Statistical Analysis of the Social Network and Discussion Threads in Slashdot: A HORRIBLE PAPER

presentation by Mykell Miller

Why study Slashdot?

- Because you haven't heard of:
  - Non-nerds
  - Youtube's comment ranking system
  - Other news sites that allow comments
  - Websites were you can't post anonymously
- Because of the important relationships on it
  - People visit Slashdot to read articles, not comments
  - Most comments are not discussions

Slashdot NEWS FOR NERDS. STUFF THAT MATTERS.
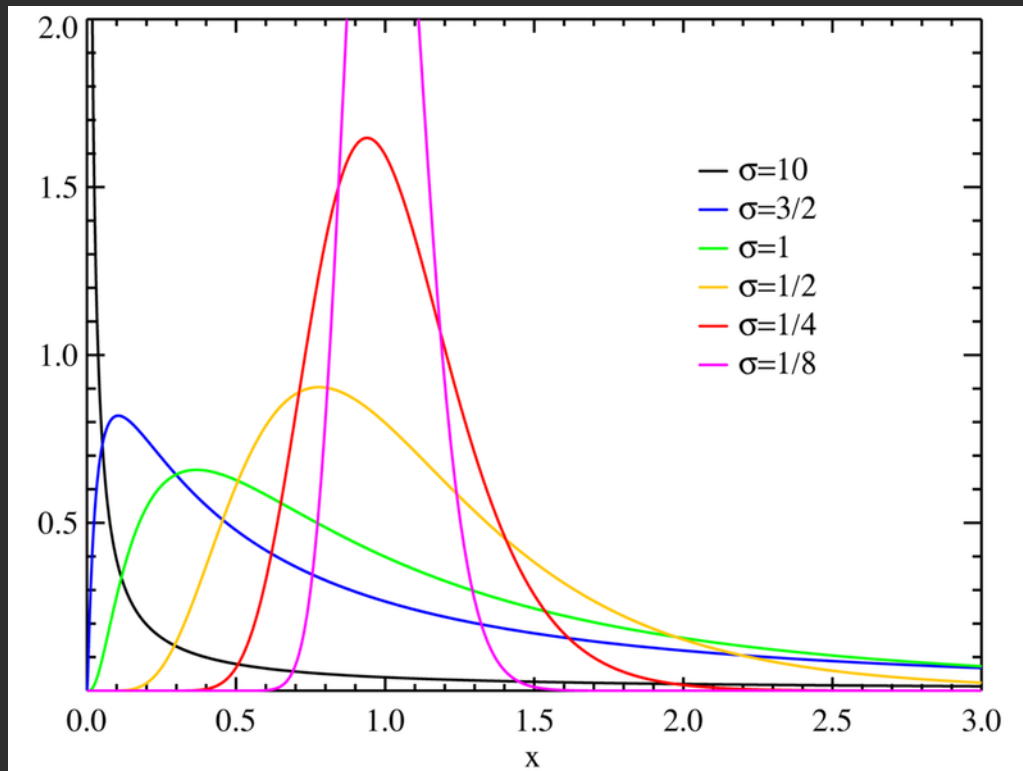
The Social Network: Building the Network

- Over 30% of comments were discarded
  - They're not studying the social network on Slashdot. They're studying their favorite part of it.
  - The discarded comments are not at all random – discarding them greatly changes the structure of the network

The Social Network: General Description

- The table on page 647 is horrible

  - What kind of random graph? Erdos-Renyi? With what probability of a link being formed?

  - So many abbreviations that it's impossible to read

The Social Network: Degree Distributions

- Power law vs lognormal distributions

The Social Network: Mixing by Score

- Discarding -1 comments
  - Don't you want to know how much spam, trolling, etc. is on Slashdot?
  - Or do you just not want to tarnish the image of your beloved Slashdot?
- Bias in favor of good writers
  - Duh

The Social Network: Community Structure

- Most clusters are of size 1

  – Most people ignore each other.

- This section is determining the strength of ties

  – But you ignored the research of Granovetter, the daddy of strength of ties

Structure of the Discussions:
Radial tree representation

- Most branches are only 1 deep

    - People usually don't reply to each other

    - It's not much of a network, it's more of an assortment

Structure of the Discussions:
The H-index as a Structural Measure of Controversy

- The worst section of a paper to ever be written

- Objective analysis of a subjective quality requires the objective measure to make sense

- Or the objective measure can be based off of something subjective but quantitative, such as someone ranking the level of controversy on a scale.

Structure of the Discussions:
The H-index as a Structural Measure of Controversy

- Good measures of controversy:

  - The frequency of swear words in comments

  - Sentiment analysis of comments

  - The results of a survey in which people ranked the level of controversy

- Barely passable measures of controversy:

  - One person ranks the level of controversy

- Horrible measure of controversy

  - The depth of the tree