TKK HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Communications and Networking

IP Traffic Measurements 2008

Mobile Internet Usage Patterns

MoMI project

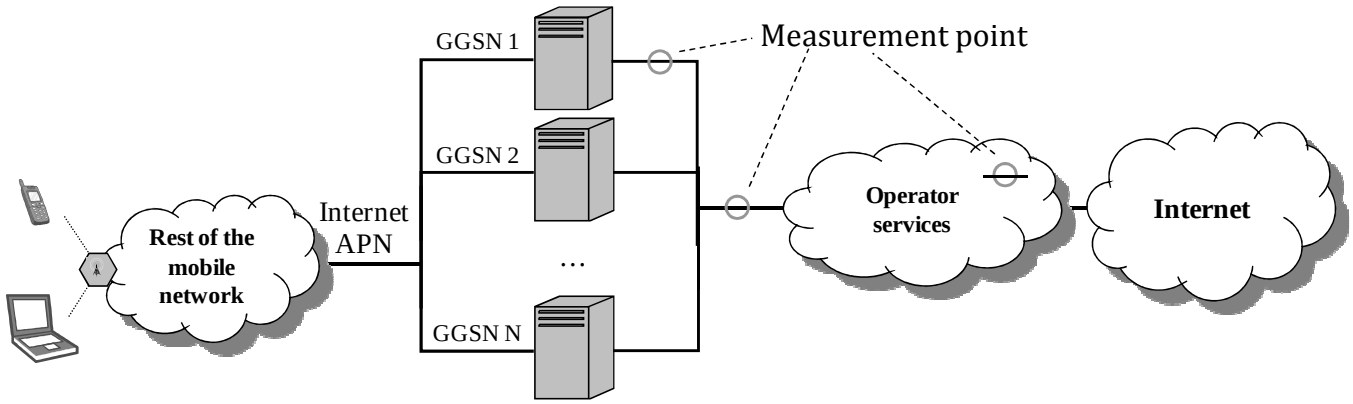
Antti Riikonen,
Antero Kivi
14.5.2009

Agenda

- Measurement Description
  - Measurement Setup & Scope
  - Measurement Analysis
- Mobile Internet Usage Patterns 2008
  - General Traffic Patterns
  - Traffic by Operating System
  - Traffic by Application Usage
  - Mobile Handset Web Traffic
    - Popular Handset Web Browsing Sites
    - Popular “Secure Web” Sites
    - Popular “Mobile Web” Sites
- Summary

Measurement Setup

- Measurements conducted annually (Oct-Dec 2005-2008) at 2-3 out of the 3 Finnish mobile network operators
- All IP traffic via the measurement point to/from Internet measured
  - Point of measurement at mobile operator Internet APN
  - Traffic generated by any terminal using mobile data connection
  - Mobile handsets/laptops/other, postpaid/prepaid, business/consumer
- Headers captured
  - Application layer: Only DNS requests, no other data
  - Network layer: IP headers
  - Transport layer: All headers (e.g. TCP, UDP)



Measurement Scope

- Trace data comparable samples of traffic from the mobile network operators
  - 2 of the 3 Finnish MNOs measured in 2008 (Elisa & DNA Finland)
  - > 90% of all mobile network packet data traffic goes via Internet APN
  - Roaming traffic routed via home network (home GGSN roaming)
    - Traffic by Finnish subscribers abroad → included
    - Traffic by foreign roamers in Finland → excluded
- Representative data of the Finnish mobile market

