Playing Hide and Seek with Mobile Dating Applications

Guojun Qin, Constantinos Patsakis, Mélanie Bouroche

Playing Hide and Seek with Mobile Dating Applications

Guojun Qin, Constantinos Patsakis and Mélanie Bouroche

Distributed Systems Group, School of Computer Science & Statistics
Trinity College, Dublin, Ireland.

Abstract. Recently, a wide range of dating applications has emerged for users of smart mobile devices. Besides allowing people to socialize with others who share the same interests, these applications use the location services of these devices to provide localized mapping of users. A user is given an approximation of his proximity to other users, making the application more attractive by increasing the chances of local interactions. While many applications provide an obfuscated location of the user, several others prefer to provide quantifiable results.
This paper illustrates that the user's location can be disclosed, with various degree of approximation, despite the obfuscation attempts. Experimenting with four of these applications, namely MoMo, WeChat, SKOUT and Plenty of Fish, we show that an attacker can easily bypass the fuzziness of the results provided, resulting in the full disclosure of a victim's location, whenever it is connected.

Keywords: Location privacy, online social networks, information revelation, geosocial networks

1 Introduction

Modern smartphones are more than just mobile phones. Due to their processing resources they are closer to mobile information systems that have access to the Internet and are location-aware, either through an embedded GPS module or through network resources. Quickly, all major social networks ported their applications to these new devices. Soon afterwards, a new species evolved, the location-based social networks, often also referred to as geosocial networks (GSNs). These applications are enriching the widely-used online social networks with location-based services. By exploiting the location awareness of users or their knowledge of proximity to points of interest, these applications are providing more fine-grained and personalised services to their users.

It is clear that this shift has not only created a whole new market, but simultaneously has drastically changed the way in which people regard their location privacy. While almost three quarters (74%) of adult smartphone owners use their phones to get directions or other information based on their current location [1], their trust in the provided privacy is not that high. This can be understood by

the number of users concerned about location sharing privacy, as the sharing of their location could be abused to disclose more sensitive personal information, such as home addresses and user identities. Similarly, another 58% of all teens have downloaded applications to their cell phone or tablet computer and 51% of teen applications users have avoided certain applications due to privacy concerns [2]. Moreover, 46% of teen applications users have turned off tracking features on their cell phone or in an application and 26% of teen applications users have uninstalled an application because they were worried about the privacy of their information.

Those research reports clearly illustrate that the privacy of location-based services is a serious concern for most smart devices users. In the past few years, many researchers have proposed several solutions to preserve users' location privacy such as location k-anonymity and cloaking granularity [3–6]. However, the location privacy threats in digital life are changing as the popularity of mobile and online dating applications is growing. According to another report, 11% of Internet users have personally used an online dating site and 7% of cell phone applications users have used a mobile dating application [7]. Additionally, 40% of online daters have used a site or an application for people with shared interests or backgrounds. The dating applications typically not only share users' pictures and interests, but also the distance between users. The latter manages to enhance even more the engagement of users to the application, as they feel that they can really meet other users, and that potentially interesting other users are in their vicinity.

The fact that users can know almost in real-time their distance to other users, motivated us to investigate whether and to what extend this feature could be used to trace other users' location, and the effort required. Our hypothesis is that these applications can provide a reliable metric, or that certain pattern would emerge which an adversary can exploit to track down a user's actual location by using simple and widely-used trilateration algorithms. It is clear that if a malicious user has more background knowledge, other users' sensitive information such as their real identities, home and work locations might be revealed as well.

The rest of this paper is organised as follows. In the next section we provide an overview of the related work, mainly focusing on attacks on online social networks and geosocial networks. In Section 3 we describe how trilateration works and Section 4 is devoted to the experimental results. We describe how we generated the experimental environment along with the individual results and impact for four of the most widely used applications of this field. In Section 5 we discuss possible counter measures that would prevent such attacks. Finally the article concludes in Section 6 with a brief summary and ideas for future work.

2 Related work

Due to the wide use of Online Social Networks (OSN), many attacks have emerged targeting their users or even the OSN infrastructure. An adversary

may try to manipulate users in many ways, either using shared information, social engineering or even by creating malicious applications [8].

