



# Sybil-resilient online content voting

Nguyen Tran, Bonan Min, Jinyang Li, and  
Lakshminarayanan Subramanian

New York University  
NSDI'09

# Website as a response to interesting marketing

The screenshot shows the Digg website interface. At the top, there is a dark blue header with the 'digg' logo on the left and navigation links for 'My Profile', 'Friends' Activity 0', and 'Submit New' on the right. Below this is a green navigation bar with dropdown menus for 'Technology', 'World & Business', 'Science', 'Gaming', 'Lifestyle', 'Entertainment', and 'Sports'. A secondary green bar contains buttons for 'Popular', 'Upcoming BETA', 'News', 'Videos', 'Images', and 'Customize'. The main content area is titled 'News, Images, Videos' and includes a 'Most Recent' filter and other time-based filters: 'Top in 24 Hr', '7 Days', '30 Days', and '365 Days'. Three articles are listed:

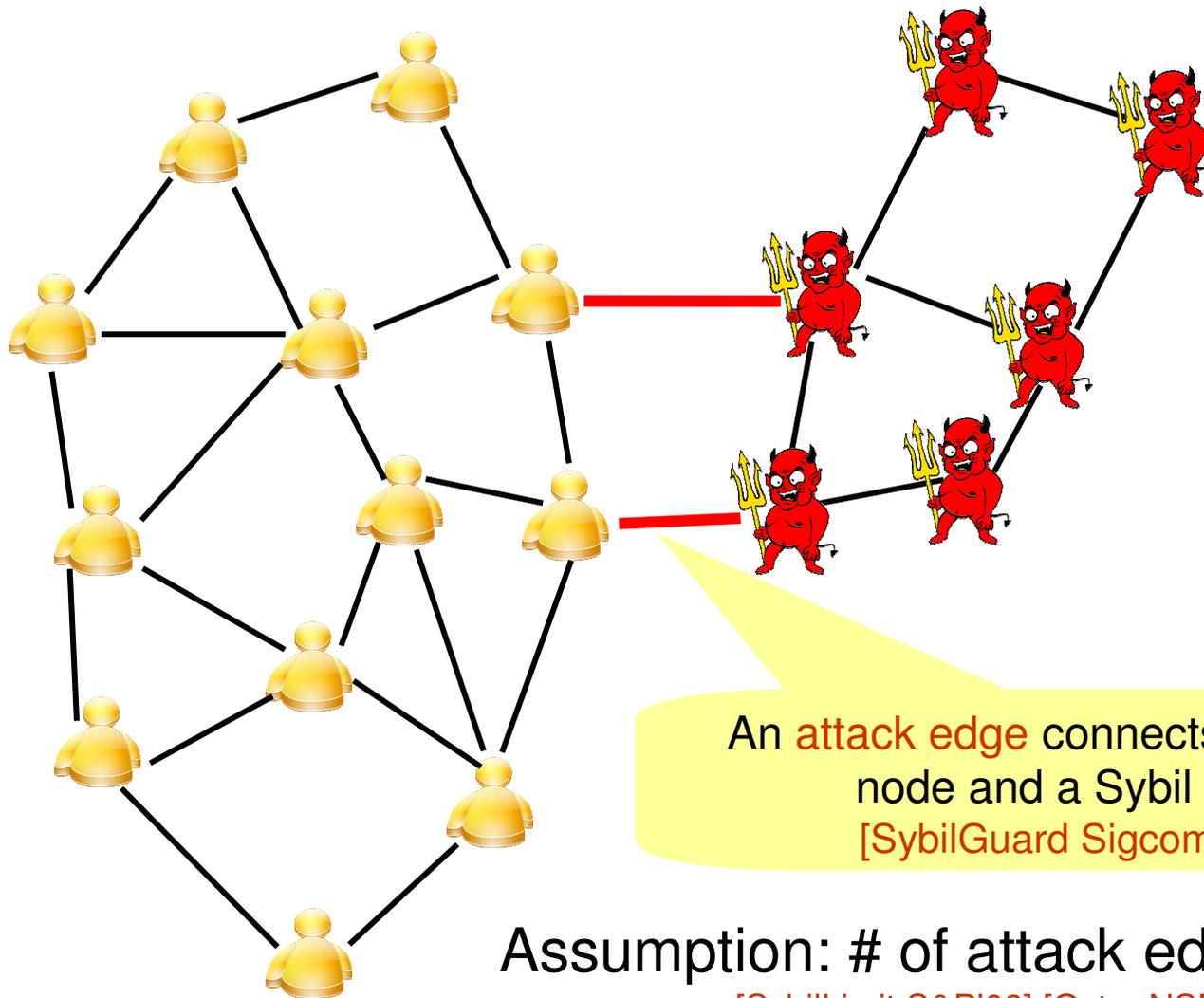
- 22 diggs** | [The Best Dam Water Money can Buy](#) | failedsuccess.com — A new product is sweeping the nation, breathing life into the dam industry. | Made popular **Jan 11, 2005**
- 14 diggs** | [REALLY Cheap Web Host](#) | mher.org — Plans starting at \$0.99; 8000 MBs of space and 150 GBs of bandwidth. Very fast and reliable. They never oversell. You can see examples of how fast at xeniac.com and slackpacks.org. | Made popular **Jun 18, 2005**
- 20 diggs** | [Free Ipod Shuffle?](#) | FreeiPodShuffle.com — MAC has just introduced a new Ipod to the bunch. The Ipod Shuffle is going to be a HOT item this year! | Made popular **Feb 2, 2005**

On the right side, there is a vertical sidebar with a blue header, a red 'ESPN' logo, and a section titled 'Top in All Time' with three items: 3307 diggs, 2034 diggs, and 1555 diggs, each with a small thumbnail image.

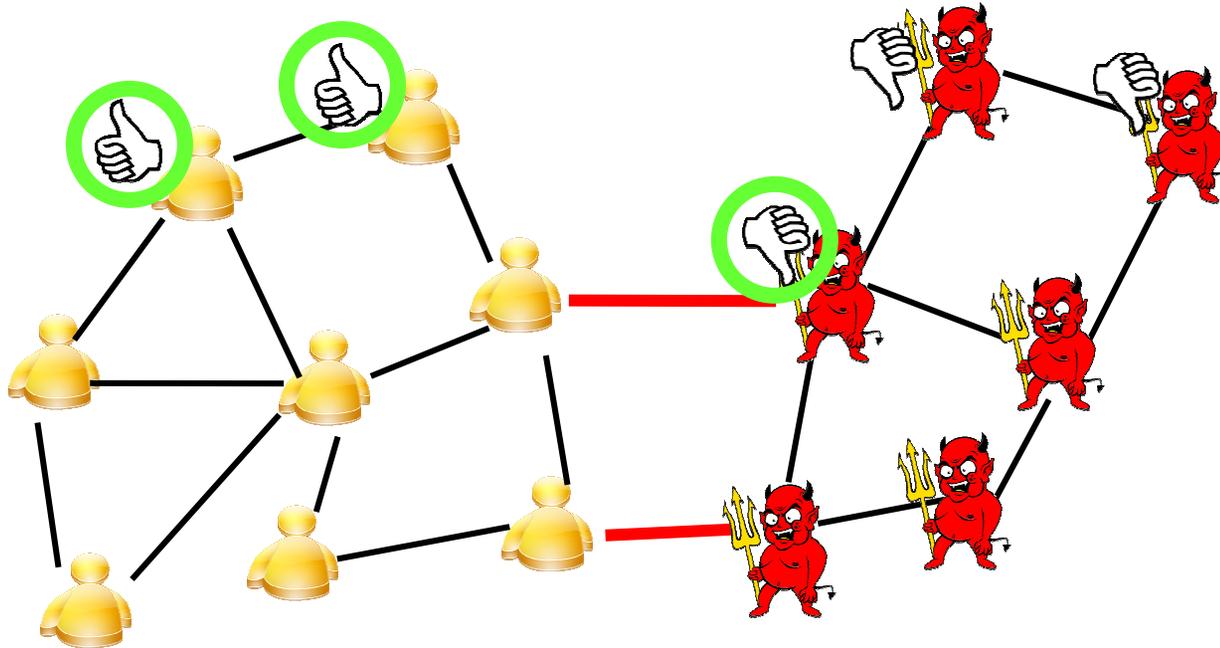
# Sybil defense is hard

- Open system → an attacker can join easily
- Lack of strong identity → an attacker can join with many fake accounts
- Need some resource that cannot be acquired in abundance
  - Links in a social network?

