Understanding the Privacy Risk of Campus-wide Wireless Geolocation Information

Joseph W. Beckman and Filipo Sharevski

Center for Education and Research in Information Assurance and Security,
Purdue University
655 Oval Drive, West Lafayette, IN, USA
{beckmanj,fsharevs}@purdue.edu
http://www.cerias.purdue.edu/

1 Introduction

In the United States, colleges and universities that receive federal funding must comply with federal regulations that grant the student, or the parents of students under the age of 18, control of the disclosure of his or her personally identifiable information from education records (Federal Education Rights and Privacy Act, 34 CFR Part 99 (1974)). Similar protections extend to student health information under the Health Insurance Portability and Accountability Act (HIPAA), 45 CFR 160-16, (1996). Data obtained through the connection of devices to campus networks is not similarly regulated.

The advent of large campus Wireless Local Area Networks (WLAN)s that are able to integrate many network devices and through software, in conjunction with the explosion of portable devices that connect to these WLANs create a new and unregulated source of personal data about students. Location data is one such data stream that can result from ubiquitous campus WLAN connectivity to personal mobile devices. Because regulation of campus wireless services in the United States falls to the universities themselves, universities can use some of students' most sensitive data without governmental oversight. Student location data is sometimes transferred by universities to local police in order to help locate individuals connected to (or recently connected to) the campus network. A lack of regulation of such uses leaves universities free to collect and choose how much data to share with other organizations, and under what circumstances they will do so. In order to fully understand the scope of the threat to student privacy posed by portable device connectivity to campus-wide WLANs, it is important to understand the regulations governing the use of geolocation data obtained from campus WLANs, the extent of the accuracy of that data, and the potential applications it may serve.

2 The Regulatory Environment Governing Campus Wireless LANs

Use of WLAN networks and the data that they generate are the purview of the network owner, including unversities that own campus-wide wireless connectivity

infrastructure. This principle can be well illustrated for purposes of this inquiry by dissecting recent judicial rulings in lawsuits attempted to leverage existing federal and state laws to enforce privacy protections against telecoms or device manufacturers. This section will show how the Computer Fraud and Abuse Act, and the Stored Communications Act and Wiretap Act sections of the Electronic Communications Privacy Act, all fail to regulate the use of geolocation data obtained by telecoms including wifi network infrastructures.

The Computer Fraud and Abuse Act (CFAA) prohibits unauthorized access to computers. In lawsuits citing the CFAA as protection from Apple's collection of geolocation data from iOS devices, the courts have found that users gave Apple the authority to collect geolocation data from their phones when they install software that uses geolocating features or simply by buying a device on which geolocation functionality was natively installed[1]. As such, Apple's access in these is not unauthorized. Further, unless claimants can prove that the collection of geolocation data harmed them tangibly[2], courts will not allow their claims to stand. Under this interpretation of the law, the courts effectively leave users with the option not to connect their mobile devices to telecom or WLAN infrastructures under whatever terms cellular or WLAN providers offer, or simply not connect their devices at all.

The Stored Communications Act (SCA) prohibits unauthorized access to a facility that provides telecommunications service including intentionally exceeding authorized access in order to obtain, change, or make unavailable to authorized users electronic communications that are in electronic storage[3]. This Act does not apply to university WLANs because in curcumstances of normal operation the university owns the communications "facilities"; so, access to geolocation data is not being accessed without authorization unless it is hacked or stolen from the university. Also part of the larger Electronic Communications Privacy Act, the Wiretap Act states that anyone who has their electronic communications intercepted, disclosed, or intentionally used in ways that violate the terms of the Wiretap Act has standing to make a civil claim against the violating entity[4]. However, U.S. federal courts have decided that only intercepts that allow the intercepting party to learn the content of communications are violations of the Act; so, gathering of geolocation data is not covered by the Wiretap Act regardless of who is gathering the data or how it is obtained [5].

