Mobile Data Service Usage Measurements

Results 2005-2006

Antero Kivi
2.4.2007
Sources of data on mobile service usage

OUR RESEARCH
- Surveys on handset panel participants
- Handset monitoring panels (SP360)
- Mobile operator charging and billing systems
- Traffic measurements at operator Internet APN

Sample of devices
- 2G/3G mobile networks
- Usage accounting systems
- TCP/IP traffic

Sample of users
- WLAN hot spots
- Other wireless access networks (WiMAX…)

Other wireless access networks
- Intranets
- Internet
- Server(s)

Source: Kivi, 2007
Major findings

- Mobile operator reporting system-based data collection
  (80-90% of Finnish mobile subscribers/terminals in fall 2005 and 2006)
  - Finnish mobile terminal base renewing rapidly, average age down 4 months (from 2.9 to 2.6 years)
  - Terminal features enabling data usage are spreading, 3G penetration up from 0.5% (2005) to 8% (2006)
  - Total mobile network packet data traffic increased almost 4x, consumer sub. traffic almost 5x
  - Only 3% unit (≈100,000) increase in number of weekly packet data users, growth in average traffic per active subscriber (>3x for consumer subs.) almost covers growth in total traffic volume
  - Lower price/MB ↔ more frequent and higher volume packet data usage
  - 3G capability ↔ more frequent and higher volume packet data usage, with all terminal types
  - Traffic predominantly to/from the Internet (Internet APN: 89% of all traffic)
  - PC traffic dominates handset traffic, data cards generate at least 35% of traffic

- TCP/IP traffic measurements
  (80-90% of Finnish mobile network packet data traffic on one week in fall 2005 and 2006)
  - Windows generates 70% of traffic in mobile networks → handset traffic profile largely hidden by Windows
  - At least 16% of traffic actually made with Symbian handsets, while Symbian devices act as GSM/UMTS terminals for 32% of traffic (handsets as modems)
  - Web dominates (>50%) mobile network traffic, P2P small (6%) but growing
  - Symbian traffic profile differs from Windows, share of email is higher (3x) and mobile P2P marginal (1/10)
  - Mobile operator sites (12%), information (9%) and entertainment (6%) significant browsing categories
  - In Symbian browsing mobile operator sites in a big (22%) and growing role
Agenda

- Operator Reporting System –Based Measurements
- TCP/IP Traffic Measurements
- APPENDIX
Operator Reporting System – Based Measurements

Mobile Data Service Usage Measurements
Operator reporting system data

Contents

• Measurement description
• Mobile subscriber base
  – Description of source data
  – Mobile subscribers by type of subscription
  – Consumer subscriber roaming
• Mobile terminal installed base
  – Description of source data
  – Mobile terminals by model
  – Mobile terminals by feature
  – Mobile terminals by manufacturer and smart phone type
  – Mobile terminals by year of introduction
• Mobile packet data traffic
  – Mobile subscriber packet data traffic
  – Mobile subscriber packet data traffic per subscriber
  – Consumer subscriber packet data traffic by pricing
  – Packet data traffic by terminal type
  – Packet data traffic by service
• Summary and conclusions
  – Summary
  – Conclusions
Operator reporting system data

Contents

• Measurement description
  • Mobile subscriber base
    – Description of source data
    – Mobile subscribers by type of subscription
    – Consumer subscriber roaming
  • Mobile terminal installed base
    – Description of source data
    – Mobile terminals by model
    – Mobile terminals by feature
    – Mobile terminals by manufacturer and smart phone type
    – Mobile terminals by year of introduction
  • Mobile packet data traffic
    – Mobile subscriber packet data traffic
    – Mobile subscriber packet data traffic per subscriber
    – Consumer subscriber packet data traffic by pricing
    – Packet data traffic by terminal type
    – Packet data traffic by service
• Summary and conclusions
  – Summary
  – Conclusions
Operator reporting system data

Measurement description

• Data collected using mobile operators’ charging-oriented reporting systems in fall 2005 and 2006
  – Ticket (CDR) and subscriber information systems of Finnish GSM/UMTS operators
  – Data primarily from 2 weeks or 1 month in Sep – Oct, 2005 and 2006

• About 80-90% of Finnish mobile subscribers/terminals included
  – Operators included in 2005 and 2006: Sonera, Elisa (+Kolumbus), DNA
    • No data on: Saunalahti, TeleFinland, others
  – Very comprehensive sample of over 4 000 000
    • Survey studies with similar results commonly with max 10^3 respondents
  – Most data from all three operators
    • In some rare cases 2005 data was available only from two operators
Contents

- Measurement description
- Mobile subscriber base
  - Description of source data
  - Mobile subscribers by type of subscription
  - Consumer subscriber roaming
- Mobile terminal installed base
  - Description of source data
  - Mobile terminals by model
  - Mobile terminals by feature
  - Mobile terminals by manufacturer and smart phone type
  - Mobile terminals by year of introduction
- Mobile packet data traffic
  - Mobile subscriber packet data traffic
  - Mobile subscriber packet data traffic per subscriber
  - Consumer subscriber packet data traffic by pricing
  - Packet data traffic by terminal type
  - Packet data traffic by service
- Conclusions
  - Summary
  - Conclusions
Description of source data

- **Active subscriber** is a subscriber that has generated packet data traffic during the studied week
- **Packet data traffic** (in subscriber context) includes all basic chargeable packet data traffic transfer by postpaid subscribers
  - Includes: basic packet data transfer
  - Excludes: roaming, MMS, and other separately chargeable traffic
  - Some error due to differences in operator-specific data sets (2006)
  - Data collected in 2005 concerning subscribers comparably included ”basic packet data transfer” only

- See appendix A for more details on collected data and estimates on potential sources error
Operator reporting system data: Mobile subscriber base

Mobile subscribers by type of subscription

- Fixed fee data packets have become more popular...
  - Increase of 3 percentage points, i.e. about 100 000 more subscribers
  - Still only 4% of consumer subs. with a data packet, i.e. about 150 000 subscribers in Finland

- …resulting from major changes in offered packet data price schemes since Q1/2006
  - Introduction of various flat-rate alternatives
  - Related to the beginning of handset bundling in 4/2006

Other remarks
- Finland still >90% postpaid country
- ¾ of postpaid subscribers are consumer subscribers
- 5-10% of all consumer subscribers with a bundling contract

---

Helsinki University of Technology
Networking Laboratory

Antero Kivi
02.04.2007
Operator reporting system data: Mobile subscriber base

Consumer subscriber roaming

- 3-4% of all consumer subscribers use their mobile phone abroad every week
  - Made or received voice calls, sent SMSs

- Packet data roaming activity increased to correspond usage at home network
  - Consumer packet data usage at home network: 11% use weekly for avg. of 2.7 MB per packet data using sub.

