HELSINKI UNIVERSITY OF TECHNOLOGY
Networking Laboratory

Measuring Mobile User Behavior and Service Usage:

Methods, Measurement Points, and Future Outlook

L.A. Global Mobility Roundtable 2007

Antero Kivi

2.6.2007

Outline

- Introduction

- Sources of data on mobile service usage

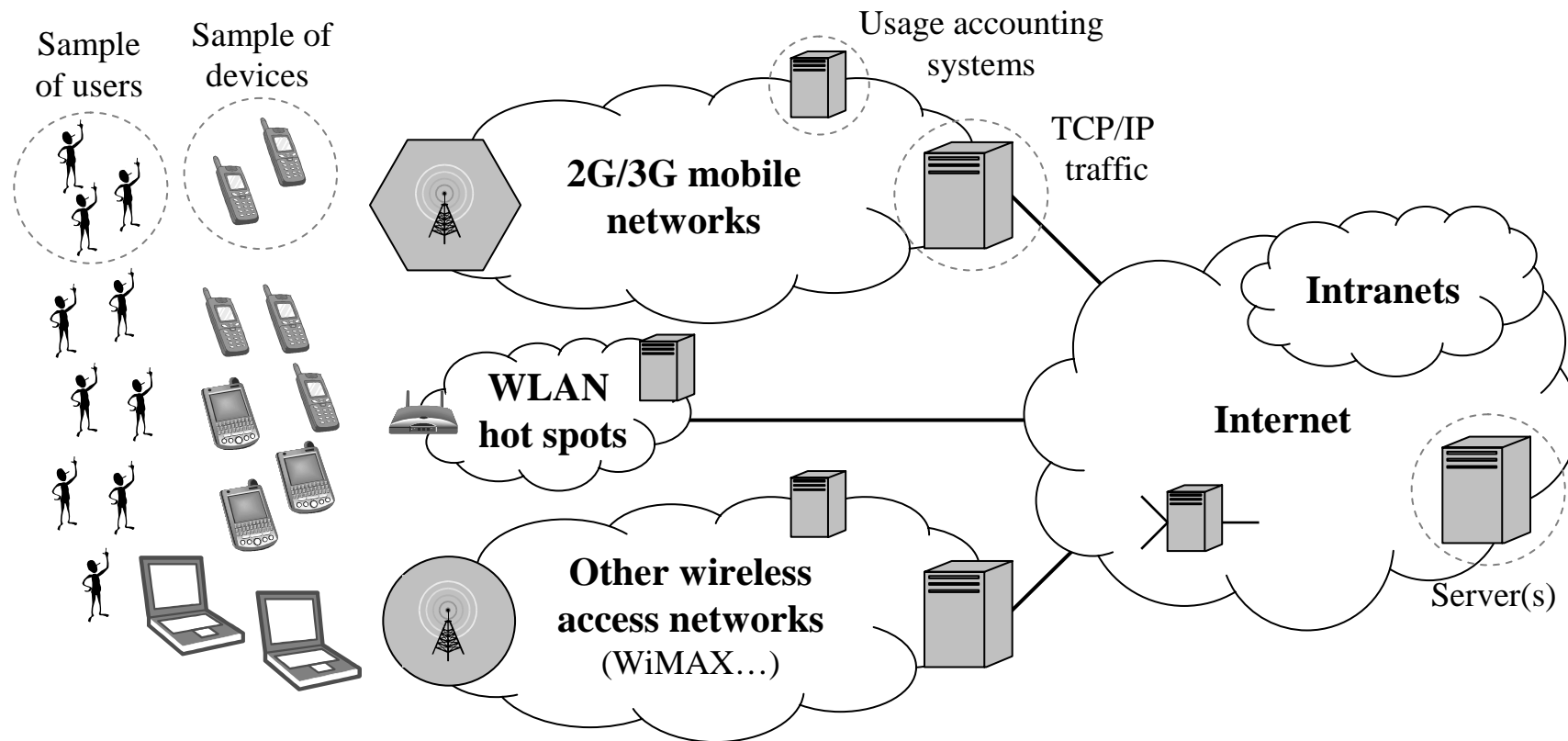- Comparison of alternative methods

- Future outlook

- Conclusions

Introduction

- Why *mobile data service usage*?
  - Reliable and transparent information on end-user behavior valuable to many stakeholders (e.g. marketing, business & product development, academics)
  - …but not often available

- Why compare alternative methods?
  - Lots of research using "real" or "empiric" data on usage of "mobile" services, but with little common ground and comparability
  - Paper purpose to highlight the fundamental differences in data collection methods, to be able to position research in proper context