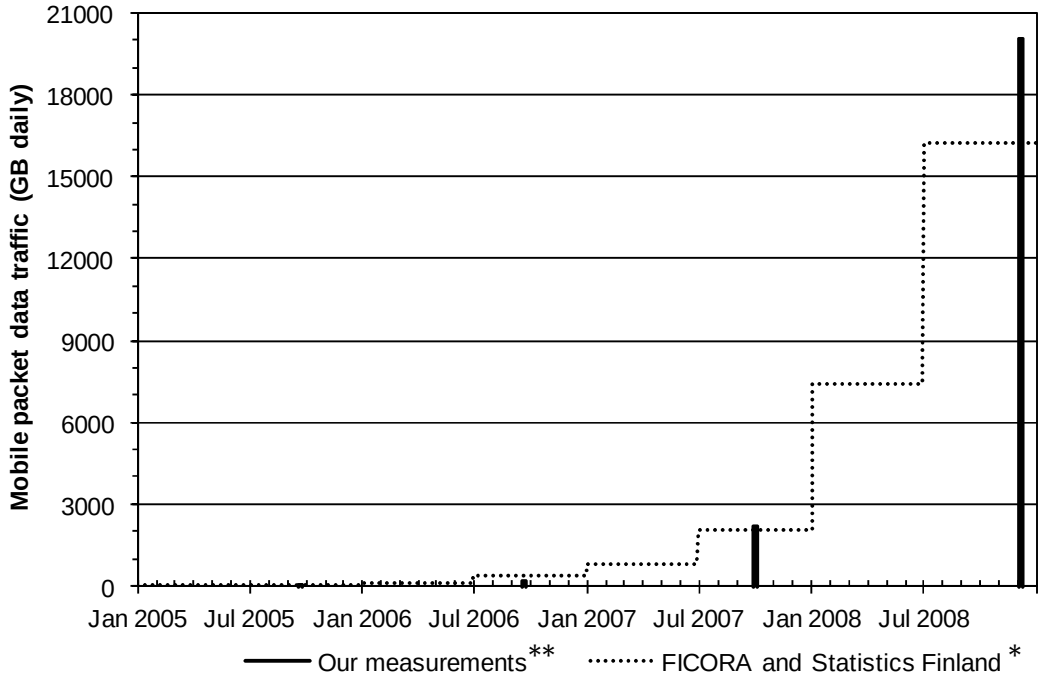
Measurement Analysis

- *Mobile terminals* identified using known IP address spaces of the mobile operators
- *Operating Systems* identified by TCP Fingerprinting
  - Using p0f* tool, not HTTP user agent or IMEI
  - Traffic traces compared to the fingerprints of previously identified OSs
  - Fingerprint database updated by TKK in the beginning of 2009
- *Applications* identified from server TCP/UDP port numbers
  - E.g. 80 = HTTP, 25 = Mail, etc.
  - Straightforward and easy to implement, but includes uncertainties
- *Popular web sites* discovered from DNS data
  - Domain names with over 40 requests during the measurement period recorded and mapped to IP addresses

* See References

Traffic in Mobile Networks Multiplied in 2008

- During 2008
  - From under 145 000 to almost 480 000 mobile broadband subscriptions*
  - Almost 1.7 million subscriptions used mobile data services*