- Some reliability issues with source data
  - Data from 2005 particularly heterogeneous

<table>
<thead>
<tr>
<th>Consumer subscriber roaming on study week</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of voice and/or SMS roamers</td>
<td>3 – 4%</td>
<td>3,6 %</td>
</tr>
<tr>
<td>Share of packet data roamers</td>
<td>0,1%</td>
<td>0,4%</td>
</tr>
<tr>
<td>Ratio of packet data roamers to voice roamers</td>
<td>2 – 6%</td>
<td>10%</td>
</tr>
<tr>
<td>Average traffic volume per packet data roaming consumer sub.</td>
<td>0.4 – 0.8 MB / week</td>
<td>2.1 MB / week</td>
</tr>
</tbody>
</table>

N > 4 000 000
Contents

• Measurement description
• Mobile subscriber base
  – Description of source data
  – Mobile subscribers by type of subscription
  – Consumer subscriber roaming
• Mobile terminal installed base
  – Description of source data
  – Mobile terminals by model
  – Mobile terminals by feature
  – Mobile terminals by manufacturer and smart phone type
  – Mobile terminals by year of introduction
• Mobile packet data traffic
  – Mobile subscriber packet data traffic
  – Mobile subscriber packet data traffic per subscriber
  – Consumer subscriber packet data traffic by pricing
  – Packet data traffic by terminal type
  – Packet data traffic by service
• Summary and conclusions
  – Summary
  – Conclusions
Operator reporting system data: Mobile terminal installed base

Description of source data

- **Terminals** include all mobile terminals observed at the network during a measurement week
  - Includes: all postpaid/prepaid subscribers’ terminals with any transaction (voice call, SMS, packet/circuit switched data traffic…)
  - Data on terminal features collected primarily from manufacturer web sites
  - Some error due to churn and differences in data sets
  - Some error due to unidentified terminals and terminal features

- **Active terminal** is a terminal that has generated packet data traffic during a studied week

- **Packet data traffic** (in terminal context) includes all mobile network packet data traffic by the terminals of all mobile subscribers (2006)
  - Includes: basic packet data transfer, roaming, MMS and other separately chargeable traffic
  - Some error due to differences in operator-specific data sets concerning included traffic (in 2006)

- See appendix B for more details on collected data and estimates on potential sources error
Operator reporting system data: Mobile terminal installed base

Mobile terminals by model

- Terminal base is less concentrated in 2006 as in 2005
  - Top 50: from 84% to 73%
  - Top 100: from 93% to 83%
  - Broader handset offering?
  - Mixture of ”old” and “new” models cause this temporarily?

- Nokia 3310 is still the most popular terminal
  - From 14% (2005) to 8% (2006)

- Nokia N70 is the most popular ”high end” handset (ranked 4th)
  - First WCDMA terminal, camera phone, smart phone…

- Other remarks
  - Roughly 1000 different terminal models identified in total
  - First WLAN terminal ranked 37th
  - Unidentified terminals likely increases concentration, but not to 2005 level
Operator reporting system data: Mobile terminal installed base

Mobile terminals by feature

• Key features for packet data usage spreading
  – Packet data 51 % → 60 %
  – Java 46 % → 56 %
  – EDGE 11 % → 25 %
  – Smart phones 6 % → 12 %
  – WCDMA 0,5 % → 8 %
  – WLAN 0,7 % → 2 %
  – HSDPA 0 % → 0,1 %

• Unidentified terminals (T) somewhat increase the figures of all features
  – 10-11% in 2006, 5-6% in 2005
  – Unknown profile likely somewhat more advanced than identified terminal base

• Other remarks
  – Share of non-handset terminals up from 0,6% (2005) to 1,3% (2006)
Operator reporting system data: Mobile terminal installed base

**Mobile terminals by manufacturer and smart phone type**

- **Nokia’s over 80% market share remarkable, but slightly decreasing**
  - First non-Nokia ranked 36th
  - Samsung has taken the 2nd place
  - Sales of bundled 3G handsets by manufacturer reflect this development

- **Nokia Symbian dominates smart phone market even clearer**
  - Smart phone base is uniforming
    - S60 platform replacing Series 80 (Nokia communicators), and new communicators are based on S60
  - Other smart phone types seem to be marginal

- **Smartphone shares not entirely reliable**
  - Smart phones among unknown terminals could change shares of different smart phone types
  - Symbian terminals still the clear majority of smart phones
Operator reporting system data: Mobile terminal installed base

Mobile terminals by year of introduction

- Mobile terminal base renewed by more than one year in 2005-2006
  - Average “age” down from 2.9 (2005) to 2.6 (2006), i.e. by 4 months
    (measurement year – year of introduction)
  - Avg. year of intro. from 2002 to 2004
- Bell curve (product life cycle) broken by 2005 and 2001
  - Many terminals from 2005 (in 2006 data) due to changed market focus towards advanced handsets
  - Very few terminals from 2001
- Reliability issues
  - Year of introduction not well defined
    - Official or accidental introduction?
    - Delay from terminal introduction to start of sales not stable, depends on e.g. manufacturer and market
    - Data from manufacturer press releases, but also from the Internet
  - Considerable 11% of "Unknown", likely with somewhat newer profile than identified terminal base
Operator reporting system data