In many instances, OSNs are used to harvest user email addresses and send them spam messages [9–12]. Going a step further than spam, malicious users might launch phishing attacks, which have better click-through rates than typical spamming as reported in [13]. The "freemium" model under which the vast majority of OSNs operate, allows users to easily create multiple accounts, launching what is known as sybil attacks [14]. The goals of the adversary typically vary and range from a simple voting scenario to a de-anonymization attack [15]. A malicious user can also launch an attack to the reputation of a user [16], usually anonymously, or try to extort the victim with the gathered information.

Overall, information about the user location can be inferred from OSNs and be exploited in many malicious ways[1], however, the location awareness opens up the possibility for even more attacks. For instance, based on collected location data, the home and work location of users or even their identities can be recovered [17–20].

Similar attacks can be launched from geospacial networks [21–23]. Nevertheless, even if some solutions have already been proposed [24], they have not been adopted. The interested reader may also refer to [25–27].

3 Trilateration attack

The trilateration attack is the application of the geometric process of trilateration which determines the location of an object based on its distance from other known points. Therefore, in the trilateration attack an adversary tries to find some points from which the distance to the target is known.

To understand the attack, we assume that the attacker co-ordinates with two more entities or that he can impersonate as two other entities. To succeed, the attacker has to select three points A_1, A_2, A_3 that are not collinear and manage to trick his victim to disclose his distance from these points (d_1, d_2, d_3). The attacker then finds the exact location of his target, as the victim V, will reside on the intersection of three circles with centers A_1, A_2, A_3 and radii d_1, d_2, d_3 respectively, as illustrated in Figure 1.

Following the same methodology, even if the distances are not exact, the location of the victim can be very well bounded. Let us assume that the accuracy of the measurement is τ, then the actual location of the user is not known, however, it resides within the area of the intersection of the three circles, see Figure 2.

Indeed, even using only the distance of three known points, the victim's location can by approximated with an error bounded by roughly by $\tau/2$. However, this bound can be further improved if more measurements are made.

As seen, the accuracy of the positioning depends on the accuracy of the distance to the known points. The experiments described below focus on investi-

[1] `seeforexamplehttp://www.pleaserobme.com/`

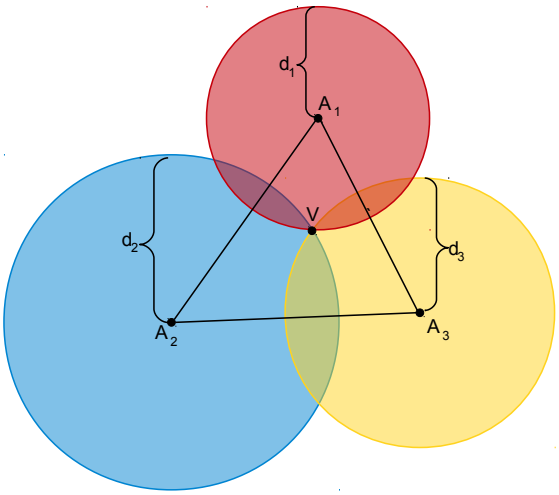


Fig. 1: Trilateration attack with exact distances.

gating the accuracy with which we can determine the distance from an attacker. This can then be used to derive the achievable accuracy for the positioning.

4 Experiments

In what follows, we describe analytically how the experiments were made, their findings and the impact for each of the applications.

4.1 Experimental methodology

In order to conduct our experiments, we needed to create a well-constrained environment for the applications, given that access to their internals or reverse engineering is not possible or legal. To avoid privacy issues that can be triggered by trying to trace individuals, we created some fake accounts on the applications that would be tested. From now on, we will only refer to two of these accounts that are going to be used, one representing the victim and the other the attacker. For convenience, we will refer to them as Alice and Bob respectively. An additional problem was that the measurements should be generic and replicable. In addition, the measurements should be independent of any kind of external noise. All these requirements can simply be met by using fake location. By setting the exact locations of Alice and Bob, one can

