



Understanding Online Social Network Usage from a Network Perspective

Fabian Schneider

Anja Feldman

Balachander Krishnamurthy

Walter Willinger

Presented by Greg Peaker

Goals (Part 1, Introduction)

- Which features of OSNS are popular and capture the users attention?
- What is the impact of OSNS on the network?
- What needs to be considered when designing future OSNs?
- Is the user's behavior homogenous?
- My Thesis: too much time on sessions, not enough analysis useful for OSN or ISP

Next Set of Goals

- Influence & Improve Service Offerings
- Researchers can propose improvements or simplifications for existing OSNs or design OSNs with novel features
- From ISP perspective, OSNs might add features that increase per-user bandwidth demand such as video or live streaming

How is this done?

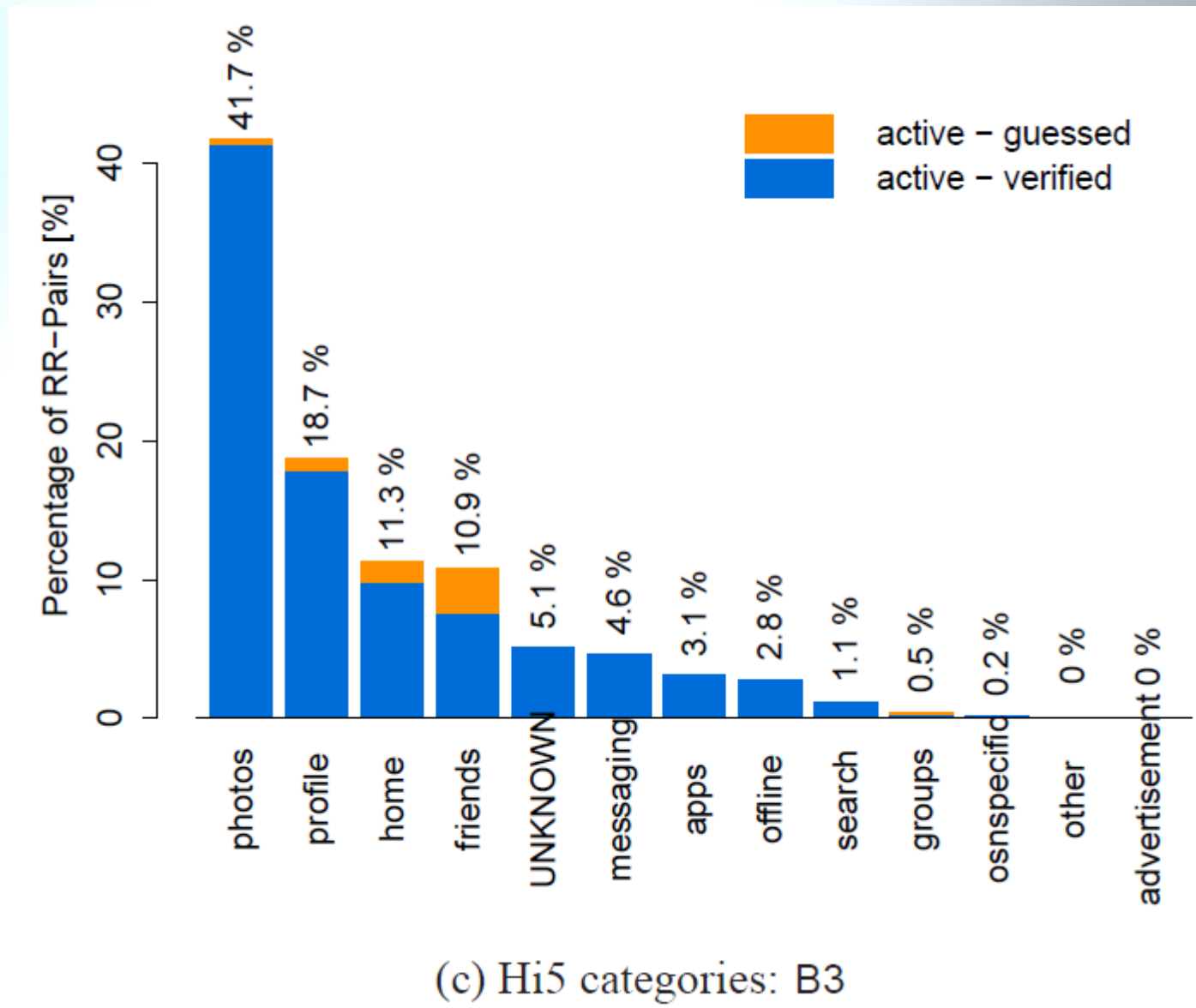
- **Extract clickstreams from passive monitor networks, first of its kind**
 - **Different vantage points within large ISPs across two continents**
 - **Facebook, LinkedIn, Hi5, StudiVZ**
 - **Feature popularity**
 - **Sessions characteristics**
 - **Dynamics within OSN sessions**
 - **“lots of time, small byte contribution”**