Contents

- Measurement description
- Mobile subscriber base
  - Description of source data
  - Mobile subscribers by type of subscription
  - Consumer subscriber roaming
- Mobile terminal installed base
  - Description of source data
  - Mobile terminals by model
  - Mobile terminals by feature
  - Mobile terminals by manufacturer and smart phone type
  - Mobile terminals by year of introduction
- Mobile packet data traffic
  - Mobile subscriber packet data traffic
  - Mobile subscriber packet data traffic per subscriber
  - Consumer subscriber packet data traffic by pricing
  - Packet data traffic by terminal type
  - Packet data traffic by service
- Summary and conclusions
  - Summary
  - Conclusions
Mobile subscriber packet data traffic

- Mobile subscriber packet data usage grown almost 4x
  - Statistics Finland: total mobile network data traffic 34 000 GB in 2005 (650 GB/week), which corresponds to the traffic volumes measured in 2005

- Consumer subscriber packet data usage up almost 5x
  - More users?
  - More usage per user?
Operator reporting system data: Mobile packet data traffic

Mobile subscriber packet data traffic per subscriber

- Only 3 percentage point increase in share of weekly users
  - 39% more (~100 000) packet data using consumer subscribers

- High growth results from increased average traffic per subscriber
  - Traffic per active consumer subscriber >3x
  - Business subs. still generate 3x as much traffic per subscriber

- Why is average usage growing?
  - Lower price/MB?
  - More capable terminals and networks (3G, HSDPA)?
  - More laptop usage?
  - More data services?
  - New data-intensive services?
  - Better usability?
  - More marketing?
  - …?

Helsinki University of Technology
Networking Laboratory

Antero Kivi
02.04.2007

N > 4 000 000

Share of consumer subscribers using packet data

Weekly packet data traffic per subscriber
Operator reporting system data: Mobile packet data traffic

Consumer subscriber packet data traffic by pricing

- Fixed fee packets very actively used
  - ≈50% use weekly
  - >70% of all traffic

- Lower price/MB ↔ higher usage
  - More frequent and voluminous usage
  - Holds also for almost all individual tariffs, except the highest (premium)
Operator reporting system data: Mobile packet data traffic

Packet data traffic by terminal type

- 3G capability ↔ higher usage for all types of terminals
  - >5x traffic per data card
  - ≈2x traffic per smart phone
  - >4x traffic per “other” handset
- No difference between 3G smart phones and other 3G handsets
  - Other 3G handsets also quite “smart” (Nokia Series 40, SonyEricsson Java platform…)
- Other remarks
  - Data cards actively used for data
    - Data cards account for 35% of all packet data traffic
    - >25x as much traffic per terminal than for handsets
  - Non-3G smart phones more used than other non-3G handsets
    - Smart phones also have other features correlating with packet data usage (e.g. large display, advanced browser, Bluetooth)
  - 14% of total traffic volume by unidentified terminals

Weekly packet data traffic per terminal (2006)

<table>
<thead>
<tr>
<th>Traffic volume (MB)</th>
<th>All terminals</th>
<th>Handsets</th>
<th>Data cards</th>
<th>Data cards (3G)</th>
<th>Data cards (non 3G)</th>
<th>Smart phones</th>
<th>Smart phones (3G)</th>
<th>Smart phones (non 3G)</th>
<th>Other handsets</th>
<th>Other handsets (3G)</th>
<th>Other handsets (non 3G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>0.3 3.2</td>
<td>0.2 1.8</td>
<td>65.2</td>
<td>47.1</td>
<td>9.5</td>
<td>1.2</td>
<td>3.4</td>
<td>0.6 2.1</td>
<td>1.1 3.4</td>
<td>0.1 0.8</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N > 4 000 000


<table>
<thead>
<tr>
<th>All terminals</th>
<th>Handsets</th>
<th>Data cards</th>
<th>Data cards (3G)</th>
<th>Data cards (non 3G)</th>
<th>Smart phones</th>
<th>Smart phones (3G)</th>
<th>Smart phones (non 3G)</th>
<th>Other handsets</th>
<th>Other handsets (3G)</th>
<th>Other handsets (non 3G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11%</td>
<td>11%</td>
<td>77%</td>
<td>87%</td>
<td>68%</td>
<td>36%</td>
<td>30%</td>
<td>34%</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Operator reporting system data: Mobile packet data traffic

Packet data traffic by service

- No actual data on service usage besides packet data traffic volumes

- Traffic predominantly to/from the Internet
  - Internet APN 89% of total packet data traffic volume (90% in 2005)
  - WAP APN ≈7% of total packet data traffic volume
  - MMS APN <1% of total packet data traffic volume
  - Corporate APNs ≈3% of total packet data traffic volume

- PC traffic dominates handset traffic
  - Data cards generate at least 35% of traffic
  - Windows OS actually creates 70% of all Internet APN traffic (see 2\textsuperscript{nd} part of the presentation)

- Effect of new data services not known
  - No detailed data on WAP/web, MMS or email usage
  - No data on other operator-provisioned data services, such as mobile TV streaming and music downloading

→ See TCP/IP traffic measurements (2\textsuperscript{nd} part of the presentation) for more information on the contents of mobile packet data traffic
Operator reporting system data

Contents

• Measurement description
• Mobile subscriber base
  – Description of source data
  – Mobile subscribers by type of subscription
  – Consumer subscriber roaming
• Mobile terminal installed base
  – Description of source data
  – Mobile terminals by model
  – Mobile terminals by feature
  – Mobile terminals by manufacturer and smart phone type
  – Mobile terminals by year of introduction
• Mobile packet data traffic
  – Mobile subscriber packet data traffic
  – Mobile subscriber packet data traffic per subscriber
  – Consumer subscriber packet data traffic by pricing
  – Packet data traffic by terminal type
  – Packet data traffic by service
• Summary and conclusions
  – Summary
  – Conclusions
Operator reporting system data: Summary and conclusions