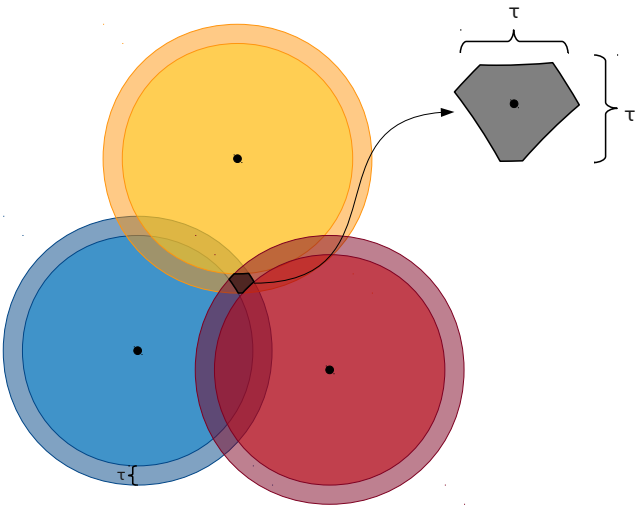


Fig. 2: Trilateration attack with approximate distances.

– replicate the exact same measurements,
– the mobile phones report always the same location and are not subject to GPS skewness or faults imposed by other antennas or lack of signal,
– the true distance between Alice and Bob can be easily recovered.

The rationale of the experiments is the following: Bob selects Alice as his target and every time he notes the distance to Alice D' as reported by the application. This is compared to the actual distance D (known from the use of the fake location), to derive a patterns between D and D' that enable Bob to deduce the actual distance from an unknown reported distance.

We found that all the applications try to obfuscate the results by reporting either rounded or randomized distances. To extract the required information, we examine the actual distance when the reported distances change. More precisely, we assume that Alice is at point A and Bob at B_0, so their distance is D_0 and the reported distance is D'_0. Bob chooses another point B_1, closer to Alice so their distance is D_1 and checks the reported distance if it is still D'_0 or it is a new value D'_1. The tests are repeated with new points A and B_0, but with the same actual initial distance D_0.

The assumption that the experiments aim to verify is the following:

Assumption: Bob can always find a points B so that when his reported distance with Alice is D', then the actual distance is always $D \pm \epsilon$.

The attack is then very straight forward: Bob records his initial distance from his target, and he starts moving until he finds that the reported distance

approximates the target D'. Then performing small steps, Bob finds a turning point, where the application reports D'' rather than D' that it was previously reporting. Bob now has found a point whose actual distance is $D \pm \epsilon$.

4.2 Experimental environment

The tests were made using two iOS devices running on a jailbroken version of iOS7. For faking the location, we have used the "LocationFaker" application from Cydia version 1.5-2. The applications that were tested are:

- MoMo version 4.8,
- WeChat version 5.1.0.6,
- SKOUT version 4.0.2,
- Plenty of Fish (POF) version 1.71

In Table 1 we summarize several characteristics of the applications. These characteristics are the distance range, how often the location is updated, whether the application displays when the last update was made and finally whether they detect usage of mock GPS location when used in Android. While in some cases the delay for location update was significant, it could be trivially bypassed. The attacker had to log off the application so his new location was used to calculate the distance with the victim.

Application Characteristics				
	MoMo	**WeChat**	**SKOUT**	**POF**
Distance Range	0.01Km	0.1Km	0.8Km	1Km
Minimum distance boundary	0.01Km	0.1Km	0.8Km	0.5Km
Location update frequency	Run time	Run time	10 mins	30 mins
Last update time	✓	✗	✗	✗
Mock GPS location detected	✓	✗	✗	✓

Table 1: Dating mobile application

4.3 Experimental results

To calculate the true distance between Alice and Bob we have used the well-known haversine formula, where the radius of the Earth is set to 6371Km. In the following paragraphs we analyse the findings and their implications for each application specifically.

MoMo Findings: The experimental results, an example of which is depicted in Figure 3, clearly indicate that MoMo is reporting the actual distance to the users, in groups of 10 meters. The formula that MoMo seams to use in order to report the distances is the following:

$$10 \left\lceil \frac{d_{True}}{10} \right\rceil$$

This means that the distances of the users are bounded by an error of 5 meters.

