An Empirical Study of Valuation and User Behavior in Social Networking Services (SNS)

Martin Falck-Ytter

Infrastructure Engineer

marf@steria.no

NTNU – Trondheim
Norwegian University of
Science and Technology

Harald Øverby

Associate Professor

haraldov@item.ntnu.no

Motivation

- Is each network connection of equal value?
- Does content productivity increase with network size?
- What generates value in a network and how do you model it?

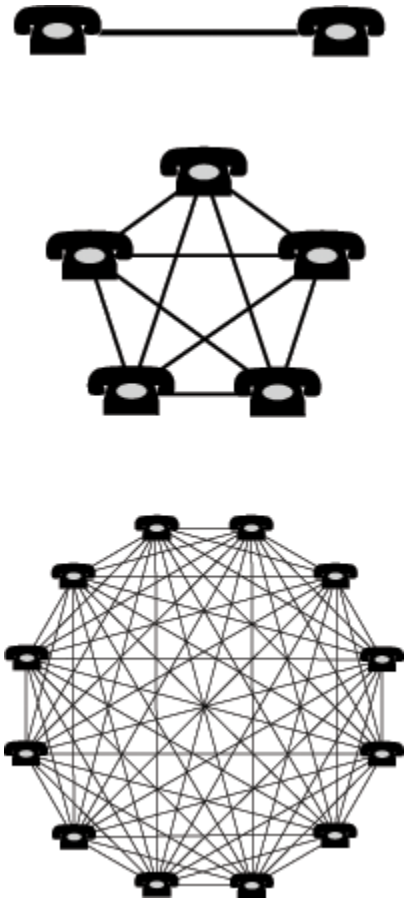
Agenda

I. **Introduction**
   - A. **Network Effects**
   - B. **Network Laws:**
     - **- Sarnoff's law**
     - **- Metcalfe's law**
     - **- Zipf's law**

II. Our Study
   - A. Content popularity: SNS and Zipf's Law
   - B. Correlation between productivity and network size in SNS
   - C. Proposed model for SNS valuation

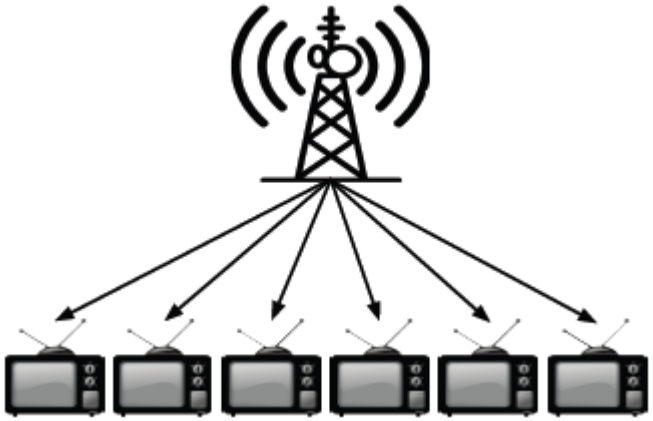
Network Effects

- Utility of consumption is affected by the number of other users using the same or compatible products
- Examples include telephone networks and social networking services



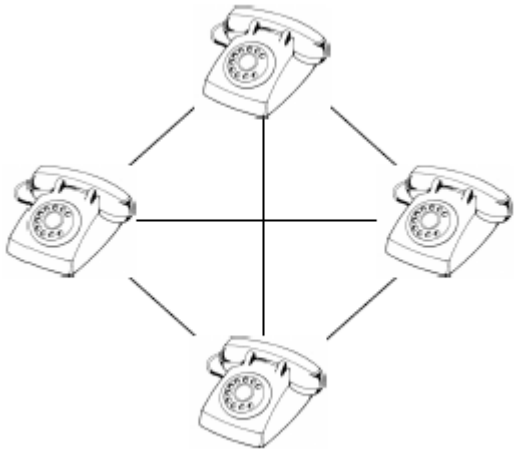
Sarnoff's law

- Applicable to broadcast networks
- Value proportional to number of subscribers
- $S(n) = n$



