Poking Facebook:
Characterization of OSN Applications

Minas Gjoka, Michael Sirivianos,
Athina Markopoulou, Xiaowei Yang

University of California, Irvine

Outline

- Motivation and contributions
- Datasets description
- Data analysis
- User Coverage
- Conclusion

Motivation

- **Very popular online social networks**
  - Facebook – 70 million users
  - overall estimated 270 million users in all OSNs

- **In May 2007, Facebook opened their platform to third-party developers for online applications**
  - in mid-February 2008, 866M installations of 16.7K distinct Facebook applications, 200K developers

- **Application popularity and adoption dynamics**
  - engineering and marketing reasons

Contributions

- First study of applications popularity and user reach in online social networks
  - Aggregate Application Popularity.
  - Popularity of Individual Applications.

- Simple and intuitive method
  - simulates the application installation process
  - captures user coverage from the popularity of applications

Outline

- Motivation and contributions
- Datasets description
- Data analysis
- User Coverage
- Conclusion

Data sets

- ## Data Set I, crawled from Adonomics
  - (day, application, #installations, #daily active users)
  - 170-day period until mid-February.

- ## Data Set II, crawled directly from Facebook
  - a subset of Facebook user profiles (300K)
  - (user ID, list of installed applications)

- ## Crawling/analysis scripts publicly available:
  - http://www.ics.uci.edu/~mgjoka/facebook

Outline

- Motivation and contributions
- Datasets description
- Data analysis
    - Aggregate Facebook application statistics
    - Popularity of individual applications
    - Application categories
- User Coverage
- Conclusion

