Exploring and Mitigating Privacy Threats of HTML5 Geolocation API

Hyungsub Kim, Sangho Lee, Jong Kim
Pohang University of Science and Technology (POSTECH)

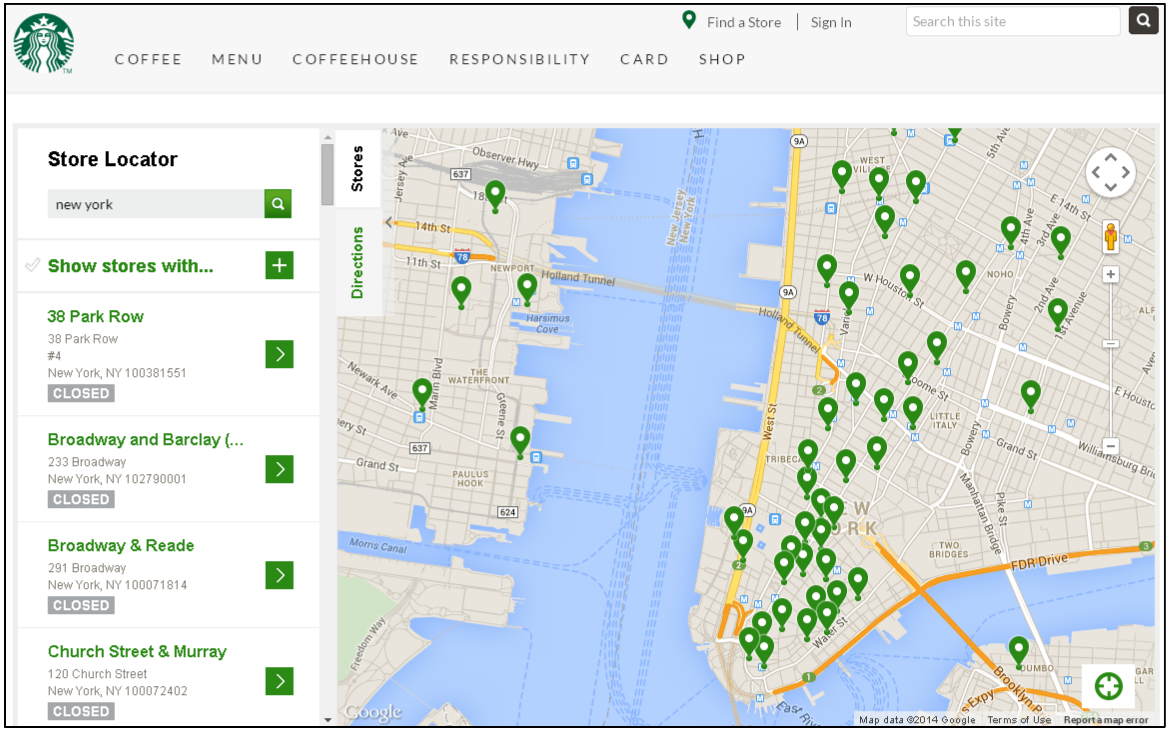
Annual Computer Security Applications Conference (ACSAC) 2014

Presenter: Hyungsub Kim

2014-12-11

Location-based service (LBS)

- **Use real-time geo-data to provide information**
  - Discovering the nearest coffee chains
- **Are used in a variety of contexts**
  - Travel information, proximity-based marketing, fraud prevention



Store Locator

new york

Show stores with...

38 Park Row
38 Park Row
#4
New York, NY 100381551
CLOSED

Broadway and Barclay (...
233 Broadway
New York, NY 102790001
CLOSED

Broadway & Reade
291 Broadway
New York, NY 100071814
CLOSED

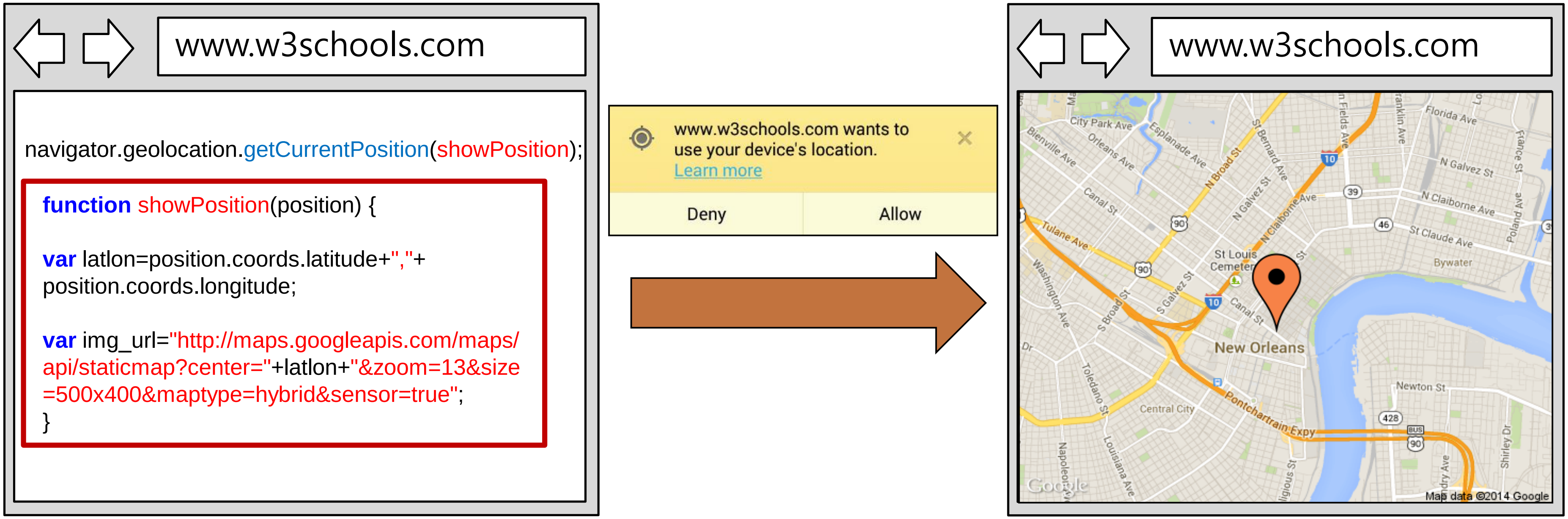
Church Street & Murray
120 Church Street
New York, NY 100072402
CLOSED

http://www.starbucks.com/store-locator/search

Web (HTML5 Geolocation API)

- **Realize LBSs via the Web**
- **Obtain latitude, longitude, altitude, heading, speed**
- **Need to obtain permissions from users**
- **Methods: ① *getCurrentPosition()* ② *watchPosition()***

