Targeted Advertising and Privacy

Yuan Xia

November 7, 2013

CPSC 457/557

Targeted Advertising

❖ Behavioral targeting

- Monitor and track users' behavior to deliver ads that are expected to increase the effectiveness of advertising
- Considered to be the most effective way to reach targeted audience

❖ Touted as a "Win-Win-Win" for Advertisers, Websites, and Users

Targeting Advertising: http://en.wikipedia.org/wiki/Targeted_advertising

Controversial with respect to Privacy

❖ Obvious "Win" for Advertisers and Websites

```
┌────────────┐      ┌────────────┐      ┌────────────┐
│  Targeting │ ───▶ │     Ad     │ ───▶ │ Profitable │
│            │      │Effectiveness│     │            │
└────────────┘      └────────────┘      └────────────┘
```

❖ Users: Win or Lose?

- Win: free tools, software, etc.

- Lose: 1) More and more personal data will be collected

  - (more personal = more valuable, competing nature of the business means aggressive is a must: Google+ for Facebook)

  2) Data collection process is almost invisible to users

  3) Data can be sold in various forms

Privacy Regulation: EU (1)

❖ Web Bugs

- Intentionally invisible – designed to be invisible to users and not stored on user's computer
- Track move and scroll, widely used on commercial websites

- **The use of such devices should be allowed only for <u>legitimate purposes</u>, with the knowledge of the users concerned**

❖ Cookies

- Stored on user's computer; tracking cookies contain records of individual's browsing history

- **should be allowed on the condition that users are provided with clear and precise information in accordance with the Data Protection Directive about the purposes of cookies**

Goldfarb, Avi and Tucker, Catherine, Privacy Regulation and Online Advertising (August 4, 2010).

Privacy Regulation: EU (2)

❖ Clickstream data

- 'Clickstream data' are data that record the webpages a user viewed at a website, how long the user spent on each webpage, the visitor's path through the site (including her points of entry and exit), the visitor's IP address, and the webpage the user viewed immediately before arriving at the website.

- **Clickstream data not explicitly restricted, but have been interpreted as potentially personally identifiable**
  - Static IP
  - Possible to inadvertently collect sensitive data (religion, medical issues, etc.)

Privacy Regulation: EU (3)

❖ Data

- 3.3 million survey responses for 9,596 different online display advertising campaigns conducted on different websites from 2001 to 2008
- 400 different kinds of products on 40 different categories
- 10 different countries represented, 894 campaigns being EU based, and 8,792 representing non-EU data

❖ Result

- Decrease in advertising effectiveness: on average a 65% reduction in Europe, no similar decline in non-European countries
- Impact was higher for "general content" websites
- Much less effect on ads with multimedia components