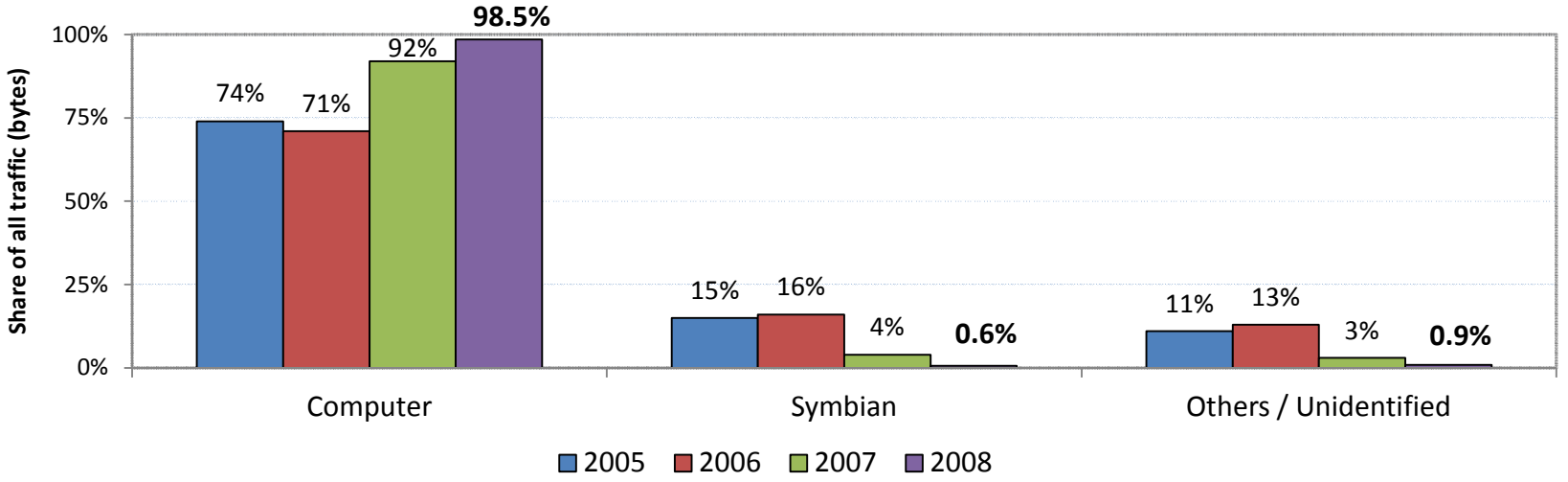
* Ficora (Finnish Communications Regulatory Authority) ** Kivi, 2009

General Traffic Patterns 2008

- Traffic dominantly towards mobile terminals (downlink)
  - 2008: 75% of total traffic downlink (Symbian: 85%)
  - 2005: 84%, 2006: 73%, 2007: 63%
- Traffic dominantly TCP:
  - TCP 94.9% of total traffic volume
  - UDP 4.8%
  - Other protocols 0.3% (e.g. control traffic)
    - Other protocols excluded from the rest of the analyses

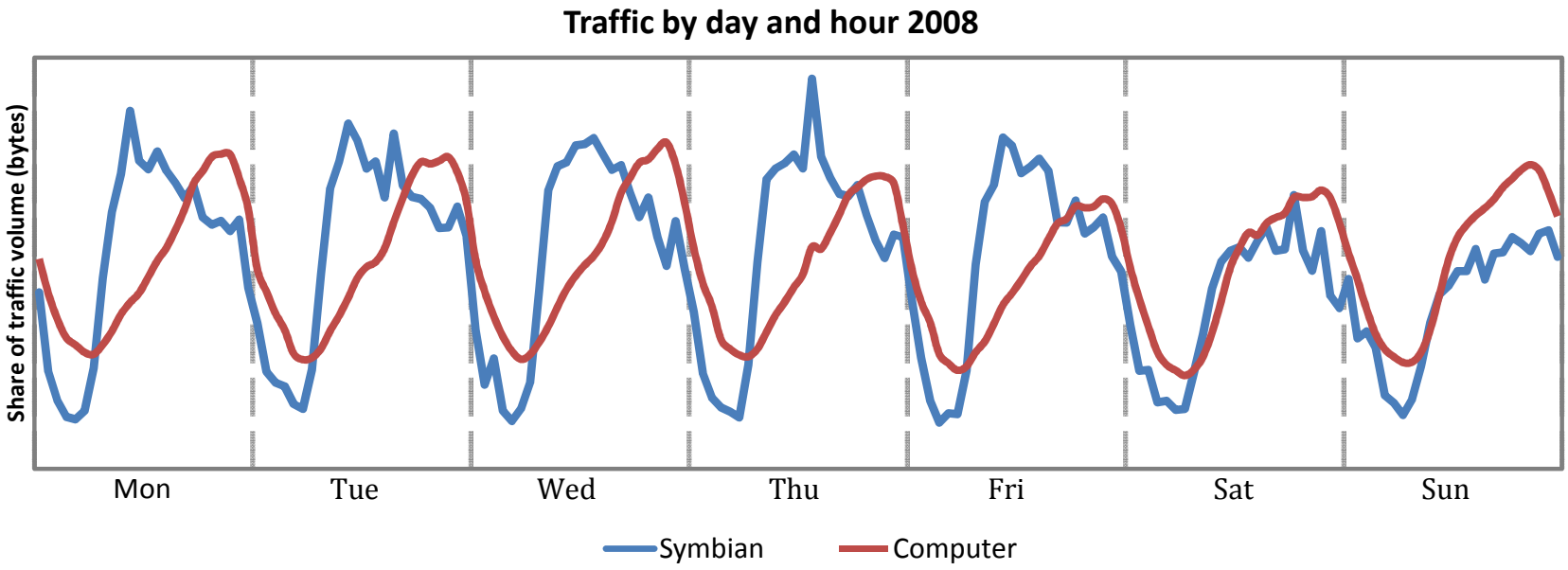
Traffic by Mobile Device Operating System