Metcalfe's law

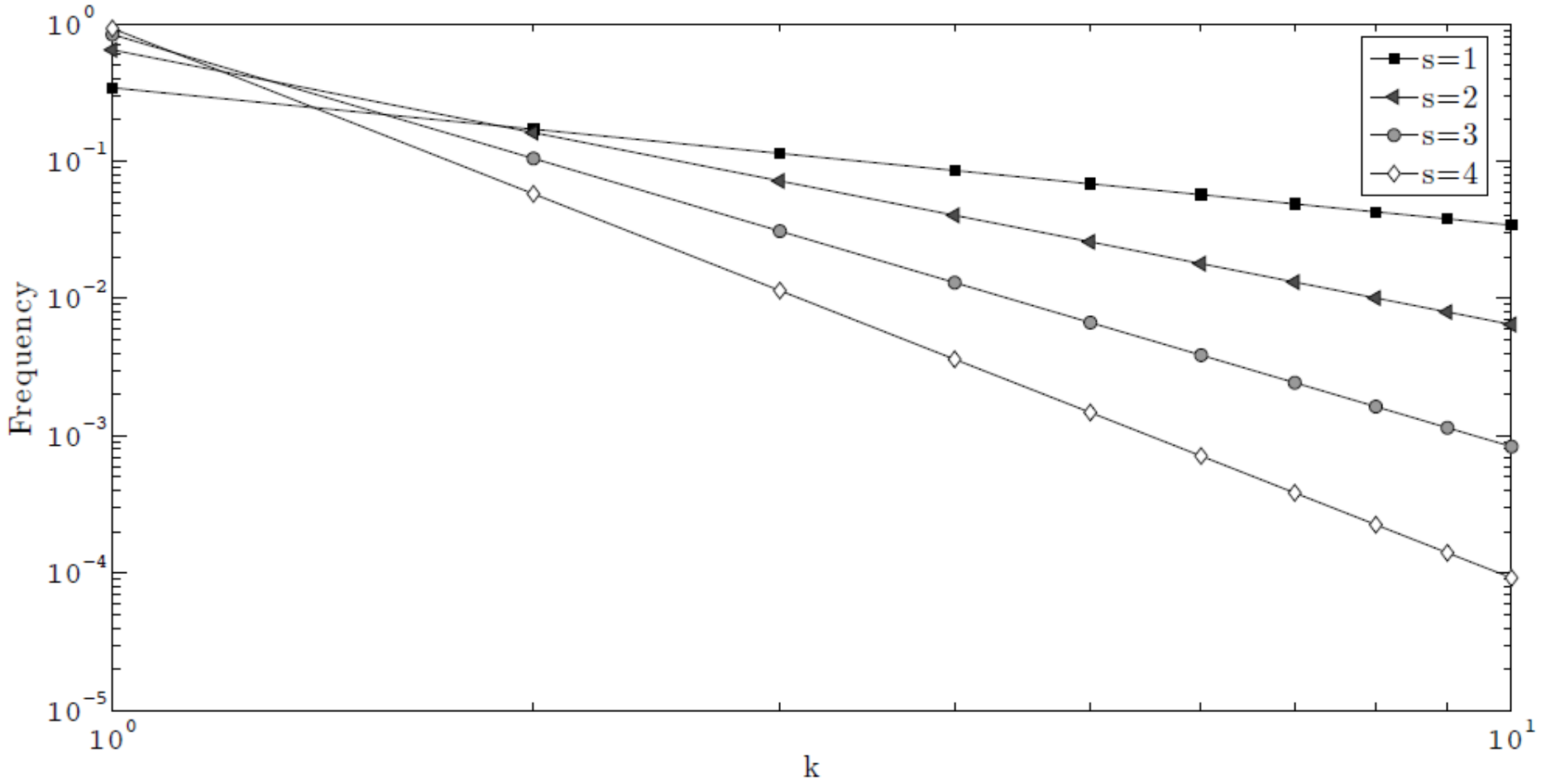
- Applicable to communication networks
  - Telephone
  - Fax
  - Email
- Value proportional to number of members squared
- $M(n) = \frac{(n-1)n}{2}$
- $M(n) \approx n^2$



Zipf's law

- Several types of data follow a Zipfian distribution.
- Zipf's law has proven to be very accurate for modeling popularity of data, such as words in the English language and sizes of large cities
- $f(k, s, N) = \frac{1/k^s}{\sum_{n=1}^{N} 1/n^s}, s > 0 \in \mathbb{R}, n \in \mathbb{I}$
  - *N* is the number of elements
  - *k* is their popularity rank
  - *s* is the value of the exponent characterizing the distribution