Implications: Using the trilateration attack, Bob can trace Alice with an accuracy of around 2.5 meters.

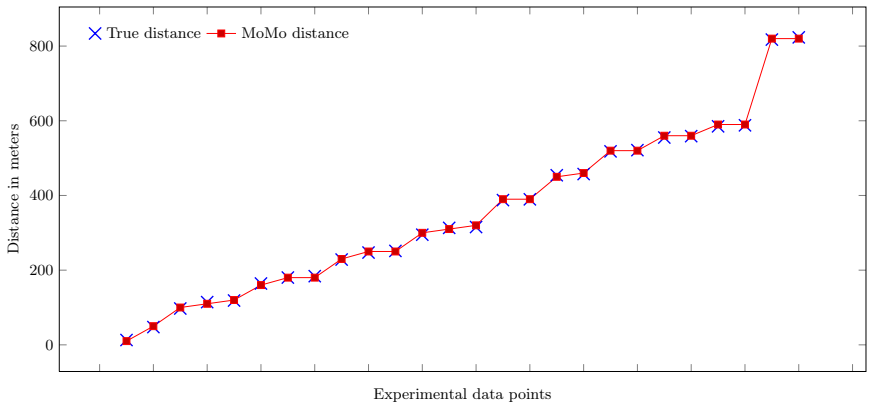
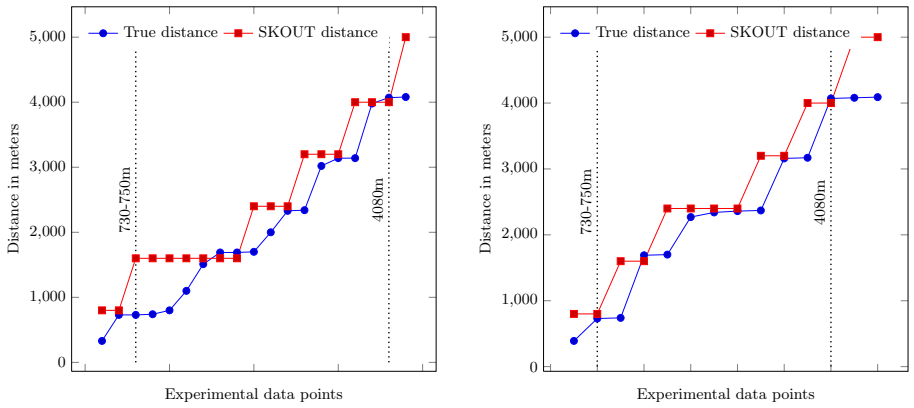


Fig. 3: A typical experiment for MoMo, the actual and reported distances coincide.

SKOUT Findings: The experimental results for SKOUT at first glance indicate that the reported distances are not correct. The application attempts to obfuscate the results, probably to provide some additional security to the users from such attacks. However, as shown in Figure 4, some patterns emerge. More precisely, as Bob moves in the range of 730-750 meters from Alice, he will see that the reported distance in the application will change from 800 to 1600 meters. The same behavior is repeated in other cases as well. Figure 4 illustrates these patterns, by showing the results of two typical experiments, comparing the reported from the application distance to the actual.

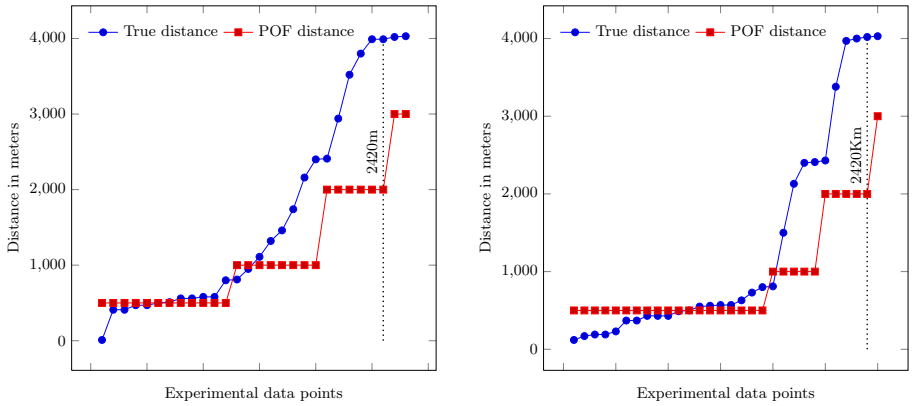
Implications: Bob can easily find a point which reports a distance of 800m. Moving around this point, Bob can find when the reported distance changes to 1600m. At that point, Bob will know that Alice's true position is 730-750 meters. It becomes apparent, that Bob can find another such point, thus the trilateration attack can be performed, tracking Alice with an accuracy of 10 meters. It is