- Methods: literature study + practical experiences
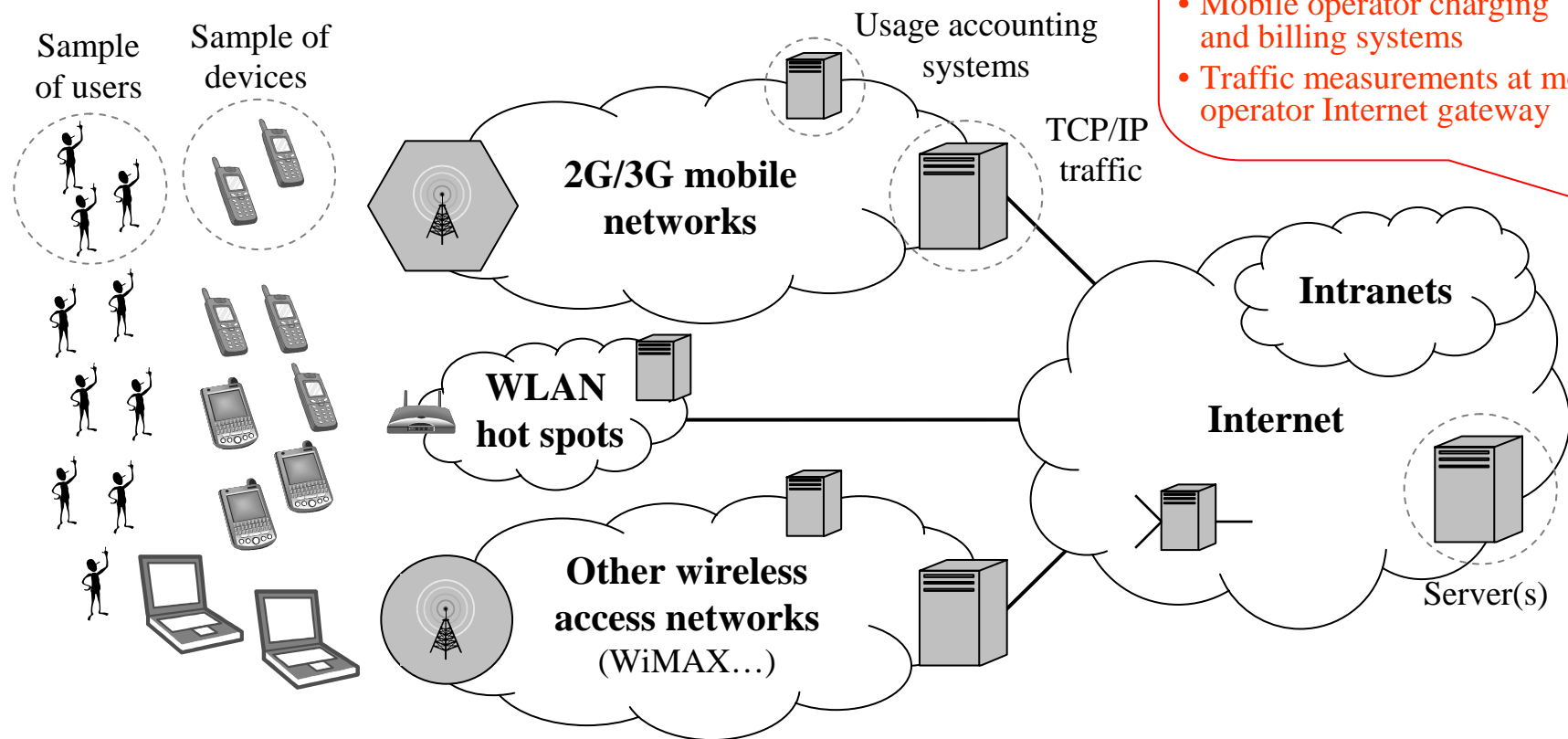
Sources of data on mobile service usage

Sample of users

Sample of devices

Usage accounting systems

2G/3G mobile networks

TCP/IP traffic

WLAN hot spots

Intranets

Internet

Other wireless access networks (WiMAX…)

Server(s)

Sources of data on mobile service usage

Sample of users

Sample of devices

Usage accounting systems

OUR RESEARCH
- Surveys on handset panel participants
- Handset monitoring panels
- Mobile operator charging and billing systems
- Traffic measurements at mobile operator Internet gateway

2G/3G mobile networks

TCP/IP traffic

WLAN hot spots

Intranets

Internet

Other wireless access networks (WiMAX…)

Server(s)

Which methods capture mobile usage?

- **Surveys and panels**
  - Most widely used data collection method
  - Ask a sample of people (telephone, postal mail, email, web, face-to-face…)
- **Mobile handset monitoring**
  - Recruit a group of people (panel)
  - Monitoring software at the handset logs of usage of applications/features
- **Traffic measurements at wireless access networks**
  - Capturing TCP/IP traffic where traffic converges (routers/gateways)
- **Usage accounting systems of wireless access networks**
  - Built-in data collection function, data on usage of chargeable services
- **Some server measurements**
  - Many alternatives (browser cookies, web server logs, search engines, proxy server logs, server usage accounting)
  - Must be able to separate mobile from fixed/PC usage

