel.

## 2. Review of the Tools

There are many tools that facilitate the harvesting of Twitter information. These are used by organizations to target users for advertising. These tools can also be subverted to stalk tweeters and even to access personal information. With this in mind a number of tools have been identified which facilitate this investigation and three have been selected. These are streamd.in, Twitonomy and Creepy and a brief overview of each is presented here.

### *2.1. Streamd.in Application*

Streamd.in is a mobile application for both android and iOS that displays tweets on Google Maps by using the geo-location details attached to each tweet [54]. This application gives the option to search any

part of the world using Google maps or just casually browse around different areas of interest. Each tweet on the map is represented by the user's profile picture and these are grouped together by location to enable anyone to focus on one specific place and look through the numerous real-time tweets coming in. The application has a number of features, such as the update interval, which gives control over how often the application refreshes. It also allows filtering of tweets on the map by picture, user or specific keywords, which enables the user to narrow searches to avoid being overloaded with tweets that are of no interest [55]. One feature that was discovered while testing the application on an android phone was that when the GPS has found someone's exact location it will then track their movements. For example if they were travelling on a train then this application would follow them along the tracks. Figure 2. shows the streamd.in application in action, tracking the GPS signal of a random person using public transport. This demonstrates what happens when someone tweets while traveling. The phone that is being tracked along the DLR link is shown as the blue dots in Figure 2.

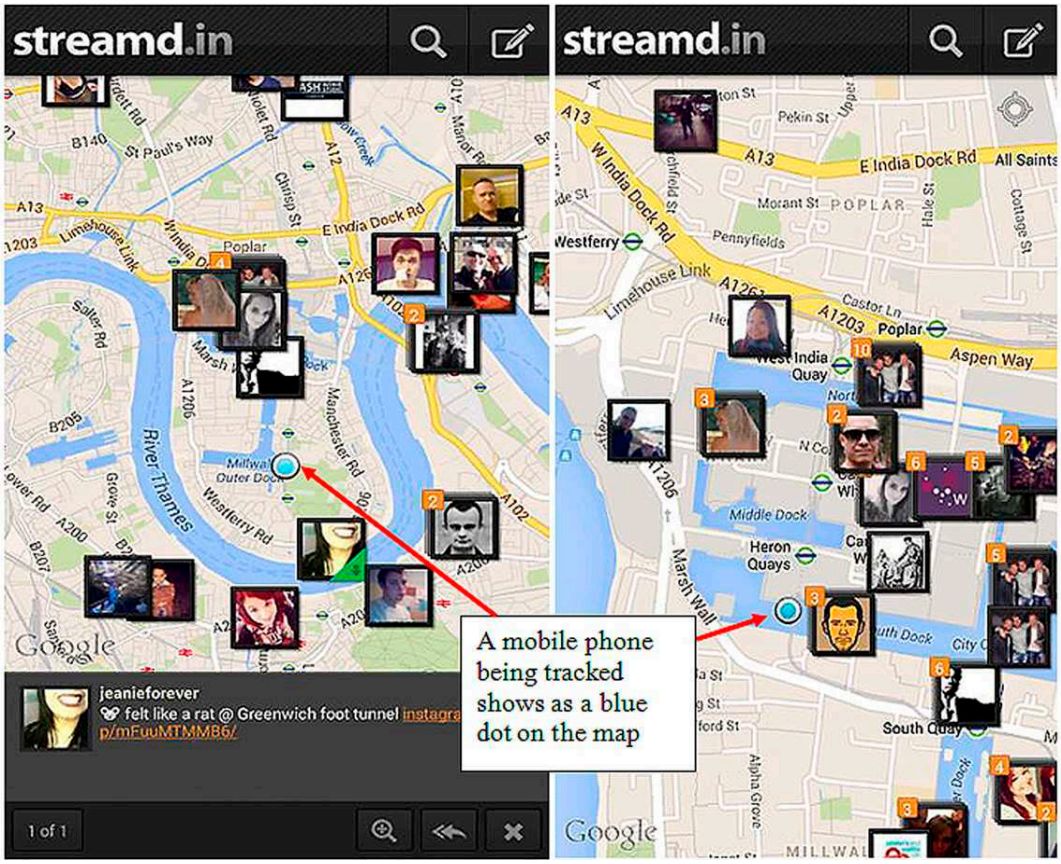


**Figure 2.** Tracking the Author's Mobile Phone Using Streamd.in.

Streamd.in does allow posting from within and the ability to view any user's timeline just by clicking on their profile picture. An indication of how many people were tweeting at the time that this test was taking place, can be seen by all the small individual pictures present in Figure 2.

The streamd.in application was found to be accurate and quick at retrieving the data when tested on the author's own Twitter account. With this application it was very easy to track someone's location just using the geo-location information on their tweets. Once a target user was identified the application could be filtered to show only their tweets. The result of this was an application that readily provides the information a cyber-stalker would find useful.

### *2.2. Twitonomy Application*

Twitonomy is a web based analytics tool that provides a vast amount of statistics on a user's Twitter profile. The user has to be registered with Twitter as the tool has to be authenticated with Twitter's API each time the search function is used. Without a subscription it only analyses the last four months' worth of tweets. With a subscription tweets can be downloaded from a specific user in excel or pdf format, the date range of the tweets can be customized and a followers report can be obtained which gives location details (if the location setting is enabled on the user's account) making it possible to find out which follower is geographically closest.

Twitonomy [56] presents the user with a number of tables and graphs offering a range of statistics which give valuable insight to anyone wanting to understand someone's tweeting habits, the "who's and what" that gets them engaged in the social media site.

An available feature is the mentions map which visually shows where in the world the most mentions are coming from and who by. On first inspection of the mentions map it was thought that it used the geo-location setting on Twitter to gather this data and plot the tweets. However, while testing reliability it was found that it actually uses the location given in the user's profile, so if they were tweeting from the other side of the world but it would appear as if they were still in London, for example.

One clear pattern that emerges when looking at the information gathered is the time of day that a Twitter user tweets the most. This can give an indication of their routine such as when that user is at work, by analyzing whether they have reduced their tweeting significantly during typical working hours, e.g., 9 to 5, as seen in Figure 3.

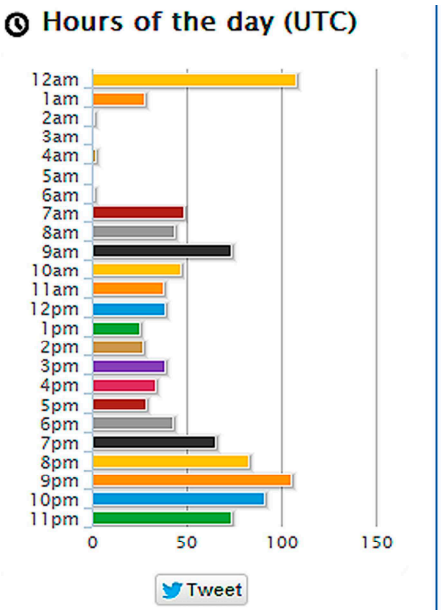


**Figure 3.** A Twitonomy User's Profile Showing Tweeting Patterns.

### *2.3. Creepy Application*

Creepy is a social media aggregation program that gathers geo-location information from the social network platforms Twitter, Instagram and Flickr. Use of Creepy requires the user to authenticate their account with each social networking site supported, which means that the user must create an account

with those sites. Once the user has authenticated themselves with Twitter and Instagram *etc.* they can add Twitter users to a target list. Other users can also be added and their geo-location data can be retrieved, regardless of whether that user follows them or even if they do not have public accounts.