The dataset

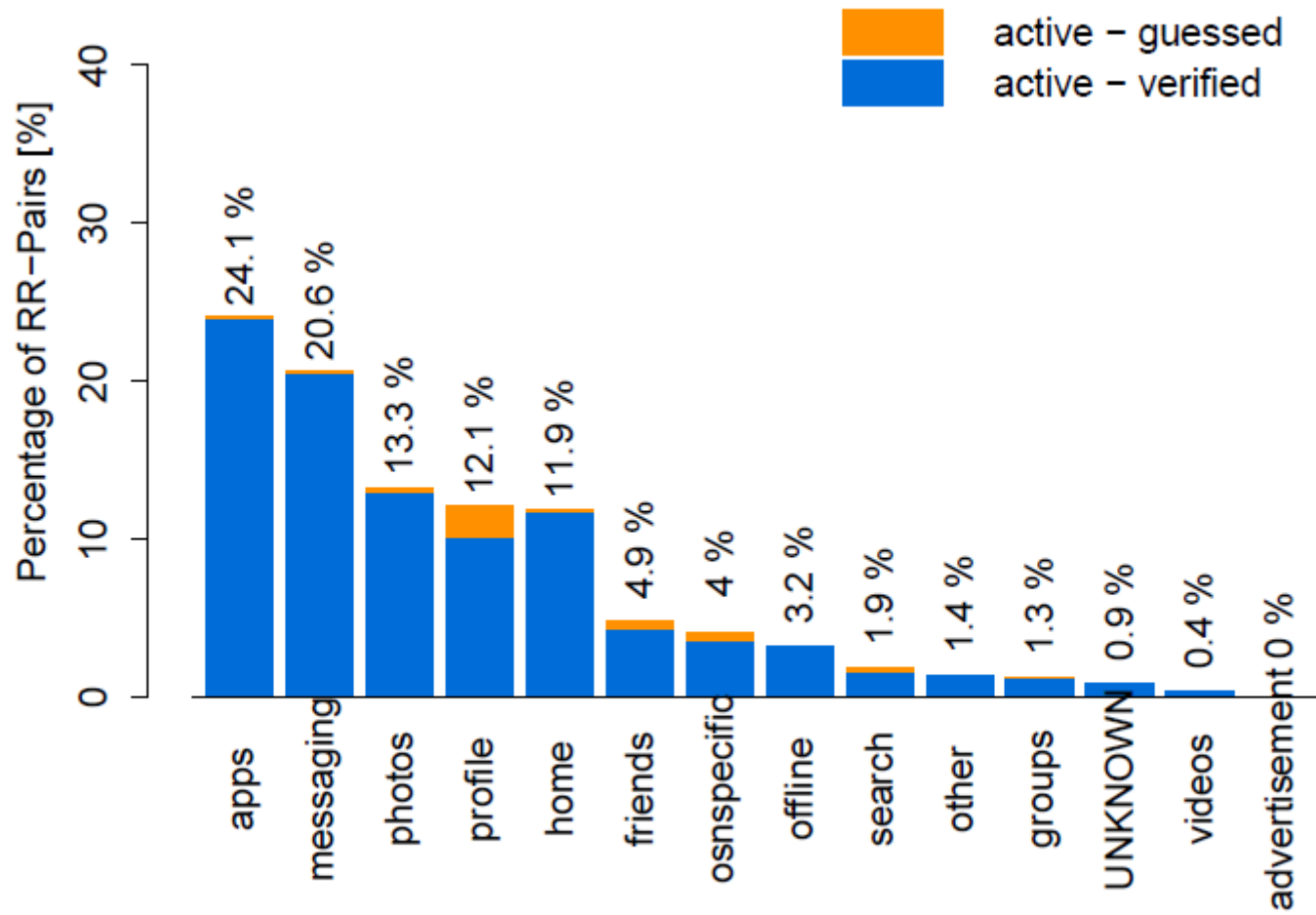
Table 2: Overview of anonymized HTTP header traces.