Zipf's law



Agenda

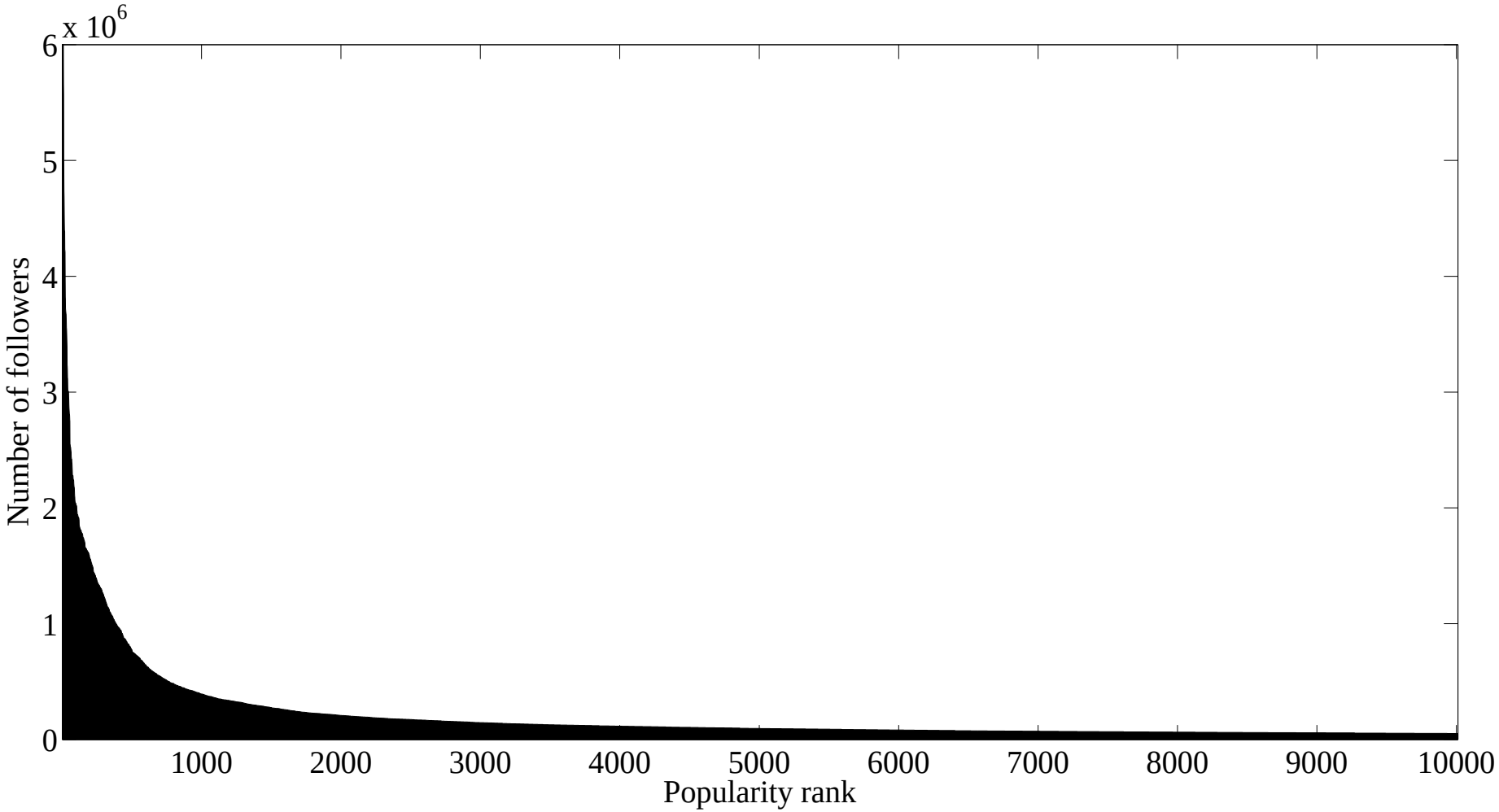
I. Introduction
   A. Network Effects
   B. Network Laws:
      - Sarnoff's law
      - Metcalfe's law
      - Zipf's law

II. Our Study
   A. Content popularity: SNS and Zipf's Law
   B. Correlation between productivity and network size in SNS
   C. Proposed model for SNS valuation

Content popularity - Twitter and Zipf's Law

- Zipf's law was used to model popularity of Twitter users
- An Internet page containing statistics for the 10 020 most popular Twitter users was used as data basis
- The value of the exponent, s, in Zipf's law was optimized to find the best-fit Zipf probability mass function