The program is freely available online to download for Linux and Windows based systems [57]. Creepy was chosen for this investigation as it provides all the geographical information needed to target a user on Google maps. This enabled us to identify where each subject works, lives and if there are particular routes that they take each day, just by using their geo-location data and tweet contents. The collected data is presented on a Google map with pins representing the geo-location information attached to each tweet posted.

One of the panels on the tool is called "Current Location Details" and shows the social media platform that the pin came from, the time and date, location of the tweet and the context of the tweet, see Figure 4. Using this feature it is possible to locate an individual user by double-clicking on that user to identify their current location on the map [58]. However, Creepy did not produce real time tweets so each time it was loaded the locations had to be reanalyzed. Although Creepy as a standalone program offers very useful results for anyone wanting to track someone's movements and understand their routines when used in conjunction with the other programs previously mentioned improved results can be obtained.

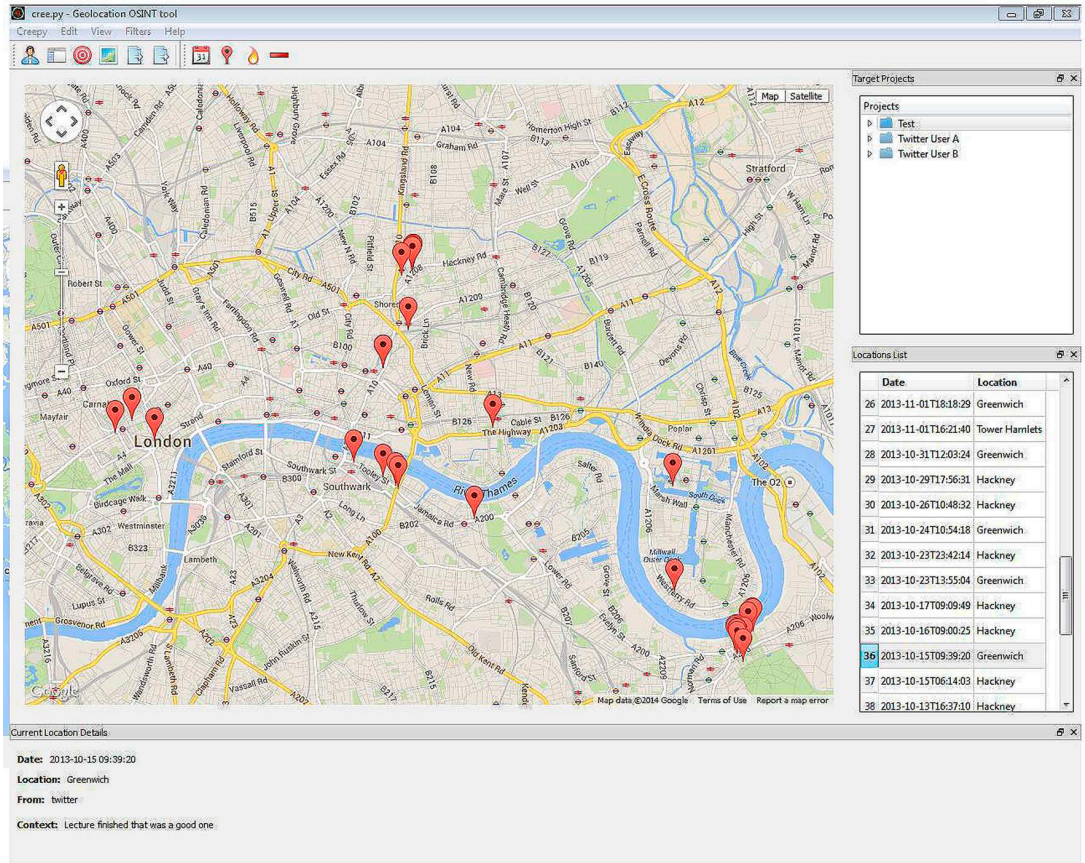


**Figure 4.** Creepy Identifying the Author's Locations.

## 3. The Experiments

A series of experiments have been conducted for this work. Initially, we have undertaken a large scale sample to test feasibility and this is reported in section 3.1. This was followed by four small scale

experiments where we selected three typical Twitter users, who agreed to participate and gave their consent for their tweets to be harvested in order to determine how much information could be gathered about them using only social networks. The information gathered in this work was obtained entirely from their social networking sites, starting with a picture that each had previously uploaded to Twitter with the geo-location setting turned on, prior to the start of this investigation. For these investigations the participants identities will be protected and they will be referred to as User A, User B and User C. The users were asked to continue to tweet normally and to keep their geo-location settings turned on for the duration of this investigation. Using geo-location tagging for tracking each individual's tweets and using their Twitter profiles and other social media, a lot of information about each of the volunteers was gathered. The objective was to see how much information could be found about each of these users, such as where they lived, where they worked, any routes that they regularly used, any other patterns of behavior and other information using social media and the results are presented.

### *3.1. Preliminary Study*

A preliminary study has been undertaken to determine the extent to which users upload inappropriate and revealing information on social networking sites. A random sample of 90 Twitter users has shown that 65.5% revealed their real names and the majority of them (94.4%) had posted pictures of themselves, see Table 1. Of these users 43.3% had geo-location turned on in their posts, which means that it would be possible to track them to their home or to their work place. 43% had links to other social networking sites, which meant it would be much easier to track these users through cyber space. Three quarters of these users had over 500 followers and were also following over 500 others. Followers are people who tend to retweet tweets from those that they follow.

**Table 1.** Data from Twitter (90 random users).

| Vulnerabilities | Number of Twitter Users from Sample |
|---|---|
| Displayed real name | 59 users (65.5%) |
| Real name as username | 33 users (36.7%) |
| Real name and geotag posts | 47 users (43.3%) |
| Shared self-images | 85 users (94.4%) |
| Specified link to other site | 39 users (43.3%) |
| Same username on multiple platforms | 30 users (33.3%) |
| Above five hundred followers | 61 users (67.8%) |
| Above five hundred following | 54 users (60%) |

There was not a huge difference when it came to gender within the sample, with 59% being female and 41% being male. We also investigated the age range of these users and found that the biggest group were in the 21 to 25 year age range, being 39% of the sample. Teenagers, aged 15 to 20 formed 27% of the sample, while the 26 to 30 year olds accounted for only 20%. The age group 31 to 35 was 7% of the sample and the 41 to 45 year olds were only 5%. There were 2% of the sample whose age could not be determined. Within our sample it would appear that the 21 to 25 year old group are the biggest users of Twitter and this is actually the age range that our three users fell within.

A survey completed using Instagram is shown in Table 2. Over 7 million users had publically identified their locations. Other personal and private information easily visible included pictures of their credit/debit cards with the 16 digit number clearly visible and various types of tickets with names, addresses and dates showing. Over 39,000 users posted a picture of their hospital band with their details showing their name, DoB and hospital ID.

**Table 2.** Data from Instagram.

| Type of information leaked | Number of users |
|---|---|
| Whereabouts information | 7,000,000 + users |
| Credit/Debit Card photos | 1,186 users |
| Flight Tickets | 2,514 users |
| Other Tickets with sensitive data | 1,808,801 |
| Hospital ID band with full name | Approximately 39,503 users |
| Using hashtag #Homesweethome | Approximately 4 million posts |
| Using hashtag #Offtowork | Approximately 350,000 posts |

### *3.2. Experiment 1—Streamd.in*

Using streamd.in the results were filtered to display each of the subject's tweets on the map in turn. The individual results from both User A and User B showed no tweets on the streamd.in map in the test period which meant that no information could be gathered about these users.