# Social links are hard to acquire in abundance

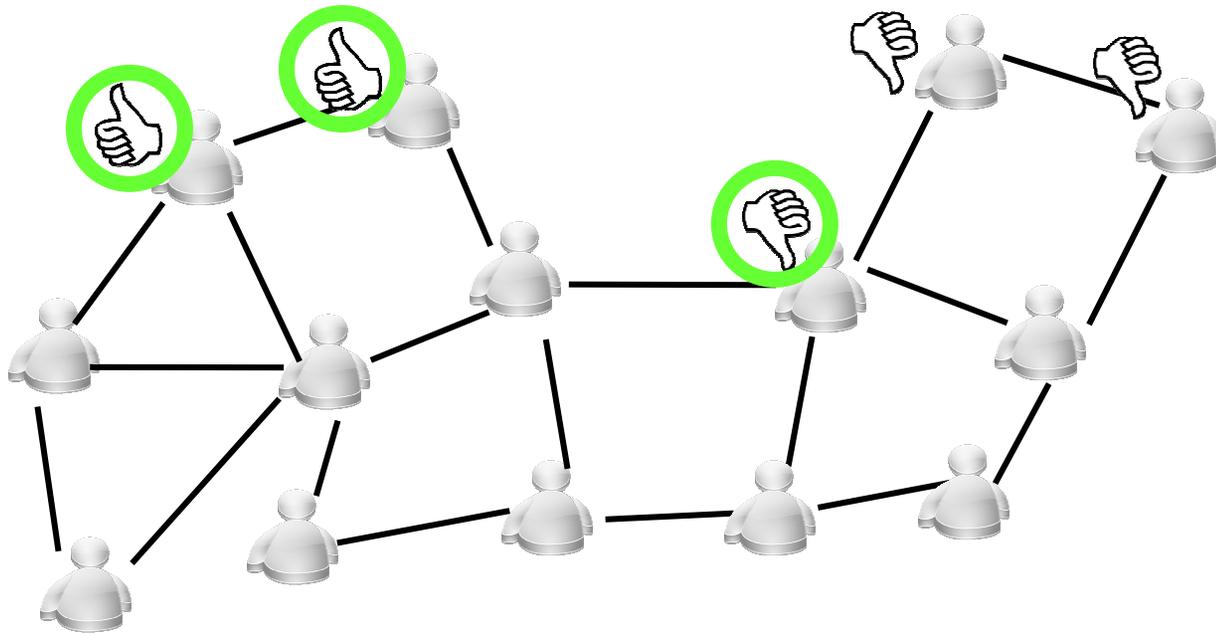


# SumUp: a Sybil-resilient vote aggregation system



- SumUp's setting: A central party collects all votes and the social graph
- Goal: extract a subset of votes
  - Include few votes from Sybils
  - Include most votes from honest users

# SumUp: a Sybil-resilient vote aggregation system

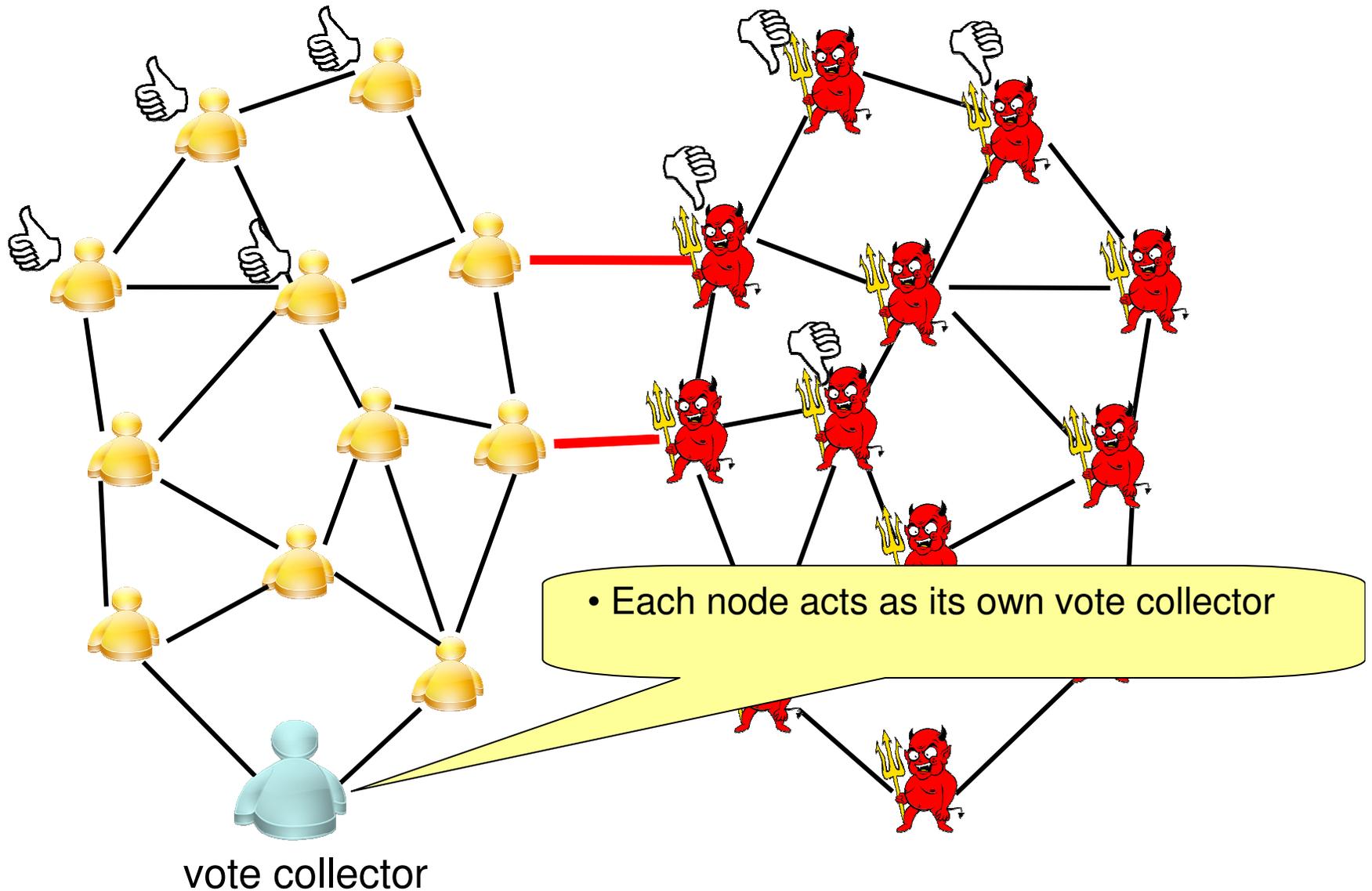


- SumUp's setting: A central party collects all votes and the social graph
- Goal: extract a subset of votes
  - Include few votes from Sybils
  - Include most votes from honest users