Obtain user's location



```
navigator.geolocation.getCurrentPosition(showPosition);

function showPosition(position) {

var latlon=position.coords.latitude+","+
position.coords.longitude;

var img_url="http://maps.googleapis.com/maps/
api/staticmap?center="+latlon+"&zoom=13&size
=500x400&maptype=hybrid&sensor=true";
}
```

www.w3schools.com wants to use your device's location.
Learn more

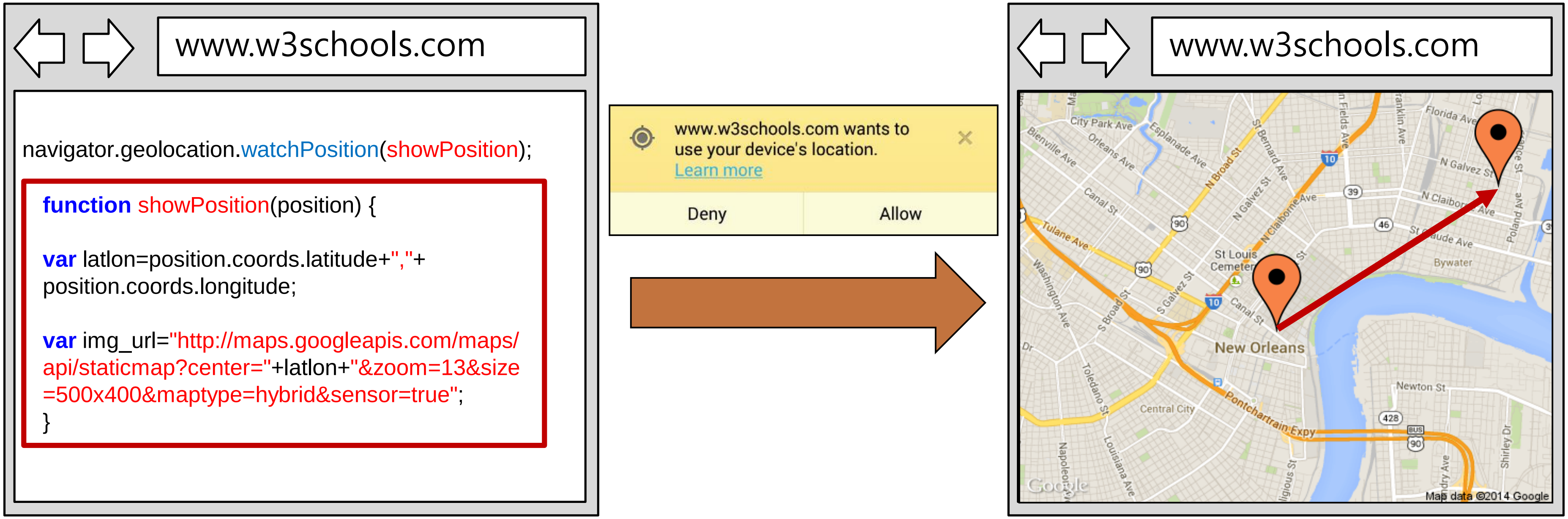
Deny | Allow

<Obtain a Google map image based on the current geolocation>

Web (HTML5 Geolocation API)

- Realize LBSs via the Web
- Obtain latitude, longitude, altitude, heading, speed
- Need to obtain permissions from users
- Methods: ① *getCurrentPosition()* ② *watchPosition()*

Track user’s location

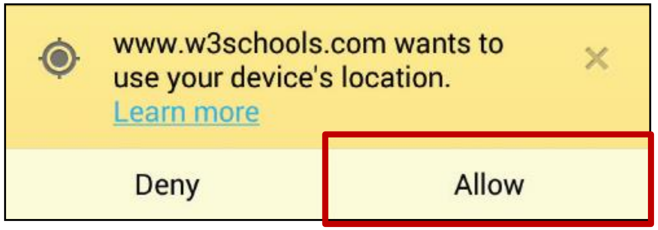


```
navigator.geolocation.watchPosition(showPosition);

function showPosition(position) {

var latlon=position.coords.latitude+","+
position.coords.longitude;

var img_url="http://maps.googleapis.com/maps/
api/staticmap?center="+latlon+"&zoom=13&size
=500x400&maptype=hybrid&sensor=true";
}
```

www.w3schools.com wants to use your device's location.
Learn more

<Obtain a Google map image based on the current geolocation>

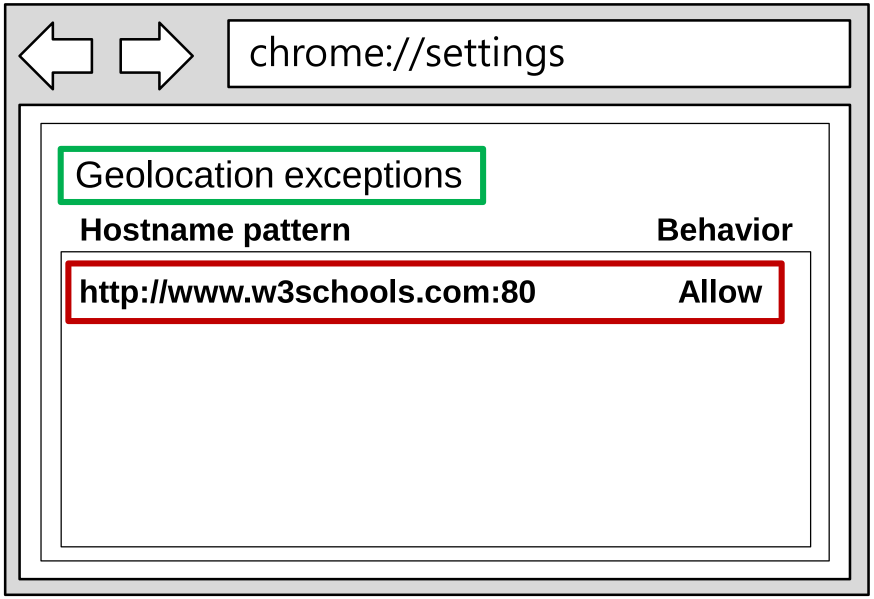
Permission model



www.w3schools.com wants to use your device's location.
Learn more

When a user grants permissions to a web site

All web pages of the web site can access the user's location.



Geolocation exceptions

Hostname pattern	**Behavior**
http://www.w3schools.com:80	**Allow**

<A list of Geolocation permissions on Chrome browser>

<For example>

http://w3schools.com/a.html
http://w3schools.com/b.html
http://w3schools.com/c.html
http://w3schools.com/dir/d.html
.
.
.

Employ a per-domain permission model

Motivation & goal

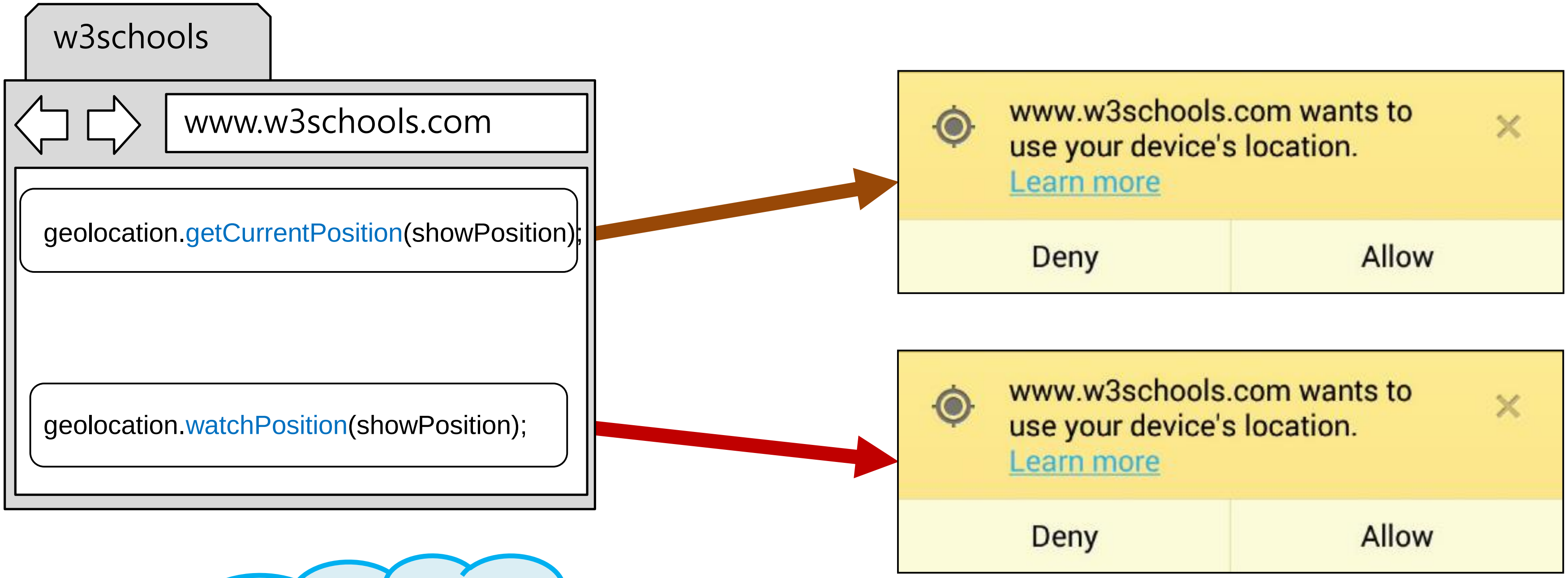
- **Motivation**
  - HTML5 Geolocation API can violate a user’s location privacy.
    - Due to its coarse-grained permission and location models