However, User C tweeted more frequently and their timeline on streamd.in showed their profile picture in many locations. Figure 5 shows the route that User C travelled along a road (shown in yellow). From the content of the tweets it can be seen that they went to get something to eat. This demonstrates just how easy it is to locate and track someone through their tweets. It would also be possible to follow them home if they continued to tweet.

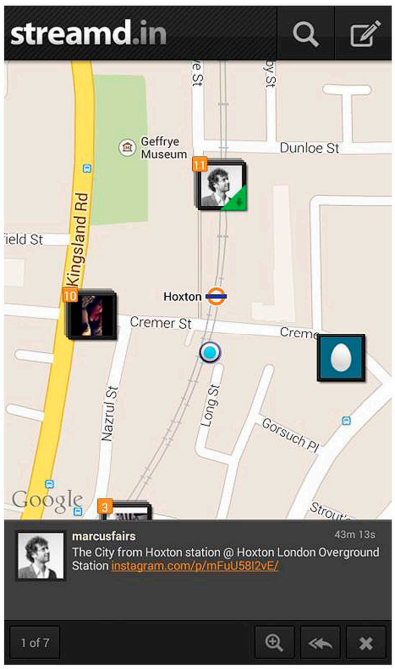


**Figure 5.** User C's Movements Shown in Yellow on the Map.

### *3.3. Experiment 2—Twitonomy*

User A's profile was entered into the Twitonomy search engine and in approximately thirty seconds it had analyzed all of their tweets posted during the last four months. After evaluating the data and graphs it was concluded that these results offered no indication as to where User A worked or lived but did offer information about when they tweet. User A mostly uses Twitter to re-tweet or reply rather than posting their own statuses. They are most active during the winter months. It was also possible to identify the days of the week User A tweets the most, as well as the time of day. From these graphs there was no indication whether User A had a job, as it was expected to see a significant drop in tweets during the traditional working hours of 9 to 5 but this pattern was not found, see Figure 6.

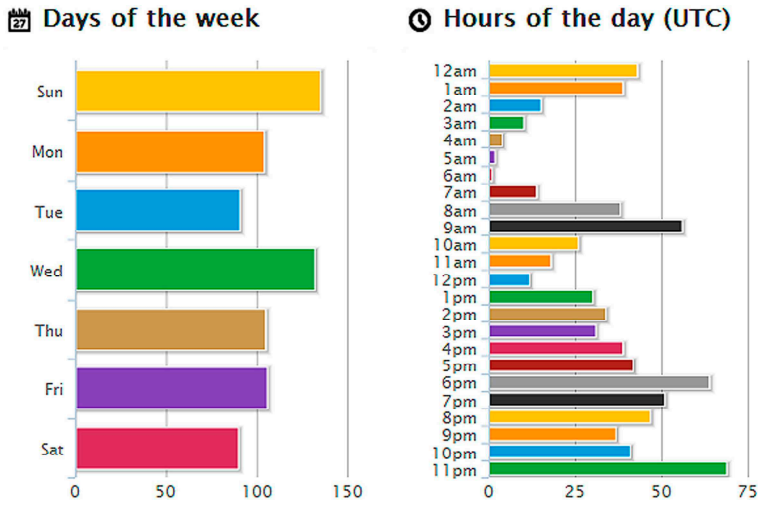


**Figure 6.** User A's Tweeting Average Number of Tweets.

The analysis of User B's profile using Twitonomy did not reveal any indication of where User B worked or lived but did offer information about when they tweet. Like User A, they use Twitter more to re-tweet and reply then actually tweeting their own statuses. It also showed sixteen links to other social networking sites.

Using the information identified by streamd.in, User C's profile name was entered into the search engine of Twitonomy. User C had a very distinctive pattern of usage related to the days of the week. There was a drop in the number of tweets sent during the week and the tweets increased at the weekends, which suggested that User C had a Monday to Friday job. The second pattern noticed was that the hours of the day that User C tweets most often are outside of the traditional working hours of 9 to 5. The last pattern worth noting was the platforms that User C uses. As seen in see Figure 7, a combination of the Twitter website and Foursquare are the most used, with 1656 and 1226 tweets respectively. The combination of the two social platforms indicates that this user checks in with Foursquare wherever they go, which actually gives even more details about the places User C goes to regularly and even where they travel from. Also it can be clearly seen that this person has an iPhone as they use the "Twitter for iPhone" application, circled in Figure 7.

Twitonomy was the least successful tool for the inappropriate tracking of User A. For User B it was able to identify their working hours, which could be used to wait for them at the end of their working day once their work place had been identified in order to orchestrate an "accidental" meeting. However,

Twitonomy was most successful with User C, as it was able to identify several distinct patterns of behavior, which would improve a stalker's opportunities for following them.


**Figure 7.** Tweeting Platforms Used by User C.

### *3.4. Experiment 3—Creepy*

Using the tool Creepy, User A's profile showed hundreds of location pins on the map each representing a tweet with associated geo-location data, see Figure 8. The dense clusters of tweets show locations where User A tweets most frequently from. It also shows individual tweets which indicate journeys between the clusters. To identify where User A lives and works the tweet contents and the geo-location data were analyzed to identify each location. If the clusters of pins did not provide any results then the single pins were individually analyzed for textual content to determine any missing information.

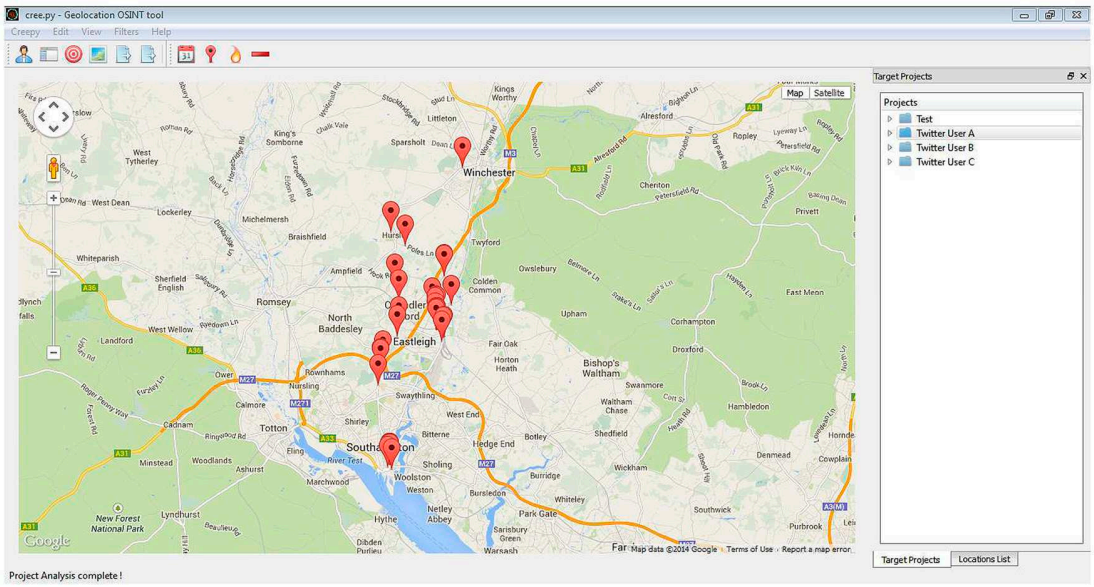


**Figure 8.** User A's Movements Tracked Using Geo-location.

Finally, the home address of User A was identified by simply reading the content of a single tweet, see Figure 9. Using Google street view the house and street were easily found. The tweet content not only identified where this user lived but also revealed that every Monday they attended "Movie Night" at the same time and in the same place. This information could be used to orchestrate a meeting with them. Their place of work was found using another one of the single pins further away from their home address. This location was again identified by the tweet content which demonstrates that it only takes one tweet to expose important information.