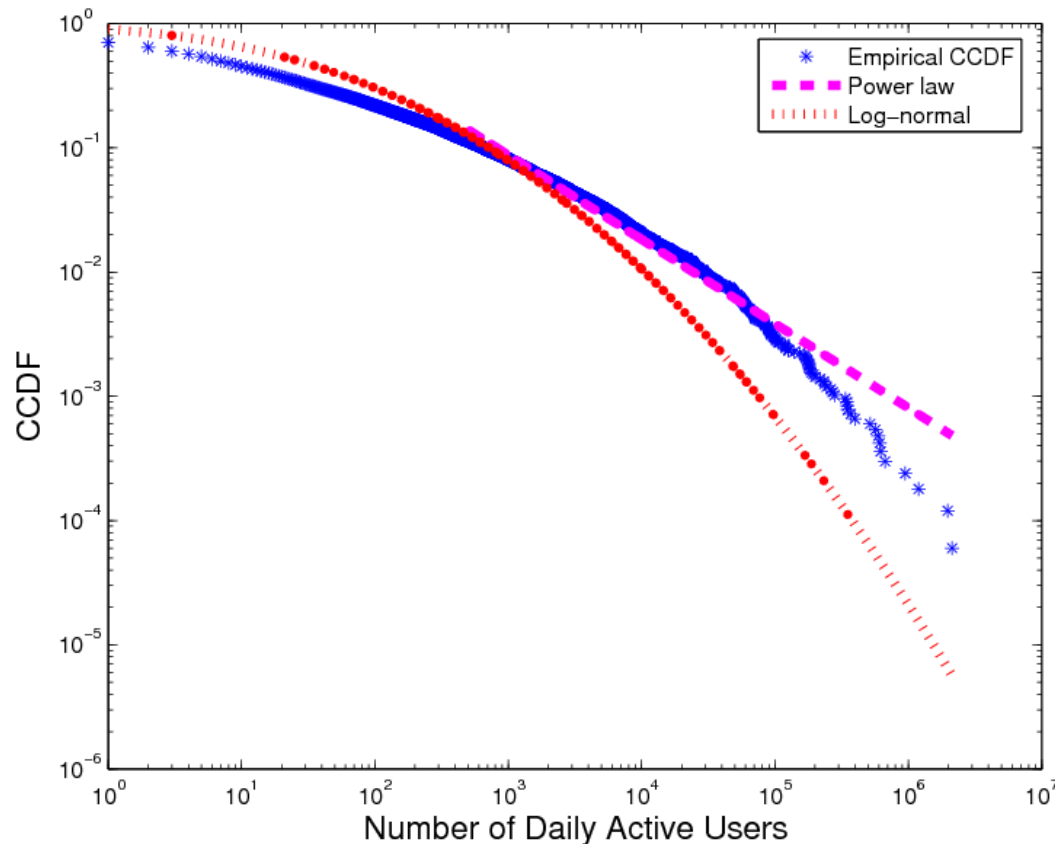# Facebook Applications
Aggregate Installation and Usage
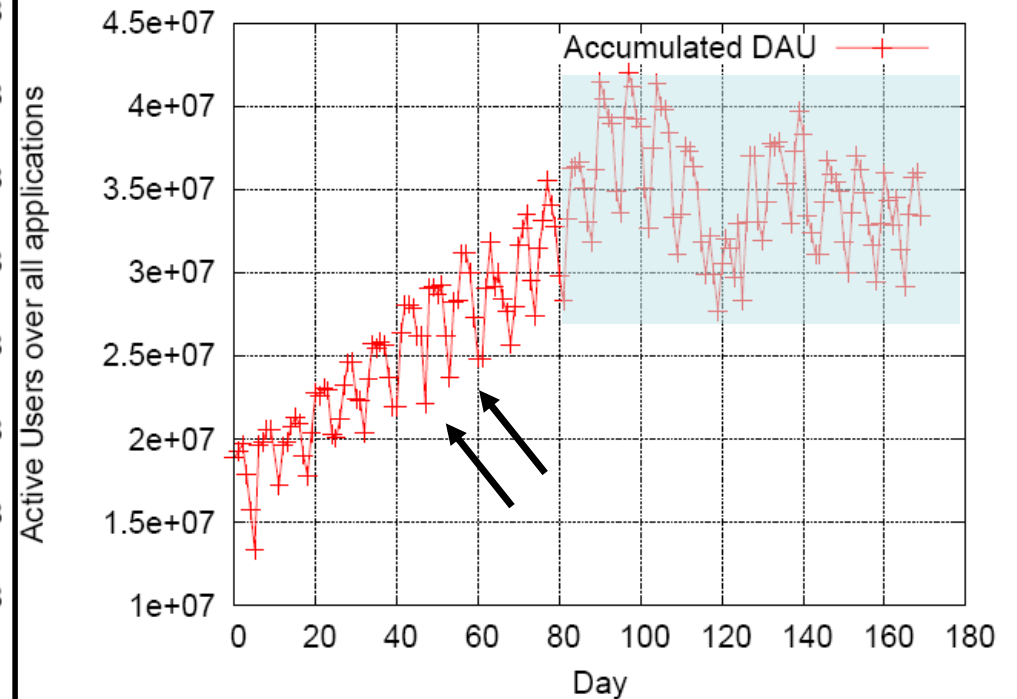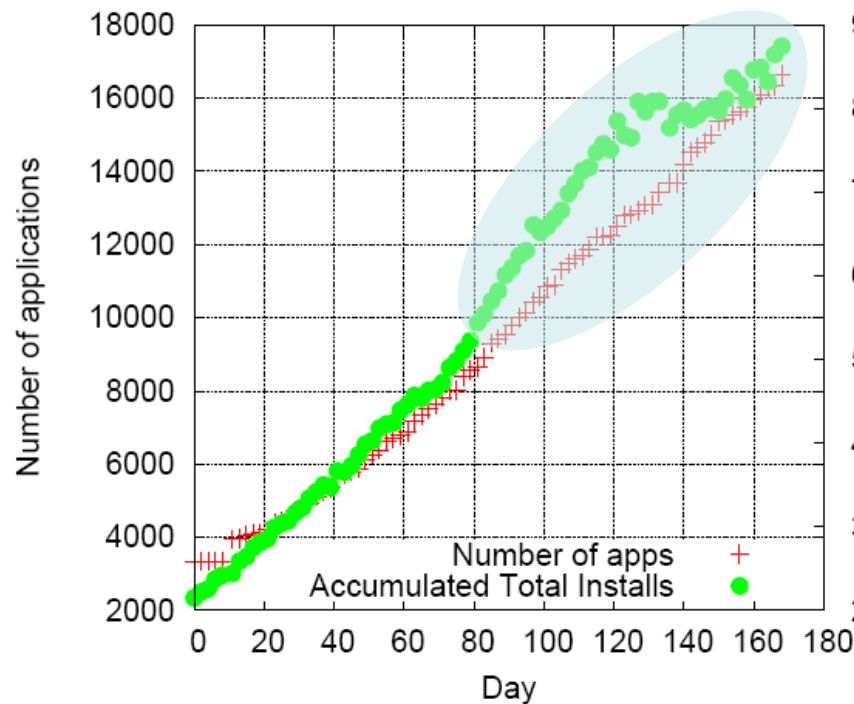
Weekly usage pattern
Average user activity decreases