Summary

- Data from mobile operators’ CDR and subscriber information systems with 80-90% (> 4 000 000) of Finnish mobile subscribers/terminals in fall 2005 and 2006
- Finnish mobile terminal base renewing rapidly, average age down 4 months (from 2.9 to 2.6 years)
- Terminal features enabling data usage are spreading, 3G penetration up from 0.5% (2005) to 8% (2006)
- Nokia handsets (>80%) and Nokia Symbian smart phones (>99% of smart phones) dominate the terminal markets
- Finland still a >90% postpaid country, ¾ of postpaid subscribers are consumers
- >4% of consumer subscribers have some monthly fixed fee tariff on packet data usage (2005: 1%)
- Total mobile network packet data traffic increased almost 4x, consumer sub. traffic almost 5x
- Only 3% unit (~100 000) increase in number of weekly packet data users
- Avg. traffic per active subscriber almost covers traffic volume growth, >3x for consumer subs.
- Lower price/MB ↔ more frequent and higher volume packet data usage
- 3G capability ↔ more frequent and higher volume packet data usage, regardless of terminal type (data card, handset, smart phone)
- Traffic predominantly to/from the Internet (Internet APN: 89% of all traffic)
- PC traffic dominates handset traffic, data cards generate at least 35% of traffic
Conclusions

• Finnish mobile market has some peculiarities
  – Dominance of Nokia and Nokia/Symbian
  – High share of postpaid subscribers

• Mobile terminal base has improved due to changed market focus towards advanced handsets

• Consumer masses have not started using mobile data services, despite improved 3G penetration

• Existing users have acquired more capable terminals and price/MB has decreased

• Critical mass for data service adoption is not yet achieved in Finland
  – Improvements in terminal base nevertheless lay the enabling conditions for mass market adoption in the (near?) future
TCP/IP Traffic Measurements

Mobile Data Service Usage Measurements
TCP/IP Traffic Measurements

Contents

• Measurement description
  – Measurement scope
  – Measurement setup
  – Identification of terminal operating systems
• General packet data traffic patterns
  – General traffic patterns
  – Traffic distribution by operating system
• Packet data traffic by application protocol
  – Most popular application protocols
  – Traffic by application protocol category
  – Traffic by application protocol category and operating system
• Packet data traffic by application and time
  – Traffic distribution by day and hour
  – Traffic distribution by day and operating system
  – Traffic distribution by hour and operating system
• Web protocol traffic patterns
  – Most popular web sites
  – Most popular web sites by category
  – Most popular web sites by category and operating system
  – Most popular web sites by category, day, and hour
• Summary and conclusions
  – Summary
  – Conclusions
TCP/IP Traffic Measurements

Contents

- Measurement description
  - Measurement scope
  - Measurement setup
  - Identification of terminal operating systems

- General packet data traffic patterns
  - General traffic patterns
  - Traffic distribution by operating system

- Packet data traffic by application protocol
  - Most popular application protocols
  - Traffic by application protocol category
  - Traffic by application protocol category and operating system

- Packet data traffic by application and time
  - Traffic distribution by day and hour
  - Traffic distribution by day and operating system
  - Traffic distribution by hour and operating system

- Web protocol traffic patterns
  - Most popular web sites
  - Most popular web sites by category
  - Most popular web sites by category and operating system
  - Most popular web sites by category, day, and hour

- Summary and conclusions
  - Summary
  - Conclusions
TCP/IP traffic: Measurement description

Measurement scope

• Packet data traffic at mobile operator Internet APN measured for a week in fall 2005 and 2006
  – TCP, UDP and IP protocol headers captured
    • In 2006 also headers of other protocols on top of IP
  – ≈90% of all packet data traffic (all APNs) goes via Internet APN
  – Measurements for two weeks in Sep – Oct 2006, one week in 2005
    • Measurements not completely simultaneous

• About 80-90% of Finnish mobile network packet data traffic included
  – Operators included Sonera, Elisa (+Kolumbus), and DNA (2006)
    • No data on: Saunalahti, TeleFinland, others
  – In 2005, measurements only at Sonera and DNA → 50-60% of all traffic
  – All traffic to/from Internet by all mobile subscribers
    (postpaid and prepaid subscribers, business and consumer subscribers)
TCP/IP traffic: Measurement description

Measurement setup

- **Measurement points comparable**
  - Traffic quantities (bytes, flows) of measurement 1 multiplied by the actual number of GGSNs in order to have proper weight for the operator’s traffic

- **Measured traffic not influenced by roaming**, as home GGSN roaming is used by both operators
  - All roaming traffic by operators’ subscribers routed via home network GGSN → all packet data roaming traffic by operators’ subscribers included, no foreign roamers’ traffic included

- **Client and server “roles” identified using terminal IP addresses**
  - Subscriber (client) terminals always on one side of traffic, all other IP addresses considered servers
  - Problem: public IP addresses for mobiles → client-server roles sometimes reversed
TCP/IP traffic: Measurement description

Identification of terminal operating systems

• Terminal operating system (OS) identified using *TCP fingerprinting*
  – Differences in implementation of TCP/IP stack in different OSs → distinct TCP “fingerprints”
  – Traffic traces are compared to the fingerprints of previously identified OSs
  – Common PC and smartphone OSs can be identified with reasonable accuracy

• Operating system identification process includes some possible bias
  – Only mobile terminal OSs identified
  – OS identification based on uplink TCP traffic only (57% of flows, 23% of bytes)
    • OS of uplink TCP flows identified → that OS resides at a certain IP address at a certain time frame
    • Downlink TCP flows, and all UDP flows accounted for different OSs based on this information
  – What is the effect of the 43% of flows not used in OS identification?
    • OS identified correctly after the first uplink TCP flow for as long as the terminal has the same IP address
    • OS identification of TCP based application protocols (e.g. web, email) is more reliable
TCP/IP Traffic Measurements

Contents

• Measurement description
  – Measurement scope
  – Measurement setup
  – Identification of terminal operating systems

• General packet data traffic patterns
  – General traffic patterns
  – Traffic distribution by operating system

