# DSybil: Optimal Sybil-Resistance for Recommendation Systems

**Presenter: Ennan Zhai**

* Announcement: thanks for the authors' helps, e.g., useful comments and referencing their talk slides.

# Roadmap

I Background and existing efforts ?

II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# Roadmap

I Background and existing efforts ?

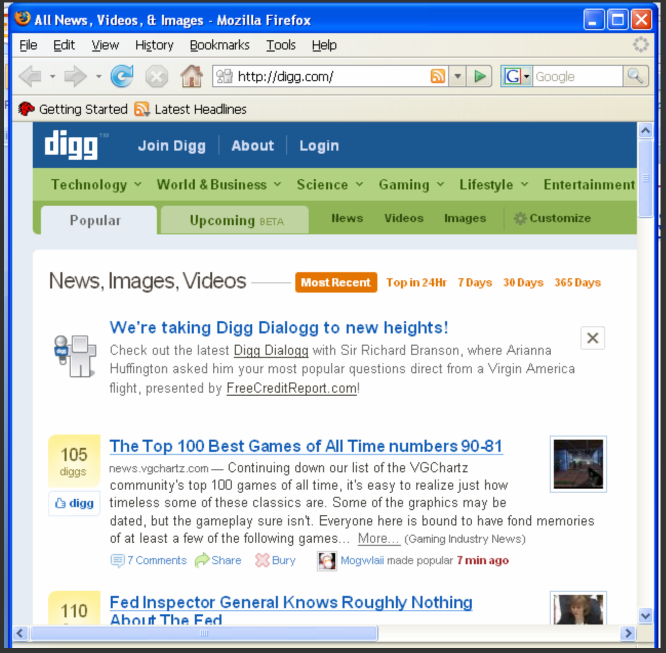
II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# Background: A Story … …

There are many popular recommendation systems in our world such as Digg, Amazon, Razor, Netflix, YouTube, Credence … …



News, Images, Videos

We're taking Digg Dialogg to new heights!

Check out the latest Digg Dialogg with Sir Richard Branson, where Arianna Huffington asked him your most popular questions direct from a Virgin America flight, presented by FreeCreditReport.com!

The Top 100 Best Games of All Time numbers 90-81

news.vgchartz.com — Continuing down our list of the VGChartz community's top 100 games of all time, it's easy to realize just how timeless some of these classics are. Some of the graphics may be dated, but the gameplay sure isn't. Everyone here is bound to have fond memories of at least a few of the following games… More… (Gaming Industry News)