$$\frac{DAU}{Total\ Installs}$$
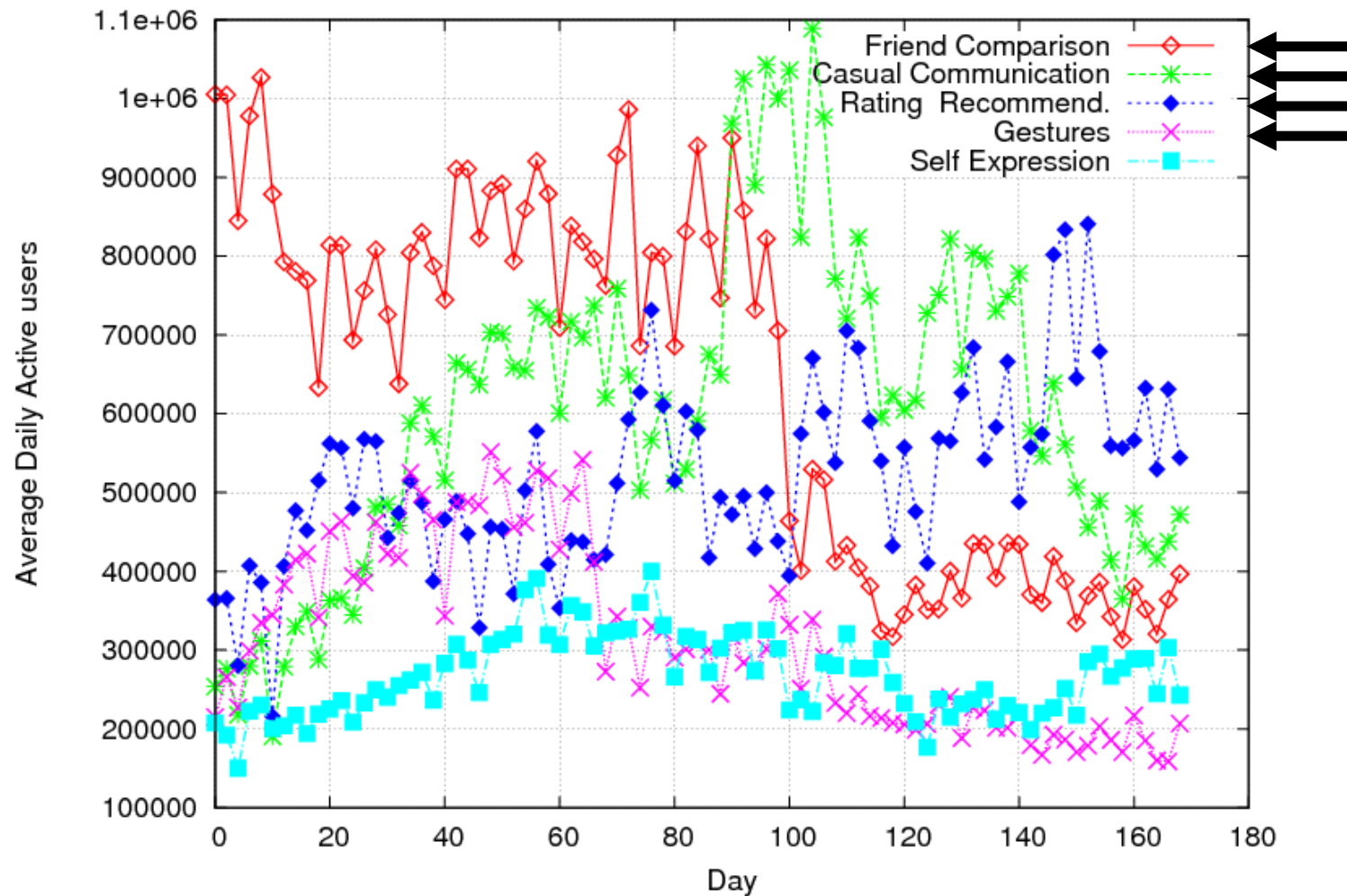
Application Popularity
Individual Applications: Daily Active Users

