Location and Geocoding

Web Architecture and Information Management
Spring 2009 — INFO 190-02 (CCN 42509)

Erik Wilde, UC Berkeley School of Information
2009-04-13

Contents

• Abstract ........................................................................ 2
• Location and the Web .................................................... 3
• 1 Location on the Web .................................................. 5
  ◦ Location and the Internet ............................................. 5
  ◦ Determining Location ................................................ 6
  ◦ Wi-Fi vs. Cell Location ............................................... 7
  ◦ Eye-Fi Dissected ....................................................... 8
• 2 GeoRSS ......................................................................... 10
  ◦ Feeds are Containers ................................................ 10
  ◦ GeoRSS (and other geolocation Formats) ..................... 11
• 3 GeoRSS Applications ................................................ 13
  ◦ Flickr and Geocoding ............................................... 13
  ◦ Exif in Camera Images .............................................. 14
  ◦ Regular Flickr Image ............................................... 15
  ◦ Geocoding in Flickr ................................................ 16
  ◦ Geocoding Problems ............................................... 17
  ◦ Getting GeoRSS from Flickr ..................................... 18
  ◦ Flickr GeoRSS in Google Maps .................................. 19
• Conclusions ..................................................................... 20

This work is licensed under a CC Attribution 3.0 Unported License
Abstract

Location currently is not a concept that is supported by the Web itself, but there are many location-aware applications available on the Web. Furthermore, the increasing availability of mobile devices and mobile Internet connectivity will turn location into an increasingly important concept on the Web. This lecture looks at some of the question how location information can be obtained, and how it is represented in today’s Web standards (such as feeds) and applications (such as Flickr and Google Maps).

Location and the Web

- The Web itself (currently) has no concept of “Location”
- Many applications using the Web use location concepts
- Location is even more important for mobile access and devices
- Location information is sensitive and relevant for privacy
Location on the Web

Location and the Internet

- IP addresses ([Internet Architecture; IP Address (1)]) are globally unique (and can be geocoded) ([http://api.hostip.info/get_html.php?position=true])
- IP addresses are not a very reliable source for location information
  - campus wired or wireless access says I am in "Berkeley, CA"
  - Home access says I am in "Santa Rosa, CA"
  - iPhone Internet access says I am in "London, UK"
- Try this at home:
  1. find out your IP address ([http://www.ipaddressworld.com/]) (it's assigned by the Internet provider)
  2. IANA manages the basic scheme of IP address allocations ([http://www.iana.org /assignments/ipv4-address-space/])
  3. ARIN is in charge of the 128.32 prefix used by UC Berkeley ([http://ws.arin.net/whois/])
- Advanced Internet network configuration is more complicated
  - Network Address Translation (NAT) ([http://en.wikipedia.org/wiki/Network_address_translation]) "hides" IP addresses of devices
  - controlled networks (such as cell phone services) can be hard to figure out

Determining Location

- Location can be determined in three different ways
  1. IP addresses for wired Internet access (available from IP configuration)
  2. ID of an Wi-Fi access point (available from Wi-Fi chip)
  3. ID of a tower of a mobile telephone network (available from phone chip)
  4. (and then there is GPS on some devices)
- Skyhook ([http://www.skyhookwireless.com/]) provides an integrated solution for determining location
- Data collection by "hand" and through cross-matching
  - match GPS coordinates and Wi-Fi/cell IDs
- Apple ([http://www.skyhookwireless.com/inaction/]) and Eye-Fi ([http://www.eye.fi/]) are Skyhook's biggest customers
- Wi-Fi and cell tower IDs require a lot of data
  - 100 million Wi-Fi access points
  - 1 million cell tower IDs
Wi-Fi vs. Cell Location

Eye-Fi Dissected
GeoRSS

Feeds are Containers

- Feeds [Content Syndication] are a format for publishing collections
- The basic feed model supports a rather limited set of features
  - title, subtitle, author, description, content
  - license, links, timestamps
- GeoRSS [$\text{http://georss.org/}$] extends feeds with a simple method for geocoding
  - GeoRSS-Simple [$\text{http://georss.org/simple}$] supports points and simple shapes
  - GeoRSS-GML [$\text{http://georss.org/gml}$] supports the Geographic Markup Language (GML) [$\text{http://en.wikipedia.org/wiki/Geographic_Markup_Language}$]
  - advanced shapes (such as polygons with holes) cannot be expressed in GeoRSS-Simple

GeoRSS (and other geolocation Formats)

```xml
  <title>Uploads from dret</title>
  <link rel="self" href="http://api.flickr.com/services/feeds/photos_public.gne?id=20266194@N00&amp;lang=en-us&amp;format=atom&amp;georss=1"/>
  <link rel="alternate" type="text/html" href="http://www.flickr.com/photos/dret/">
  <id>tag:flickr.com,2005:/photos/public/626282</id>
  <icon>http://farm4.static.flickr.com/3089/buddyicons/20266194@N00.jpg?1229911061#20266194@N00</icon>
  <updated>2009-04-08T00:31:46Z</updated>
  <generator uri="http://www.flickr.com/">Flickr</generator>
  <entry>
    <link rel="enclosure" type="image/jpeg" href="http://farm4.static.flickr.com/3591/3421997673_7c747786b2_o.jpg"/>
    <georss:point>37.871499 -122.257479</georss:point>
    <geo:lat>37.871499</geo:lat>
    <geo:long>-122.257479</geo:long>
    <woe:woeid>55858022</woe:woeid>
  </entry>
</feed>
```
GeoRSS Applications

Flickr and Geocoding

- Flickr supports geocoding
  - early adopters invented their own conventions for geotagging
  - Flickr quickly realized the potential and supported geocoding
  - Flickr allows to control the disclosure of location information [http://www.flickr.com/help/map/?search=location+settings#202]

- Uploaded photos may already have embedded location information
  - location info can be added by Eye-Fi or by camera GPS

- Publishing JPEG can reveal a lot of information
  - it is possible to “cleanse” JPEG before publishing them
  - Flickr has an option to “hide” Exif data
  - Flickr does not strip the Exif data which can still be read
Regular Flickr Image

Geocoding in Flickr
GeoRSS Applications

Geocoding Problems (17)

- Camera position vs. photographed object
- Lens (tele vs. wide angle) make a big difference
- Focus setting is interesting as well

Getting GeoRSS from Flickr (18)

- Flickr by default provides non-geotagged feeds
- Adding "&georss=1" to the feed URI references a GeoRSS feed
- Google Maps reads GeoRSS feeds and displays them on the map
  - Simply enter the feed's URI in the Google Maps search field
**Flicker GeoRSS in Google Maps**

![GeoRSS in Google Maps](image)

**Conclusions**

- Geocoded information allows new ways of visualization
- Geolocation-aware applications become popular and widely used
- GeoRSS is a simple way of expressing location information
- GeoRSS feeds are a good foundation for managing geocoded datasets