**Figure 9.** User A's Give Away Tweet.

User B's profile showed that the geo-location information contained distinct clusters that could indicate a place of work or home residence. To identify these locations it was necessary to analyze the tweet attached to each of the pins. Further analysis of these clusters revealed the home residence of User B which was clearly identifiable by a number of tweets that specifically mention the word "home", see Figure 10.


**Figure 10.** User B's Give Away Tweet.

Also found within the clusters was User B's route to work which when put together with some of the single pins on the map identified a consistent route to and from work using public transport, see Figure 11. Using this regular route it was easy to find the pin that represented their work place, even though User B only tweeted once from their workplace. The tweet content clearly stated that they were at work, "*My view of the London eye and whitehall looking very dull and rainy from my office window*", see Figure 12.

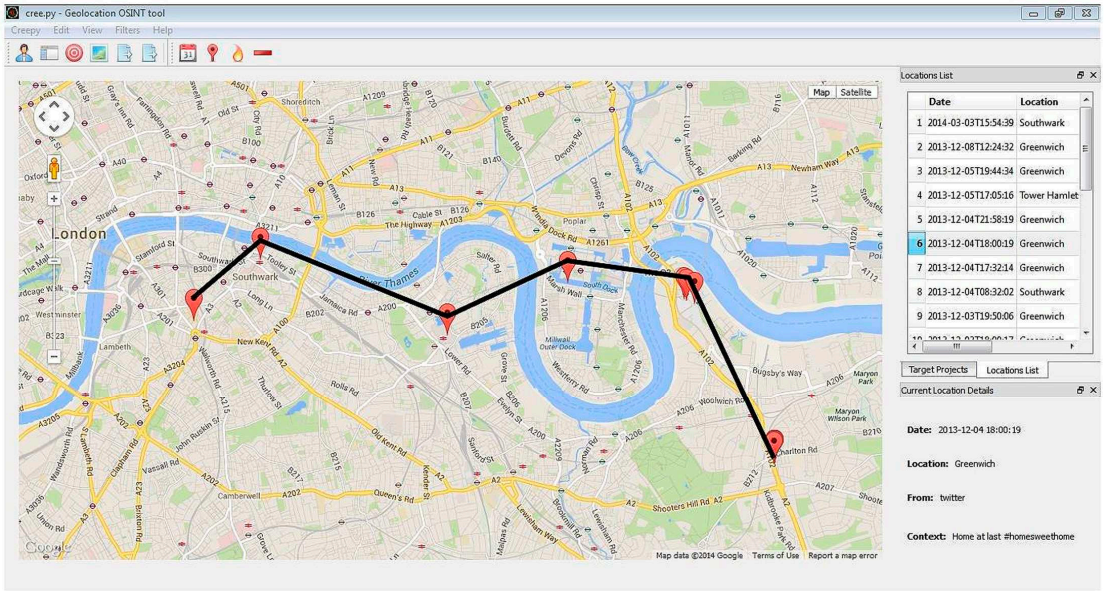


**Figure 11.** User B's Regular Route to Work.

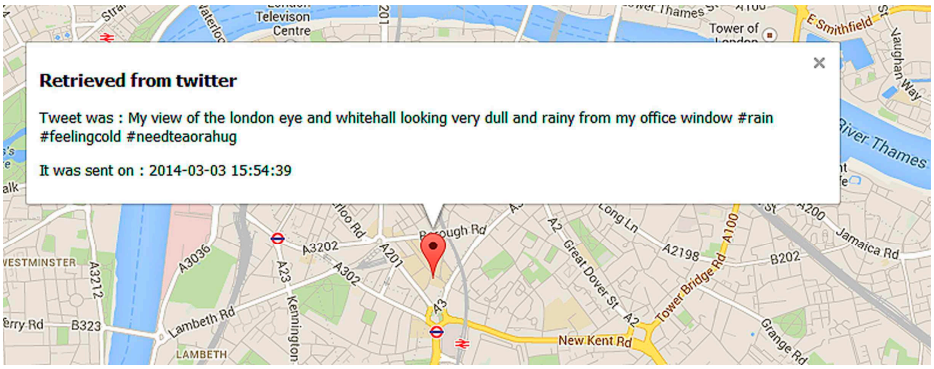


**Figure 12.** Tweet Revealing the Location of User B's Work Address.

User C was a prolific tweeter as demonstrated by the large number of pin drops on the map, see Figure 13. Analysis of the tweet contents revealed a number of patterns. The first pattern identified was that every month User C goes out of the city to visit their parents. A perfect time to burgle their home. The next pattern found was that they regularly take the same route to their work place and this was identified because at each station and bus stop on route they shared the information on Twitter, which makes them a very easy person to track. While analyzing their route to work it was discovered that the second half of their journey home may change depending on whether they needed to stop at the supermarket. To confirm that these were the correct routes going to and from their work place, the timestamps were verified and confirmed against the working hours established using Twitonomy.

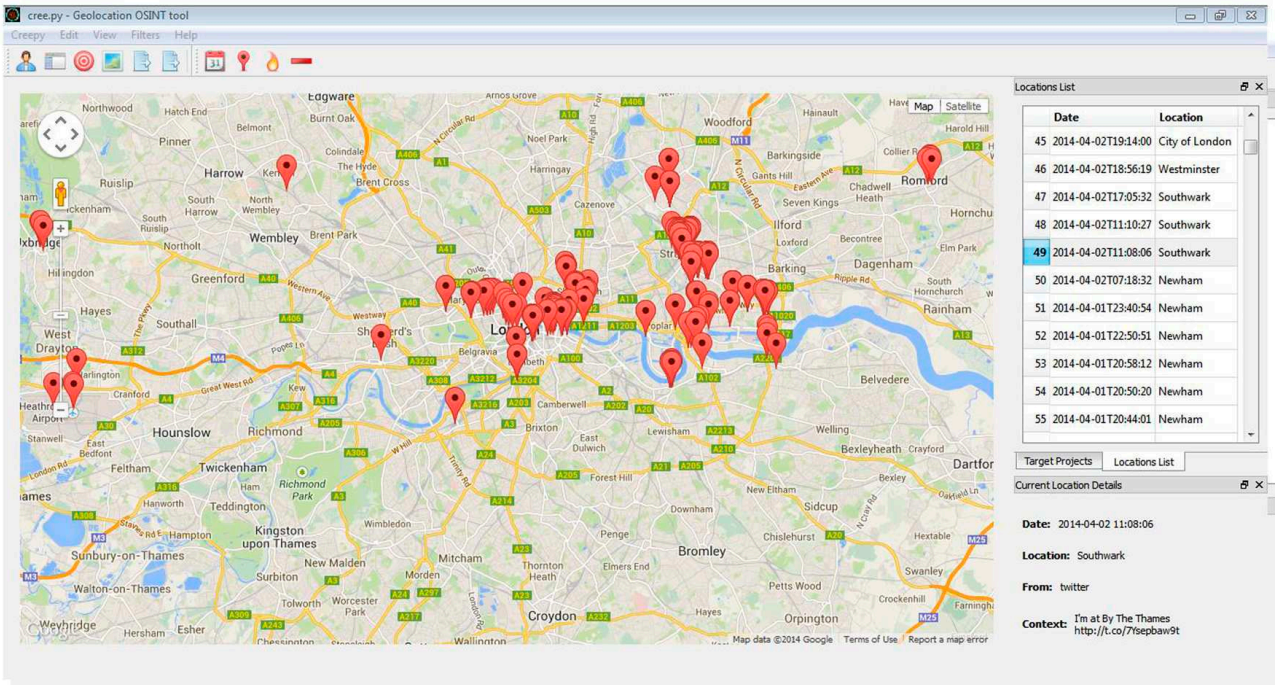


**Figure 13.** Pin Drops Identifying Locations Frequented by User C.

Another pattern in User C's habits that was identified was that on average they attend the gym three times per week. If it is a week day then they normally attend between 7 and 10 and if it is at the weekend an earlier time of between 1 and 3 was identified. This was found by analyzing their tweet content while making a note of the times and dates that they attended, which can be seen listed in Figure 14.