- Highly skewed distribution

- Not a power law

- Pick top-5 applications daily: only 17 unique apps in the 170-day period

Application Popularity
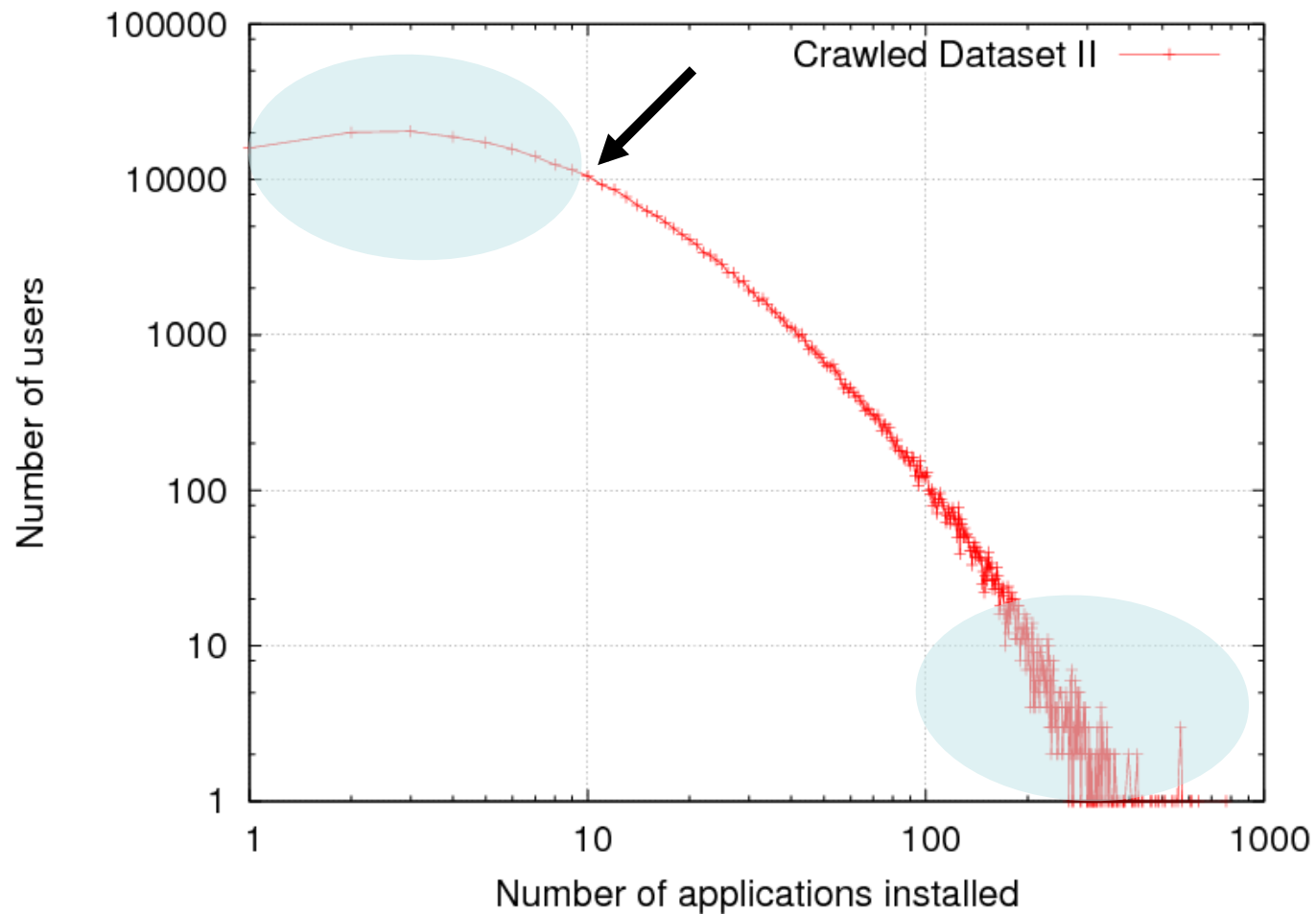## The effect of application category