- Computers originate over 98% of traffic in the mobile network
  - Computer: Mostly Windows (over 93% of total traffic)
  - USB modems, data cards
  - OS identification necessary to uncover handset traffic
- < 1% of traffic generated by handsets
  - Handset: Symbian OS, no significant amount of iPhone / other OS traffic identified
  - Exclusive distributor of iPhone (TeliaSonera) not included in the measurements
  - Symbian traffic increasing in absolute terms
  - On average one computer generates hundreds times the traffic than one mobile handset



Handset Traffic Differs from Computer Traffic

- Handset traffic has high variations
  - Reflects human activity?
  - Peaks in the morning
  - Use during weekends lower and different than during weekdays
- Computer traffic more evenly distributed
  - More continuous traffic?
  - Peaks in the evening (6-10pm)
  - Computer traffic with mobile access used at home, or more capacity available in the evening?

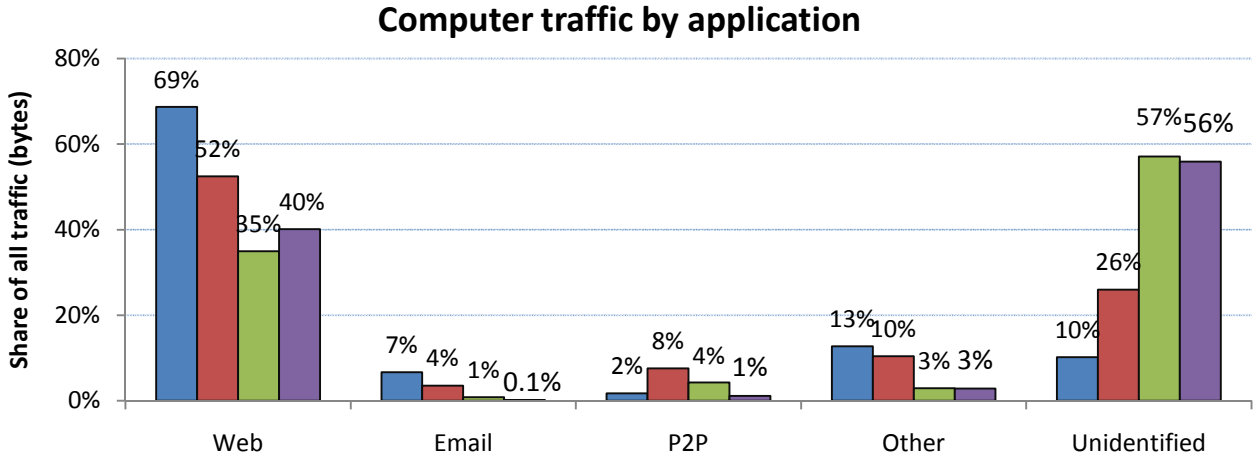


Identification of Applications

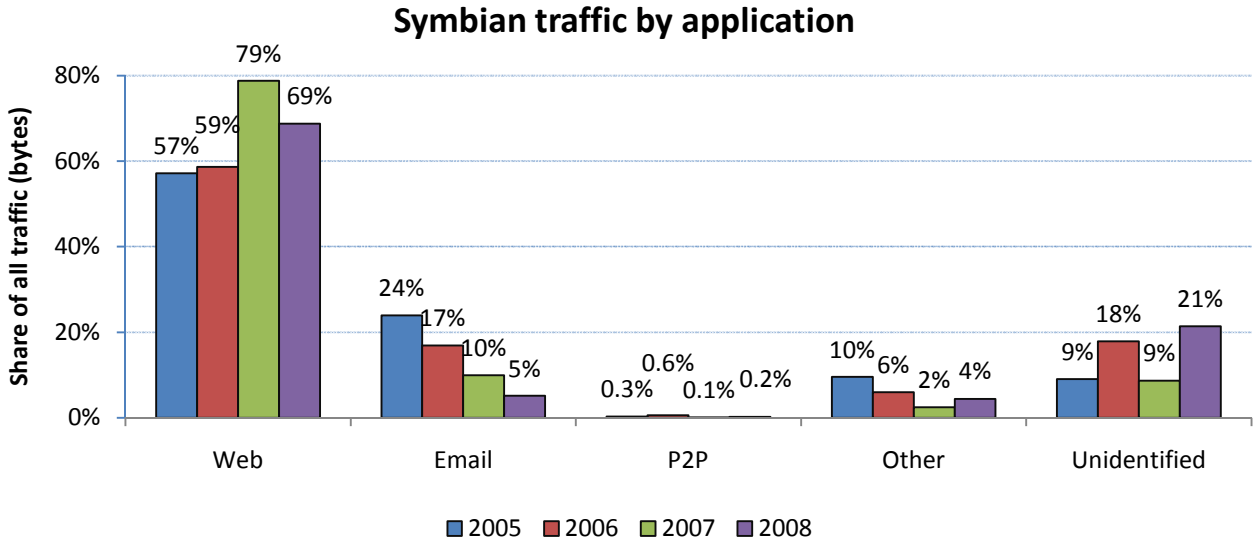
- Application protocols identified with server-side TCP and UDP port numbers
  - Port number based identification not fully accurate
    - Applications can use port space dynamically, or masquerade as other protocols (e.g. P2P, streaming)
  - Port numbers grouped into five categories
    - Web, Email, P2P, Other identified, and Unidentified

Application protocol category	Major transport protocol ports included	
Web	TCP	HTTP (80), HTTPS (443)
Email	TCP	SMTP (25), POP3 (110), IMAP (443), IMAP/SSL (993), POP3/SSL (995)
P2P	TCP	e.g. 411, 412, 1214, 1412, 4661-4662, 6346-6347, 6881-6889, 7777, 51413

Traffic by Applications

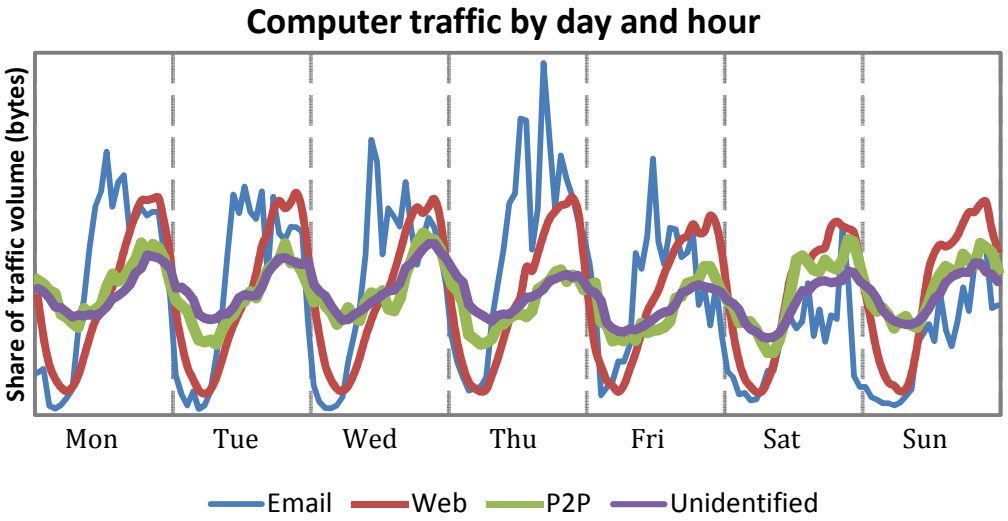


- Computer traffic dominated by Web and Unidentified applications
  - Web usage demands a lot of capacity
  - What is the share of Unidentified P2P?