Laws of individual states also do not offer de facto protection to mobile device users of geolocation information that their devices may transmit. The state constitution of California, for example, offers privacy protections against both state and non-state actors, but violations must: (1) infringe on a privacy interest protected by law; (2) petitioners must expect reasonably that they are

[1] In re iPhone Application Litigation, 844 F. Supp. 2d 1040 (N.D. Cal. 2012).
[2] Czech v. Wall Street on Demand, Inc., 674 F. Supp. 2d 1102 (D. Minn. 2009).
[3] Ibid.
[4] Ibid.
[5] In re iPhone Application Litigation, 844 F. Supp. 2d 1040 (N.D. Cal. 2012);Yunker v. Pandora Media, Inc., 2013 W.L. 1282980 (2013).

entitled to privacy; and (3) the actions of the invading party must be egregious[6]. These characterizations are broad, but courts have ruled that the gathering and use of geolocation data only violates a users privacy rights under the California Constitution if the behavior is far outside of accepted social norms [7]. State law, therefore, may apply in university cases if the university's use of personal data is egregious, but not otherwise.

U.S. laws governing the collection and use of geolocation data revolve around some common themes. First, legal claims against entities for use of geolocation data are only sustainable if such collection can be proven to harm the party that is described by the collected electronic information[8]. Even if the use of geolocation were to cause harm, law provides collectors and users of geolocation data with exemption from wrongdoing if the harm created is less than the benefit created for society [9]. Court findings that reject the argument that personal information is property also disallow the use of property laws to govern the collection of geolocation data and further restrict users' options for legal relief from collection and use of geolocation data. Finally, courts currently hold that disclaimers of liability, such as those that users sign when they activate cellular service, or that they agree to when downloading applications are valid and protective of vendors when users sue to block data collection[10]. It would appear then that universities are also exempt from liability stemming from the collection or use of geolocation data if they require users to agree to such a disclaimer before allowing them to connect to the campus WLAN.

3 A Rich Dataset Available to Universities

The amount of data that can be recovered from WLANs is well illustrated by [2]. In addition to the protocols used and types of devices used, the authors used various forms of data from 550 access points and 7,000 wireless users to determine that half of wireless users remained "...close to home 98% of the time" [2]. This characteristic of mobile device users suggests that few data points linking a mobile device and location are needed to uniquely identify a user. [3] support the claim that mobile device users can be identified from a small number of location data points by conducting an experiment in two European cities, one Swiss and the other Swedish. In the Swiss city of Borlange, for example, when the researchers sampled data from a set in which the location information was attributed to a user's work or home with p=0.9, the users identity could be determined 65% of the time from only 20 mobile location samples.

[6] In re iPhone Application Litigation, 844 F. Supp. 2d 1040 (N.D. Cal. 2012).
[7] In re iPhone Application Litigation, 844 F. Supp. 2d 1040 (N.D. Cal. 2012); Low v. LinkedIn Corp., 900 F. Supp. 2d 1010 (N.D. Cal. 2012).
[8] In Re Google, Inc. Privacy Policy Liticagation, No. 5: 12-cv-001382-PSG (N.D. Cal. July 15, 2015).
[9] In re iPhone Application Litigation, 844 F. Supp. 2d 1040 (N.D. Cal. 2012).
[10] Ibid.

In the case of campus-wide WLAN, the basic location information the university has for the devices connected on its network includes: *client username*, *IP address*, *MAC address*, *association and disassociation time*, *Access Point name* (to which the client was or is still attached), *Access Point coordinates*, *wireless protocol*, *security protocol*, *Received Signal Strength Indicator (RSSI)*, and *the session traffic volume*. Because campus wireless networks are generally for the private use of university students, staff, and faculty, they often require username and password authentication in order to connect. With the client username as a unique university authentication identifier the access point coordinates and detailed building/floor plans it is easy to determine clients' *general prevalence trace* with a coarse accuracy and precision.