• Packet data traffic by application protocol
  – Most popular application protocols
  – Traffic by application protocol category
  – Traffic by application protocol category and operating system

• Packet data traffic by application and time
  – Traffic distribution by day and hour
  – Traffic distribution by day and operating system
  – Traffic distribution by hour and operating system

• Web protocol traffic patterns
  – Most popular web sites
  – Most popular web sites by category
  – Most popular web sites by category and operating system
  – Most popular web sites by category, day, and hour

• Summary and conclusions
  – Summary
  – Conclusions
TCP/IP traffic: General packet data traffic patterns

General traffic patterns

- Traffic volume grown about 4x between the measurements in 2005 and 2006
  - Accurate absolute growth figures not available from this measurement, as only 2 operators measured in 2005
  - Statistics Finland: total mobile network data traffic 34 000 GB in 2005 (650 GB/week)

- Traffic dominantly TCP
  - 2006: 88% of byte volume
  - 2005: 84% of byte volume

- Traffic dominantly towards the mobile terminals (downlink)
  - 2006: 73% of byte volume
  - 2005: 84% of byte volume

- Other protocols <2% of total traffic volume
  - Mainly IPSec ESP traffic (VPN), >80% of other protocols
  - Excluded from the rest of the analyses
TCP/IP traffic: General packet data traffic patterns

Traffic by mobile terminal operating system

- Windows originates 70% of traffic in mobile network
  - Data cards, GPRS modems, handsets as modems via Bluetooth/cable
  - A few PCs create more traffic than many mobiles → OS identification necessary to uncover handset traffic
  - Windows Mobile, Windows CE, and Pocket PC traffic in “Others” category

- At least 16% of traffic actually made with Symbian handsets
  - 32% of traffic with Symbian device as the GSM/UMTS network terminal (CDR data) → 4–16% of this traffic from modem usage

- Unknown 12% of traffic problematic
  - All other handsets, possibly additional laptop and Symbian traffic
  - Telematics, machine-to-machine (M2M) comm., alarm terminals, remote cameras…?
  - Do intelligent modems / GPRS modules, VPN, or access network elements (SGSN/GGSN, firewall) alter the TCP fingerprint?
TCP/IP Traffic Measurements

Contents

• Measurement description
  – Measurement scope
  – Measurement setup
  – Identification of terminal operating systems

• General packet data traffic patterns
  – General traffic patterns
  – Traffic distribution by operating system

• Packet data traffic by application protocol
  – Most popular application protocols
  – Traffic by application protocol category
  – Traffic by application protocol category and operating system

• Packet data traffic by application and time
  – Traffic distribution by day and hour
  – Traffic distribution by day and operating system
  – Traffic distribution by hour and operating system

• Web protocol traffic patterns
  – Most popular web sites
  – Most popular web sites by category
  – Most popular web sites by category and operating system
  – Most popular web sites by category, day, and hour

• Summary and conclusions
  – Summary
  – Conclusions
TCP/IP traffic: Packet data traffic by application protocol

Most popular application protocols

<table>
<thead>
<tr>
<th>Rank</th>
<th>TCP port</th>
<th>Share of TCP traffic (bytes)</th>
<th>UDP port</th>
<th>Share of UDP traffic (bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>80</td>
<td>53.6%</td>
<td>2746</td>
<td>14.0%</td>
</tr>
<tr>
<td>2.</td>
<td>443</td>
<td>6.0%</td>
<td>10000</td>
<td>13.2%</td>
</tr>
<tr>
<td>3.</td>
<td>4662</td>
<td>3.0%</td>
<td>4500</td>
<td>10.2%</td>
</tr>
<tr>
<td>4.</td>
<td>110</td>
<td>2.6%</td>
<td>370</td>
<td>8.9%</td>
</tr>
<tr>
<td>5.</td>
<td>143</td>
<td>1.3%</td>
<td>500</td>
<td>4.1%</td>
</tr>
<tr>
<td>6.</td>
<td>119</td>
<td>1.2%</td>
<td>53</td>
<td>3.7%</td>
</tr>
<tr>
<td>7.</td>
<td>25</td>
<td>1.1%</td>
<td>32555</td>
<td>1.4%</td>
</tr>
<tr>
<td>8.</td>
<td>7777</td>
<td>1.1%</td>
<td>5004</td>
<td>1.4%</td>
</tr>
<tr>
<td>9.</td>
<td>6881</td>
<td>1.0%</td>
<td>434</td>
<td>1.2%</td>
</tr>
<tr>
<td>10.</td>
<td>7171</td>
<td>0.7%</td>
<td>49000</td>
<td>0.8%</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>28.4%</td>
<td>Others</td>
<td>41.2%</td>
</tr>
</tbody>
</table>

- Application protocols identified with server-side TCP/UDP port numbers
  - Nearly all 65000 TCP and UDP ports observed
  - Port number based identification not full proof
    - Applications using port space dynamically, or masquerading as other apps. (e.g. P2P, streaming…)
    - Subscriber terminals also as servers as public IP addresses used → client ports observed

- TCP traffic mainly web and email
  - Web: HTTP (80), HTTPS (443)
  - Email: POP3 (110), IMAP (143), SMTP (25)
  - News: NNTP (119)
  - P2P: eDonkey (4662), Limeware (7777), BitTorrent (6881)
  - Unknown: 7171

- UDP traffic mainly VPN
  - VPN: CheckPoint UDP Encapsulation (2746), NDMP / Cisco IPsec VPN (10000), IPsec / NAT-Traversal (4500), ISAKMP / IKE (500)
  - Others: F-Secure updates / BackWeb (370), Mobile IP agent (434), DNS (53)
  - Unknown: 5004, 32555, 49000
TCP/IP traffic: Packet data traffic by application protocol

Traffic by application protocol category

- Port numbers grouped into 6 application protocol categories
  - >30% uncategorized protocols → categories simplify but are ambiguous

- Web traffic dominates with >50% traffic share
  - Includes video streaming, possibly P2P

- ”Other” category with most growth
  - P2P and streaming becoming more mainstream also in mobile network?
  - Also client ports, self-initiated Windows traffic, client ports, malware