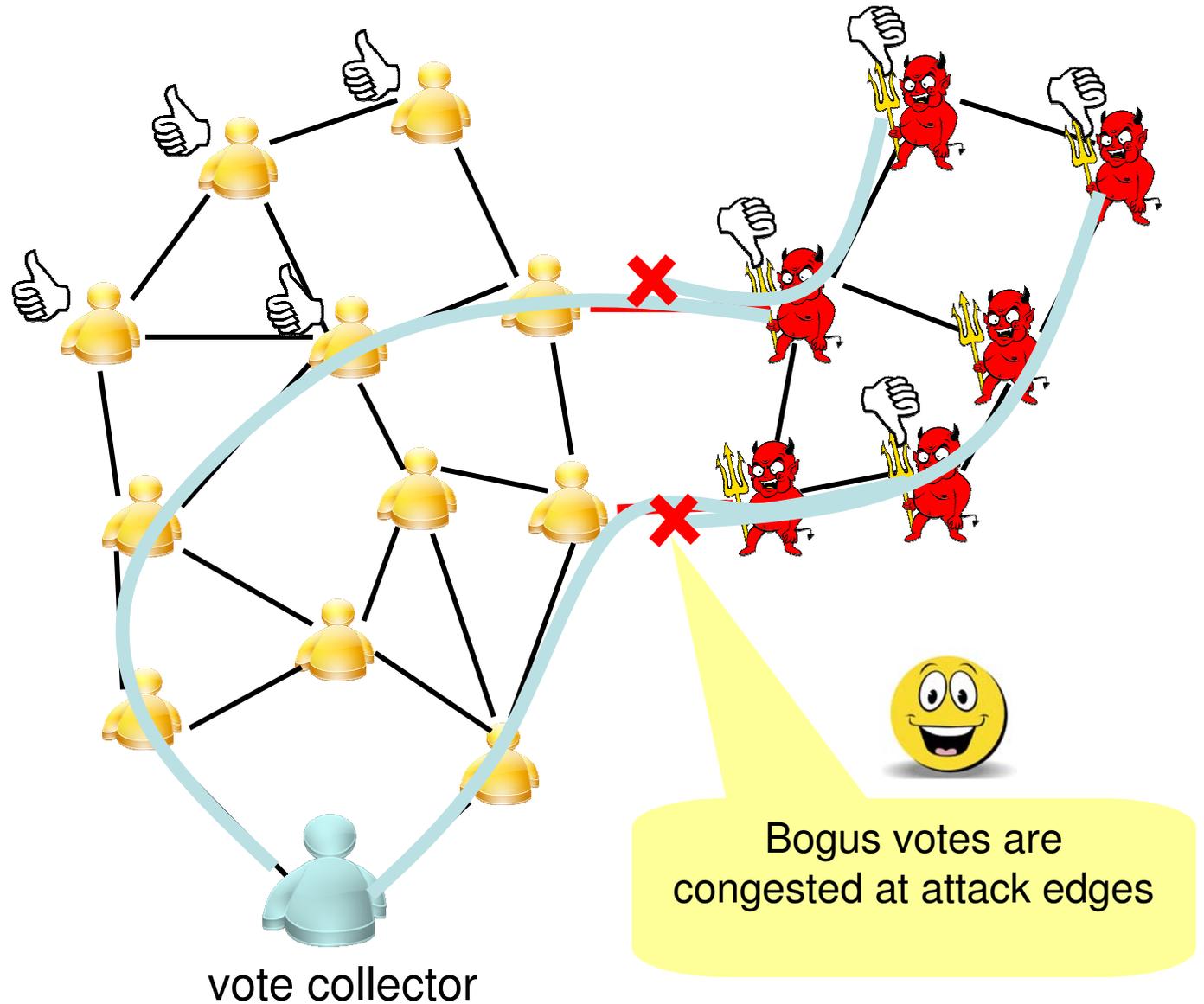
# SumUp overview

- Design insights
  1. Designate a vote collector
  2. Use max-flow to collect votes
  3. Assign appropriate link capacities