Outline

- Motivation and contributions
- Datasets description
- Data analysis
- User Coverage
  - Model
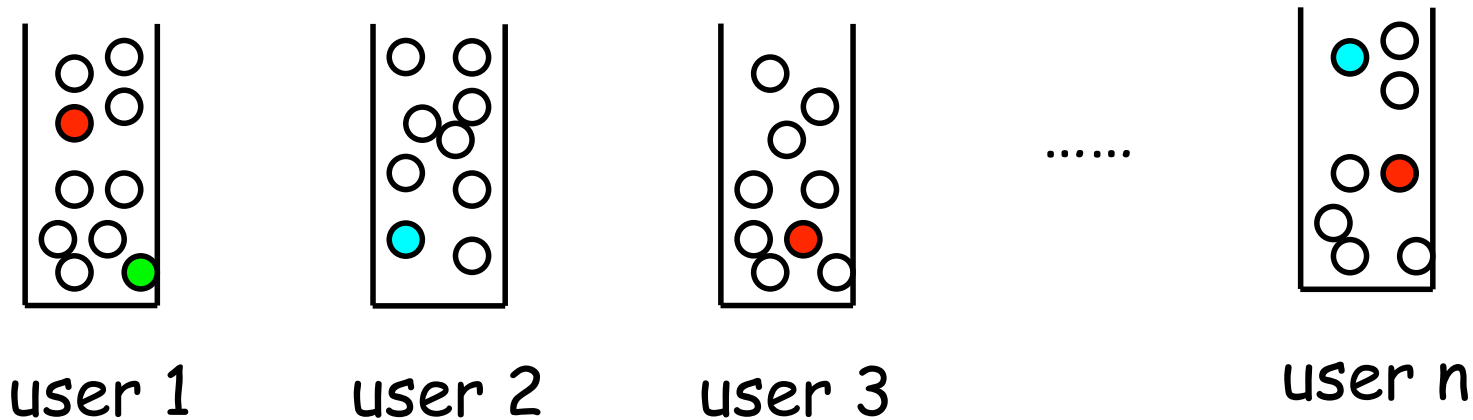 - Validation Simulations
- Conclusion