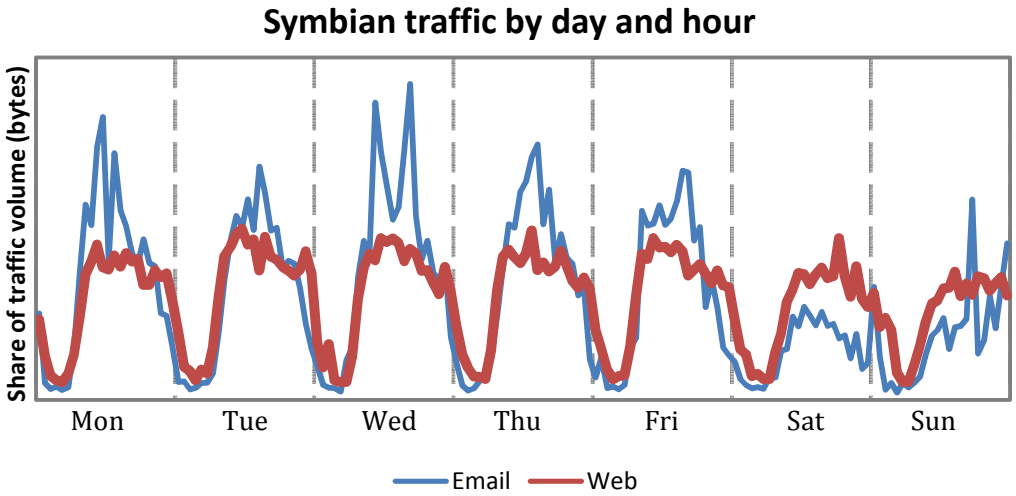


- Handset traffic more clearly dominated by Web
  - Email share decreasing
  - Unidentified traffic does not have clear correlation with other categories

Traffic by Day and Hour



- Computers:
  - Web use pattern same on both, weekends and weekdays
  - Is Unidentified traffic mostly P2P?
    - P2P and Unidentified traffic strongly correlated! Also in terms of uplink/downlink ratio



- Handsets:
  - Email more concentrated on working hours and weekdays
  - Web traffic does not have clear peak hours, slightly weekday oriented
    - Web traffic may include e.g. webmail or streaming
    - Handset web use more evenly distributed during daytime than computer Web use

Handset Web Browsing

- Includes HTTP traffic (85% of all Web traffic)
- Handset Web browsing not very concentrated
  - All domain names have less than 5% of the total HTTP traffic
- Popular sites from fixed side visited also with handsets
  - Lots of local (Finnish) content
  - Finnish media houses (Top 3)
  - Social media
  - Adult content
- Notice
  - YLE and Sanoma servers also significant, but traffic could not be identified to a specific domain name
  - Ranking based on byte volume, i.e. size of the web page matters
  - Categorization subjective
  - ”Non browsing” domains filtered manually

Rank	Domain name*	% of HTTP traffic	Information
1	iltalehti.fi	4%	Traditional media
2	kauppalehti.fi	2%	Traditional media
3	mtv3.fi	2%	Traditional media
4	suomi24.fi	1%	Social media
5	opera-mini.net	1%	Opera Mini browsing
6	tube8.com	< 1%	Adult content
7	irc-galleria.net	< 1%	Social media
8	facebook.com	< 1%	Social media
9	bigbrother.fi	< 1%	Traditional media
10	sihteeriopisto.net	< 1%	Adult content
11	hs.fi (includes oikotie.fi)	< 1%	Traditional media
12	flickr.com	< 1%	Social media
13	ilmatieteenlaitos.fi	< 1%	Information (weather)
14	wikimedia.org	< 1%	Social media (mostly uplink)
15	blogger.com	< 1%	Social media