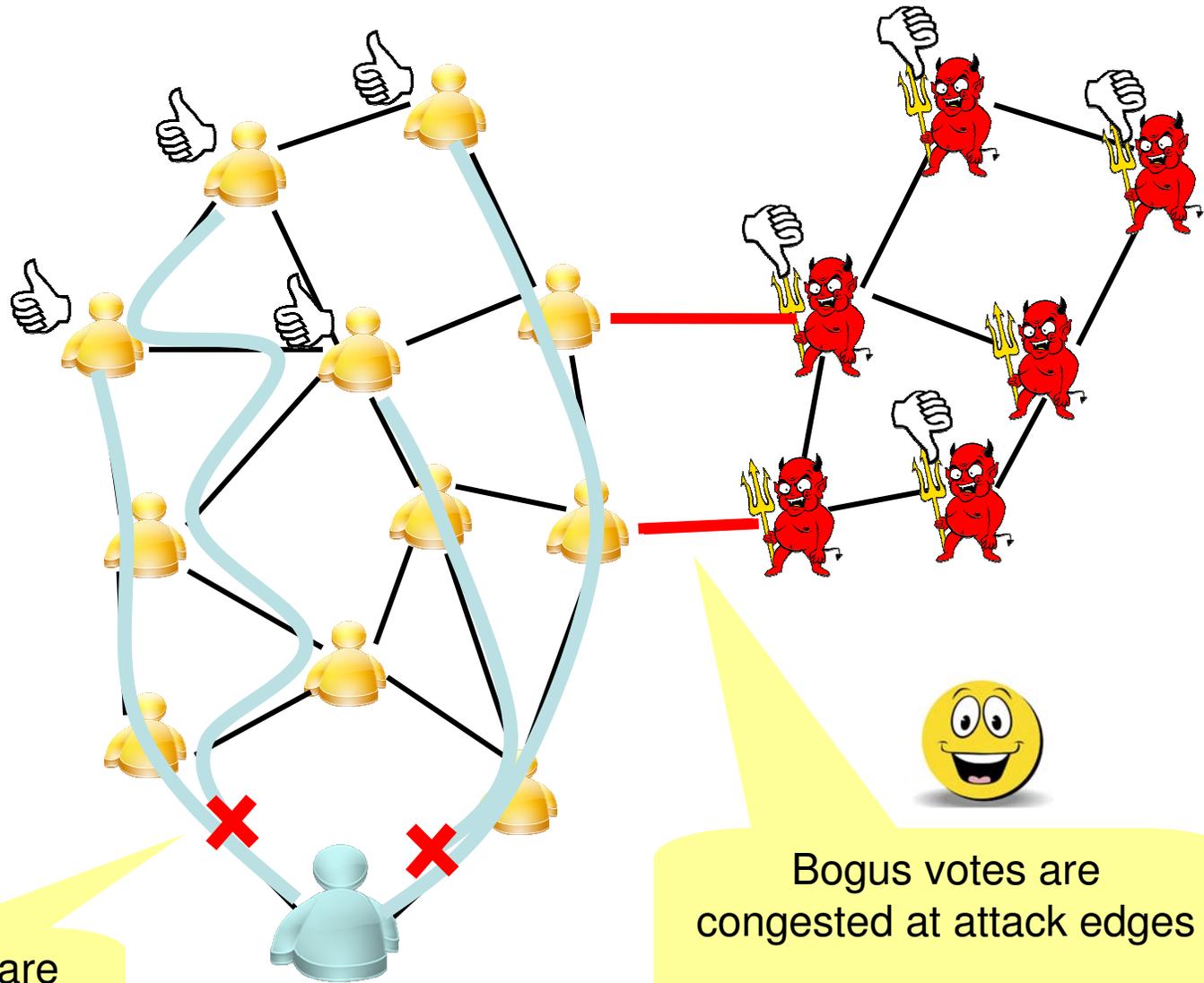
# #1: Designate a vote collector



# #2: Use max-flow to collect votes



# #2: Use max-flow to collect votes

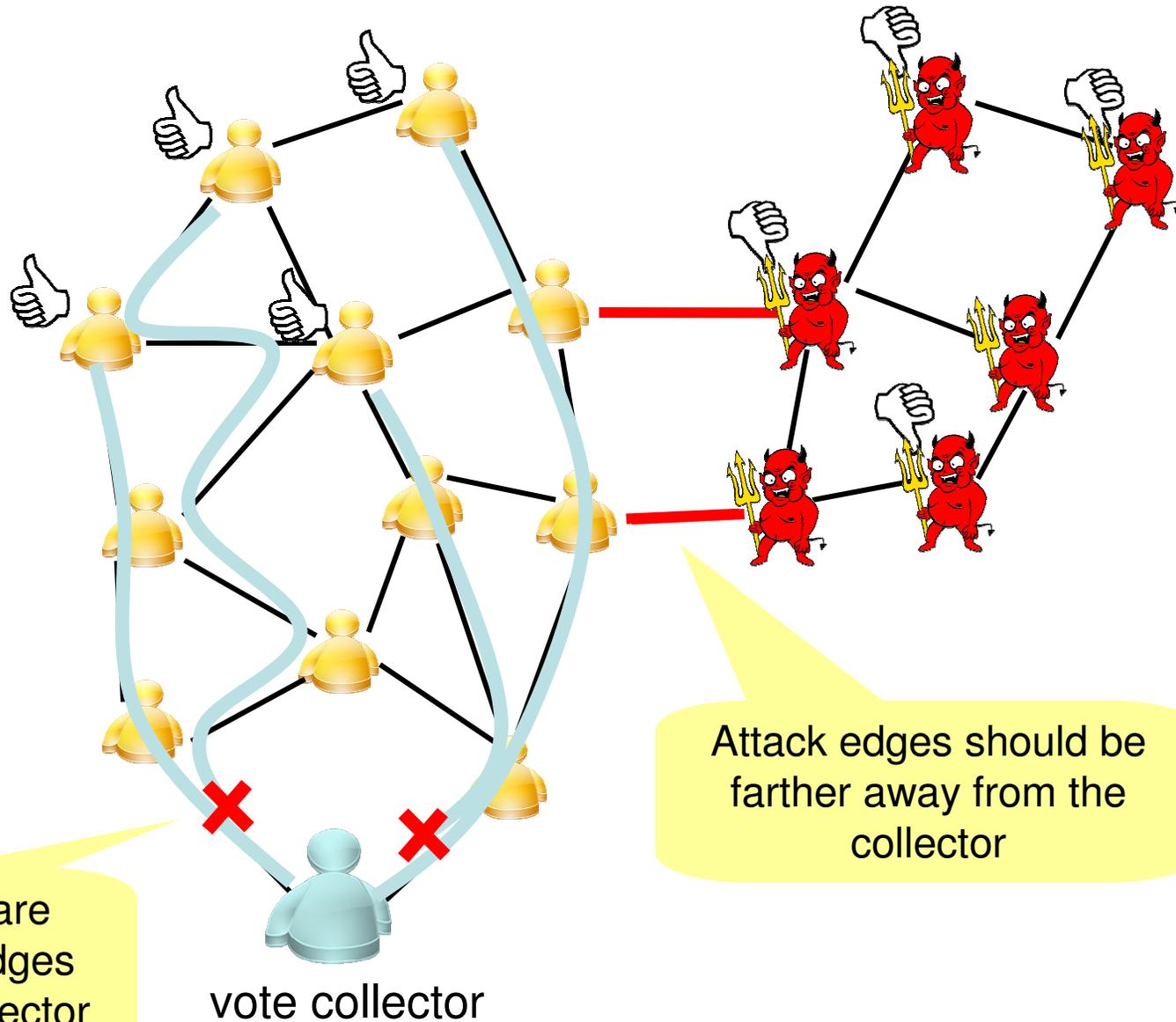


Honest votes are also congested

vote collector

Bogus votes are congested at attack edges

# #3: Assign appropriate link capacities



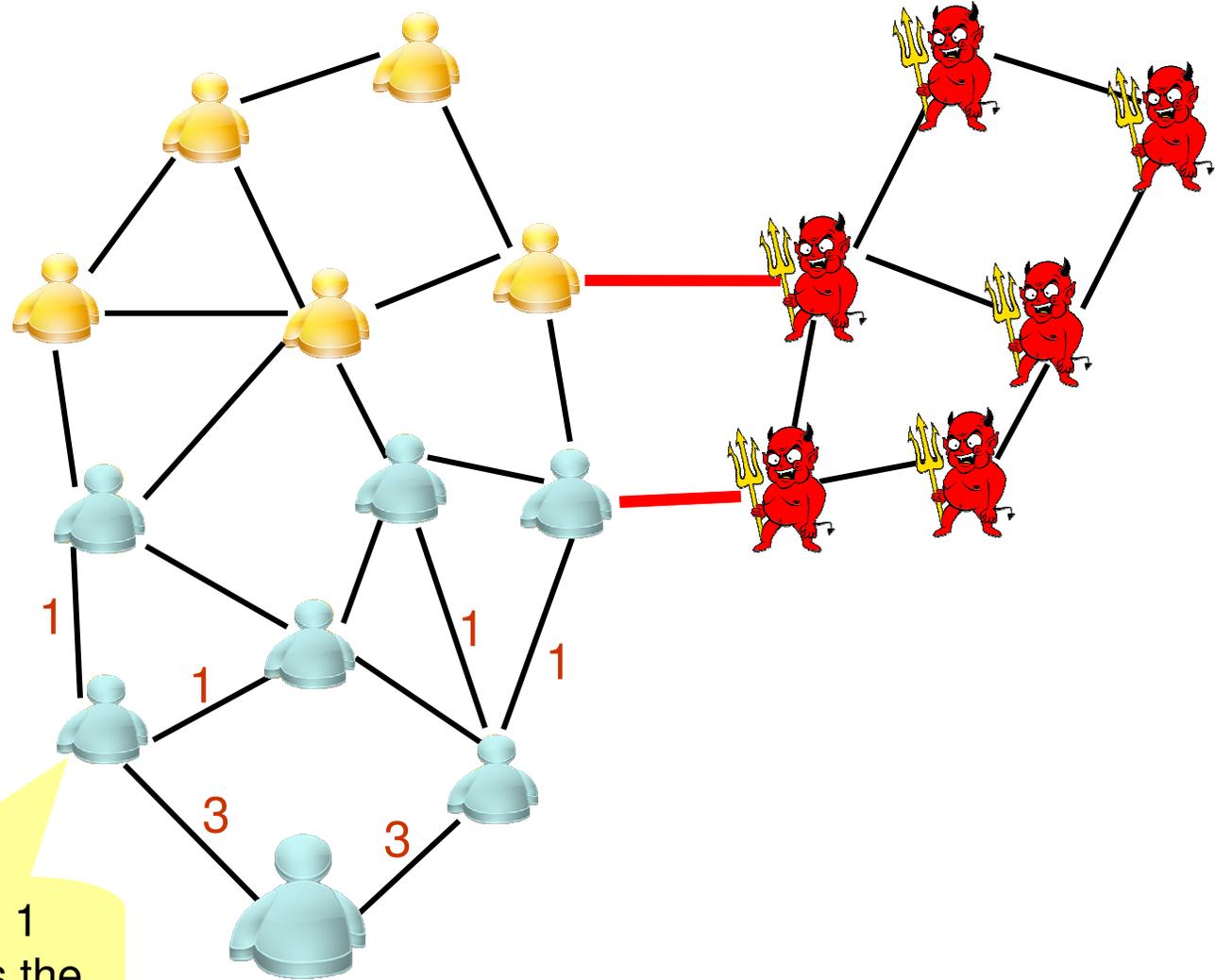
# SumUp's design details

- Capacity assignment
- Leverage user feedback to reduce bogus votes

# Capacity assignment

- Assign link capacity to collect at most  $v$  votes
  - Ideally  $v$  is the number of honest votes
- Assign greater capacity to edges that are closer to the vote collector