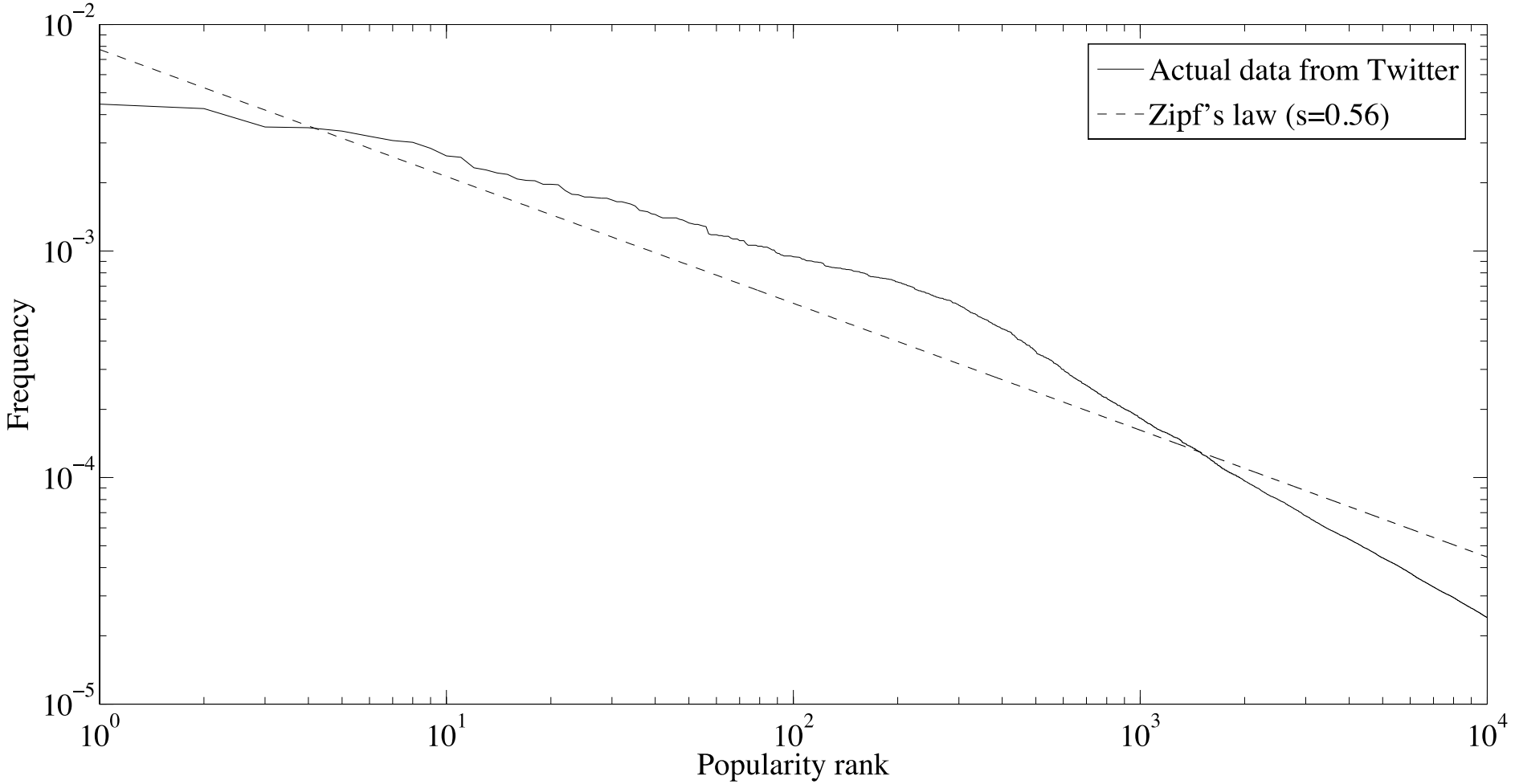
Raw data from Twitter



Optimization of the exponent, s (Twitter)

- A minimum sum of squares problem was solved to find the optimal value of the exponent, s:
- $min \sum_{k=1}^{10011} (f_k - \frac{1/k^s}{\sum_{n=1}^{10011} 1/n^s}) \; subject\ to\ s > 0$
- The optimal value of *s* was 0.56