- **Goal**
  - Explore the current status of the Geolocation API
    - Discovering overprivileged web sites
    - Finding vulnerable web browsers
  - Mitigate the privacy threats of the Geolocation API
    - Supporting fine-grained permission and location models
    - Inspecting the location sensitivity of each web page

Outline

- Privacy threats
- Real-world privacy problems
- Proposed scheme
- Evaluation

(1) No per-method permission



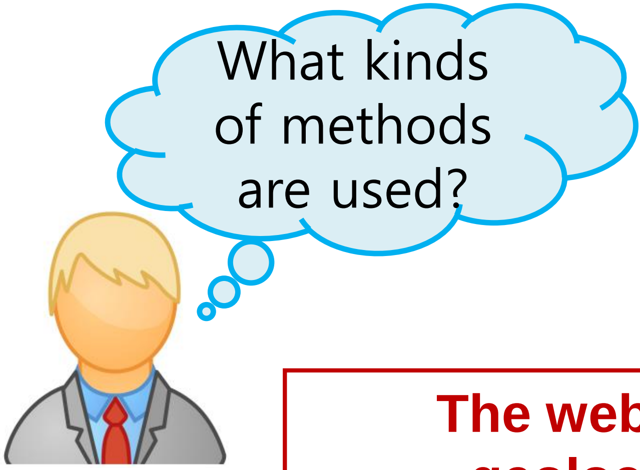
geolocation.getCurrentPosition(showPosition);

geolocation.watchPosition(showPosition);

www.w3schools.com wants to use your device's location.
Learn more

www.w3schools.com wants to use your device's location.
Learn more

< Permission dialogs >

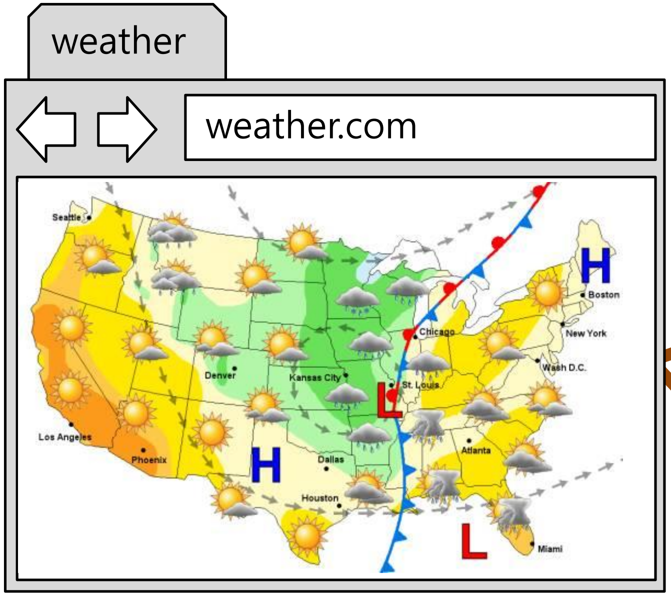


< User >

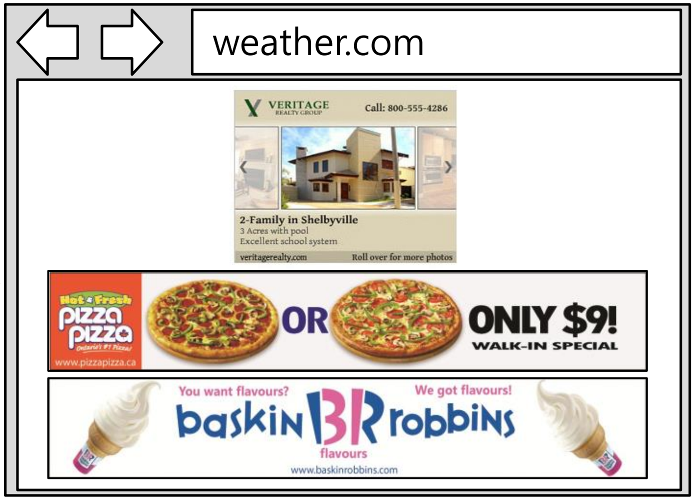
1. Problem: Cannot distinguish the two different methods

The website can silently track the geolocation changes of a user.