ID	start date	dur	sites	size	rr-pairs
A1	22 Aug'08 noon	24h	all	>5 TB	>80 M
A2	18 Sep'08 4am	48h	all	>10 TB	>200 M
A3	01 Apr'09 2am	24h	all	>6 TB	>170 M
B1	21 Feb'08 7pm	25h	OSNs	>15 GB	>2 M
B2	14 Jun'08 8pm	38h	OSNs	>50 GB	>3 M
B3	23 Jun'08 10am	>7d	OSNs	>110 GB	>7 M

Which features popular?

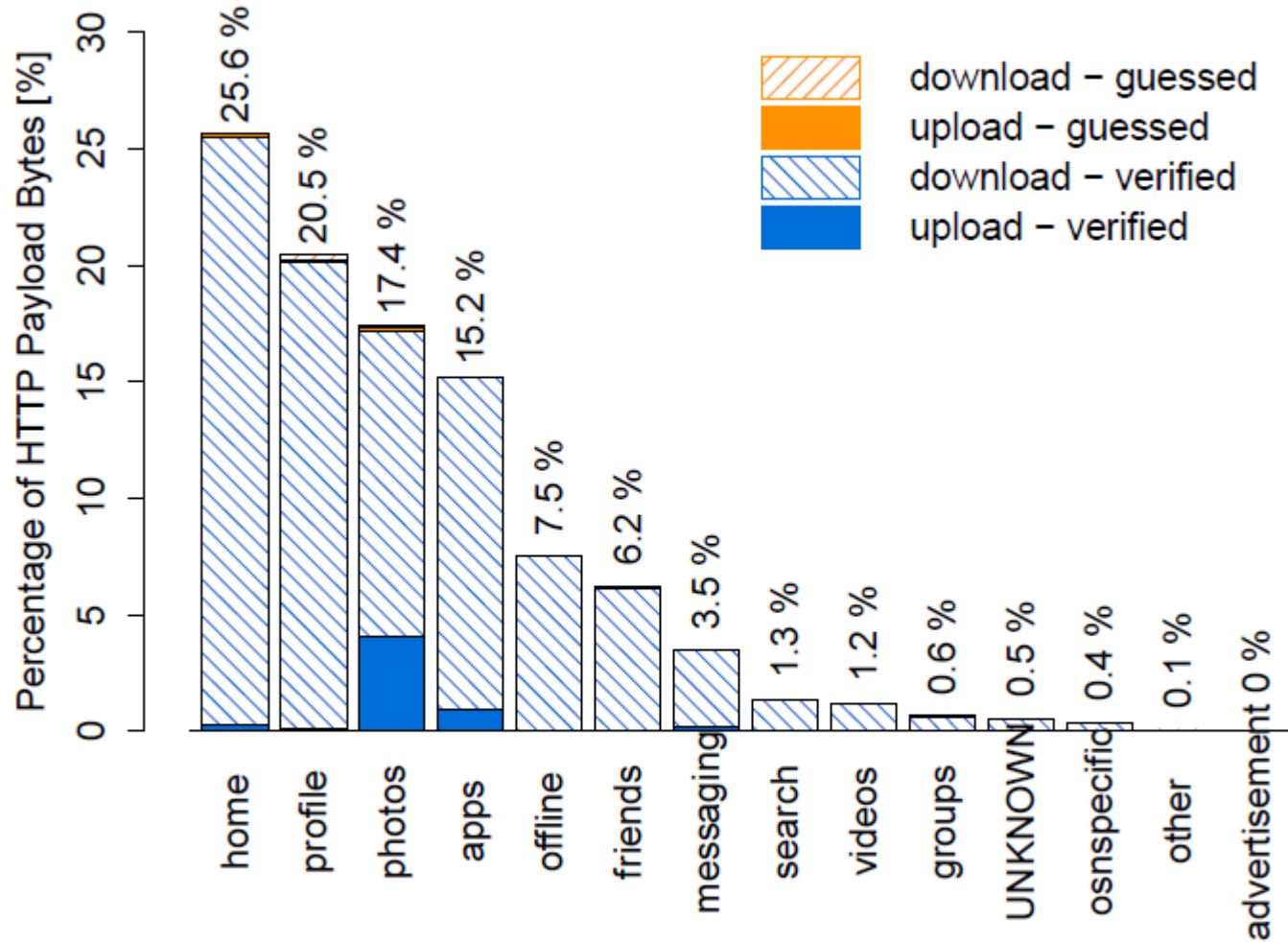


Which features popular?