Method comparison by research scope

Methods / Attributes	Surveys and panels	Handset monitoring	Wireless access network		Server measurements
			TCP/IP traffic measurements	Usage accounting systems	
Researchers with access to data	Anybody (analysts, consulting firms, investment banks, academics…)	Those with a monitoring client and ability to recruit the sample	Wireless access operators	Wireless access / service operators	Service providers (mobile operators, 3rd party providers, search engines…)
Method scope and sample size	Survey respondents: $10 - 10^5$ Panel participants: $10 - 10^3$	Sample of panelists using one handset OS and software platform: $10 - 10^3$	Access operator's subscriber / terminal base: $10^2 - 10^7$	Access operator's subscriber / terminal base: $10^2 - 10^7$	All users of the measured service(s): $10^2 - 10^7$

Method comparison by research scope

Methods / Attributes	Surveys and panels	Handset monitoring	Wireless access network		Server measurements
			TCP/IP traffic measurements	Usage accounting systems	
Researchers with access to data	An... ...ts,	Those with a moni...	Specific industry actors / service providers		S...vice providers ...ors,
	Anybody				
Method scope and sample size	Survey... ...re...	...nelists	Access ope... sub...		...s of the
	Sample		**Entire user / subscriber base**		

Method comparison by data characteristics

Methods / Attributes	Surveys and panels	Handset monitoring	Wireless access network		Server measurements
			TCP/IP traffic measurements	**Usage accounting systems**	
Nature of data	Subjective data on perceived aggregate service usage, diary method could sort out individual events	Accurate and objective data on handset application and feature usage	Quite accurate and objective profile of TCP/IP traffic in the network	Accurate and objective data on access operator network and service usage	Objective data on service-specific usage, accuracy depending on the method
Explanatory / independent variables •User •Device •Time •Location	Any background variables on the respondents Perceived time and context of usage Panels using a diary method enable more accuracy in time and location of usage	Any background variables on the panelists Handset model and access network used Time of usage Location of usage (cell ID, WLAN access point name, GPS coordinates)	No data on individual users without extra effort, different terminal operating systems identifiable Time of usage No location data	Some background variables on the subscribers (type, tariff), no data on real end-users Terminal model in GSM/UMTS Time of usage No location data	Depending on the used method (identification of individual users, background data on registered users, separation of mobile usage, and time of usage) No location data
Usage / dependent variables •Usage volume •Usage frequency	Perceived amount and frequency of service usage	Volume, frequency, and duration of handset application and feature usage	Volume of usage (bytes, flows) per application protocol and traffic destination	Volume and frequency of service usage (e.g. calling, SMS, data transfer)	Depends on the used method
			Additional measurements could 1) link usage accounting data and traffic data 2) add location data to both measurements		

Method comparison by data characteristics

Methods / Attributes	Surveys and panels	Handset monitoring	Wireless access network		Server measurements
			TCP/IP traffic measurements	Usage accounting systems	
Nature of data	**Subjective data** **Perceived usage**	Accurate and ... and ...	**Objective data** **Actual usage**		...ive data on
Explanatory / independent variables • User • Device • Time • Location	**User OK** **BG data OK** **Device NOT** **Time NOT** **Location NOT**	**User OK** **BG data OK** **Device OK** **Time OK** **Location OK**	**User NOT** **BG data NOT** **Device NOT** **Time OK** **Location NOT**	**User OK** **BG data NOT** **Device OK** **Time OK** **Location NOT**	**User OK** **BG data NOT** **Device NOT** **Time OK** **Location NOT**
Usage / dependent variables • Usage volume • Usage frequency	**Perceived usage** **Any data item**	Volume, frequency and d... b...	**Registered/logged usage** **Specific data items**		Depends on the used

Limitations on data collection

- ## Privacy legislation
  - International differences, convergence expected at EU level
  - In Finland: personal information (e.g. customer registers) vs. identification information (in telecommunication networks) → define the applied law and authority concerned
  - Different interpretations, depending on research purpose and specific data items

- ## Privacy concerns
  - Effect of data collection on participation and behavior?
  - "Big brother" concerns, what's collected and by whom?

- ## Resource limitations
  - Cost-benefit, effort from service provider always required
  - Resource requirements vary, and are implementation specific

Future outlook

- Improvements at handsets, routers, servers...
  → all automated methods to become more important

- Increased complexity, user base fragmentation
  → surveys somewhat less potent

- Convergence of devices and radio technologies, usage events (e.g. offline usage) not otherwise measurable
  → mobile handsets a promising place to measure

- Concentration of services to big players (e.g. Google)
  → lots of information on user behavior

- Divergence of traffic to other wireless access networks
  → 2G/3G mobile networks somewhat less useful

Summary and conclusions

- No single method for all situations
  - Surveys and handset monitoring → detailed sample based data
  - Wireless access network measurements → less detailed data, large scope
  - Server measurements → detailed data, focused scope

- Automated methods to complement surveys in the future
  - All automated methods to become more important
  - Surveys still provide something not directly measurable

- Mobile handsets and certain server methods seem most potential, 2G/3G networks somewhat less useful

Further information

- Related publications
  - http://www.netlab.tkk.fi/tutkimus/coin/

- Contact the author
  - antero.kivi(a)tkk.fi