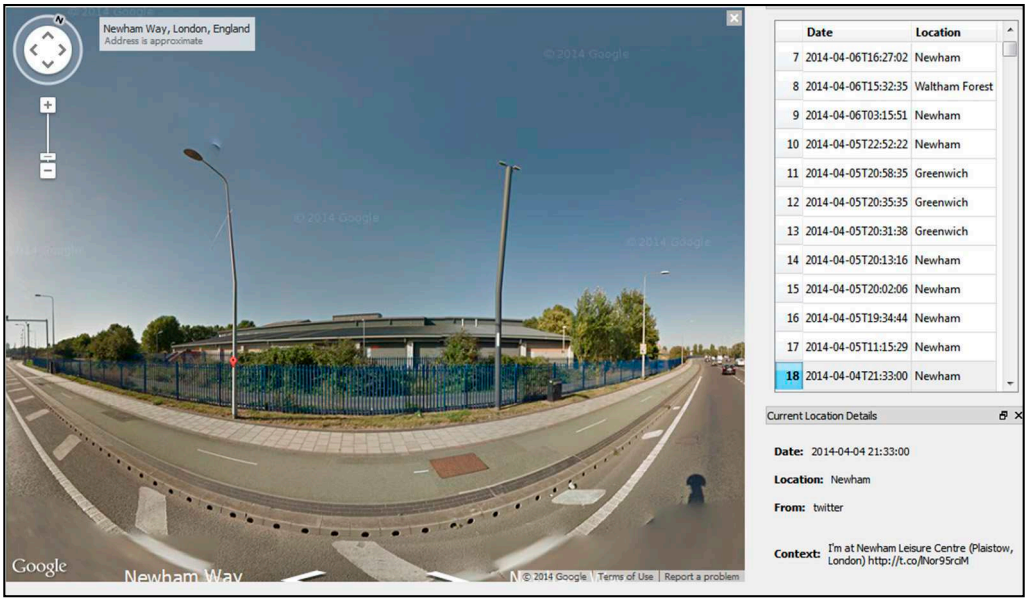


**Figure 14.** The Leisure Centre Attended Regularly By User C.

Finding a place of work or home residence for User C was particularly difficult as they never mention them in their tweets. Each tweet and pin drop had to be analyzed very thoroughly to extract this information. Eventually a distinct pattern emerged which showed that User C would be in the area of Southwark every week day, but never at weekends and always between the same times. User C checked in frequently with Foursquare, which identified a few places on the same street. It would be easy to go and wait outside these places to see which one User C actually works at. From their work place it would then be possible to track their route home to discover their residence. Again because they check in when getting on and off public transport it can be approximately determined which part of a particular street they live on but to confirm a specific dwelling it would be necessary to follow the user in person.

Overall the tool Creepy was very successful at identifying the patterns and habits of all three of the target users. Creepy gathered the least amount of information on User A, as they varied their route to work which made it harder to track them. However, because they tweeted a lot when socializing it was possible to determine a weekly routine that could be used by a stalker. Establishing where User B lived and worked was relatively easy. They only tweeted at home or on their way to work which made it very easy to track the route that they took, as well as determining both their work place and home address. The most information leaked was by User C. The reason Creepy was so successful for User C was because they tweeted on average 30 times a day, which over a long period of time led to clearly identifiable patterns, such as their trips to the supermarket and the gym. Creepy has been an excellent tool for finding information on all three subjects as it was very easy to establish set patterns in their behavior.

### *3.5. Experiment 4—Other Social Networks*

The fourth experiment was to determine how much each subject's tweeting exposed the rest of their social networking "presence". As all three were using Twitter it was expected that they would also have accounts on other social media such as Facebook, LinkedIn, Foursquare and Instagram. Instagram is a mobile based application rather than a computer based one.

User A's Twitter feed gave no indication that they had any other social media accounts. After entering their name into a Google search their Facebook page was found listed right under their Twitter account. A quick analysis of the profile pictures confirmed that this was the same person, as they had used the same picture on both sites, which is typical user behavior. Without logging into Facebook only very basic information could be viewed, such as their interests, but nothing personal was revealed. However, once we logged into a Facebook account which had been created for this work, and that had no links with User A, *i.e.*, was not "friends" with them, a small number of their pictures, mainly from the profile picture album, were visible, as well as where they were currently living. This corroborated the information previously obtained. User A also had a profile on Instagram (Instagram24.com) and using their profile name it was possible to locate their pictures, including other peoples pictures that they had "liked". Using Google Street View the exact location and door number of their home were easily identified. Figure 15 shows the front door of User A's home with their Twitter pin drops surrounding the residence. User A no longer lives at this residence.

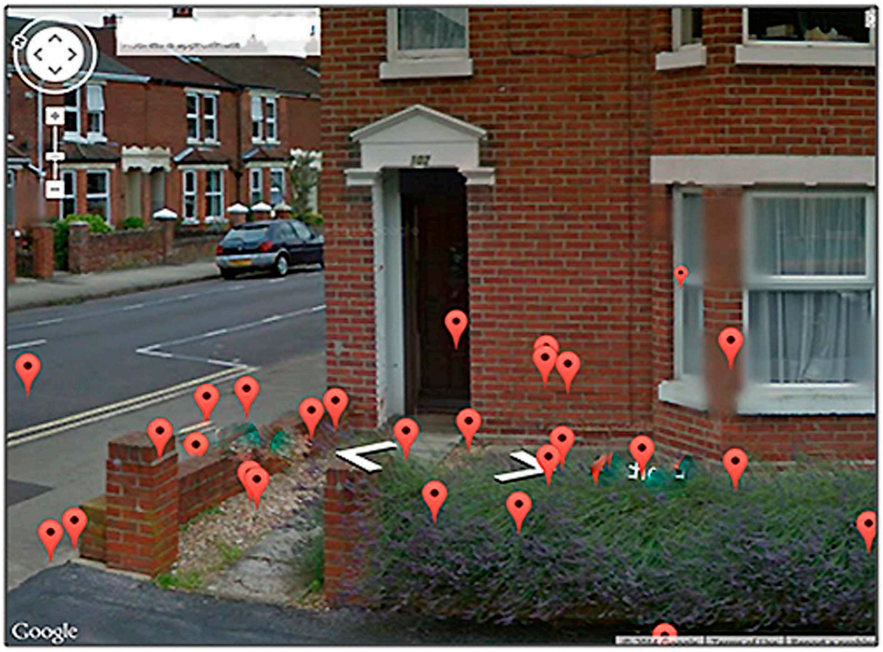


**Figure 15.** User A's Front Door (anonymized).

Since we had previously discovered where User A worked there was a good chance they would be on the professional social media site LinkedIn. An initial search using the same profile name turned up no results, so the last name was dropped and User A appeared in the search results with a different surname. Although this subject clearly uses two different surnames names depending on the situation, it was still very easy to find all of these profiles and link them back to that person.