Number of applications per user
Dataset II

Users and Applications

- Popularity of applications is publicly available.
- Unknown how applications distributed among users
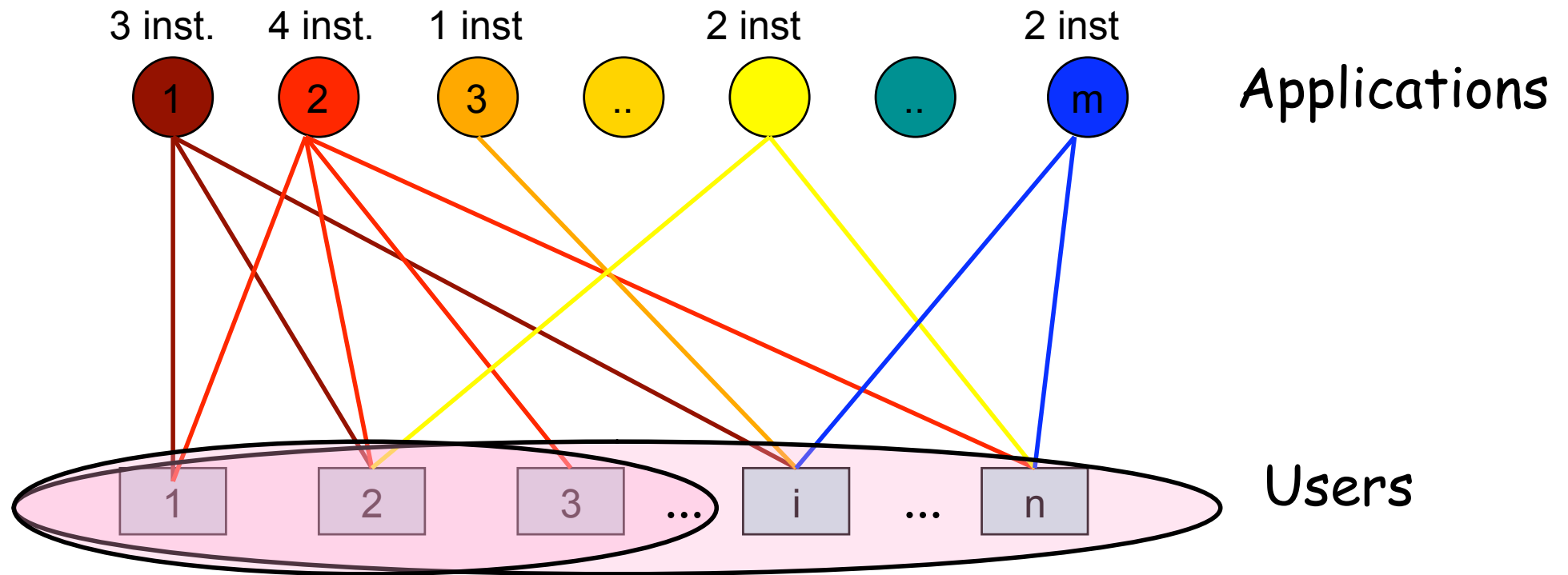- Example of usefulness: advertising