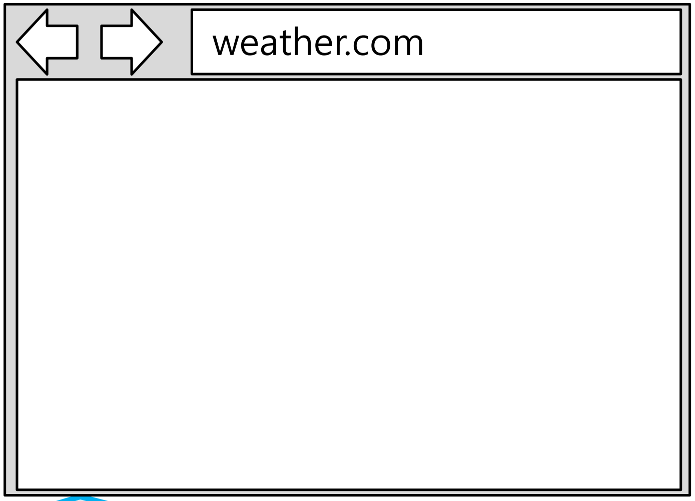
(2) No re-confirm process

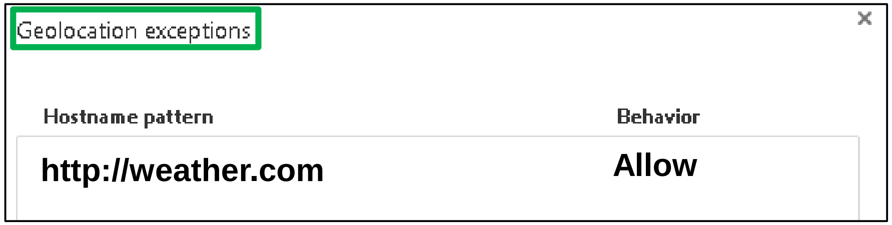


< Modified >



< Deleted >




Geolocation exceptions

Hostname pattern	Behavior
http://weather.com	Allow

2. Problem: Web sites preserve permissions even after their contents are changed.

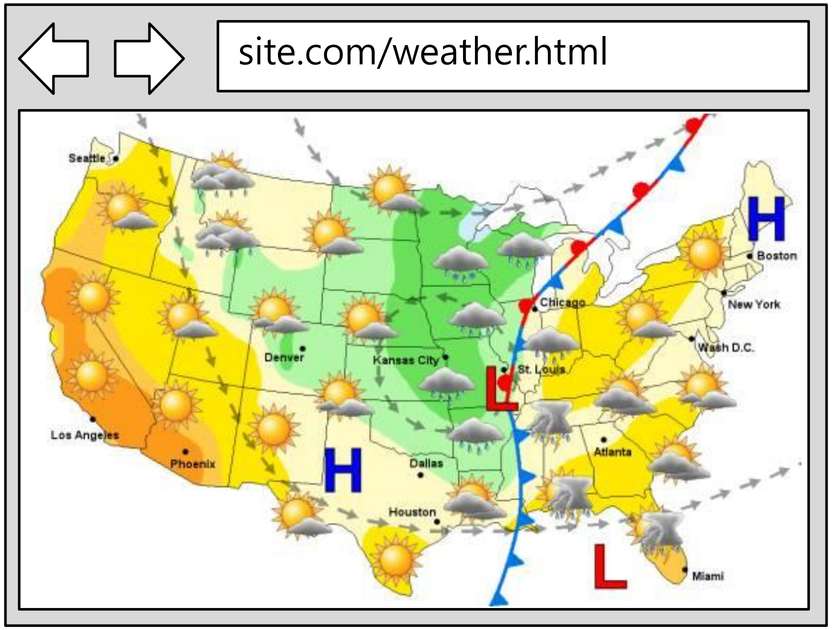
(3) Coarse-grained permission

- Some webpages do not need to access the pinpoint location of users.
  - City- or country-level location may be sufficient.

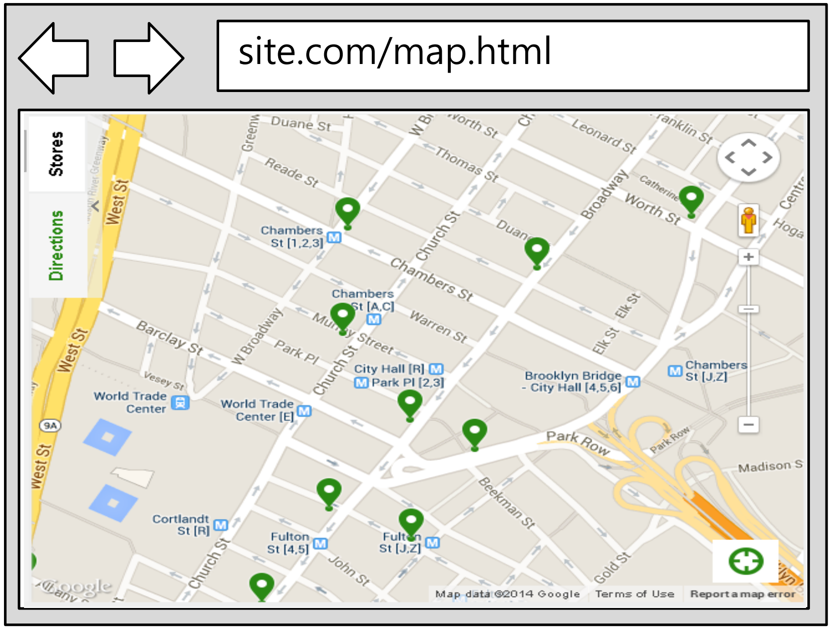
3. Problem: All webpages can access the exact location of a user.

Country-level location granularity

Pinpoint-level location granularity



site.com/weather.html



site.com/map.html

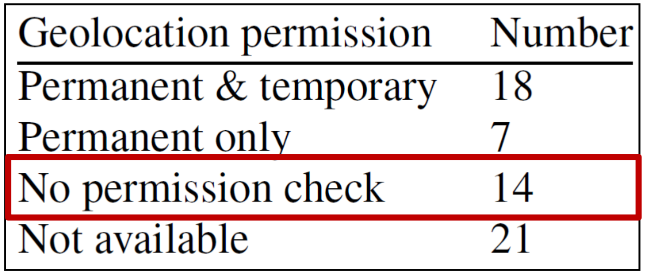
Outline

- Privacy threats
- **Real-world privacy problems**
- Proposed scheme
- Evaluation

(1) How browsers implement API

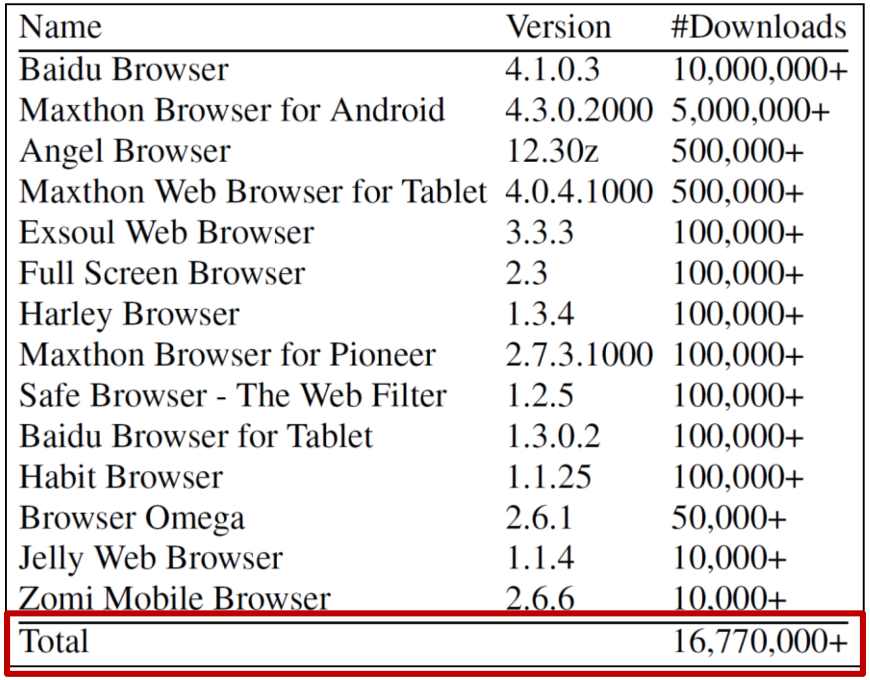
- Investigate top 60 Android web browsers
- Detect 39 web browsers that support the Geolocation API

4. Problem: Do not ask for user permissions



Geolocation permission	Number
Permanent & temporary	18
Permanent only	7
No permission check	14
Not available	21