* Operator sites not included

Handset ”Secure Web” Traffic

- Includes HTTPS traffic (14% of all Web traffic)
- Top list dominated by mail/sync traffic
  - Top 5 domains would be also in the HTTP top 15
  - ¼ of traffic to/from a single nokia.com subdomain
    - Mail for Exchange server?
  - Handset based banking and gambling also observed

Rank	Domain name*	% of HTTPS traffic	Information
1	nokia.com	25%	Mail?
2	sok.fi	7%	Mail
3	fmdm.net	3%	media management
4	logica.com	3%	Mail
5	veikkaus.fi	3%	Gambling
6	op.fi	2%	M-banking
7	turku.fi	1%	Intellisync
8	hus.fi	1%	Mail
9	eqonline.fi	1%	M-banking
10	f-secure.com	1%	F-Secure Mobile Service

* Operator sites not included

Handset ”Mobile Web” Traffic

Rank	Domain name*	% of mobile web traffic
1	nokia.mobi	50%
2	m.facebook.com	19%
3	m.hs.fi	8%
4	yle.mobi	6%
5	m.youtube.com	2%
6	www.foreca.mobi	< 1%
7	wap.jamba.fi	< 1%
8	m.volvooceanrace.org	< 1%
9	wap.sp.fi	< 1%
10	www.ovi.mobi	< 1%
11	wap.aftonbladet.se	< 1%
12	m.espn.go.com	< 1%
13	m.ebay.com	< 1%
14	m.note.nokia.com	< 1%
15	wap.eniro.fi	< 1%
16	wap.veikkaus.fi	< 1%
17	020202.mobi	< 1%
18	wap.weatherproof.fi	< 1%
19	nokia.12dld.mobi	< 1%
20	m.goal.com	< 1%

- Includes Web traffic to/from domain names:
  - ***m.*** & ***wap.*** & ***.mobi***
- In total 110 different ”mobile web” domains (with >40 DNS requests) found from DNS data
- Important groups
  - Traditional media (YLE & Sanoma)
  - Mobile (Nokia)
  - Social media (Facebook, YouTube)
- Is the high share of nokia.mobi explained by Nokia Download! application server?

* Operator sites not included

Summary

- Traffic in Finnish mobile networks multiplied in 2008
- Computers generate most of the traffic (>98%) in mobile networks
  - Use mostly Web (40%) and Unidentified (56%)
  - Unidentified traffic possibly largely P2P!
- Share of handset generated traffic only <1%
  - Still, handset traffic volume approx. doubled in 2008
  - Handset use dominated by Web (69%), though consists also of mail and streaming
- Handset and computer traffic profiles differ also by daily distribution of usage
  - Handsets more morning/working day oriented, computer use peaks in the evening
- Significant use of some ”mobile web” sites noticed
  - Mobile web sites can be 1/10 of ”normal web sites” in size
  - Still, the total amount of “mobile web” sites low

Further information

- Questions?
  - Contact:
    - Antti Riikonen,
    - Antero Kivi,
    - firstname.lastname (at) tkk.fi

- MoMI project:
  - http://www.netlab.tkk.fi/tutkimus/momi/
  - Project partners:

References

- Ficora, 2009. *Market review 3/2008*. Available at: http://www.ficora.fi/attachments/englantiav/5CXG4SC7Y/Market_Review_3_2008.pdf
- Kivi, 2008. *Mobile Data Service Usage Measurements - Results 2005-2007*. Available at: http://www.netlab.tkk.fi/%7Ejakivi/publications/Kivi_Mobile_Data_Service_Usage_2005_2007.pdf
- Kivi, 2009. Diffusion of Mobile Data in Finland. Accepted for publication at *NETNOMICS: Economic Research and Electronic Networking*, 2009.
- p0f passive operating system fingerprinting tool. Available at: http://lcamtuf.coredump.cx/p0f.shtml