1. UCE is more than annoying

- Unsolicited commercial email (UCE), a.k.a spam
  - Sent in bulk and unsolicited
  - Mostly unwanted by users: discarded in seconds
  - Among all email traffics: 55%-80% are spam
- Waste: Bandwidth, CPU and Storage, Productivity

Annoyance++

Productivity--

Money--

2. Get around effective content filters

- Content-based filter works on keywords/tokens, but:
  - Scrambled words
    - Mortgage, click to order
  - Obfuscated with HTML/CSS
- Images with excerpt from legitimate sources

3. Sender based detection

- Sender-based filter works on sender classifications:
  - Does not rely on contents
  - Sometimes faster (process only email header/log)
- 1. IP address blacklisting
- 2. Sender (Domain/IP address) reputation
  - Limited manual feedbacks from users
- How about sender behavior?

4. Social networks features

- Email social networks:
  - To represent email communication patterns
  - A directed graph
  - Email users as nodes, transaction as edge
  - A sent email to B = directed edge A→B
- Observe communication patterns
- Social networks analysis
- Human users communicate in social groups
- Spammers high fan-out, low response rate

5. A learning approach

- Email transaction logs
- Email exchanges between users
- Email social networks
- Feature vectors, one per sender
- Extract communication features and patterns
- Feature weighting
- Vectors with weighted features
- Machine Learning
- Sender scores

- Machine Learning techniques on senders
  - Give score of spammer likelihood
  - Supervised/semi-supervised learning?
  - Unsupervised learning?

6. Filtering or resisting

- With a score for each sender, get a score for each email
  - Parallel single threshold approach

7. Preliminary results

- Score histogram for legitimate senders
- Score histogram for spammers

8. Conclusions and references

- Exploit social networks features to complement content-based filters
- Components to be further explored and evaluated
  1. Feature selection and extraction
  2. Machine learning details
  3. Filtering and resisting schemes