- P2P traffic increasing
  - More P2P likely in Web and ”Other”
  - Still much smaller than in fixed Internet

- Other remarks
  - Relative shares of web, email, and VPN traffic are decreasing
    - All growing in absolute, but less than other applications
    - VPN & Web likely include more email

---

### Traffic by application protocol category

<table>
<thead>
<tr>
<th>Application protocol category</th>
<th>Transport protocol ports included*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>TCP 80, 443, 8080</td>
</tr>
<tr>
<td>Email</td>
<td>TCP 110, 143, 25, 993, 995</td>
</tr>
<tr>
<td>Control traffic</td>
<td>UDP 53, 5060, 123</td>
</tr>
<tr>
<td>P2P, file transfer</td>
<td>TCP 4662, 7777, 6881, 1412, 20, 9999, 6346, 411, 6882, 412</td>
</tr>
<tr>
<td>VPN</td>
<td>TCP 10000, 500</td>
</tr>
<tr>
<td></td>
<td>UDP 2746, 10000, 4500, 500, 1194</td>
</tr>
<tr>
<td>Others</td>
<td>TCP 119, 7171, 11469, 554, 1863, 12001, 1352, 3000, 1935, 32459, 22, 1750</td>
</tr>
<tr>
<td></td>
<td>UDP 370, 32555, 5004</td>
</tr>
</tbody>
</table>

* TCP/UDP ports with at least 0.5% of total bytes in category
TCP/IP traffic: Packet data traffic by application protocol

Traffic by application protocol category and OS

- **Windows profile as the profile for all traffic**
  - Imposes itself with 70% share of traffic

- **Symbian profile differs from Windows in some ways**
  - Email share 3x higher
  - P2P share only 1/10

- **Web and “Other” driving traffic growth on Symbian**
  - Decreasing email traffic still significant with 13% share

- **Other remarks**
  - VPN also used on Symbian
    - Nokia Mobile VPN Client, related to Check Point
  - Share of Symbian P2P traffic very low (0.6%)
TCP/IP Traffic Measurements

Contents

• Measurement description
  – Measurement scope
  – Measurement setup
  – Identification of terminal operating systems
• General packet data traffic patterns
  – General traffic patterns
  – Traffic distribution by operating system
• Packet data traffic by application protocol
  – Most popular application protocols
  – Traffic by application protocol category
  – Traffic by application protocol category and operating system
• Packet data traffic by application and time
  – Traffic distribution by day and hour
  – Traffic distribution by day and operating system
  – Traffic distribution by hour and operating system
• Web protocol traffic patterns
  – Most popular web sites
  – Most popular web sites by category
  – Most popular web sites by category and operating system
  – Most popular web sites by category, day, and hour
• Summary and conclusions
  – Summary
  – Conclusions
TCP/IP traffic: Packet data traffic by application and time

Traffic by day and hour

- Monday most active day, peak hour at 8–9PM
  - 16% of weekly traffic on Mondays
  - 6.5% of daily traffic at 8–9PM
- Business-oriented email and VPN less used on weekends
- Web is free-time oriented
  - Almost equal usage during weekdays and weekend
  - Most usage in the evening on weekdays
- Other remarks
  - VPN peaks at 8–9AM on weekdays
  - Email used quite evenly during office hours
  - P2P traffic present all day
TCP/IP traffic: Packet data traffic by application and time

Traffic by day and operating system

- Windows profile again as the profile for all traffic
- Symbian profile quite similar to Windows profile
  - Symbian P2P not representative due to negligible P2P volume (0.6% of traffic)
TCP/IP traffic: Packet data traffic by application and time

Traffic by hour and operating system

- Windows profile again as the profile for all traffic
- Symbian usage quite even from 8AM to 9 PM
  - Peak hour 12AM–1PM
  - No increase in web usage in the evening
  - VPN traffic has a 2\(^{nd}\) peak before noon
  - P2P traffic again error prone due to small volume
TCP/IP Traffic Measurements

Contents

• Measurement description
  – Measurement scope
  – Measurement setup
  – Identification of terminal operating systems

• General packet data traffic patterns
  – General traffic patterns
  – Traffic distribution by operating system

• Packet data traffic by application protocol
  – Most popular application protocols
  – Traffic by application protocol category
  – Traffic by application protocol category and operating system

• Packet data traffic by application and time
  – Traffic distribution by day and hour
  – Traffic distribution by day and operating system
  – Traffic distribution by hour and operating system

• Web protocol traffic patterns
  – Most popular web sites
  – Most popular web sites by category
  – Most popular web sites by category and operating system
  – Most popular web sites by category, day, and hour

• Summary and conclusions
  – Summary
  – Conclusions
## TCP/IP traffic: Web protocol traffic patterns