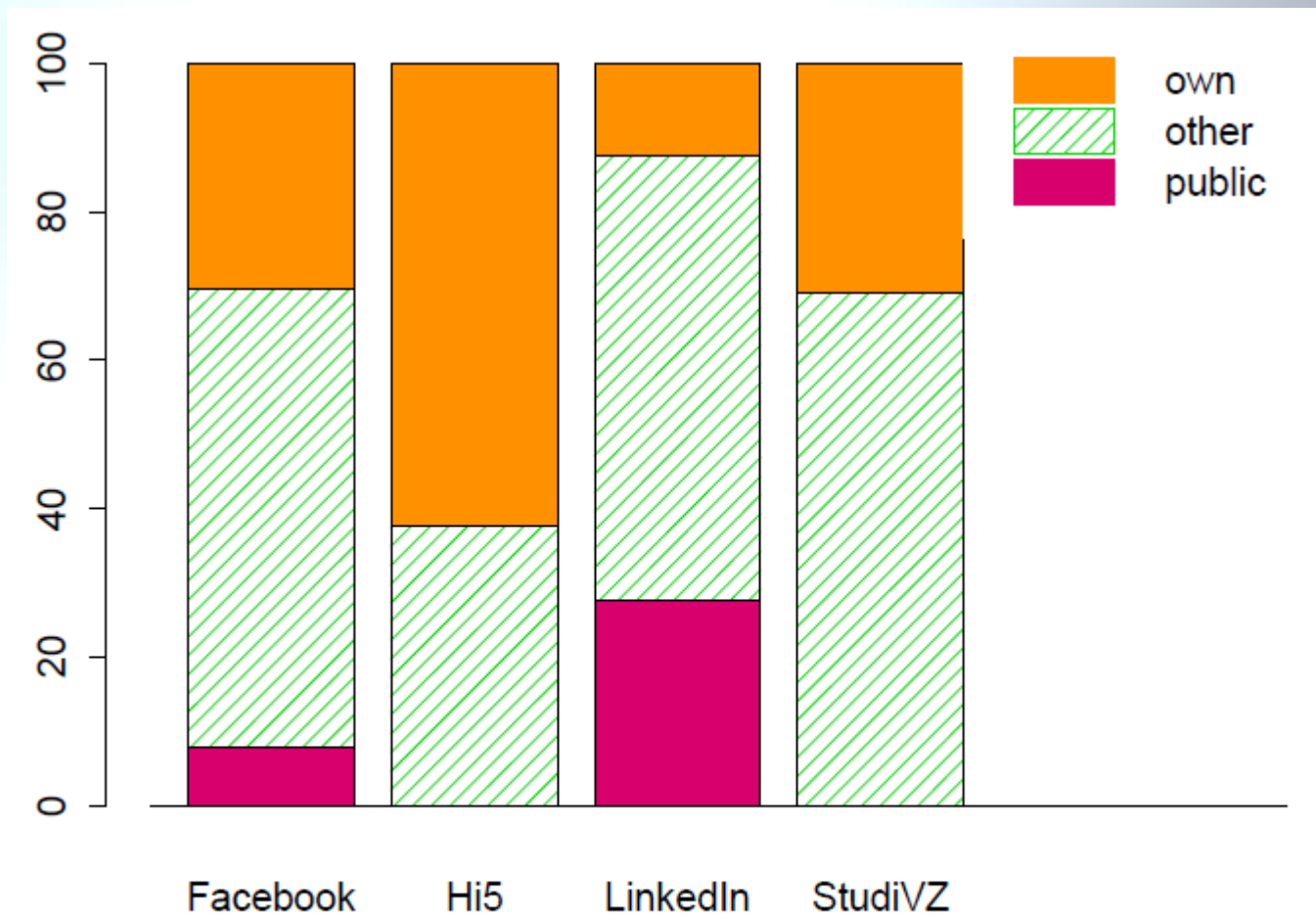
(b) Facebook categories: B3

Impact of OSNs on Network?



(b) Facebook categories by bytes: A2

User Behavior Homogenous?



(a) All OSNs: Distribution of profiles

User Behavior Homogenous?