user 1 user 2 user 3   user n

n(users) on the order of millions

m(unique apps) on the order of thousands

total installations on the order of hundreds of millions
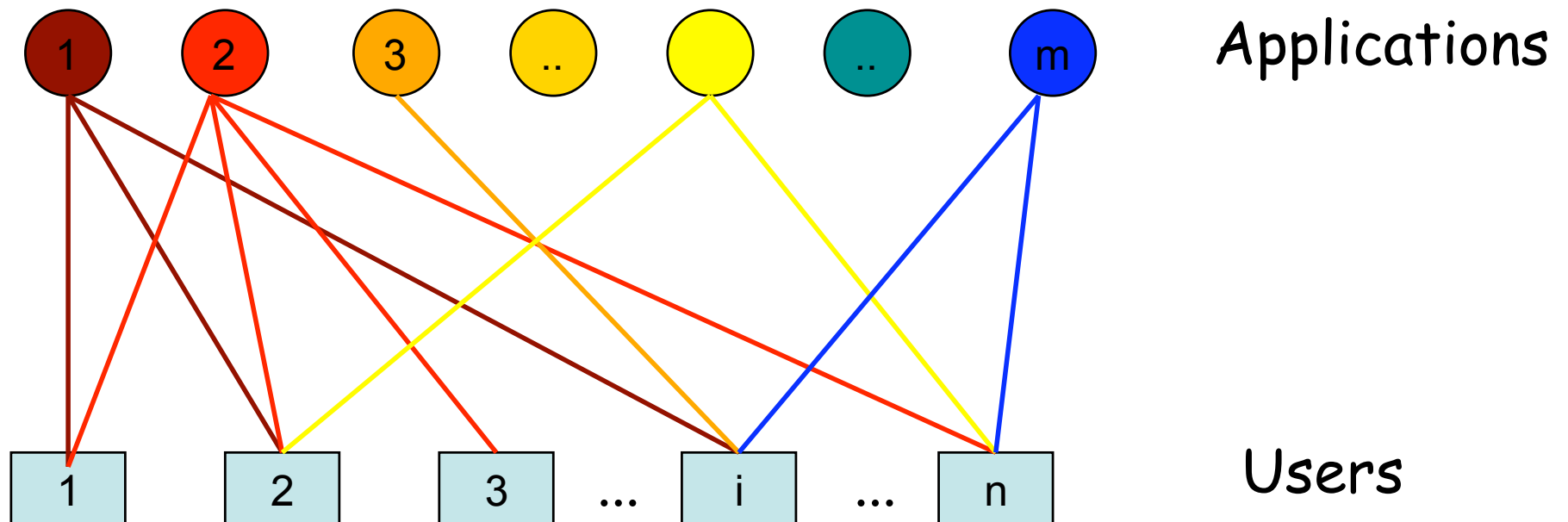
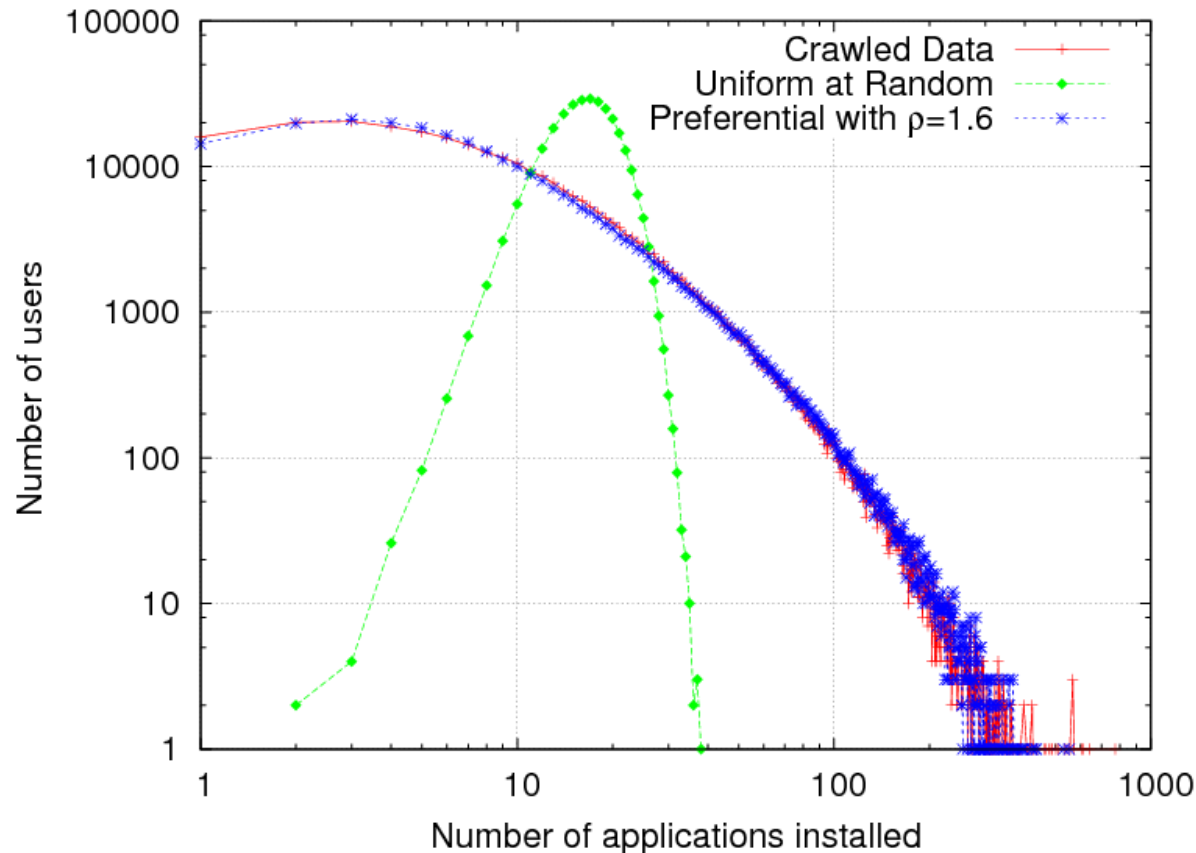# Users-Applications
Model (1)

Users-Applications
Model (2)

- Simulate a preferential installation process based on a balls and bins model:

$$P_{bin}(i) = \frac{balls(i)^\rho + init}{\sum_{j \in B}(balls(j)^\rho) + init}$$