Privacy Regulation: US

❖ Self-regulation

- Websites set their own policies on ad targeting
- FTC provides guidelines

❖ Reality

- FTC is **considering** moving to directly regulating online targeting in the US
- advertisers against the potential for an increase in privacy protections
- privacy regulation may affect the direction of innovation on the advertising-supported internet
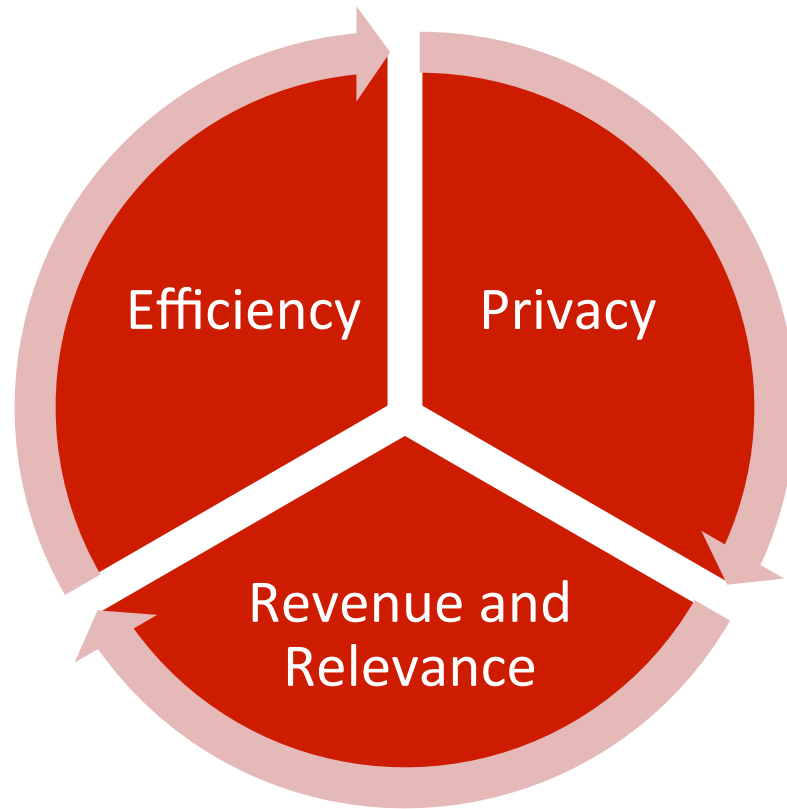
Mobile Age

❖ Mobile devices carry more personal info

❖ **Location**

- Home address, work address, occupation, income, etc.

- Schedule, age, gender, etc.

- Lifestyle
  - Drink
  - Smoke
  - Food preference
  - Race
  - Etc.

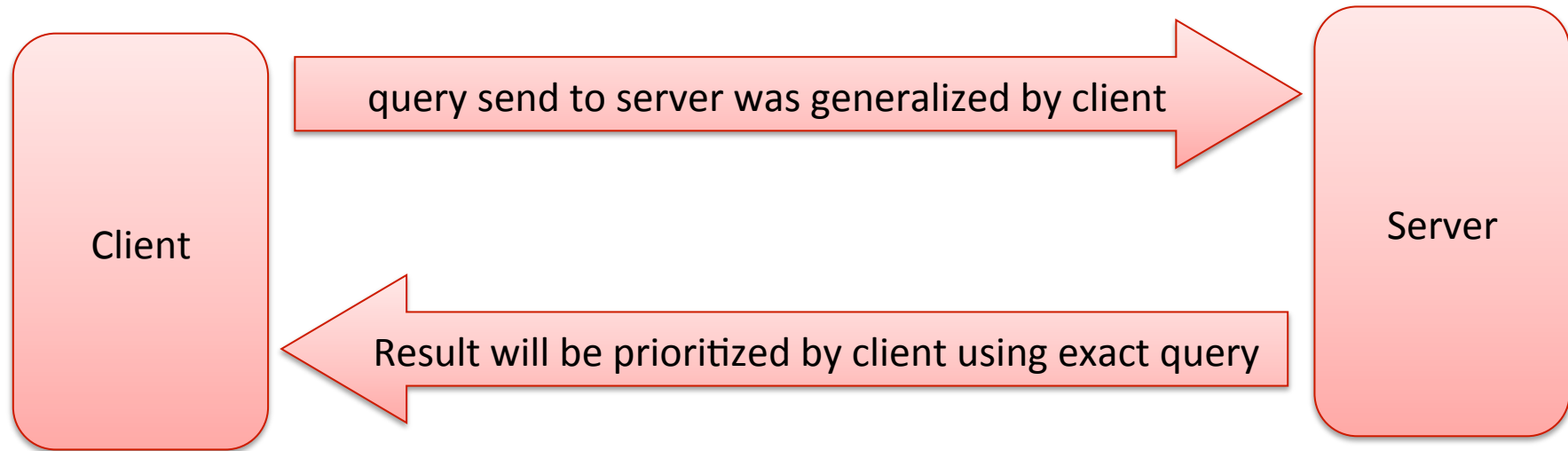# Goal

Privacy-Aware Personalization (1)