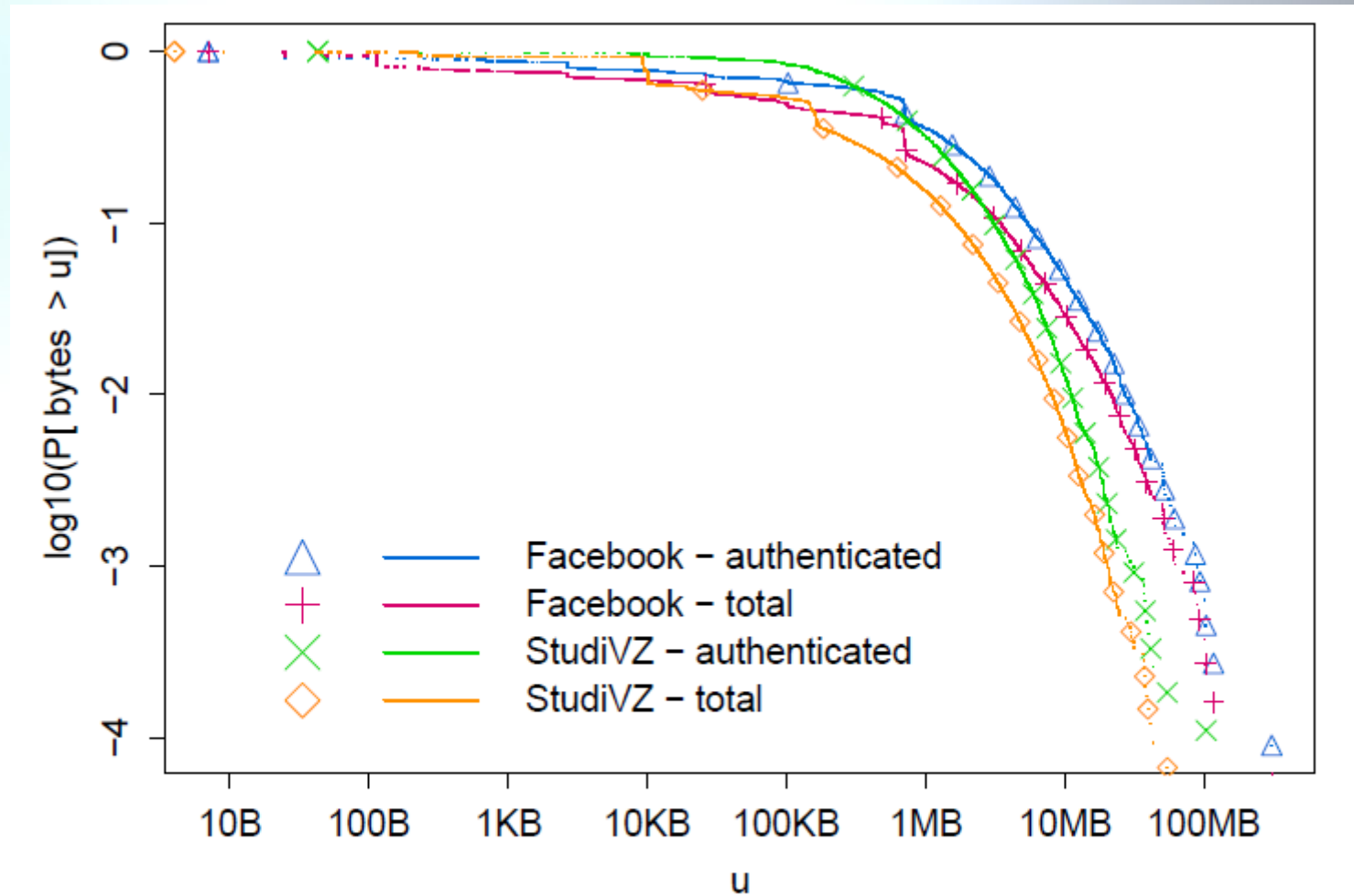


Figure 8: CCDF of Bytes per OSN subsession for Facebook and StudiVZ for A2.

For future OSNs?