<How the 60 Android web browsers support the Geolocation API>

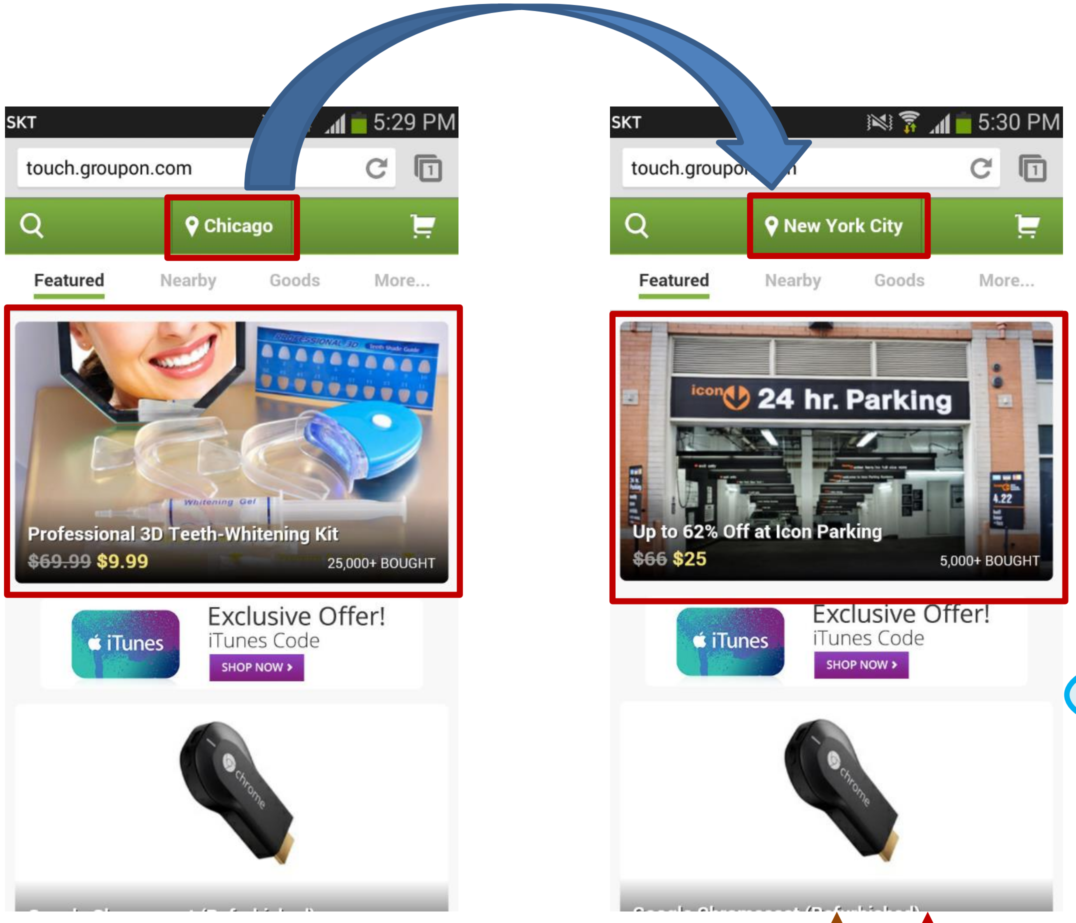


Name	Version	#Downloads
Baidu Browser	4.1.0.3	10,000,000+
Maxthon Browser for Android	4.3.0.2000	5,000,000+
Angel Browser	12.30z	500,000+
Maxthon Web Browser for Tablet	4.0.4.1000	500,000+
Exsoul Web Browser	3.3.3	100,000+
Full Screen Browser	2.3	100,000+
Harley Browser	1.3.4	100,000+
Maxthon Browser for Pioneer	2.7.3.1000	100,000+
Safe Browser - The Web Filter	1.2.5	100,000+
Baidu Browser for Tablet	1.3.0.2	100,000+
Habit Browser	1.1.25	100,000+
Browser Omega	2.6.1	50,000+
Jelly Web Browser	1.1.4	10,000+
Zomi Mobile Browser	2.6.6	10,000+
Total		16,770,000+

<Vulnerable browsers doing not ask for permissions>

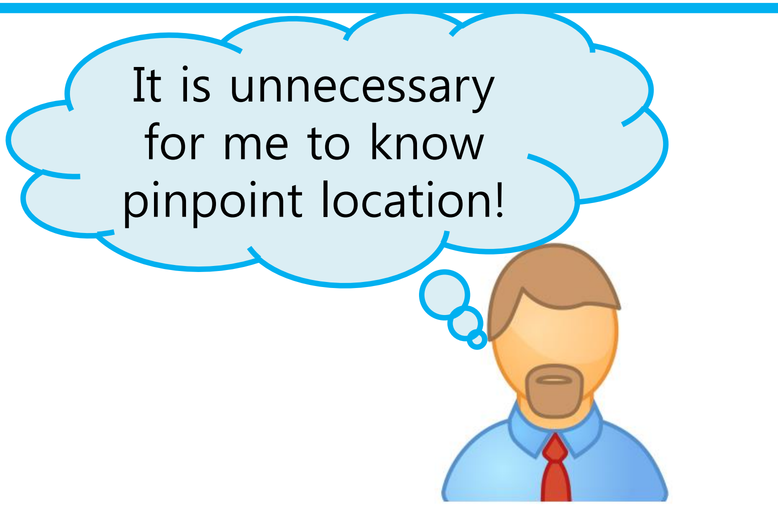
More than 16 million users have installed the 14 vulnerable browsers!

(2) How web sites utilize API



City: New York, NY 10028, USA

Pinpoint: 99-109 E 81st St New York, NY 10028, USA



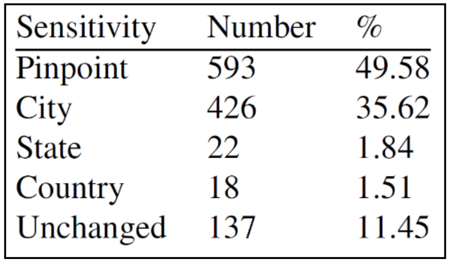
< Website owner >

< User >

(2) How web sites utilize API

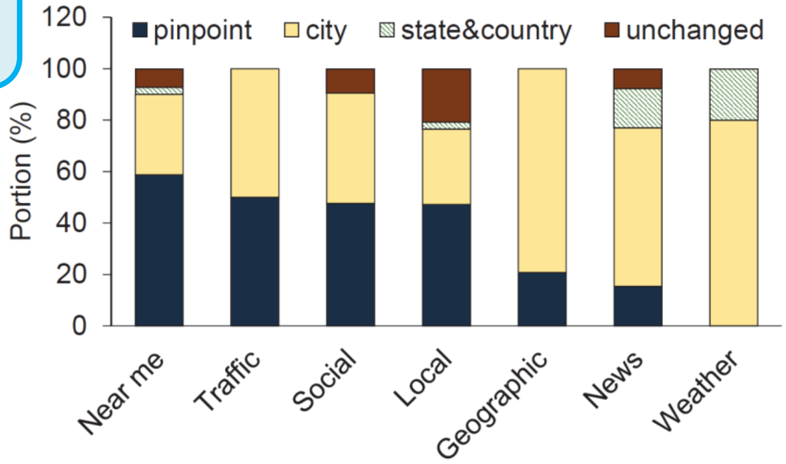
- Collect 1196 web pages that use the Geolocation API
  - Inspect web sites listed on *Alexa.com*
  - Use the Google search engine with keywords (e.g., “near me” and “around me”)
  - Use an HTML code search engine (globalogiq.com)
- Analyze categories and location sensitivity of the web pages

5. Problem: About half of web pages do not need to use exact geolocation



Sensitivity	Number	%
Pinpoint	593	49.58
City	426	35.62
State	22	1.84
Country	18	1.51
Unchanged	137	11.45

<Location sensitivity of web pages using the Geolocation API>



<Location sensitivity according to web page categories>

Outline

- Privacy threats
- Real-world privacy problems
- Proposed scheme
- Evaluation

Design goals

- **To support a fine-grained permission model**
  - Per-method permissions to separate positioning and tracking
  - Per-webpage permissions

- **To support a fine-grained location model**
  - Verify the changes of web pages

- **To recommend appropriate privacy settings**
  - Inspect the location sensitivity of each web page