### Most popular web sites

<table>
<thead>
<tr>
<th>Rank*</th>
<th>Domain name of site</th>
<th>Share of web traffic volume</th>
<th>Share of web site visits**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>akamai.net</td>
<td>4.2%</td>
<td>1.5%</td>
</tr>
<tr>
<td>2</td>
<td>irc-galleria.net</td>
<td>2.1%</td>
<td>2.9%</td>
</tr>
<tr>
<td>3</td>
<td>statistik-gallup.net</td>
<td>0.4%</td>
<td>4.4%</td>
</tr>
<tr>
<td>4</td>
<td>basefarm.net</td>
<td>1.7%</td>
<td>2.2%</td>
</tr>
<tr>
<td>5</td>
<td>mtv3.fi</td>
<td>1.6%</td>
<td>2.3%</td>
</tr>
<tr>
<td>6</td>
<td>google.com</td>
<td>1.7%</td>
<td>1.9%</td>
</tr>
<tr>
<td>7</td>
<td>sanomawsoy.fi</td>
<td>0.6%</td>
<td>1.7%</td>
</tr>
<tr>
<td>8</td>
<td>doubleclick.net</td>
<td>0.5%</td>
<td>1.5%</td>
</tr>
<tr>
<td>9</td>
<td>akamaitechnologies.com</td>
<td>1.7%</td>
<td>0.2%</td>
</tr>
<tr>
<td>10</td>
<td>nebula.fi</td>
<td>0.9%</td>
<td>0.9%</td>
</tr>
<tr>
<td>11</td>
<td>youtube.com</td>
<td>1.4%</td>
<td>0.3%</td>
</tr>
<tr>
<td>12</td>
<td>yahoo.com</td>
<td>0.6%</td>
<td>1.0%</td>
</tr>
<tr>
<td>13</td>
<td>iltalehti.fi</td>
<td>0.8%</td>
<td>0.8%</td>
</tr>
<tr>
<td>14</td>
<td>yle.fi</td>
<td>0.7%</td>
<td>0.7%</td>
</tr>
<tr>
<td>15</td>
<td>seksitreffit.fi</td>
<td>0.5%</td>
<td>0.8%</td>
</tr>
<tr>
<td></td>
<td>Other identified domains</td>
<td>51.1%</td>
<td>49.0%</td>
</tr>
<tr>
<td></td>
<td>Unknown addresses</td>
<td>19.2%</td>
<td>18.8%</td>
</tr>
<tr>
<td></td>
<td>Private addresses</td>
<td>10.2%</td>
<td>9.2%</td>
</tr>
</tbody>
</table>

* Ranked by the domain’s combined share of traffic volume and site visits
** Share of TCP flows to/from the domain

- Server IP addresses of all TCP flows with ports 80, 8080, 8000, 8888, and 443 included
  - Might include P2P or malware traffic as well (traffic to e.g. port 80 goes through firewalls)
  - \( \Rightarrow \) 225000 web server IP addresses
  - \( \Rightarrow \) 220000 domain / sub domain names
    - First PTR record listed taken into account
  - \( \Rightarrow \) 41000 domain names

- Web traffic not very concentrated to few domains
  - 21% of traffic in top 10, 47% in top 100

- Unknown addresses about 19% of traffic
  - Server (web or other) IP addresses for which no reverse DNS entry was available

- Private addresses 10% of traffic volume
  - Web/WAP servers/proxies in operator network?
TCP/IP traffic: Web protocol traffic patterns

Most popular web sites by category

<table>
<thead>
<tr>
<th>Rank*</th>
<th>Site category</th>
<th>Major domains included**</th>
<th>Share of web traffic volume</th>
<th>Share of web site visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mobile operator site</td>
<td>-</td>
<td>11.8%</td>
<td>9.9%</td>
</tr>
<tr>
<td>2.</td>
<td>Information</td>
<td>mtv3.fi, sanomawsoy.fi,</td>
<td>8.8%</td>
<td>11.8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ittalehti.fi, yle.fi,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>sanoma.fi, helsinginsanomat.fi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Entertainment</td>
<td>irc-galleria.net, youtube.com, telkku.com, veikkaus.fi</td>
<td>5.9%</td>
<td>5.7%</td>
</tr>
<tr>
<td>4.</td>
<td>Web search</td>
<td>google.com, yahoo.com</td>
<td>2.5%</td>
<td>3.0%</td>
</tr>
<tr>
<td>5.</td>
<td>Adult content</td>
<td>seksitreffit.fi, shiteeriopisto.net</td>
<td>1.8%</td>
<td>1.7%</td>
</tr>
<tr>
<td>6.</td>
<td>Messaging</td>
<td>luukku.com, hotmail.com,</td>
<td>1.4%</td>
<td>2.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>msn.com, passport.net,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gmail.com</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Banking</td>
<td>op.fi, sampo.fi, nordea.fi, eQonline.fi, huoneistokeskus.fi, alandsbanken.fi, aktia.fi, osuuspankki.fi</td>
<td>1.4%</td>
<td>1.2%</td>
</tr>
<tr>
<td>8.</td>
<td>Advertising</td>
<td>doubleclick.net, advertising.com, adtech.de, theonlinetrader.com, tradedoubler.com</td>
<td>1.2%</td>
<td>3.4%</td>
</tr>
<tr>
<td></td>
<td>Hosting / corporate site</td>
<td>akamai.net, statistik-gallup.net, basefarm.net</td>
<td>23.3%</td>
<td>22.2%</td>
</tr>
<tr>
<td></td>
<td>Other sites</td>
<td>-</td>
<td>12.5%</td>
<td>10.2%</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>-</td>
<td>19.2%</td>
<td>18.8%</td>
</tr>
<tr>
<td></td>
<td>Private</td>
<td>-</td>
<td>10.2%</td>
<td>9.2%</td>
</tr>
</tbody>
</table>

- Domain names grouped into 12 categories, despite many potential sources of error
  - About 40% share of unknown, private, and uncategorized sites
  - Domain name resolving method gives domains of hosting service providers and not the hosted services
  - Categorization based on domain name, not sub domain name
  - Categorization itself is subjective, ambiguous, and error prone
  - Overlapping categories (e.g. Information and Entertainment, Web search and Messaging)
  - Possible non-web traffic

- Mobile operator sites (12%), information (9%) and entertainment (6%) significant categories

- Lots of traffic to “infrastructure” hosts not intentionally connected by users
  - Advertising (banners, pop up windows)
  - Load sharing (e.g. akamai.net)
  - Web site analytics (e.g. statistik-gallup.net)
TCP/IP traffic: Web protocol traffic patterns

Most popular web sites by category and operating system

- Mobile operator sites in a bigger, and growing role for Symbian
  - 22% of all Symbian browsing traffic
  - More "basic" consumer users accessing operator services instead of Internet sites in 2006?