(a) Experiment 1 (b) Experiment 2

Fig. 4: Graphical representation of the actual vs the reported distances for SKOUT

worthwhile to notice that the same behavior is noticed in the transition from 4Km to 5Km, where the actual distance is 4080 meters. Therefore, the victim's location can be almost accurate.

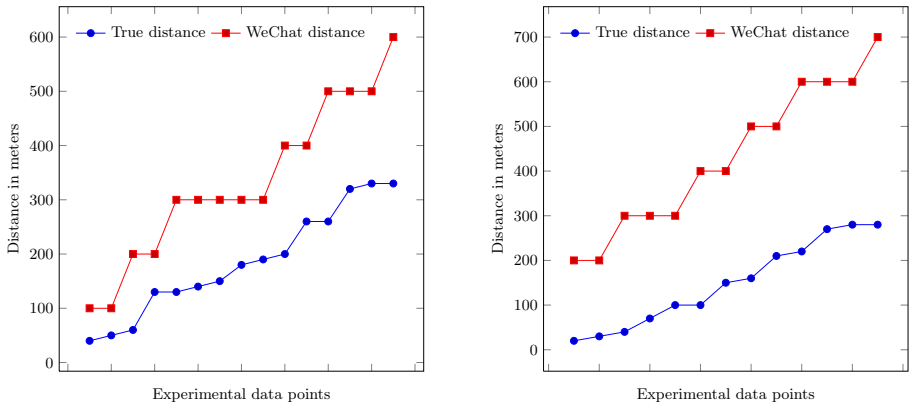


(a) Experiment 1 (b) Experiment 2

Fig. 5: Graphical representation of the actual vs the reported distances for Plenty of Fish.

Plenty of Fish Findings: This application also attempts to obfuscate the reported distances, more or less in the same way as SKOUT does. The reported distances might not reflect the actual rounded to kilometer distances, nevertheless, there are again specific patterns that emerge. For instance, the application will change the reported distance from 2Km to 1Km while Bob approaches Alice. At that point, their actual distance is 2420 meters. The results of two typical experiments, comparing the reported to the actual distance are illustrated in Figure 5, and clearly indicate the aforementioned patterns.

Implications: Using the same steps as in SKOUT, Bob can trace Alice using the trilateration attack with almost absolute accuracy.



(a) Experiment 1 (b) Experiment 2

Fig. 6: Graphical representation of the actual vs the reported distances for WeChat.

WeChat Findings: From the applications that were tested, WeChat was the one that tried to obfuscate the results the most. The initial results indicated that the application is not returning the actual distances. Moreover, the application is reporting the distance between Alice and Bob in a non standard way. For instance, the reported distance might be 500m when the actual distance might be 160m or 260m. Additionally, if Bob decides to track Alice, even if they are at the same positions as before, the reported distance might be different over time. Therefore, we may assume that WeChat is trying to detect probable attacks and stop them, by reporting inaccurate distances. Nevertheless, even if Bob cannot find an exact and stable point where his distance from Alice will change, there is some information that can be extracted. In all our experiments we noticed that when the reported distance between Alice and Bob is 200m, their actual dis-

tance is less than 150m. Figure 6 illustrates the results of two such experiments, indicating the reported behavior.

Implications: Exploiting the last comment, Bob can try to find points which are reported 200m away from Alice. This means that Alice will be at most 150 meters away from him. Exploiting this fact along with more points, Bob can accurately find Alice's location.