# Assign capacity via ticket distribution

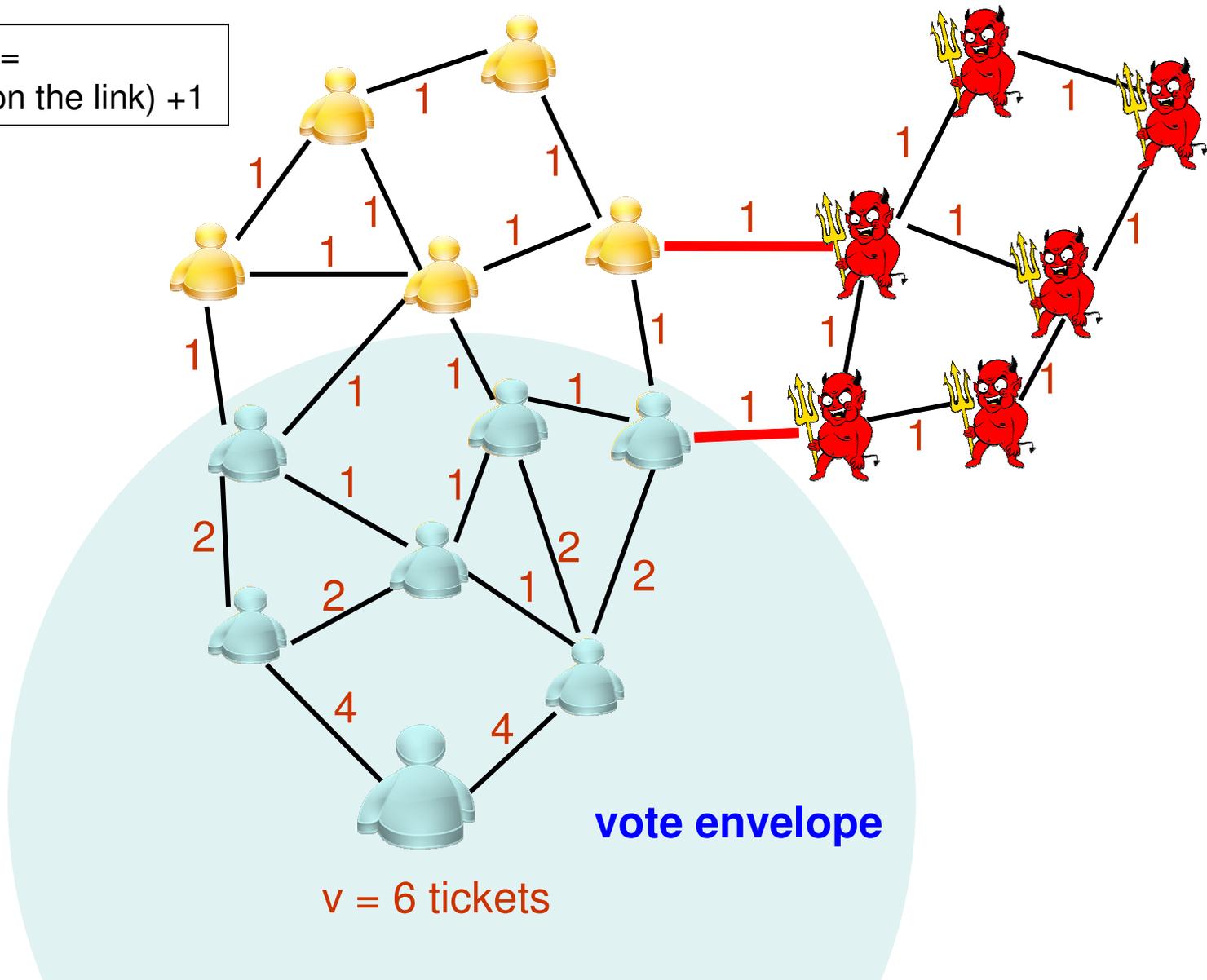


A node consumes 1 ticket and distributes the rest to out-going links

$v = 6$  tickets

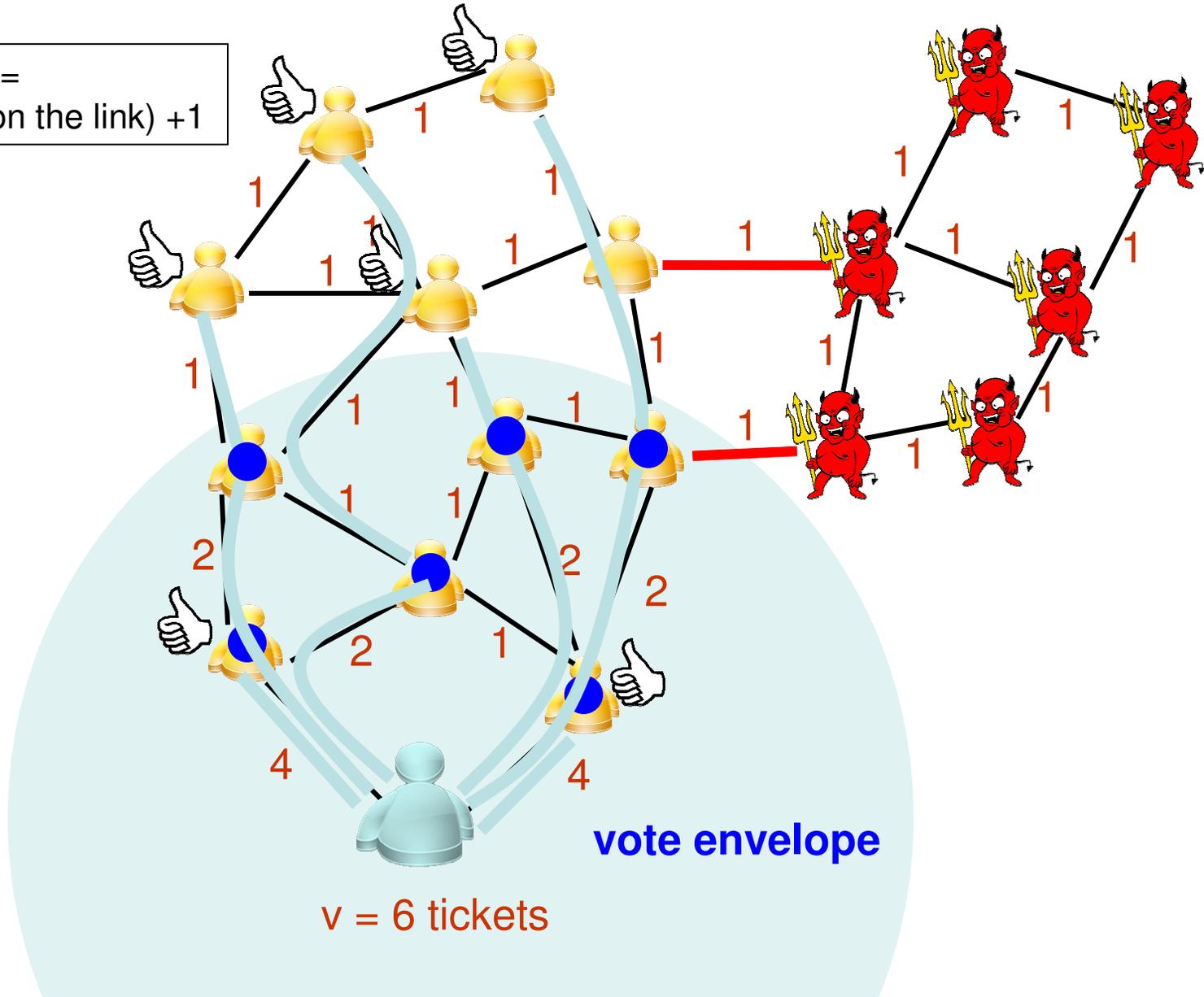
# Assign capacity via ticket distribution

Link capacity =  
(# of tickets on the link) + 1



# Assign capacity via ticket distribution

Link capacity =  
(# of tickets on the link) + 1



# Approximate $v$ to # of honest votes

- Observation

- When number of honest votes  $\gg v$ , number of collected votes  $\approx v$
- When number of honest votes  $\ll v$ , number of collected votes  $\ll v$ 
  - Not many bogus votes are collected

# Setting $v$

- Iteratively adjust  $v$ :

1. Start with a small  $v = 100$
2. Collect votes using current  $v$
3. If # of collected votes  $> 0.5 * v$ , double  $v$  and repeat step 2

- Final  $v$  approximates number of honest votes regardless of attacks

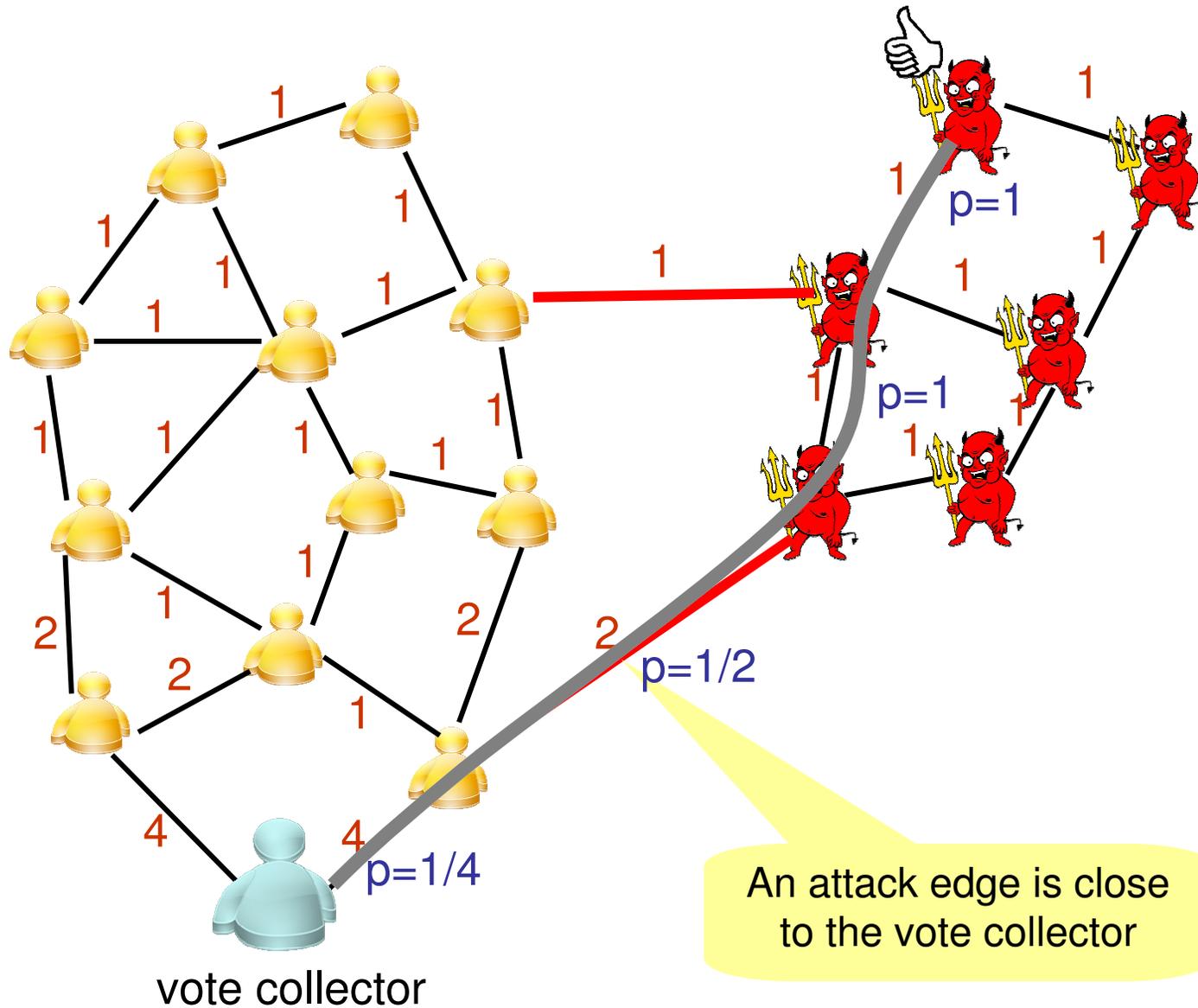
# SumUp's provable properties

- Limit bogus votes
  - When  $v \ll n$ , expected bogus votes per attack edge is  $1 + o(1)$
  - Even when  $v = \Theta(n)$ , expected bogus votes per attack edge is  $O(\log n)$  [SybilLimit IEEE S&P'08]
- Collect a large fraction of honest votes
  - On a random graph, ~100% honest votes can be collected