Applications

Users

Users-Applications
Fitting

- We use the crawled dataset to fit the parameters of the model
- Clearly not uniform
- Good fit with ρ=1.6 and init=5

$$P_{bin}(i) = \frac{balls(i)^{\rho} + init}{\sum_{j \in B}(balls(j)^{\rho} + init)}$$

User Coverage
Simulation vs. Real Data (1)

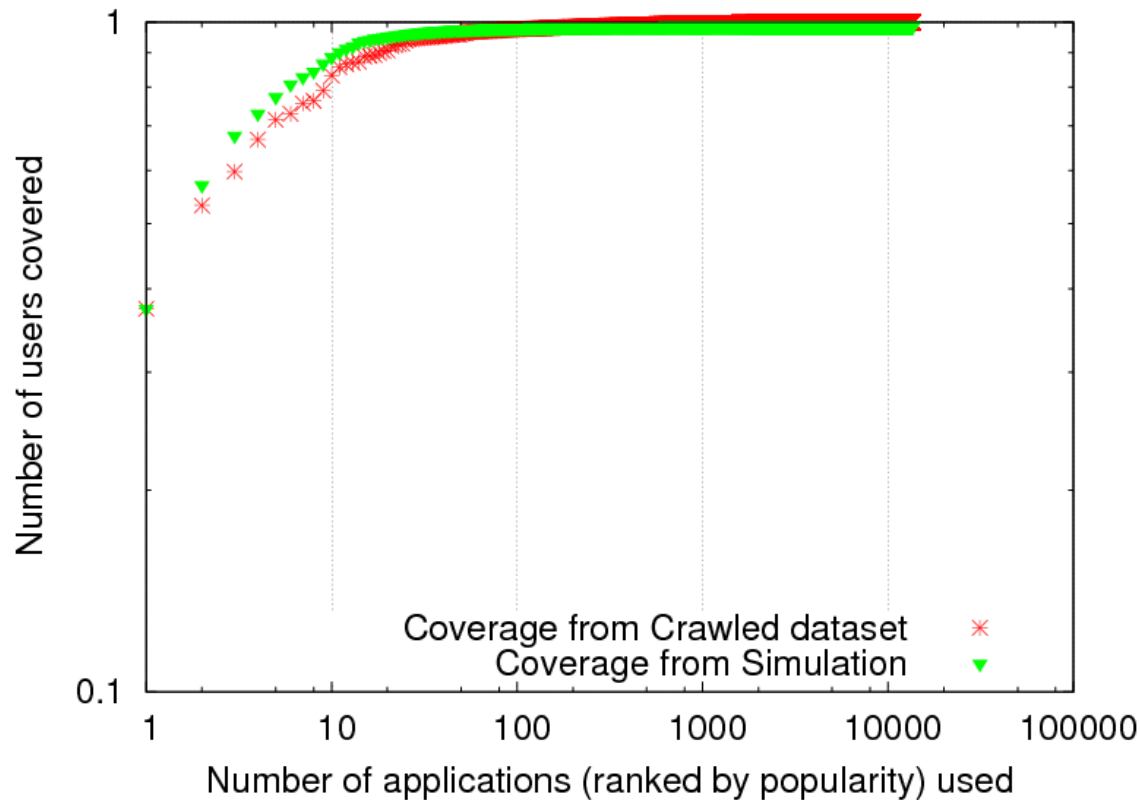| Application Name	Popul Rank	#Installations	Coverage Real(%)	Coverage Simulation(%)
Flixster	5	87609	30.2	30.2
Graffiti	15	45396	41.6	39.8
Flirtable	46	19504	43.9	42.6
Hug Me	99	9685	44.9	43.6
Nicknames	12	50825	51.5	51.1

} cumulatively

Total=213019 ← 73.5%

One instance of five apps randomly selected

User Coverage
Simulation vs. Real Data (2)

- User coverage for all applications cumulatively (taken in decreasing order of popularity)

- Simulation with fitted parameters agrees with crawled dataset

Conclusion

- A first study of FB application usage.
  - average user activity decreases
  - application installation process model

- Future extensions
  - study dynamic aspects, such as application virality.
  - further analysis through the balls and bins model

Questions?