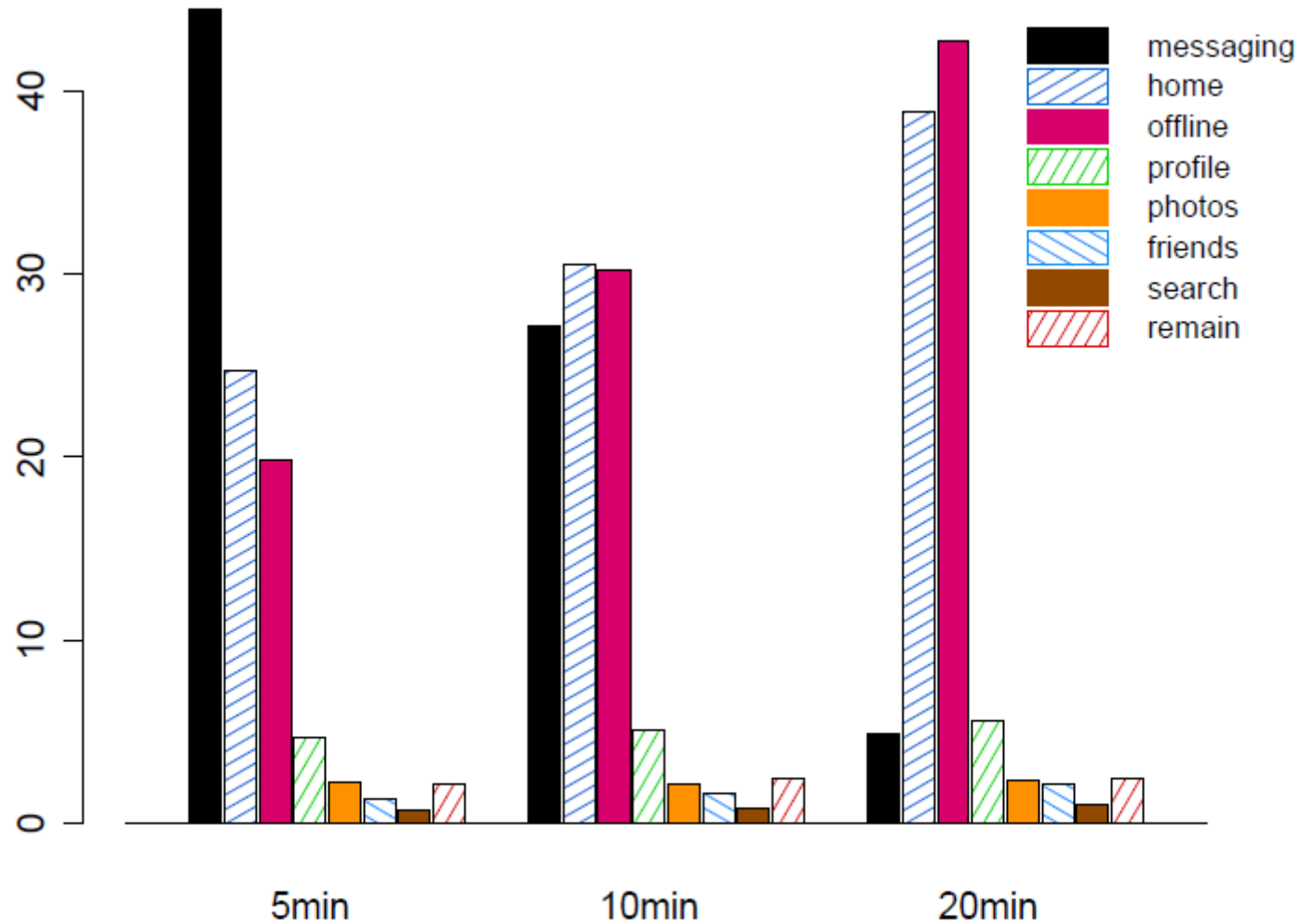


Figure 11: Actions after inactivity period for Facebook and A2.

For future OSNs?