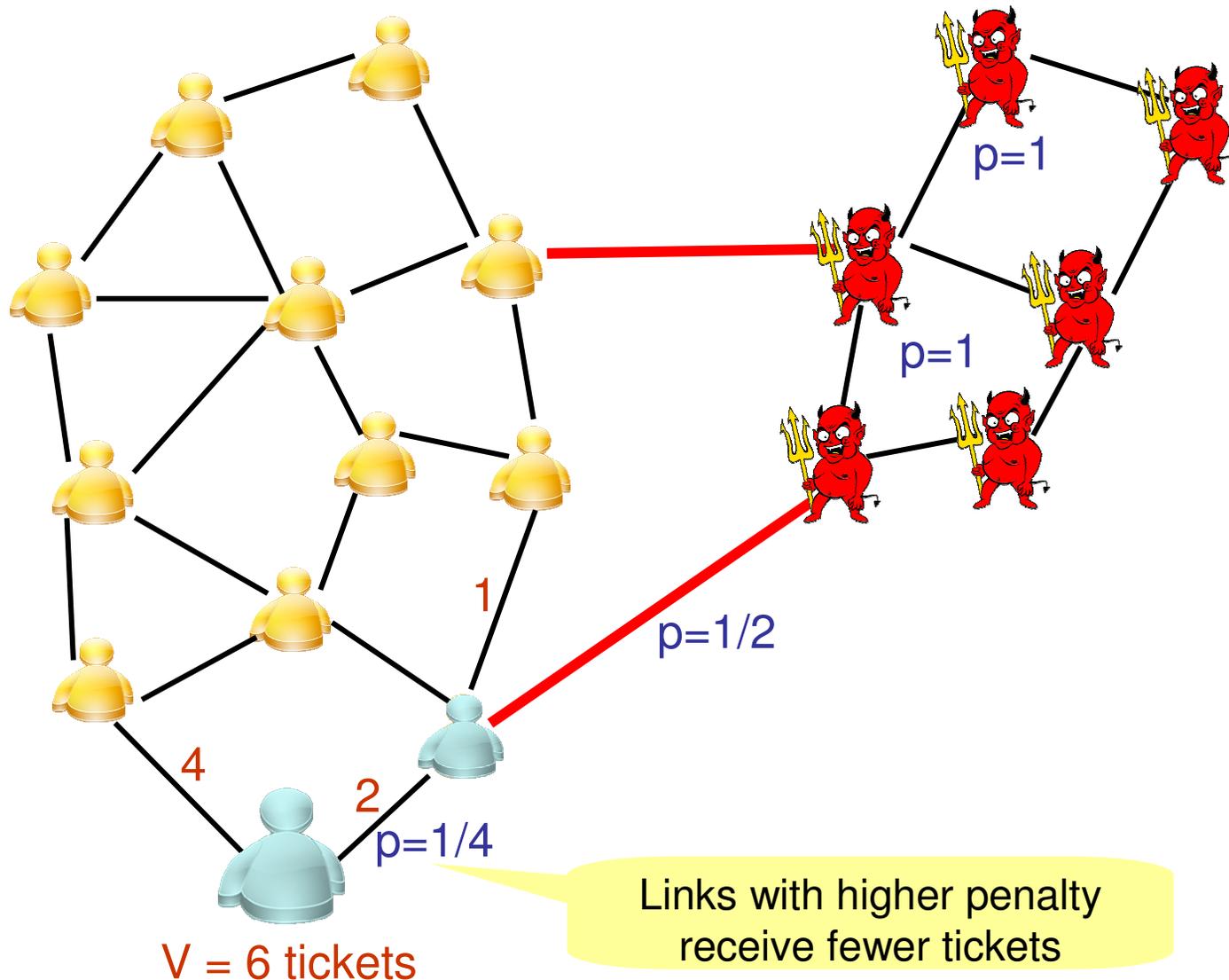
# Leverage user feedback on votes

- If vote collector can tag some votes as bogus, SumUp can do better:
  - Reduce capacity on attack edges close to the collector
  - Possibly ignore attack edges
- Idea: penalize all links along the path taken by the bogus vote [Ostra NSDI'08]