3.1 Pilot Study - A Flagship Public University Campus

In support of the general prevalence proposition, we obtained a sample of real-world WLAN records collected form the campus users of a flagship university in the United States. The sample includes all sessions recorded in a period of 72 hours for which total of around 1.5 milion records are yielded, each containing the aforementioned attributes used in determining the general prevalence. In addition to these, each records includes invormation about the *vendor* and the *device type*. The overal campus area including the interior of the facilities and buildings is covered with 8,432 access points. The average number of users conected simultaneously on the campus WLAN is 33,126 and the average session duration is 756.245 seconds. Each user produces on average 14 records a day and is uniquelly connected to an average of 3.56 access points.

3.2 Potential Applications Threatening Student Privacy

Raw WLAN data fits well the purpose of deriving the mobility paths or the users for a given time period, set of access points, or for abstracting a set of mobility profiles charaterising the user population using the campus WLAN service. Following the *mobility profiler framework* [10], the paths - although to a coarse granlarity level - are constructed from an ordered set of access point identifiers that corrpond to a user's movement between different points of connection. In the case of the WLAN access points, the accuracy and precission of the user's location is within the error margin of 10 meters [3,4,8]; however, the three dimensional overalap of wireless signals from the ordered set of access points coupled with the coordinates and detailed floor plans enable them to be clustered to remove any oscilation problems on the paths. Using the paths of WLAN access points as inputs, the a topology of all the possible paths is constructed, out of which the frequent mobility patterns for each user can be derived separately. For example, one can employ different selection criteria of the paths matching certain mobility profiles). Finally, the WLAN campus user profiles are derived from the previously discovered mobility patterns by adding the time-context information also present in the WLAN data records.

With these patterns of movement represented as an ordered set of access point identifiers and their frequency of occurance, university analysis can be furthered to yield the likelihood of occurance movement patterns on weekends/weekdays, during semsters and off-semester periods, time portions of a day of over academic years. Adding time context information, which can include but is not limited to the time period a campus event (conference, sports, or similar) occured, the acutal locaiton of the access point, or academic calendars and schedules, the moblity profiles yield a signficant amount of information on the campus users' behavior. University authorities can also learn wether users with fellowships or contracts are compliant to the contracted terms (i.e. athletes who are required to take classes), detect outstanding behaviors, or analyize various profile groups present on the campus.

Risks to the security and privacy of geolocation data within WLAN environments are significant and varied. Like regulated electronically stored data, unauthorized access resulting from hacking also puts geolocation data describing users movements at risk. However, the lack of regulation of geolocation data releases WLAN owners and administrators from formally evaluating and working to control the risks of this data in their environments. In contrast, health care entities subject to the Health Information Portability and Accountability Act (HIPAA) Security Rule [7] in the United States are required to perform regular security risk assessments. These assessments evaluate the administrative, technical, and physical controls of health care entities in the context of the personal health information they hold. A similar requirement exists for credit card processors who comply with Payment Card Industry (PCI) standards [5]. Absent a legal framework regulating geolocation data gathered from WLANs, the extent to which it is protected is unclear.

Arguably more importantly, lack of regulation gives WLAN owners the freedom to use geolocation data as they choose, and to share it with whom they choose, with few consequences. At best, the answers to these questions are codified in formal policies and procedures created by the WLAN owners. In other cases, no formal guidance may be offered at all. Anecdotes exist of university technical staff fulfilling real-time requests for information from local law enforcement personnel in order to track an individual who had threatened his own safety on a social media account. This anecdote is concerning not only because it contains a transfer of this very sensitive information to law enforcement personnel outside of the university, but also because the request is processed in real time by low-level security personnel in routine operations. While this event remains an anecdote and the legality of this type of information transfer has not been litigated in court, it appears to be legal according to the framework of U.S. laws discussed earlier. If legal, only organizational policies, procedures, and system activity monitoring would protect WLAN users from potential harm.