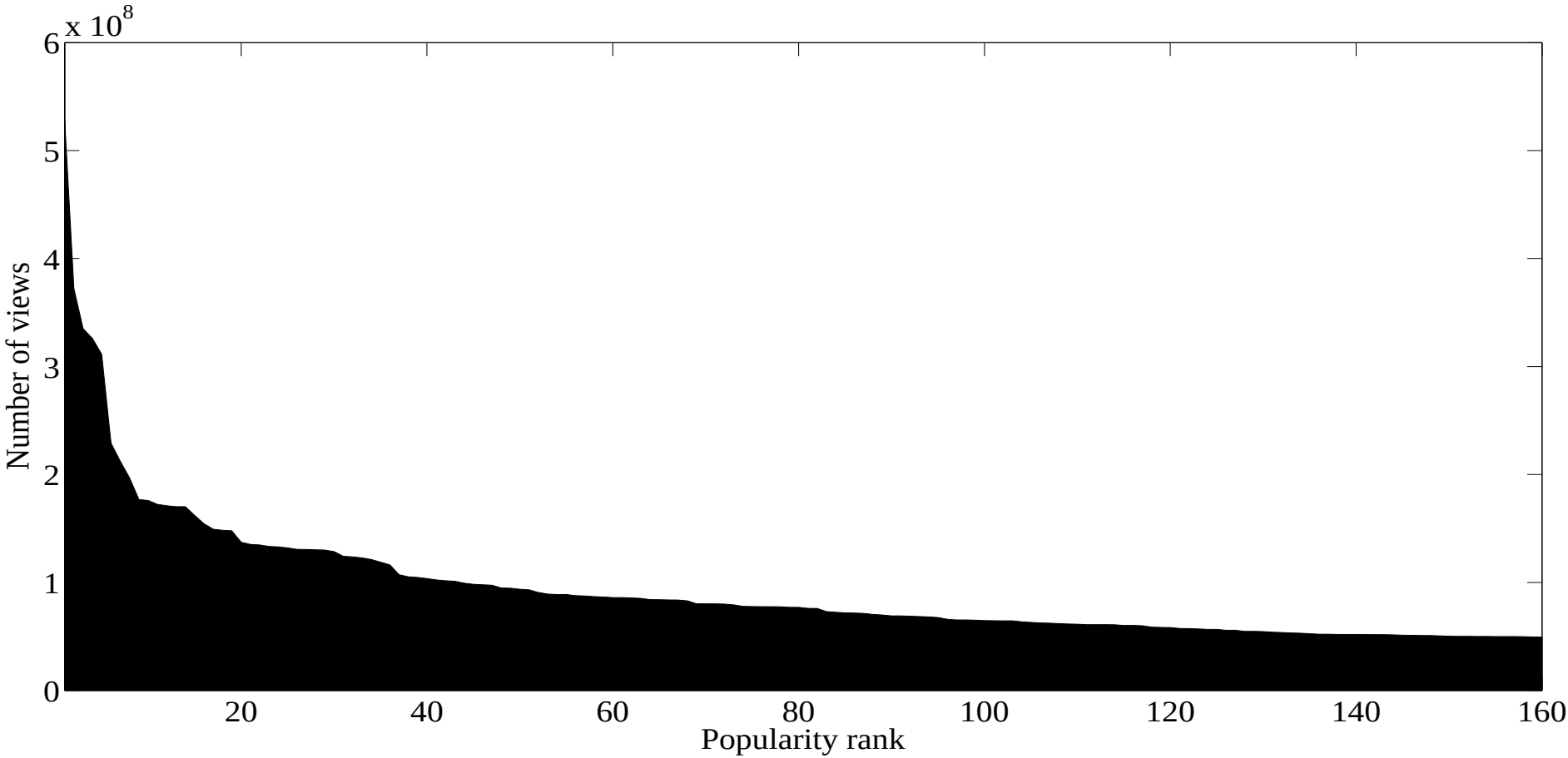
Twitter compared with Zipf's law



Content popularity - Youtube and Zipf's Law

- Zipf's law was used to model popularity of Youtube videos
- Number of views for the 160 most popular Youtube videos was retrieved
- The value of the exponent, s, in Zipf's law was optimized to find the best-fit Zipf probability mass function