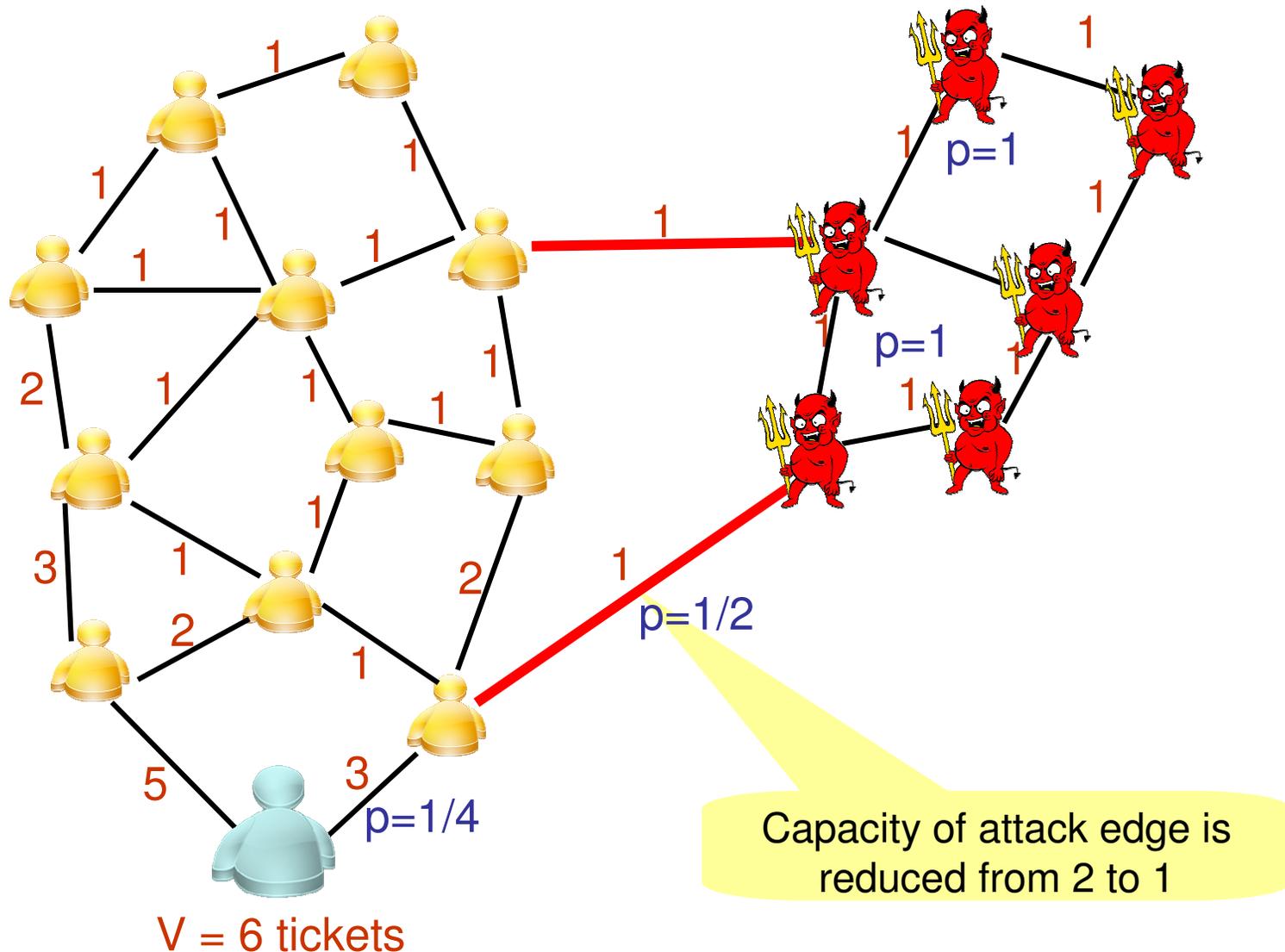
# Associate a penalty with each link



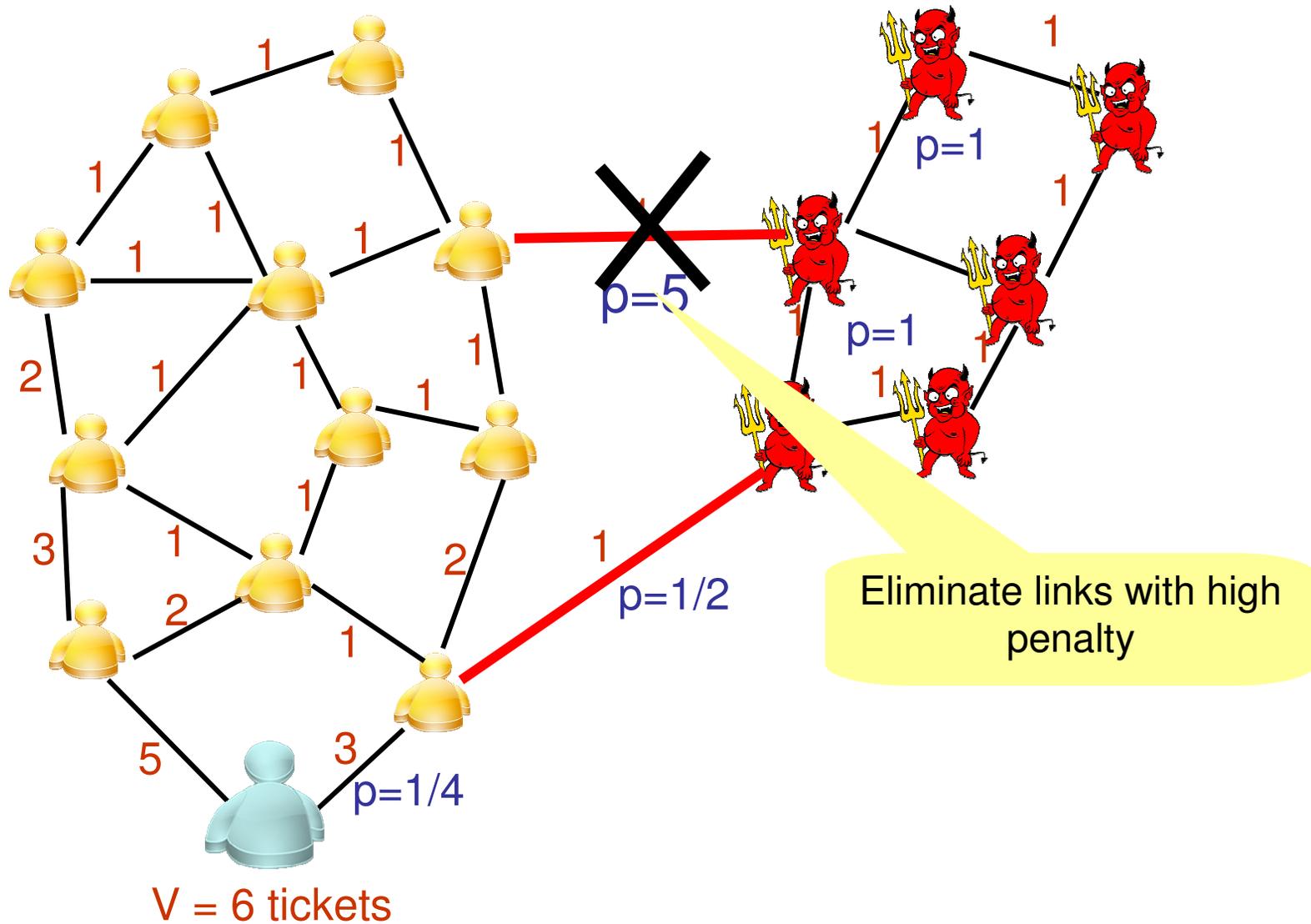
# Reassign capacity according to penalty



# Reassign capacity according to penalty



# Eliminate links with high penalty



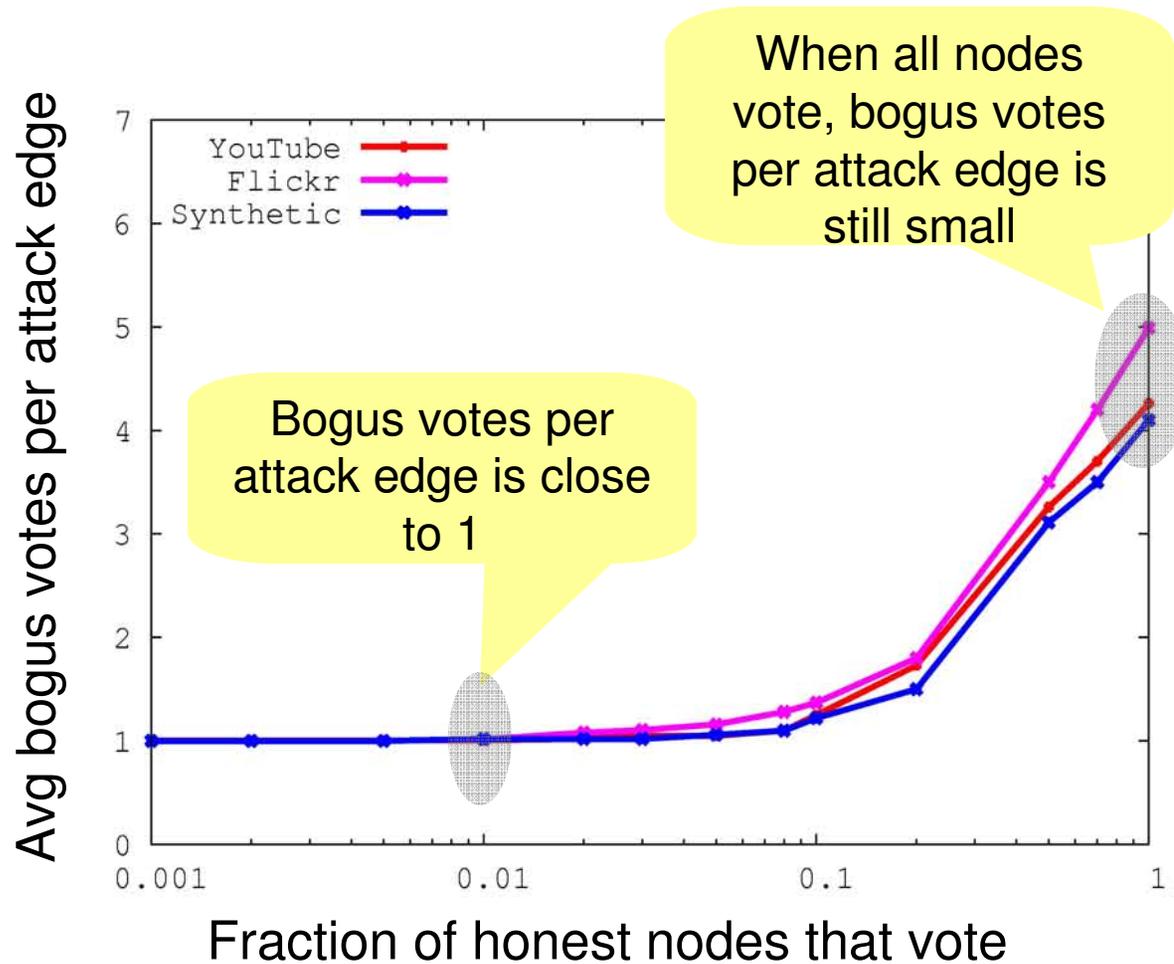
# Evaluation

1. How does SumUp perform on real social networks?
2. Can SumUp detect Sybil attacks?

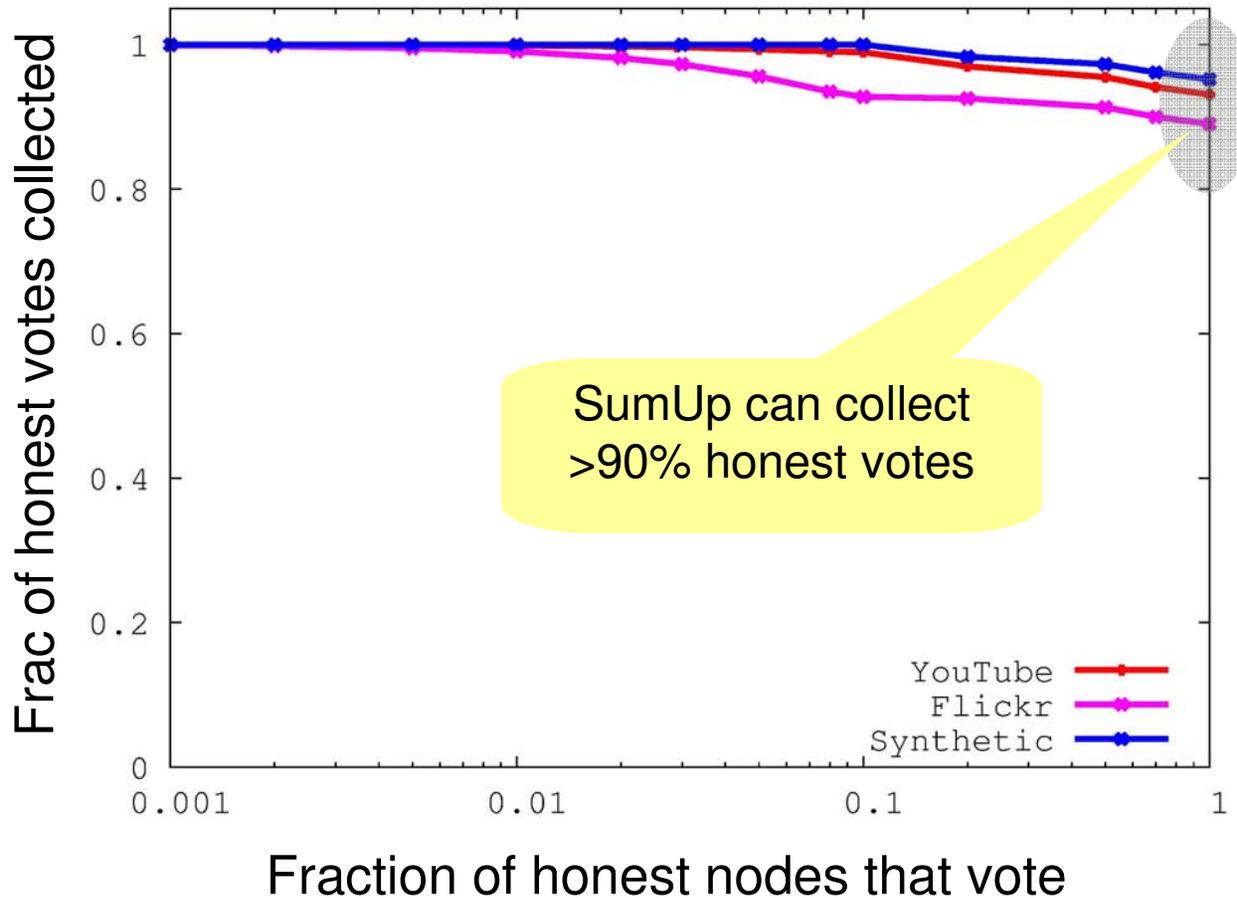
# Simulation setup

- Use 3 social networks
  - YouTube (0.5 million nodes)
  - Flickr (1.5 million nodes)
  - Synthetic (3 million nodes)
- Inject 100 attack edges randomly
- Use a random vote collector
- Choose a random set of honest voters