It is instructive to draw parallels between adequacy of controls and hypothetical uses of WLAN geolocation data with a 2014 U.S. legal case that involved HIPAA-protected data. In this case, a pharmacist working in a private-sector pharmacy gained access to the personally identifiable information (PII) about

the mother of her husbands infant child. She then distributed the information to her husband and two othersi for use in the paternity suit pending against her husband. The childs mother successfully sued the pharmacists employer for violating her rights under HIPAA, despite the employers defense that they should not be held accountable for the actions of an employee who knowingly violated their workplace policies regarding PII [11]. This case illustrates both the inadequacy of organizational policies and procedures as protection against privacy violations and the need for federal regulation of the use and distribution of sensitive personal information. In order to draw the parallel to WLAN geolocation data, consider the following hypothetical case. In violation of organizational policy, a network analyst for a university begins to pull geolocation data for a rivals WLAN-connected mobile device. The engineer notices a pattern of travel that may be indicative of illegal activity, and informs local law enforcement. In this case, no U.S. law would protect the rival from such a violation of privacy.

3.3 Application and Coverage under European Union Regulation

Clearly, the context of the arguments in this paper are not impacted by the General Data Protection Regulation [1], as the activities are not based within the European Union. Nor are these activites regulated by the US-EU Privacy Shield since the data in question is not transmitted from the EU to the U.S [6]. It is yet unclear the extent to which the General Data Prtection Regulation may apply to a European university engaging in these types of practices. In particular, the exemption for law enforcement activities may limit the amount of liability a European university may hold in the above given example.

4 Conclusion

Advancing wireless networking technology, coupled with an explosion of mobile devices, has created a new challenge to individual privacy by allowing operators of large WLANs to track the movements of those connected to the WLAN across its expanse. In particular, large university operators are able to exploit this new stream of data by forcing users to sign on to a sophisticated WLAN for necessary services and through a lack of laws preventing the use or dissemination of geolocation data. Despite increasing privacy concerns around technology in the U.S. and regulation of sensitive health and academic data, geolocation data from university WLANs remains unregulated. The private nature of a person's movements and location and the ease with which universities can access data about those movements suggest that the United States Federal government should consider regulating this data similarly to health or academic data.

References

1. European Parliament and Council. (2016).: Regulation (EU) 2016/679 General Data Protection Regulation. (2016)

[11] Walgreen Co. v. Hinchy, 21 N.E.3d 99 (Ind. Ct. App. 2014).

2. Henderson, T., Kotz, D. and Abyzov, I.: The Changing Usage of a Mature Campus-wide Wireless Network. Computer Networks, 52(14), pp.2690-2712 (2008)
3. Freudiger, J., Shokri, R., and Hubaux, J. P.: Evaluating the privacy risk of location-based services. Financial Cryptography and Data Security, pp. 31-46 (2011)
4. Lymberopoulos, D., Liu, J., Yang, X., Choudhury, R.R., Handziski, V. and Sen, S.: A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned. Proc. Intl. Conf. on Information Processing in Sensor Networks pp. 178-189. (2015)
5. Payment Card Industry. (2014).: Payment Card Industry Data Security Standards. (2014)
6. United States Department of Commerce. (2016). Fact Sheet: Overview of the EU-U.S. Privacy Shield Framework. (2016).
7. United States Department of Health and Human Services. (2013). HIPAA Security Rule Administrative Simplification. (2013)
8. Yang, C. and Shao, H.R.,: WiFi-based indoor positioning. Communications Magazine, IEEE, 53(3), pp.150-157.(2015)
9. Rea, M., Cordobs de la Calle, H., Giustiniano, D. and Lenders, V.,: Robust WiFi Time-of-Flight Positioning System. Intl. Conf. on Information Processing in Sensor Networks (2016)
10. Bayir, M., Demirbas, M., Eagle N.,: Discovering SpatioTemporal Mobility Profiles of Cellphone Users. World of Wireless, Mobile and Multimedia Networks & Workshops IEEE International Symposium on a, pp. 1-9. IEEE, 2009.