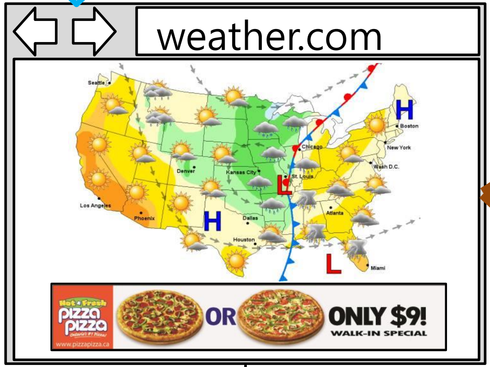
(1) Verify changes of web pages

Proposed scheme
(2/7)

Fuzzy hash value:
c98e8c29770f9

< stored hash value >

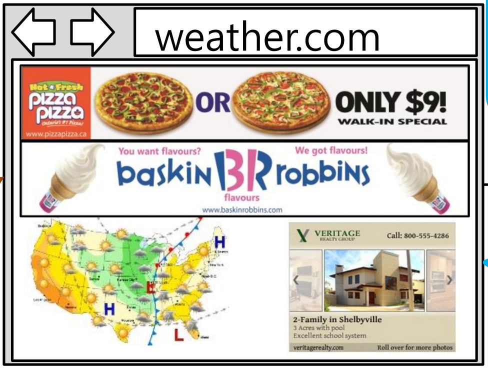
< The first visit >



< Modified >

< Deleted >

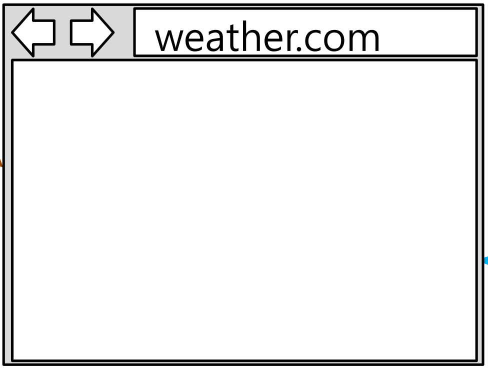
< The second visit >



4) Compare the new hash value with the stored one

Fuzzy hash value:
a1169830770b7837ef8

< new hash value >



Fuzzy hash value:
3a2290

< new hash value >

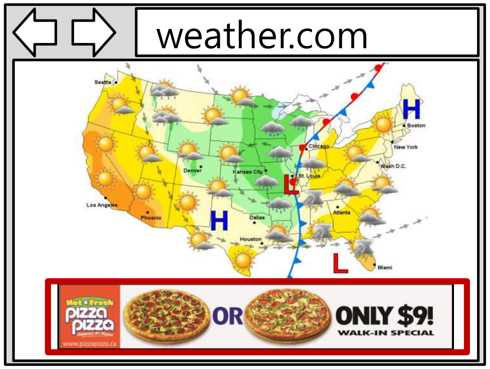
1) Extract the resources of the page
2) Compute a hash value of them
3) Store the value in DB

Browser ask for user permissions again!

How to remove dynamic contents?

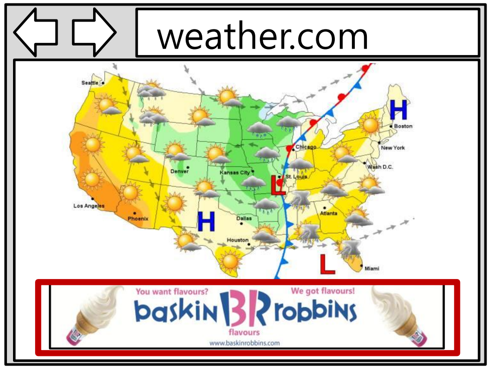
(1) Verify changes of web pages

①



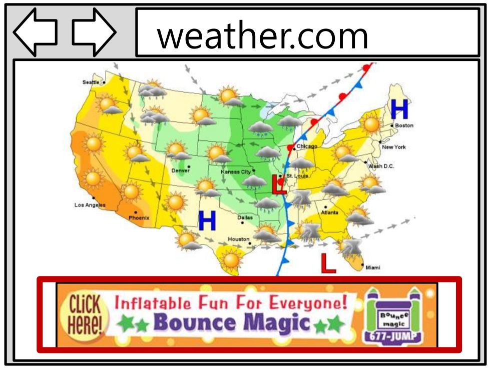
2) Inspect changed contents over the visits

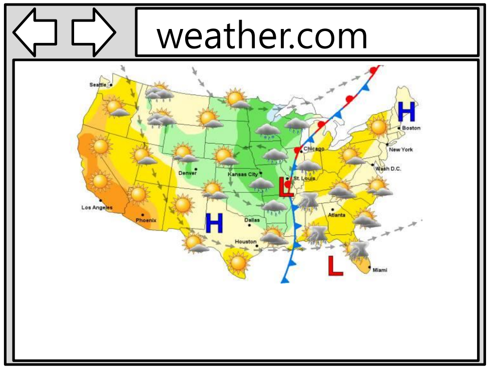
②



1) Access the same page at the same time

③





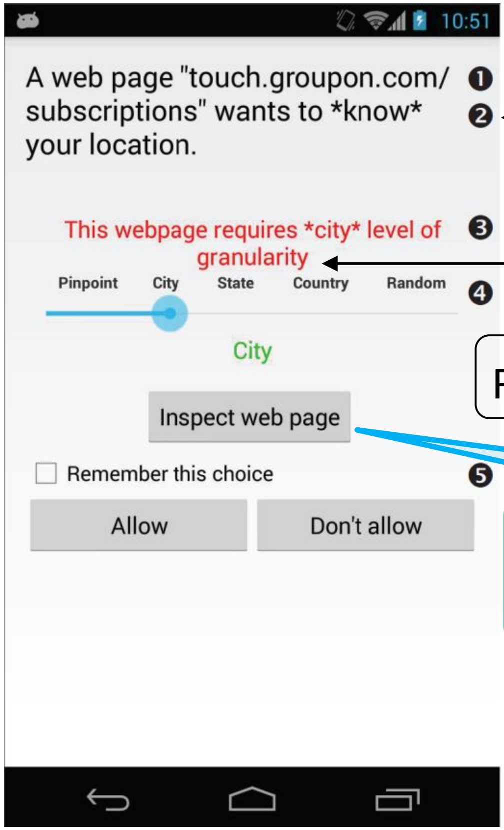
3) Compute a hash value while excluding the dynamic content

Time

New permission dialog

Per-page: touch.groupon.com/nearby
Per-domain: touch.groupon.com

Per-page permissions



A web page "touch.groupon.com/ subscriptions" wants to *know* your location.