# SumUp limits number of bogus votes



# SumUp collects most honest votes



# Evaluate SumUp on Digg

## Digg marks “popular” articles

- 130,000 popular articles among 7 million articles submitted

## Digg’s social network

- 3 million nodes
- 0.5 million nodes in a connected component
- 80% of votes are from the connected component

**Users cast positive  
or negative votes**

The image shows a screenshot of the Digg website. The main content area displays a list of articles with their digg counts and submission times. The first article is 'World's Strongest Pope' with 75 diggs, submitted 2 minutes ago. The second is 'Space-Age Solar-Powered Catamaran to Circle the Globe' with 83 diggs, submitted 12 minutes ago. The third is 'Napolitano 'Clarifies' But the Dye is Cast' with 113 diggs, submitted 12 minutes ago. The interface includes navigation links like 'Submit News', 'Logout', and 'Search Digg'. A 'Friends' section is visible on the right, showing a list of users with their profiles and relationship types (e.g., Mutual Friend, Fan). Red arrows point from the text boxes to the 'digg' buttons on the articles and the 'Friends' section.

# Evaluating SumUp on Digg

- Kevin Rose (Digg founder) → vote collector
- Run SumUp for all votes cast before the article is marked as “popular”
- Normal articles → fraction of votes collected  $> 0.7$
- Suspicious articles → fraction of votes collected  $\ll 0.7$

## Suspicious articles have evidence of attack

- ~800 suspicious articles have less than 50% votes collected by SumUp
- Manually sampled 30 articles
- Found (subjective) evidence of attacks in 15 the articles
  - 1 article is an advertisement
  - 10 articles have many newly registered voters
  - 4 articles receive < 50 votes after marked “popular”

# Examples of suspicious articles

22  
diggs

## [The Best Dam Water Money can Buy](#)

failedsuccess.com — A new product is sweeping the nation, breathing life into the industry.

 digg

14  
diggs

## [REALLY Cheap Web Host](#)

mher.org — Plans starting at \$0.99; 8000 MBs of space and 150 GBs of bandwidth. Very fast and reliable. They never oversell. You can see examples of how fast.

 digg

18  
diggs

## [Access Your PC Anywhere and Anytime for FREE](#)

webex.com — With MyWebEx PC, you can access your files, programs and email from anywhere to your PC at the home or office with ease, and it's 100% Free and No Credit Card Required.

 digg

20  
diggs

## [Free Ipod Shuffle?](#)

FreeIpodShuffle.com — MAC has just introduced a new Ipod to the bunch. The new Ipod Shuffle is going to be a HOT item this year!

 digg

 Share  Bury  Who dugg this? Made popular **Feb 2, 2005**

# An example of suspicious articles with no evidence of attack

206  
diggs

## [New iBook, Mac mini part numbers appear](#)

thinksecret.com — Apple's forthcoming iBook update will also usher in revised Mac mini models, according to databases.



Share



Bury



2 Comments

Made popular **Jul 22, 2005**

### Who Dugg This?



[redawgts](#)



[ShyGuy29201](#)



[mcpaige](#)



[HadjiQuest](#)



[filmer](#)



[NDFord005](#)



[cheesegod](#)



[DiZASTiX](#)



[incu\\_vamp](#)



[producer](#)



[inactive](#)



[brando\\_ub](#)



[npacheco](#)



[garbanzo](#)



[teddysetgo](#)



[produce](#)



[filmpro](#)



[kkeeter](#)



[HANABI](#)



[hardcoredj](#)



[WorfoSAUR](#)



[filmaker\\_pro](#)

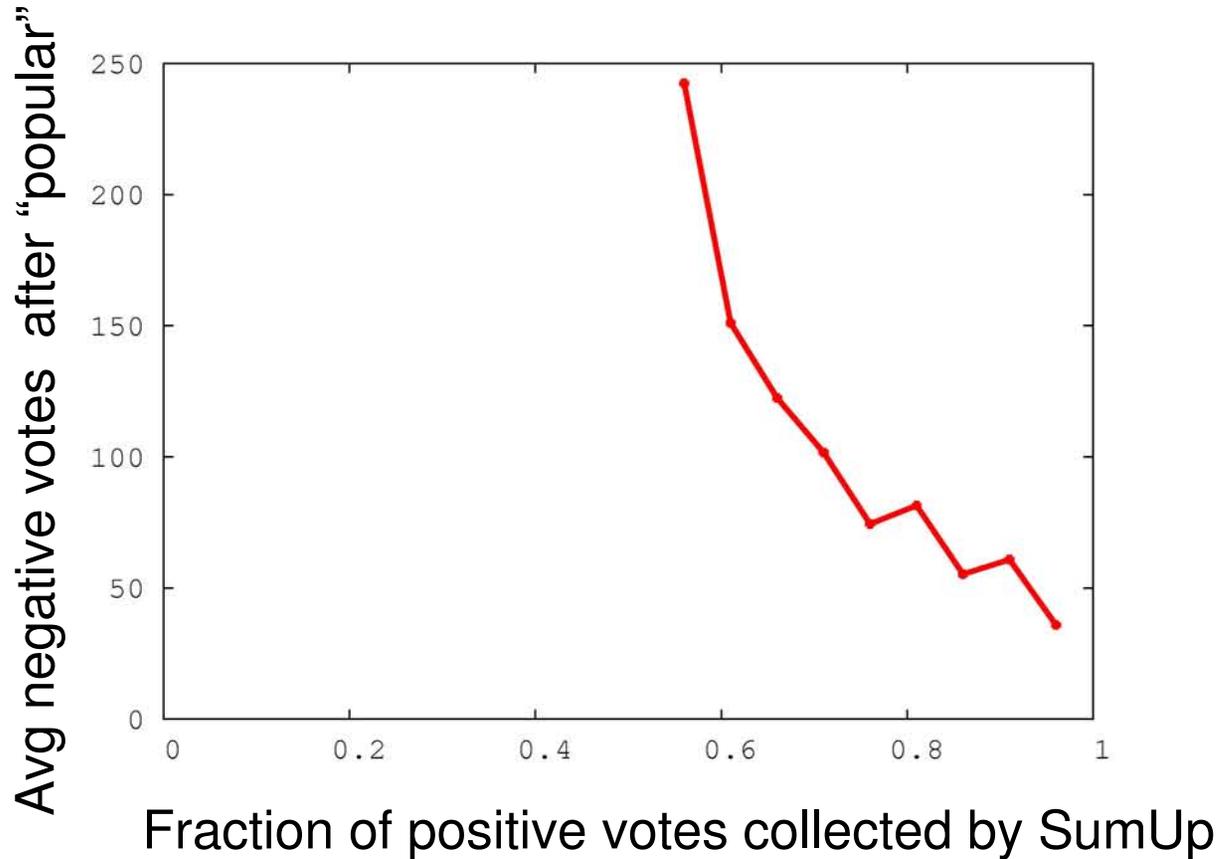


[merlin](#)



[peterapokotos](#)

# Suspicious articles receive more negative votes



- Obtained negative votes for 5794 “popular” articles from 08/2008 to 09/2008

# Related work

- Node admission
  - SybilGuard [Sigcomm'06], SybilLimit [IEEE S&P'08], SybilInfer [NDSS'09]
- Fighting spam
  - Ostra [NSDI'08]
- User reputation systems
  - SybilProof [P2PEcon'05], Appleseed [ISF'05], Advogato [SSYM'98]
- Content voting systems
  - Credence [NSDI'06]

# Conclusion

- Defending against Sybil attacks is important for content voting systems
- SumUp vote aggregation:
  1. Limit # of bogus votes by # of attack edges
  2. Collect many votes from honest users
  3. Ignore votes from attackers that repetitively cast bogus votes