❖ Privacy-Aware Ad Delivery

- Keeps personal information on the client device and performs personalization on the client
- Query Generalization

❖ Privacy-Preserving Statistics Gathering

- Add noise

Michaela Hardt and Suman Nath, Privacy-Aware Personalization for Mobile Advertising, 2011

Privacy-Aware Personalization (2)

Client

Server

query send to server was generalized by client

Result will be prioritized by client using exact query

Experiments

❖ Data

- Using a trace of location-aware searches in Microsoft Bing for mobile, consists of 1,519,307 records, more than 116,432 unique user-IDs

- focus on clicks to "Food & Dining" businesses (largest in the trace)

- the first 90% of the trace as training data and the remainder to evaluate the framework and to compute targeted ads

❖ Search generalization

- 5 levels of location generalization (precision of latitude and longitude)

- 3 levels user interest generalization (cheese -> groceries -> food & dining)

- 3 levels of query generalization

Evaluation

❖ Precision: The fraction of targeted ads in our framework on which users actually click. Precision is an indicator of relevance

❖ Coverage: The fraction of contexts for which our framework computes and displays a targeted business
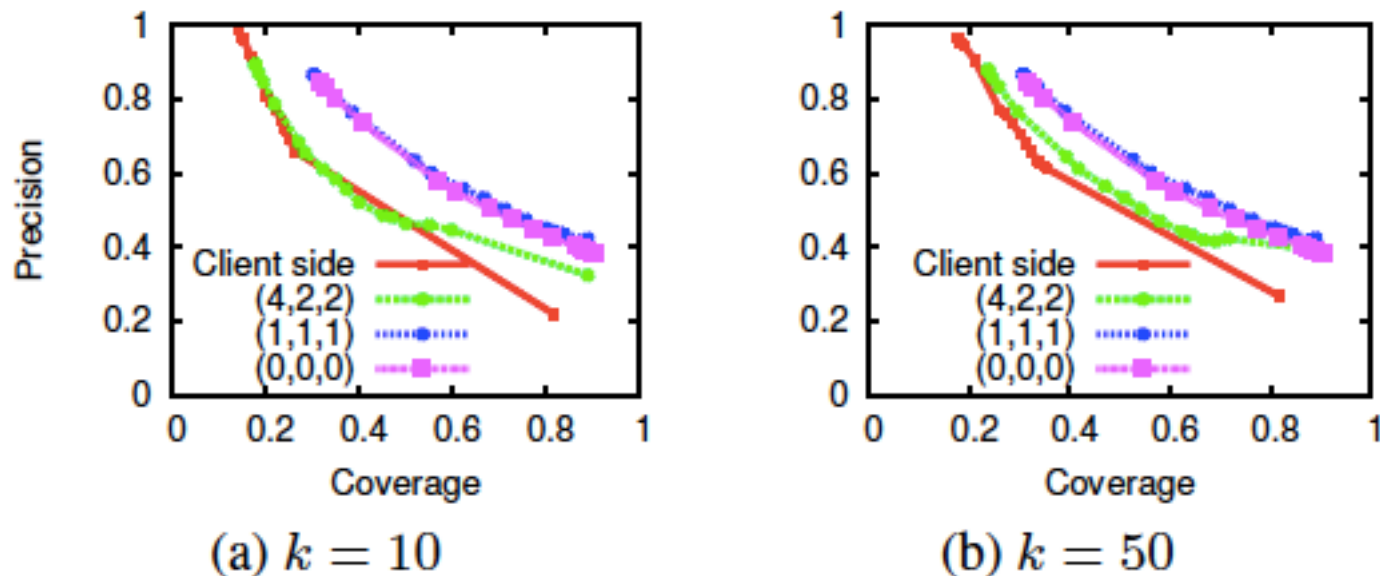
(a) $k = 10$ (b) $k = 50$

Figure 6: Varying information disclosure.

Jawbone Up

Google Glass

Smart Car

Smart TV

Smart Kitchen – Internet of Things

Discussion