User B also gave no indication on Twitter that they had any other social media profiles but again their name was entered into a Google search, which found their LinkedIn, Facebook and Google+ accounts. Using each of these profiles for the different sites, it was possible to confirm where they worked, the city they live in and where they were studying.

Although User C had been the easiest to identify using Twitter, they were the most difficult to locate on other social media sites, as they only used their first name on their Twitter account. With only their

first name a search returned hundreds of thousands of hits. The only social media site that was linked to their Twitter account was Foursquare. Further research to determine the full name of User C required an exhaustive search of their Twitter account, which revealed two tweets with pictures. The first picture, see Figure 16, shows that they had posted an image of their ID bracelet while they were in hospital. After enhancing this picture using a photo editing program their last name and their NHS ID number could be clearly read. As we already had their first name from their Twitter profile, this meant that we now had User C's full name.

**Figure 16.** User C's Hospital ID Bracelet (anonymized).

The second picture found in User C's tweets was even more revealing as it showed their e-ticket and their full name (including a middle name), which airports they will be passing through and how long they will be stopping at each airport, see Figure 17. It details their flight plans to Mexico via Madrid and Miami and returning via JFK airport, New York. The dates of travel can also be clearly seen, so anyone can know how long they will be away. Looking at the flight schedule we can determine that User C will be out of the UK from 20th September 2014 until at least the 2nd October 2014. Any burglar can assume that they have more than a week when their home will be unoccupied.

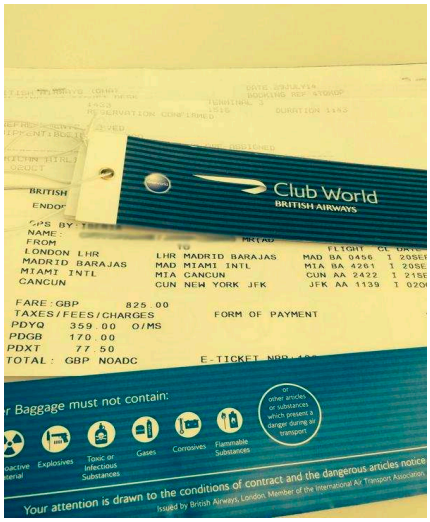


**Figure 17.** User C's E-tickets (anonymised).

Having discovered both of these pictures and obtained the missing information, we were then able to locate User C's Facebook account, which was set to private so we could only see basic information. After searching other social networking sites such as LinkedIn and Instagram we were unsuccessful in finding any other accounts linked to User C, so we can assume they only have Twitter, Foursquare and Facebook. Even though Facebook did not give us any new information on User C it did confirm a lot of the details that we were only speculating about, such as their work place, which they shared on Facebook. With this new research we now have solid information about User C, all of which has been leaked by the user themselves in one way or another.

### *3.6. Evaluation*

This research initially set out to demonstrate how anyone with a computer could use the data gathered from social networking sites to commit identify theft. While undertaking the investigations it quickly became clear that this was not the only criminal activity that could be conducted. Creepy was the most successful at obtaining results using the geo-location tagging on tweets because it enabled the identification of all three test subject's home addresses, work addresses and in most cases their routes to and from work. It was a combination of the geo-location data and the tweet contents that gave each of these locations away. The tool streamd.in was able to accurately plot real time tweets onto Google maps using the geo-location data provided with the tweets. This tool was invaluable for identifying User C but it could just as easily be used to track anyone as they travelled around or enjoyed social events.

With the information that has been gathered on all three target subjects it is quite clear that through the use of geo-location tagging they are at risk of being a victim of stalking or even burglary. It was relatively easy to obtain addresses for their places of work and for their home residences and also to determine their individual routines throughout the week. For example, User A goes to film night every Monday evening which would be the perfect opportunity to either rob their home or to orchestrate a meeting. Similarly, User B clearly identified their home address, work place and their route to work. User C was very easy to track on Twitter, although tracking them across other web sites was more challenging as they only used their first name on their Twitter account.

Throughout this work we have seen that the contents of individual tweets often gives away a lot more information than the tweeter intended. For example, they would not want strangers to know where they lived, but they had inadvertently given this information away in their tweet contents. It was the two individual tweets that finally enabled us to identify User C's surname and from there it was easy to track them on Facebook. This highlights how much users leak personal information willingly and without realizing that anyone can read them, not just their friends or followers, even though they may be very careful when using Twitter.

## 4. Recommendations

With the many risks highlighted during this work it is necessary to include recommendations to provide the reader with some guidelines on how to reduce their risk when using Twitter. Awareness is the key to protecting your personal information. Below are a number of points that every Twitter user should follow:

1. If using Twitter is important to limit the amount of information added to an individual's profile, so do not include where you live even if it is only the city.
2. Avoid using your full name. This will reduce the chance of identity theft from Twitter and other social networks. Use an alias or at the very most only initials.
3. Avoid using an actual profile picture of yourself. Use a cartoon or a picture of yourself when very young which will make it more difficult to identify you in person.
4. Set your profile to private "Protect my Tweets" which is located in the "Security and privacy" settings. This will allow only your permitted followers to view your tweets, as followers have to be accepted by the user, unlike on public accounts.
5. Remove geo-location tagging on tweets. If some tweets already have geo-location data attached to them Twitter has a function to delete this data.
6. Do not provide your phone number.
7. Remove "Let others find me by my email address" as this is another piece of information that can lead to the disclosure of personal information.
8. Do not connect your Twitter account with any other social media sites, such as Facebook to avoid unintentional sharing information.
9. Limit the amount of Apps that have access to your profile. Always question whether it is necessary to give access.
10. Be very selective about what you put in your tweets and always check to make sure you are not accidentally giving away personal information.

As has been identified during the course of this work, it is very easy to inadvertently leak personal information without realizing. We have identified that when sharing information using Twitter it is important to be very selective about what you tweet as any small detail could give away personal information. Profiles must be protected by only allowing friends to view them, which means being extra careful when confirming "follow" requests by only accepting people you know personally.

## 5. Conclusions

A number of experiments have been conducted using the three freely available tools, streamd.in, Twitonomy and Creepy. Of the three tools used for this work Creepy provided the most results using the geo-location tagging on tweets. However, at times the combination of all three tools were used to obtain the required information.

These experiments initially set out to demonstrate how one could use the data gathered to commit crimes such as identify theft. However, it quickly became apparent that this wasn't the only criminal activity that could be conducted. With the information gathered on all three target subjects it was clear that through the use of geo-location tagging they were at risk of being a victim of crime, such as burglary or online stalking which could turn into physical stalking or even harassment. With all the information each user provided to Twitter, they were easy to track. They had their privacy settings on default which meant that their tweets and pictures were public. All the subjects were given help and advice on how to protect their information and identities online at the conclusion of this work.

This work clearly high-lights the importance of online privacy and security and the need for Twitter users to be more careful about what they share. These experiments have demonstrated how easily the

information obtained using Twitter could be used to track a "victim" through cyber space to harvest even more personal data about them as they have no idea who is reading their tweets. The data gathered led to further harvesting of private information which users had not intended to reveal, such as their home address or where they work. The geo-location data was determined to be very accurate, which highlights the importance of using this correctly or by turning it off all together.