This webpage requires *city* level of granularity

☐ Remember this choice

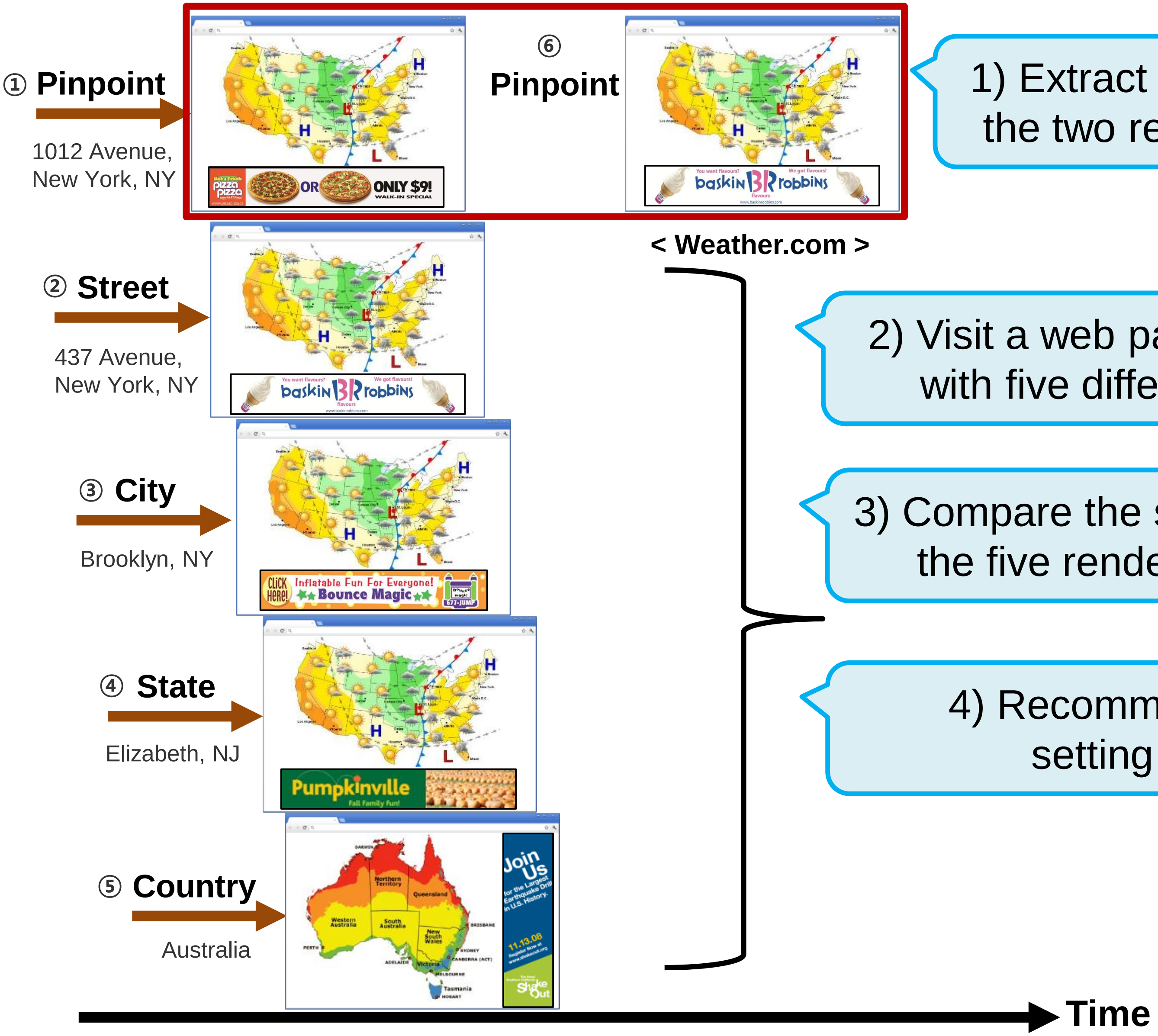
know: geolocation.getCurrentPosition();
track: geolocation.watchPosition();

Per-method permissions

Recommend appropriate privacy setting

Inspect location sensitivity of the webpage

(2) Inspect location sensitivity



① Pinpoint

1012 Avenue,
New York, NY

⑥ Pinpoint

< Weather.com >

② Street

437 Avenue,
New York, NY

③ City

Brooklyn, NY

④ State

Elizabeth, NJ

⑤ Country

Australia

Time

1) Extract static contents from the two rendered web pages

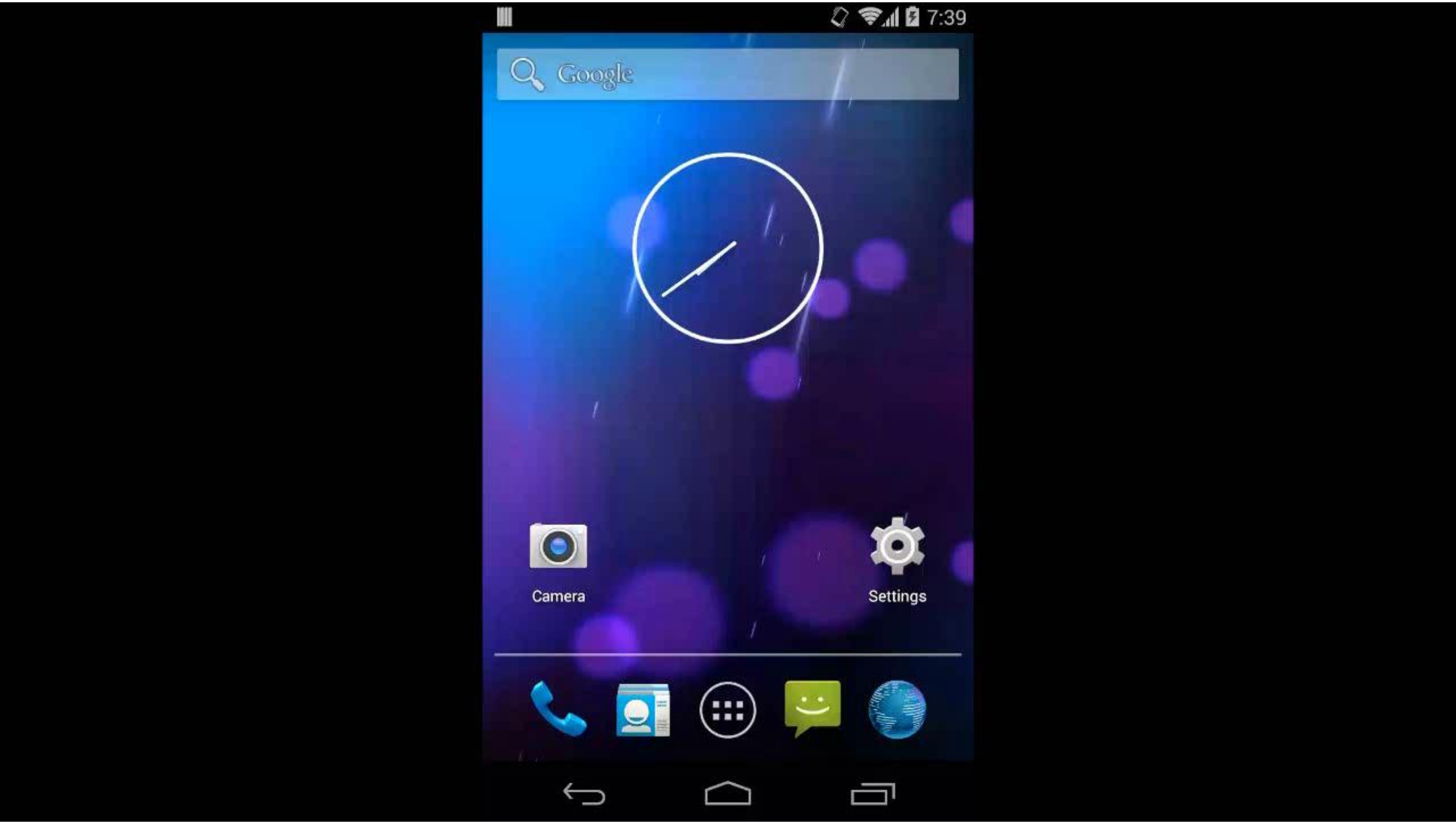
2) Visit a web page multiple times with five different addresses