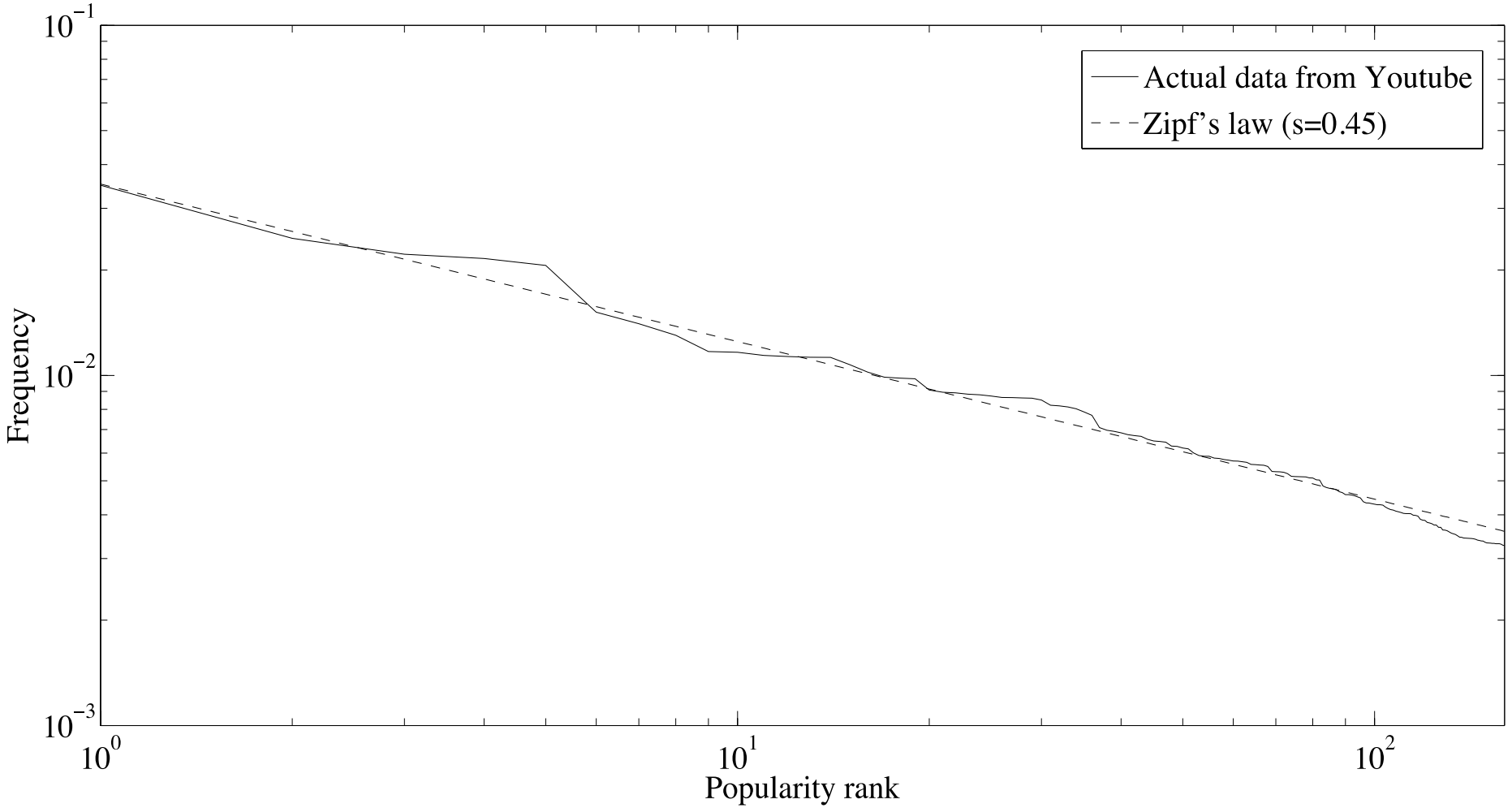
Raw data from Youtube



Optimization of the exponent, s (Youtube)

- The procedure performed was the same as with the fitting of Zipf's law with Twitter.
- The optimal value of the exponent this time, *s*, was 0.45