5 Discussion and counter measures

The results from the previous section demonstrate that users' position can be estimated very accurately from the distances provided by the applications. Those experiments, however, assume ideal GPS positioning (i.e. the GPS position reported to the application by both users is completely accurate). Due to the urban morphology, as well as the fact that users are often indoors, GPS is actually inaccurate, with an an error that can reach several meters in dense urban areas. The official study from the US government[2] clearly indicates this fact. In rural settings, however, the positioning is expected to be very accurate, and even in urban settings, using map information about the area (residential vs office building, public spaces etc.), would allow to improve the estimation. In addition, since these users are sharing their photographs, they can be identified amongst a set of people in a public space. While this attack requires several steps, the use of the fake location by the attacker means that it is fairly fast, and users typically stay at some places for a significant amount of time (e.g., office, restaurant, home), rendering the attack highly practical. However, it should be highlighted that due to this well known GPS inefficiency, many mobile OSes are using GPS in combination with wifi networks or even the signal from mobile carriers to further improve the position accuracy.

As we have shown, allowing users to arbitrarily test their distances with other users is not a good policy, even if the results are somehow obfuscated. The optimal, in terms of privacy, would be not to disclose any kind of information about the location of the users. However, this would cripple the user engagement that these applications are trying to get via displaying it. A first measure is definitely generalizing the results in terms of "far", "close" etc, without quantifying the actual distance. However, even this measure is not sufficient. An adversary could exploit the change from one category to the other, just as discussed in the previous section. Therefore, the best policy would be to fuzz the results in order to report more random distances between the users. If users could decide on the amount of fuzziness, they could provide their desired level of privacy and create safe "areas" or a specific radius of tracing.

Private proximity schemes, such as [28, 24], could also be considered. These schemes allow two parties to exchange privately whether they are close or not, without disclosing any further information to each other, the server or any eavesdropper. The adoption of these schemes is not very straight forward, as these

[2] `http://www.gps.gov/systems/gps/performance/accuracy/`

schemes require the two parties to have some sort of trust to each other, which translates to key exchange. Therefore, dating sites that operate with arbitrary users that do not already know each other and just want to flirt does not fit well within this application scenario. Nevertheless, it would be worthwhile to consider escalating information, so that proximity for instance is only disclosed to authenticated "friends" and not to all subscribed users. This approach could certainly provide more privacy to the users while limiting the computational effort significantly.

6 Conclusions

The quest for finding one's other half leads many people to use online dating applications. While this engages people to another way of communication, they are exposed in many ways, mainly due to the nature of Internet. However, as we highlighted in this work, users are exposed to an additional risk due to the location awareness of the smart phone applications. By spoofing his location, a malicious user can manipulate these applications in order to disclose the actual location of an honest user using simple and well known methods. This way, not only private information can be leaked, but cyber-stalking can become real-life, automated stalking exposing users even physically. It is shown that currently applied methods, even if they attempt to somehow obfuscate the results, fail and that a user's actual location can be disclosed very accurately.

Acknowledgments

This work was supported by Science Foundation Ireland under the Principal Investigator research program 10/IN.1/I2980 "Self-organizing Architectures for Autonomic Management of Smart Cities".

References

1. Zickuhr, K.: Location-based services. Pew Internet and American Life Project (2013)
2. Madden, M., Lenhart, A., Cortesi, S., Gasser, U.: Teens and mobile apps privacy. Pew Internet and American Life Project (2013)
3. Gedik, B., Liu, L.: Protecting location privacy with personalized k-anonymity: Architecture and algorithms. Mobile Computing, IEEE Transactions on **7**(1) (2008) 1–18
4. Beresford, A.R., Stajano, F.: Location privacy in pervasive computing. Pervasive Computing, IEEE **2**(1) (2003) 46–55
5. Cheng, R., Zhang, Y., Bertino, E., Prabhakar, S.: Preserving user location privacy in mobile data management infrastructures. In: Privacy Enhancing Technologies. (2006) 393–412
6. Damiani, M.L., Silvestri, C., Bertino, E.: Fine-grained cloaking of sensitive positions in location-sharing applications. IEEE Pervasive Computing **10**(4) (2011) 64–72