3) Compare the static contents with the five rendered web pages

4) Recommend a privacy setting to a user

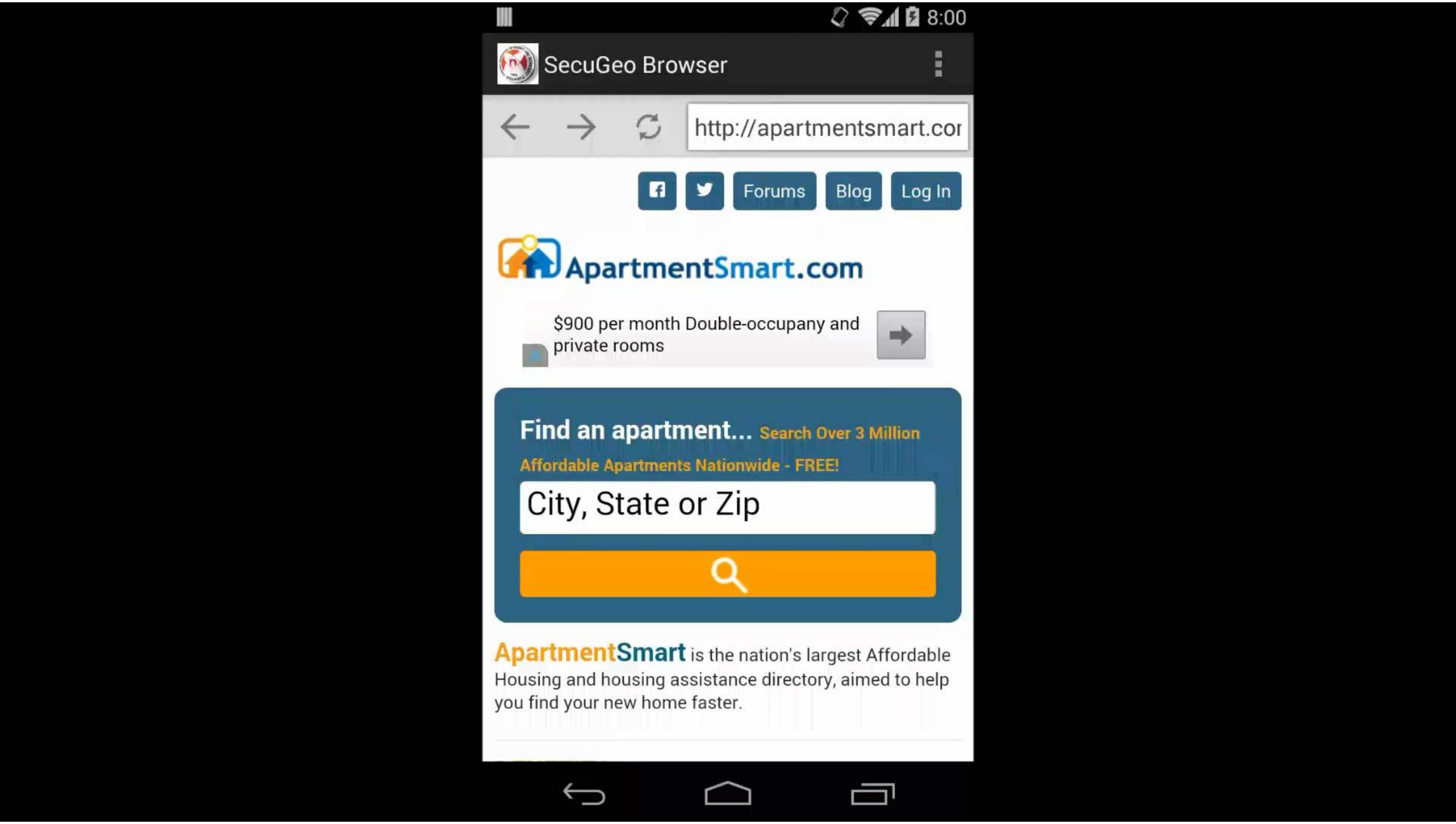
Demonstration videos

1. http://www.tgifridays.com



Demonstration videos

2. http://apartmentsmart.com



$900 per month Double-occupany and private rooms

Find an apartment... Search Over 3 Million Affordable Apartments Nationwide - FREE!

City, State or Zip

ApartmentSmart is the nation's largest Affordable Housing and housing assistance directory, aimed to help you find your new home faster.

Outline

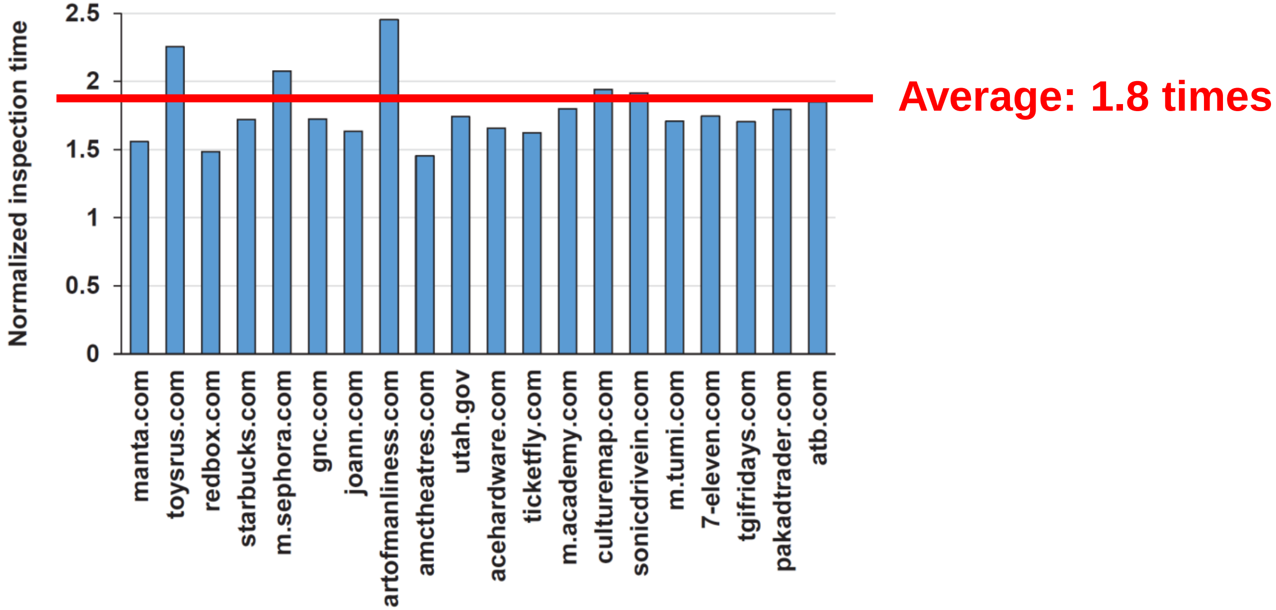
- Privacy threats
- Real-world privacy problems
- Proposed scheme
- Evaluation

Accuracy

- **Randomly choose 200 web pages**
- **Correctly estimate the location sensitivity**
  - Comparing the results of our scheme with manual inspection
  - Accuracy: **93.5%**

Time overhead

- Choose Top 20 web pages using the Geolocation API from *Alexa.com*
- Measure time to inspect location sensitivity
  - About 1.8 times longer than the page loading time
- Is negligible because it happened just one time.
  - Only when the browser visits the web page for the first time.



<Inspection time normalized to page loading time>

Conclusion

- **Consider the privacy problems of the Geolocation API**
  - Point out coarse-grained permission and location models
  - Detect vulnerable web browsers and overprivileged web sites
- **Propose a scheme to enhance the Geolocation API**
  - Support fine-grained permission and location models
  - Recommend appropriate privacy settings according to automatic location sensitivity inspection

Thank you

Are there any questions?