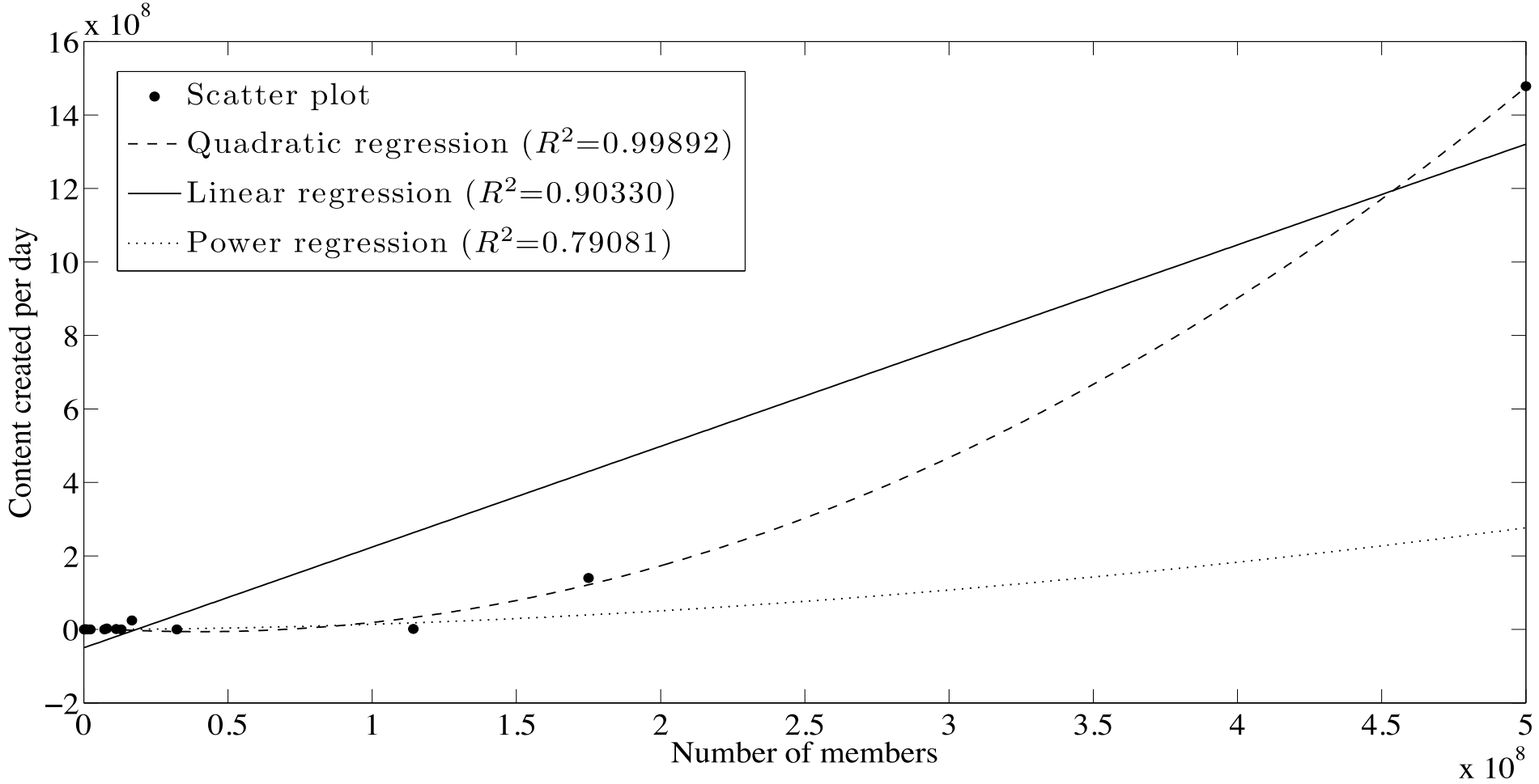
Youtube compared with Zipf's law



Correlation between productivity and network size in SNS

- The relationship between network size and content created in SNS was studied to see whether content productivity increases with network size
- 15 social networking services provided information about network size and content productivity
- Various best-fit functions were calculated and tested

Correlation between productivity and network size in SNS



Correlation between productivity and network size in SNS

- The quadratic model fitted the data significantly better than the linear model
- Consequently, average productivity increased with network size for SNS studied

