W3C Geolocation API calls for Better User Privacy Protection

Ioannis Krontiris, Andreas Albers and Kai Rannenberg

Chair of Mobile Business and Multilateral Security
Goethe University, Frankfurt, Germany

1 Introduction

The W3C's Geolocation API will be able to standardize rapidly the transmission of location information of users on the Web. However, such sensitive information raises serious privacy concerns - especially in the mobile Internet. Our position is that the introduction of this API has to be complemented with additional means in order to prevent privacy violations originating from combining location data with unique identifying information in web browsers (e.g. cookies).

For this, first we are briefly enumerating classical and new methods used to identify users via browsers. Subsequently, the W3C's Geolocation API is introduced and potential threats to the privacy of mobile users with respect to these identification methods discussed. The paper concludes by offering two distinct solution approaches for the outlined situation.

The issues that we identify fall outside the scope of the API itself, but we believe that privacy should be addressed and highlighted in a more general frame. The large number of implicated stakeholders creates a complicated setting and in order to offer privacy protection to the end user, we need to see also threats that go beyond the responsibilities of each party separately.

2 Traditional and new Privacy Issues of Web Browsers

The following methods and means allow unattended acquisition of personal information about online users without their consent from browsers:

- Cookies: Web Browser Cookies storing a unique user ID.
- Local Shared Objects: Local Shared Objects are also known as flash cookies and offer basically the same functionality as regular cookies. However, flash cookies are typically unaffected by the privacy settings and mechanisms of web browsers (e.g. regular deletion of existing cookies) and therefore become more persistent.
- Document Object Model (DOM) Storage: The DOM Storage concept constitutes the advanced version of regular cookies and was introduced with HTML 5. Basically, it provides the same functionality as regular and flash cookies. However, it offers website

owners a simplified, but sophisticated handling of the stored data as well as increased storage space.

Whereas the outlined means above are traditionally used to distinctly identify a mobile user, recent research in this context shows the possibilities of a new concept:

- Web Browser Footprint: The generation of a web browser footprint consisting of the user agent information (e.g. browsers version, language), installed plug-ins and fonts, etc. allows a pretty accurate tracking of mobile users across multiple visited websites [1]. Consequently, this footprint is unaffected by the privacy setting and mechanism and will be in any case disclosed to services providers.

3 Location Privacy Threats

The W3C Geolocation API provides a JavaScript API to allow web sites to request location information (latitude and longitude coordinates) from web browsers. The browser determines the current location by contacting a third-party location provider and then passes the answer to the requesting web site. In particular, the following stakeholders are implicated in determining and reporting location information through the Geolocation API:

(a) The web site, which requests and receives the location information.

(b) The third-party location provider, which computes the location information.

(c) The web browser, which represents the user.

In general the location information of the user is currently accumulated at two different points (b and c above). Since currently there are only a couple of third-party location providers, location information from all users is concentrated in their servers. Users' location information is also accumulated at the service providers who requested the information in the first place. The use of the same services (e.g. Google Maps) multiple times from the same person and from different locations corresponds to location logs on the service provider.

Even though the specification tries to remain agnostic concerning user's identity, significant risks emerge by the centralized approach taken. What we find critical is the use of unique identifying information to link subsequent location requests back to the same user and build the *mobility path*, *mobility pattern* and *mobility profile* of a particular user [2]. This could allow the extraction of endpoints like home or work place of and individual and eventually lead to his identification.

Web sites can distinguish between clients submitting location information by using identifying information, which we discussed in the previous section. Especially web browser footprints allow the distinguishing of clients even if cookies, local shared objects and DOM storage objects get deleted frequently. In addition, third-party location providers can use the unique identifier of the client to link location information of users. This ID is sent with each request from the browser to the location provider and it remains the same even for two weeks [3].

4 The Privacy by Policies Approach

Doty et al. [4] present a collection of web sites implementing the Geolocation API and notice that most of them do not clearly state the purpose and practices of the collected location information and provide to the users little (if any) control on that information.

Both Google Location Services and Skyhook Wireless Location Services state in their privacy policy that they do not use collected information to identify the user. However, these policies are subject to change at any time without notification and they call for a certain amount of trust from the user's side.

The W3C specification states strict requirements on notice, consent and usage of location information for web sites using the Geolocation API. However, these requirements are not imposed by the way the API works, allowing third parties to use it without conforming to them.

There are several suggestions that the API should make the above requirements functional and make the transmission of privacy preferences from the user and notification from the requesting websites part of the API function calls. While we support the adoption of these proposals in future drafts of the specification, our position is that these measures can only solve parts of the problem and that a different approach is required to address more severe threats.

In general, privacy by policy cannot protect from stronger attackers, who would not be deterred by policies and regulations. A consensus has not been reached in the privacy research community on how realistic these stronger attacker models are. Cryptography researchers and privacy rights organizations tend to agree that we should protect access to location information, making it tamper-proof against both

- malicious hackers with the desire to intrude on other people's privacy, and
- against companies profiling and accumulating users' location information for profit maximization.

Towards this direction we need to incorporate techniques that provide provable privacy guarantees, meaning that even if an attacker has access to the necessary information, no personally identifiable data can be created or recreated with reasonable effort.

5 Suggested Directions: Privacy by Tools

Obfuscation of location data is a technique that has already been proposed to the W3C Geolocation Working Group and is being considered for the second version of the API. This solution, however, can be applied only in cases where the web site does not need the precise latitude and longitude in order to provide its service. In the rest of the cases the user remains unprotected. Moreover, this solution does not solve the problem of the accumulation of location information at the third-party location providers.

We believe that we need to build more tools that emphasize on the control of user's privacy locally, on the mobile phone. Privacy implications begin by sending sensitive data from the mobile phone to third parties, therefore we can better control our own privacy, by monitoring and controlling this information before it is sent out.

Monitor: A background process in the browser of the user, which keeps track of the information sent from the mobile phone, can be turned into a monitor of the privacy "exposure"

of the user. With non-intrusive user interfaces, this tool could warn the user, when his or her privacy is about to be exposed, taking under consideration the particular context and some predefined privacy preferences. Context in this case could be for example, whether the user is at home or work at the time the location is determined, the frequency at which the information is revealed, etc.

Control: An example of control could be on what we described in the above sections, i.e. the possibility of website owners to combine a browser footprint with the location information of mobile users. This could be accomplished by suppressing all or at least parts of unnecessary browser information when at the same time location information is requested by a website. Since mobile users cannot detect or deal with this situation manually, the browser has to provide a (standardized) process for this task.

6 Conclusions

In this paper we have argued that incorporating privacy mechanisms into the Geolocation API itself is not sufficient to protect the privacy of mobile users. Instead, Geolocation API specifications has to be flanked by additional means and requirements for browsers, which support the API.

References

[1] B. Krishnamurthy and C. Wills, "Privacy Diffusion on the Web: A Longitudinal Perspective," in *Proceedings of the 18th International World Wide Web Conference (WWW 2009)*, pp. 541–541, April 2009.

[2] M. A. Bayir, M. Demirbas, and N. Eagle, "Discovering SpatioTemporal Mobility Profiles of Cellphone Users," in *Proceedings of the 10th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2009)*, pp. 1–9, June 2009.

[3] Mozilla Firefox. `http://www.mozilla.com/en-US/firefox/geolocation/`.

[4] N. Doty, D. K. Mulligan, and E. Wilde, "Privacy Issues of the W3C Geolocation API," Tech. Rep. Report 2010-038, UC Berkley School of Information, February 2010.