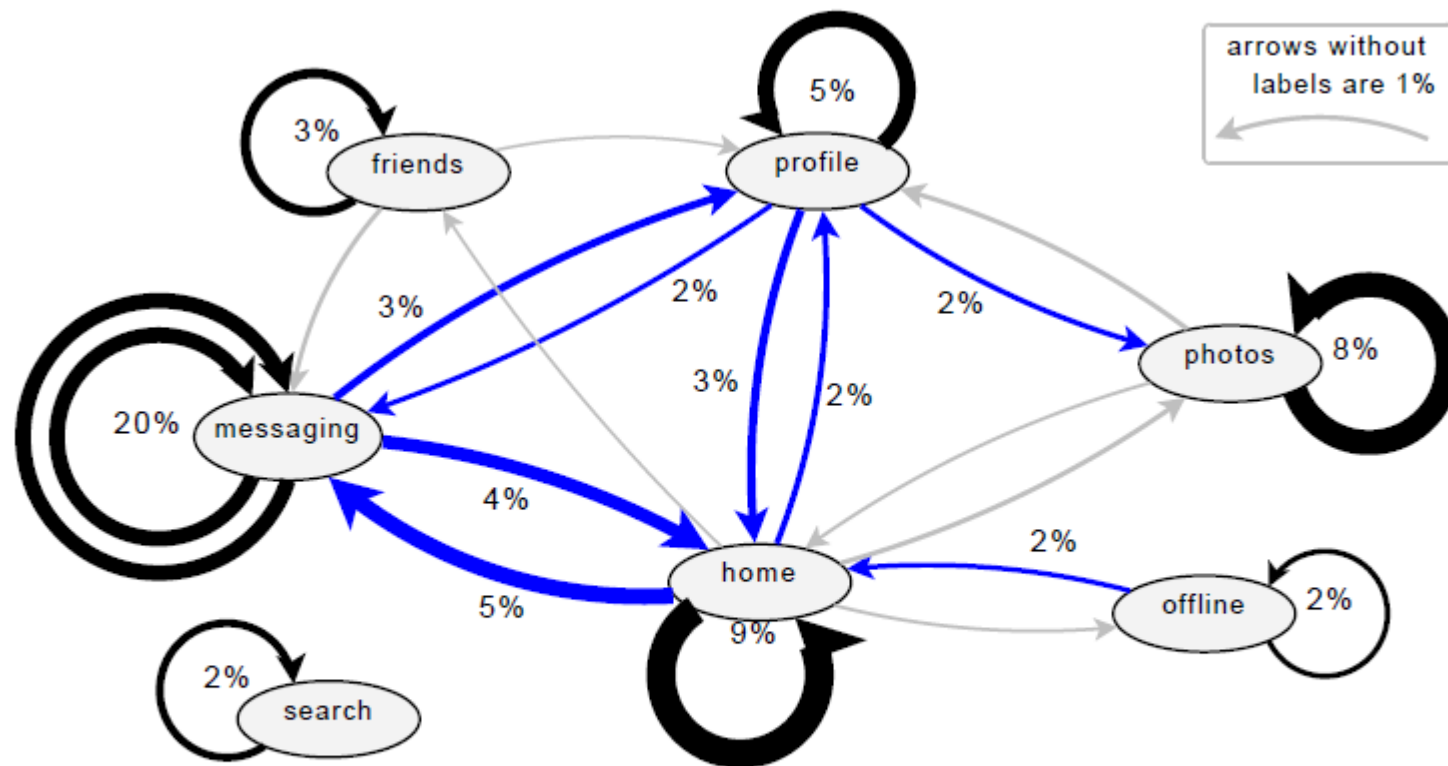


Figure 12: Typical click sequences for Facebook and A2 (only transitions $> 1\%$ are shown).

- **Profiles most popular, then photos & exchanging messages**
- **Most popular clicks doesn't represent most volume**
 - **Photos use most bandwidth**
- **Some sessions lots of accesses, most do "handful"**

Network Features

- **Typical Sessions:**
 - Facebook: 200 KB – 10 MB
 - StudiVZ: 50 KB – 5 MB
 - LinkedIn: 10 KB – 1 MB
- **Take step further; break down into bytes being used in upload/download and then for what feature categories.**
- **Too much time on sessions, not enough on analysis useful for ISP**

Summary

- **Successfully reconstruct OSN clickstreams**
 - Customizable methodology identify OSN sessions
 - Extract OSN usage from coarse information like session duration to minute details
- **Users stay same feature categories**
- **Long Sessions, users don't continuously interact OSN**
- **Next Steps:**
 - Customize larger set OSNs
 - Collaborate with social scientists

Goals (Part 1, Introduction)

- Which features of OSNS are popular and capture the users attention?
 - Yes
- What is the impact of OSNS on the network?
 - No, Not Enough
- What needs to be considered when designing future OSNs?
 - Maybe
- Is the user's behavior homogenous?
 - Yes

Next Set of Goals

- Influence & Improve Service Offerings
 - No
- Researchers can propose improvements or simplifications for existing OSNs or design OSNs with novel features
 - No
- From ISP perspective, OSNs might add features that increase per-user bandwidth demand such as video or live streaming
 - Maybe
 - Messaging increases during lunch time & early evening
 - Photos and apps more popular in afternoon to early evening

What I would do...

- Better analyze this data from ISP and OSN perspective
- I believe current data is too much from neat ideas learned using clickstream
- Better structure results
- ISP: I care about minimizing cost, done through routing patterns and contracts
 - What & how much data is used at times
- OSN: I care about better user experience
 - What are user experiences like and how can I make this faster?