With all the privacy issues highlighted during this work the natural suggestion would be to stop using social networking sites but in order to maximize job prospects for graduating students, social media is a must. Not using social networks is not an option for many people. The risks can be reduced by promoting the safety recommendations in this work, especially to young people. There are however issues of crime prevention and public safety which need to be considered at the level of public policy and the social network providers should take a more active role to improve safety for their users. Social networks are a way of life for many and there is no getting away from it, so protecting your privacy and preventing yourself from becoming a target for crime should be top of the agenda for all users.

## Acknowledgments

The authors would like to thank Isa Usman Wakili who provided the statistics presented in Tables 1 and 2. We would also like to thank the three subjects of the experiments for their co-operation in allowing us to stalk them and to publish their information.

## Author Contributions

All authors were involved in drafting the article and revising it critically for important intellectual content, and all authors approved the final version to be published. The initial concept was conceived by Diane Gan. Lily R. Jenkins and Diane Gan designed the experiments and Lily R. Jenkins collected the data and collated the results.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Discover Twitter—What is Twitter and How to Use It. Available online: https://discover.twitter.com/ (accessed on 4 February 2014).
2. Ball, J.; Lewis, P. Twitter and the Riots: How the News Spread. The Guardian, 7 December 2011. Available online: http://www.theguardian.com/uk/2011/dec/07/twitter-riots-how-news-spread (accessed on 4 August 2014).
3. Proctera, R.; Crump, J.; Karstedt, S.; Voss, A.; Cantijoch, M. Reading the riots: what were the police doing on Twitter? *Polic. Soc.: Int. J. Res. Policy* **2013**, *23*, 413–436.
4. Twitter Statistics. Available online: http://www.statisticbrain.com/twitter-statistics (accessed on 18 February 2014).
5. Stone, B. Location, Location, Location, August 20, 2009. Available online: https://blog.twitter.com/2009/location-location-location (accessed on 20 February 2014).

6. Vicente, C.R.; Freni, D.; Bettini, C.; Jensen, C.S. Location-related privacy in geo-social networks. *IEEE Internet Comput.* **2011**, *15*, 20–27. doi:10.1109/MIC.2011.29.
7. Jenkins, L.R.; Gan, D.E. Investigation into the privacy issues of using social media. In Proceedings of the CFET 2014—7th International Conference on Cybercrime, Forensics, Education and Training, Christ Church Canterbury, UK, 10–11 July 2014; ISBN:97801909067158.
8. Luo, W.; Liu, J.; Liu, J.; Fan, C. An analysis of security in social networks. In Proceedings of the 2009 8th IEEE International Conference on Dependable, Autonomic and Secure Networking, Chengdu, China, 12–14 December 2009.
9. Abdulhamid, S.M.; Ahmad, S.; Waziri, V.O.; Jibril, F.N. Privacy and National Security Issues in Social Networks: The Challenges. *Int. J. Comput. Internet Manag.* **2011**, *19*, 14–20.
10. Hajli, N.; Lin, X. Exploring the security of information sharing on social networking sites: The role of perceived control of information. *J. Bus. Ethics* **2014**, doi:10.1007/s10551-014-2346-x.
11. Edwards, A. Teenage Youth Crime Commissioner Who Quit over Offensive Tweets is Questioned by Special Branch. Available online: http://www.dailymail.co.uk/news/article-2312044/Paris-Brown-Foul-mouthed-youth-commissioner-quit-offensive-tweets-questioned-police-caution.htm (accessed on 20 April 2013).
12. Reuters. Twitter Told to Improve Security. Available online: http://www.pcpro.co.uk/news/security/368542/twitter-told-to-improve-security (accessed on 8 July 2011).
13. Brinkmann, M. Twitter Improves Account Security, Improves Password Reset. Available online: http://www.ghacks.net/2014/05/09/twitter-improves-account-security-improves-password-reset/ (accessed on 9 May 2014).
14. Srinivasan, S. Lack of Privacy Awareness in Social Networks. *ISACA J.* **2012**, *6*. Available online: http://www.isaca.org/Journal/Past-Issues/2012/Volume-6/Pages/Lack-of-Privacy-Awareness-in-Social-Networks.aspx (accessed on 20 February 2014).
15. Twitter Privacy Policy 2014. Available online: https://twitter.com/privacy (accessed on 4 February 2014).
16. Madden, M. Privacy management on social media sites. In *Pew Internet Report*; Pew Research Center: Washington, DC, USA, 2012; pp. 1–20. Available online: http://pewinternet.org/Reports/2012/Privacy-management-on-social-media.aspx (accessed on 23 November 2014).
17. Bennett, S. Social Media And Crime—How Secure Is Your Information? Covering the World of Social Media. Available online: http://www.adweek.com/socialtimes/social-media-crime/491227?red=at (accessed on 25 September 2013).
18. Welter, A. Social Media and Crime, Crime Wire. Available online: http://www.instantcheckmate.com/crimewire/social-media-and-crime-2/#prettyPhoto (accessed on 21 August 2013).
19. Hardwick, L. How to Improve Your Twitter Security and Privacy. Available online: https://nakedsecurity.sophos.com/2014/08/26/how-to-improve-your-twitter-security-and-privacy/ (accessed on 26 August 2013).
20. Neagu, A. Twitter Security Tips: How to Improve your Twitter Security and Privacy in 10 Easy Steps. Available online: https://heimdalsecurity.com/blog/twitter-security-privacy-essential-guide (accessed on 13 November 2014).

21. Hannay, P.; Baatard, G. GeoIntelligence: Data mining locational social media content for profiling and information gathering. In Proceedings of the 2nd International Cyber Resilience Conference, School of Computer and Information Science, Security Research Centre, Edith Cowan University, Perth, Western Australia, 1–2 August 2011. Available online: http://data.openduck.com/wp-posts/2012/09/paper-geo/geointelf.pdf (accessed on 10 December 2013).
22. Valli, C.; Hannay, P. Geotagging Where Cyberspace Come to Your Place. *Openduck* **2010**, 1–4. Available online: http://data.openduck.com/wp-posts/2010/07/paper-geotagging/valli-hannay-geotagging.pdf (accessed on 10 December 2013).
23. Paske, B.; Lyle, J. Follow you follow me: Using location tracking to mitigate multi-device privacy threats. In Proceedings of the Workshop on Multi-device App Middleware, Montreal, QC, Canada, 3–7 December 2012; Article No. 4, pp. 1–6. doi:10.1145/2405172.2405176.
24. Thomas, L.; Briggs, P.; Little, L. Location tracking via social networking sites. In Proceedings of the 5th Annual ACM Web Science Conference (ACM WebSci14), Paris, France, 2–4 May 2013; pp. 405–412. doi:10.1145/2464464.2501852.
25. Brownlee, J. This Creepy App Isn't Just Stalking Women without Their Knowledge, It's A Wake-up Call about Facebook Privacy Available online: http://www.cultofmac.com/157641/this-creepy-app-isnt-just-stalking-women-without-their-knowledge-its-a-wake-up-call-about-facebook-privacy/ (accessed on 20 February 2014).
26. Palmer, S. How Twitter is being used by Australian engineering academic units. In Proceedings of the 24th 2013 Australasian Association for Engineering Education Conference, Griffith School of Engineering, Griffith University, Brisbane, Australia, 8–11 December 2013.
27. Forgie, S.E.; Duff, J.P.; Ross, S. Twelve tips for using Twitter as a learning tool in medical education. *US Natl. Lib. Med.* **2013**, *35*, 8–14. doi:10.3109/0142159X.2012.746448. Available online: http://www.ncbi.nlm.nih.gov/pubmed/23259608 (accessed on 12 August 2014).
28. Visser, R.D.; Evering, L.C.; Barrett, D.E. #TwitterforTeachers: The Implications of Twitter as a Self-Directed Professional Development Tool for K–12 Teachers. *J. Res. Technol. Educ.* **2014**, *46*, 396–413. doi:10.1080/15391523.2014.925694.
29. Carpenter, J.P.; Krutka, D.G. How and why educators use twitter: A survey of the field. *J. Res. Technol. Educ.* **2014**, *46*, 414–434. doi:10.1080/15391523.2014.925701
30. Prestridge, S. A focus on students' use of Twitter—Their interactions with each other, content and interface. *Act. Learn. High. Educ.* **2014**, *15*, 101–115. doi:10.1177/1469787414527394.
31. Junco1, R.; Elavsky, C.M.; Heiberger, G. Putting twitter to the test: Assessing outcomes for student collaboration, engagement and success. *Br. J. Educ. Technol.* **2013**, *44*, 273–287. doi:10.1111/j.1467-8535.2012.01284.x.
32. Pentina, I.; Zhang, L.; Basmanova, O. Antecedents and consequences of trust in a social media brand: A cross-cultural study of Twitter. *Comput. Hum. Behav.* **2013**, *29*, 1546–1555. doi:10.1016/j.chb.2013.01.045.
33. Visa, F. Twitter as a Reporting Tool for Breaking News—Journalists tweeting the 2011 UK riots. *Digit. Journal.* **2013**, *1*, 27–47. doi:10.1080/21670811.2012.741316.
34. Hermida, A.; Lewis, S.C.; Zamith, R. Sourcing the Arab Spring: A Case Study of Andy Carvin's Sources on Twitter during the Tunisian and Egyptian Revolutions. *J. Comput. Mediat. Commun.* **2014**, *19*, 479–499. doi:10.1111/jcc4.12074.