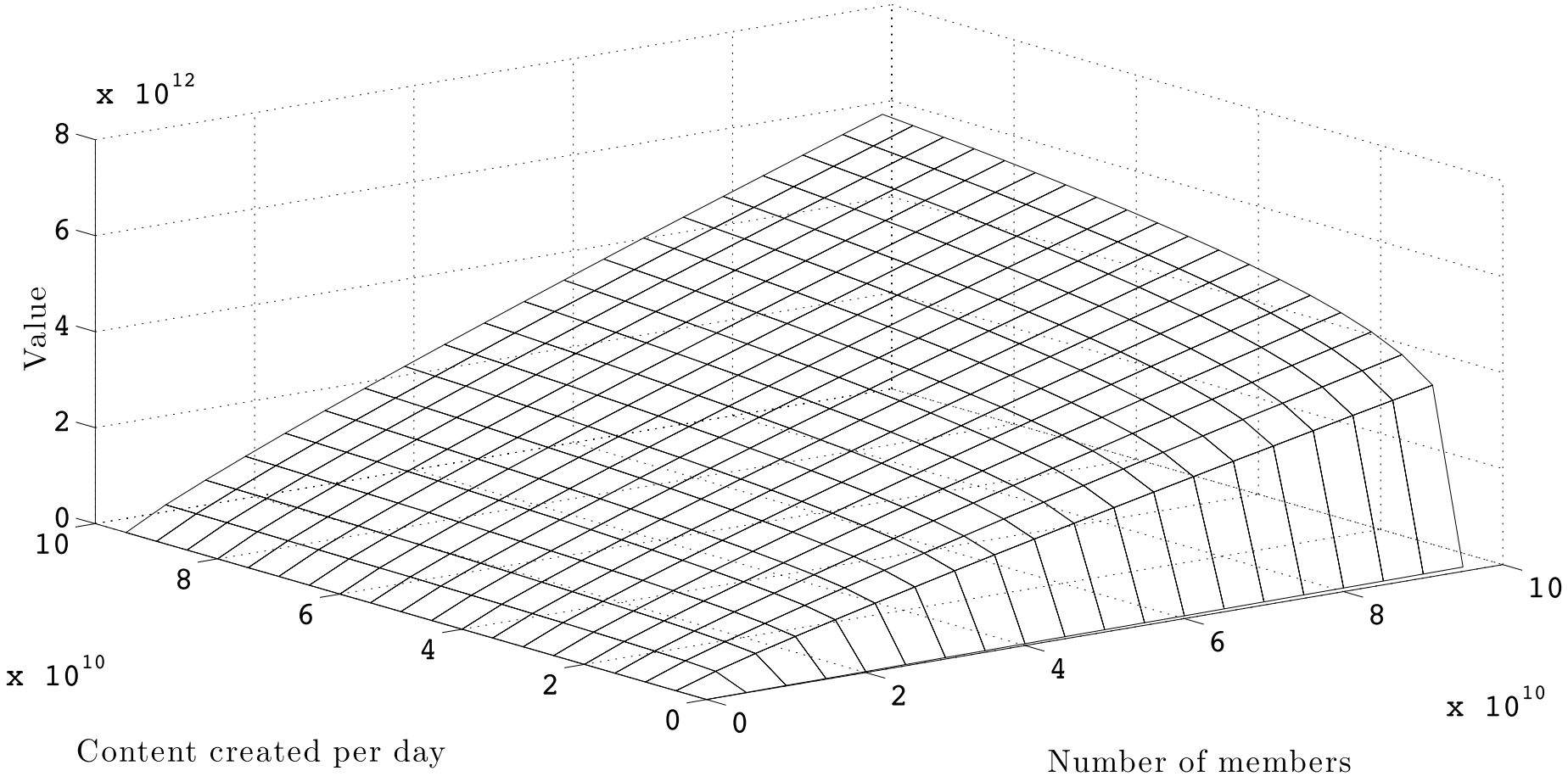
Estimated value of Social Networking Services

- Three alternative response surface models for valuation of SNS were based on network size, average content created per day and actual market value in United States dollar
- Only five social networks were able to provide the information needed for our valuation model
- The software Mathematica 8 was used to calculate a best-fit linear, quadratic and power response surface

Estimated value of Social Networking Services

- The estimated best-fit linear and quadratic response surfaces had both undesirable properties. These models were therefore considered inappropriate
- The best-fit power response surface was:
- $V_{prs}(n, c) = 14.1514 n^{0.892437} c^{0.167022}$
  - where *n* is the network size and *c* average content created per day

Estimated value of Social Networking Services



Conclusions

- Zipf's law with *s=0.56* was not an accurate describer of popularity of Twitter members
- Zipf's law with *s=0.45* was a good describer of popularity of Youtube videos
- Content productivity increases with network size for SNS studied
- An empirical model for SNS valuation was proposed based on two variables: network size (n) and average content created per day (c). The best-fit response surface was the following power function:
  - $V_{prs}(n, c) = 14.1514 n^{0.892437} c^{0.167022}$

Questions?

Martin Falck-Ytter

Infrastructure Engineer

marf@steria.no