7. Smith, A., Duggan, M.: Online dating and relationships. Pew Internet and American Life Project (2013)
8. Patsakis, C., Asthenidis, A., Chatzidimitriou, A.: Social networks as an attack platform: Facebook case study. In: ICN. (2009) 245–247
9. Brown, G., Howe, T., Ihbe, M., Prakash, A., Borders, K.: Social networks and context-aware spam. In: Proceedings of the 2008 ACM conference on Computer supported cooperative work. CSCW '08 (2008) 403–412
10. Huber, M., Mulazzani, M., Weippl, E., Kitzler, G., Goluch, S.: Friend-in-the-middle attacks: Exploiting social networking sites for spam. Internet Computing, IEEE **15**(3) (2011) 28–34
11. Abu-Nimeh, S., Chen, T., Alzubi, O.: Malicious and spam posts in online social networks. Computer **44**(9) (2011) 23–28
12. Huber, M., Mulazzani, M., Weippl, E., Kitzler, G., Goluch, S.: Exploiting social networking sites for spam. In: Proceedings of the 17th ACM conference on Computer and communications security. CCS '10 (2010) 693–695
13. Jagatic, T.N., Johnson, N.A., Jakobsson, M., Menczer, F.: Social phishing. Communications of the ACM **50**(10) (2007) 94–100
14. Douceur, J.R.: The sybil attack. In: Peer-to-peer Systems. Springer (2002) 251–260
15. Backstrom, L., Dwork, C., Kleinberg, J.: Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. In: Proceedings of the 16th international conference on World Wide Web, ACM (2007) 181–190
16. Hoffman, K., Zage, D., Nita-Rotaru, C.: A survey of attack and defense techniques for reputation systems. ACM Computing Surveys (CSUR) **42**(1) (2009) 1
17. Krumm, J.: Inference attacks on location tracks. In: Pervasive Computing. Springer (2007) 127–143
18. Gambs, S., Killijian, M.O., del Prado Cortez, M.N.: Show me how you move and i will tell you who you are. Transactions on Data Privacy **4**(2) (2011) 103–126
19. Gambs, S., Killijian, M.O., del Prado Cortez, M.N.: De-anonymization attack on geolocated data. In: TrustCom/ISPA/IUCC. (2013) 789–797
20. Golle, P., Partridge, K.: On the anonymity of home/work location pairs. In: Pervasive Computing. Springer (2009) 390–397
21. Pontes, T., Vasconcelos, M.A., Almeida, J.M., Kumaraguru, P., Almeida, V.: We know where you live: privacy characterization of foursquare behavior. In: UbiComp. (2012) 898–905
22. Kostakos, V., Venkatanathan, J., Reynolds, B., Sadeh, N., Toch, E., Shaikh, S.A., Jones, S.: Who's your best friend?: Targeted privacy attacks in location-sharing social networks. In: Proceedings of the 13th International Conference on Ubiquitous Computing. UbiComp '11 (2011) 177–186
23. He, W., Liu, X., Ren, M.: Location cheating: A security challenge to location-based social network services. In: Distributed Computing Systems (ICDCS), 2011 31st International Conference on, IEEE (2011) 740–749
24. Mascetti, S., Freni, D., Bettini, C., Wang, X.S., Jajodia, S.: Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies. The VLDB journal **20**(4) (2011) 541–566
25. Zheleva, E., Getoor, L.: To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In: Proceedings of the 18th international conference on World wide web, ACM (2009) 531–540
26. Carbunar, B., Rahman, M., Pissinou, N., Vasilakos, A.: A survey of privacy vulnerabilities and defenses in geosocial networks. Communications Magazine, IEEE **51**(11) (November 2013) 114–119

27. Ruiz Vicente, C., Freni, D., Bettini, C., Jensen, C.S.: Location-related privacy in geo-social networks. Internet Computing, IEEE **15**(3) (2011) 20–27
28. Narayanan, A., Thiagarajan, N., Lakhani, M., Hamburg, M., Boneh, D.: Location privacy via private proximity testing. In: Network & Distributed System Security Symposium. (2011)