35. Malleson, N.; Andresen, M.A. The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartogr. Geogr. Inf. Sci.* **2015**, *42*, 112–121. doi:10.1080/15230406.2014.905756.
36. Gabielkov, M.; Rao, A.; Legout, A. Studying social networks at scale: Macroscopic anatomy of the Twitter Social Graph. In Proceedings of the 2014 ACM International Conference on Measurement and Modeling of Computer Systems, (SIGMETRICS '14), Austin, TX, USA, 16–20 June 2014; Volume 42, Issue 1, pp. 277–288. doi:10.1145/2591971.2591985.
37. Graham, M.; Hale, S.A.; Gaffney, D. Where in the World Are You? Geo-location and Language Identification in Twitter. *Prof. Geogr.* **2014**, *42*, 277–288. doi:10.1080/00330124.2014.907699.
38. Jeong, Y.; Coyle, E. What Are You Worrying About on Facebook and Twitter? An Empirical Investigation of Young Social Network Site Users' Privacy Perceptions and Behaviors. *J. Interact. Advertising* **2014**, doi:10.1080/15252019.2014.930678.
39. Davis, K.; James, C. Tweens' conceptions of privacy online: implications for educators. *Learn. Media Technol.* **2013**, *38*, 4–25.
40. Zhang, H.; Choudury, M.D.; Grudin, J. Creepy but Inevitable? In Proceedings of the Evolution of Social Networking, CSCW '14, Baltimore, ML, USA, 15–19 February 2014. Available online: http://dx.doi.org/10.1145/2531602.2531643 (accessed on 15 August 2014).
41. Mao, H.; Shuai, X.; Kapadia, A. Loose Tweets: An Analysis of Privacy Leaks on Twitter. In Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society, Chicago, IL, USA, 17–21 October 2011; pp. 1–12, ISBN:978-1-4503-1002-4.
42. Mearns, G.; Simmonds, R.; Richardson, R.; Turner, M.; Watson, P.; Missier, P. Tweet My Street: A Cross-Disciplinary Collaboration for the Analysis of Local Twitter Data. *Future Internet* **2014**, *6*, 378–396. doi:10.3390/fi6020378.
43. Ryoo, K.; Moon, S. Inferring Twitter user locations with 10 km accuracy. In Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web, Seoul, Korea, 7–11 April 2014; pp. 643–648, ISBN:978-1-4503-2745-9.
44. Jin, L.; Chen, Y.; Wang, T.; Hui, P.; Vasilakos, A.V. Understanding User Behavior in Online Social Networks: A Survey. *IEEE Commun. Mag.* **2013**, *51*, 144–150.
45. Shi, Z.; Ru, H.; Whinston, A.B. Content Sharing In a Social Broadcasting Environment: Evidence from Twitter. *MIS Q.* **2014**, *38*, 123–142.
46. Yang, Z.; Guo, J.; Cai, K.; Tang, J.; Li, J.; Zhang, L.; Su, Z. Understanding retweeting behaviors in social networks, CIKM'10, Proceedings of the 19th ACM International Conference on Information and Knowledge Management, Toronto, ON, Canada, 26–30 October 2010.
47. Zhang, J.; Liu, B.; Tang, J.; Chen, T.; Li, J. Social Influence Locality for Modeling Retweeting Behaviors. In Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, Beijing, China, 3–9 August 2013; pp. 2761–2767.
48. Wang, J.; Wang, L.; Wu, W. Predicting information popularity degree in microblogging diffusion networks. *Int. J. Multimed. Ubiquitous Eng.* **2014**, *9*, 21–30.
49. Lee, K.; Mahmud, J.; Chen, J.; Zhou, M.; Nichols, J. Who will retweet this? Automatically Identifying and Engaging Strangers on Twitter to Spread Information. In Proceedings of the 19th International Conference on Intelligent User Interfaces, New York, NY, USA, 24–27 February 2014; pp. 247–256.

50. Pervin, N.; Takeda, H.; Toriumi, F. Factors Affecting Retweetability: An Event-Centric Analysis on Twitter. In Proceedings of the International Conference on Information Systems (ICIS) 2014, Auckland, New Zealand, 14–17 December 2014.
51. Lia, J.; Pengb, W.; Lia, T.; Sunb, T.; Lic, Q.; Xuc, J. Social network user influence sense-making and dynamics prediction. *Expert Syst. Appl.* **2014**, *41*, 5115–5124
52. Liu, G.; Shi, C.; Chen, Q.; Wu, B.; Qi, J. A Two-Phase Model for Retweet Number Prediction. In Proceedings of the Web-Age Information Management—15th International Conference, WAIM 2014, Macau, China, 16–18 June 2014; Volume 8485, pp. 781–792.
53. Macskassy, S.A.; Michelson, M. Why do People Retweet? Anti-Homophily Wins the Day? In Proceedings of the International Conference on Weblogs and Social Media (ICWSM), Barcelona, Spain, 17–21 July 2011.
54. StreamdIn API. Available online: http://download.cnet.com/StreamdIn/3000–12941_4–75695954.html (accessed on 20 February 2014).
55. Shetty, N. Streamd.in—Tweets on Google maps in an Awesome Way. 2010. Available online: http://www.twi5.com/streamd-in-tweets-on-google-maps-in-an-awesome-way/8666/ (accessed on 1 February 2015).
56. Twitonomy: Twitter #Analytics and Much More. Available online: http://www.twitonomy.com/index.php Twitonomy.com (accessed on 12 February 2014).
57. Creepy API. Available online: http://ilektrojohn.github.io/creepy (accessed on 12 February 2014).
58. Kakavas, Y. Creepy, the Geo-location Information Aggregator, 2011—InfoSec Institute. Available online: http://resources.infosecinstitute.com/creepy/ (accessed on 31 January 2014).