- Share of Information high for Symbian
  - Lots of mobile-adapted content (e.g. mtv3, YLE, HS)

- Lots of uncertainty
  - Very high share of the last 4 non-specific categories
  - Higher share of these for Windows (70%) could imply presence of non-web protocols
TCP/IP traffic: Web protocol traffic patterns

Most popular web sites by category, day and hour

- Entertainment and Web search free time oriented
  - More traffic on weekends and in the evening...
  - … as was for web traffic in general

- Banking most active on weekdays, peak at 9–10AM
  - Banking only 1.4% of all web traffic
TCP/IP Traffic Measurements

Contents

- **Measurement description**
  - Measurement scope
  - Measurement setup
  - Identification of terminal operating systems

- **General packet data traffic patterns**
  - General traffic patterns
  - Traffic distribution by operating system

- **Packet data traffic by application protocol**
  - Most popular application protocols
  - Traffic by application protocol category
  - Traffic by application protocol category and operating system

- **Packet data traffic by application and time**
  - Traffic distribution by day and hour
  - Traffic distribution by day and operating system
  - Traffic distribution by hour and operating system

- **Web protocol traffic patterns**
  - Most popular web sites
  - Most popular web sites by category
  - Most popular web sites by category and operating system
  - Most popular web sites by category, day, and hour

- **Summary and conclusions**
  - Summary
  - Conclusions
Summary

- 80-90% of Finnish mobile network packet data traffic captured at three mobile operators’ Internet APN for a week in fall 2005 and 2006
- Windows originates 70% of traffic in mobile networks → handset traffic profile largely hidden by Windows traffic
- Web dominates (>50%) mobile network traffic, P2P small (6%) but growing
- Most traffic generated on Mondays (16% of weekly traffic), peak-hour at 8-9PM (6.5% of daily traffic)
- Mobile operator sites (12%), information (9%) and entertainment (6%) significant browsing categories
- At least 16% of traffic actually made with Symbian handsets, while Symbian devices act as GSM/UMTS terminals for 32% of traffic (handsets as modems)
- Symbian traffic profile differs from Windows, share of email is higher (3x) and mobile P2P traffic marginal (1/10)
- Mobile operator sites in a big (22%) and growing role in Symbian browsing, ”information” sites also actively accessed
TCP/IP traffic: Summary and conclusions

Conclusions

• Symbian profile differs from general traffic profile in many ways
  – Separating terminal operating system is essential for understanding handset-based mobile service usage from traffic measurements

• Consumer-oriented usage is growing faster than business usage
  – Some P2P traffic has emerged, while the relative share of email and VPN is decreasing

• Mobile operator portal usage share increasing
  – Because (some) consumer subscribers have started using mobile data?
  – Could easily usable operator portals compete with Internet as average skill level of users drops when other user segments start adopting mobile data?
APPENDIX

Mobile Data Service Usage Measurements
APPENDIX A

Description of source data: subscribers

- **Packet data traffic** includes all basic chargeable packet data traffic transfer by postpaid subscribers
  - Includes: basic packet data transfer
  - Excludes: roaming, MMS, and other separately chargeable traffic
  - Error due to differences in operator-specific data sets (2006)
    1. Includes: basic packet data transfer
       Excludes: roaming, MMS and other separately chargeable traffic
    2. Includes: basic packet data transfer, roaming, MMS and other separately chargeable traffic
       → share of extra other separately chargeable traffic (e.g. mobile TV streaming) unknown
       → share of extra MMS traffic <0,3% of all traffic, unknown contribution to active users
       → share of extra roaming traffic <0,1% of all traffic, and <1% of active users
    3. Includes: basic packet data transfer
       Excludes: roaming, MMS and other separately chargeable traffic
  - Data collected in 2005 concerning subscribers comparably included ”basic packet data transfer” only
APPENDIX B

Description of source data: terminals

- Terminals include all terminals of all mobile subscribers observed at the network during the measurement week
  - Data sets from each operator in 2005
    1. All terminals observed at the network in week 34, no particular transactions nor operator subscription required
    2. All postpaid/prepaid subscribers’ terminals with made voice calls / sent SMSs in September
    3. Top 100 postpaid/prepaid subscribers’ terminals with chargeable transactions (phone call, SMS, other) in September
  - Data sets from each operator in 2006
    1. All postpaid/prepaid subscribers’ terminals observed at the network on weeks 37/39, no particular transactions required
    2. All postpaid subscribers’ terminals with chargeable transactions (voice calls, SMSs, packet/circuit-switched data traffic) in weeks 37/39
    3. Top 150 terminals (with chargeable transactions) and top 100 data generating terminals of postpaid and prepaid subscribers in September
  - Error due to churn and other differences in data sets
    - Missing prepaid subscribers (2006 only): <5% , who supposedly have less advanced handsets than postpaid subs.
    - Max. 2.5% excess in measured terminal base due to foreign roamers and emergence call readiness (2005 only)
  - Error due to unidentified terminals and terminal features
    - 9.5% (2006) and 4.5% (2005) of terminals were unidentified → somewhat more advanced terminals than identified terminal base in general
      - TAC codes not identified: Manufacturers don’t deliver TAC–model data to TAC allocating organizations in real time → no data for most recent terminals
      - Terminal models outside top 100 / top 150 models: most recent features not evenly represented among top 100
    - 0.5% -1.5% of identified terminals without some specific information on terminal features

- Packet data traffic includes all mobile network packet data traffic by mobile subscribers (2006)
  - Includes: basic packet data transfer, roaming, MMS and other separately chargeable traffic
  - Error due to differences in operator-specific data sets concerning included traffic (2006 only)
    1. Includes: basic packet data transfer, MMS and other separately chargeable traffic. Excludes: roaming
      → share of missing roaming traffic 2%, contribution to active users <2%
    2. Includes: basic packet data transfer and roaming traffic of postpaid subscribers.
      Excludes: MMS and other separately chargeable traffic of postpaid subscribers, and any traffic of prepaid subscribers
      → share of missing prepaid subscriber traffic unknown, likely <1% of traffic
      → share of missing MMS traffic <0.1% of all traffic, unknown contribution to active users
      → share of missing separately chargeable traffic unknown, likely negligible
    3. Includes: basic packet data transfer, roaming, MMS and other separately chargeable traffic
Further information

• Publications on the topic

• COIN project web site
  – http://www.netlab.tkk.fi/tutkimus/coin/

• Contact antero.kivi(at)tkk.fi