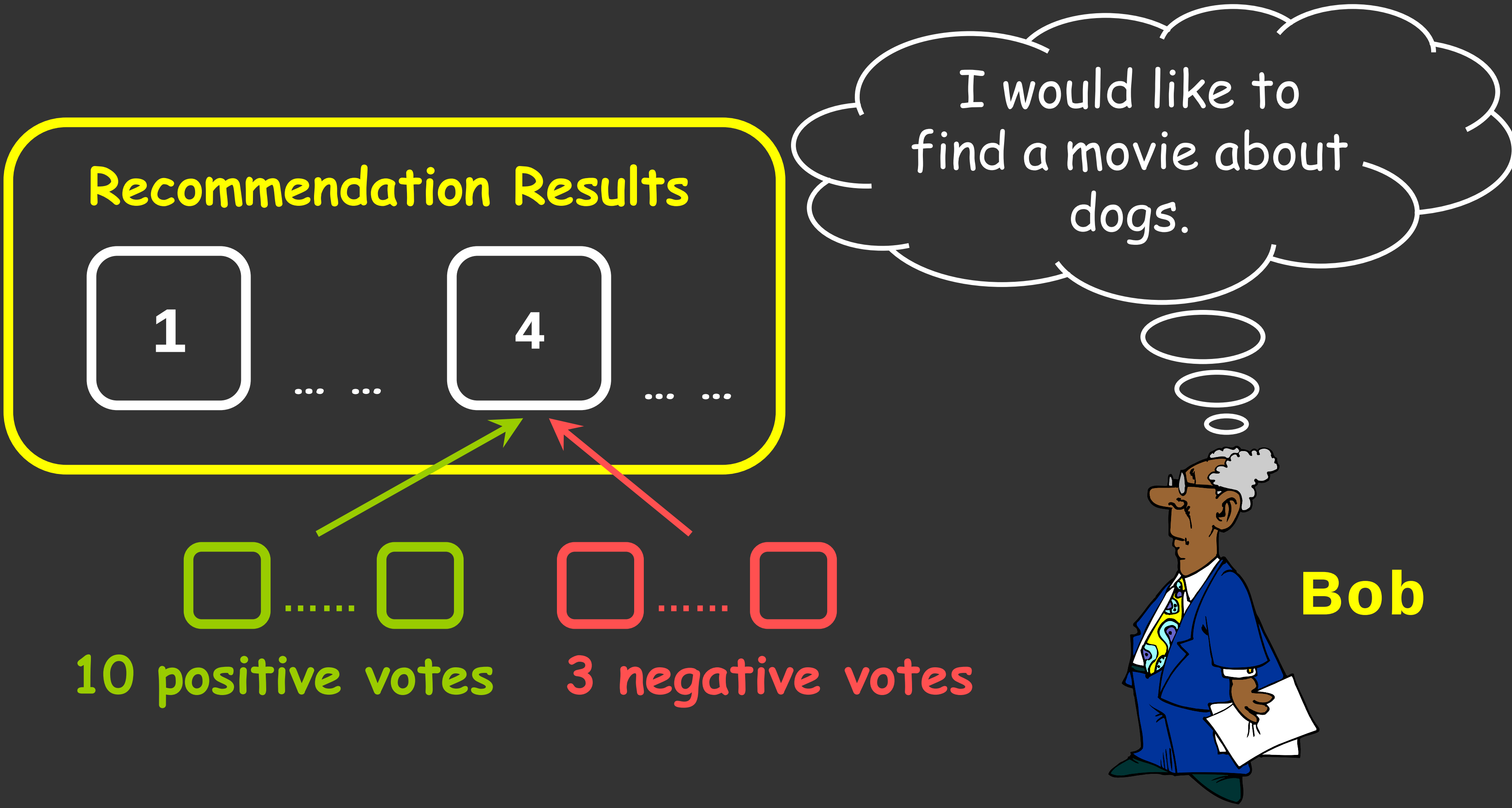
Mogwlaii made popular 7 min ago

Fed Inspector General Knows Roughly Nothing About The Fed

# Background: A Story … …

In these recommendation systems, attackers are able to cast misleading votes … …

# Background: A Story … …



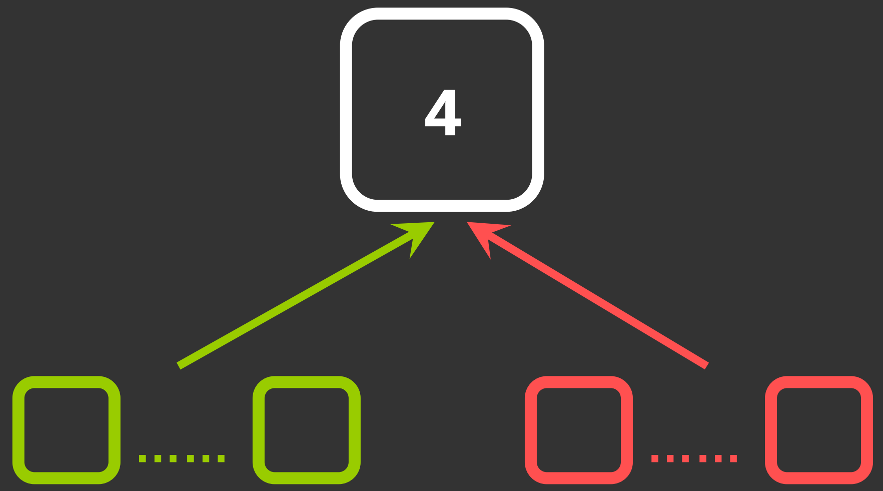
Recommendation Results

1 … … 4 … …

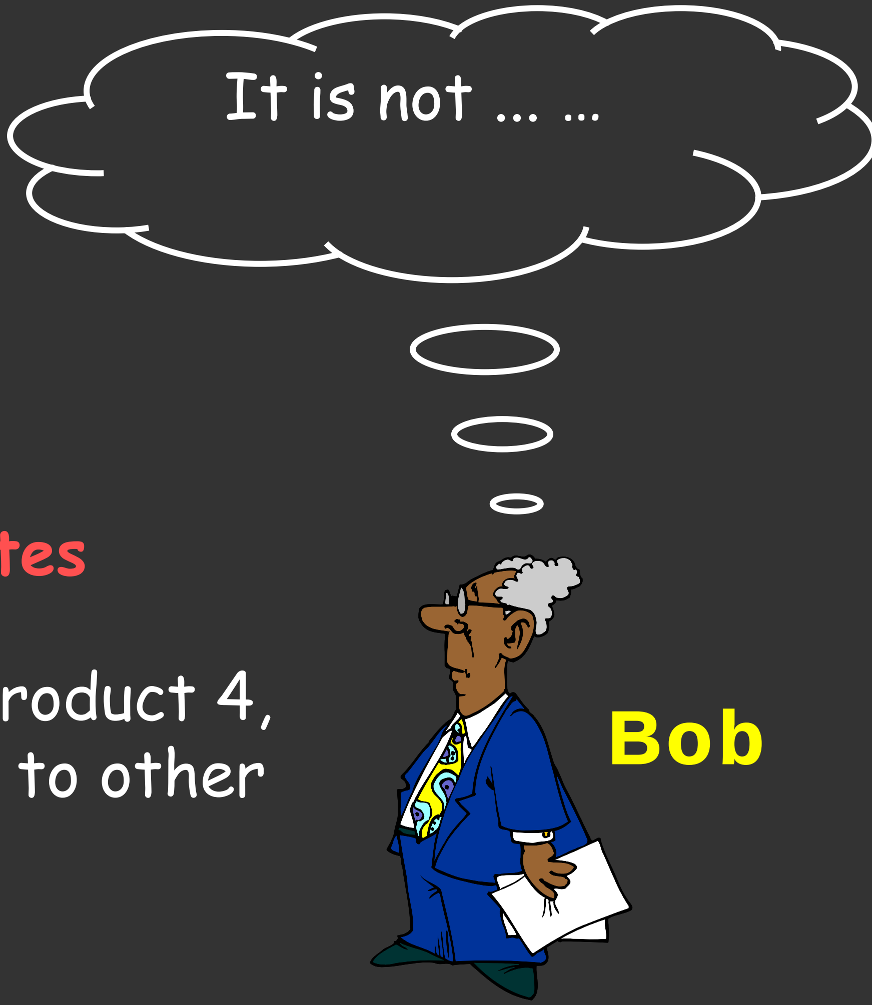
10 positive votes

3 negative votes

# Background: A Story … …



10 positive votes 3 negative votes

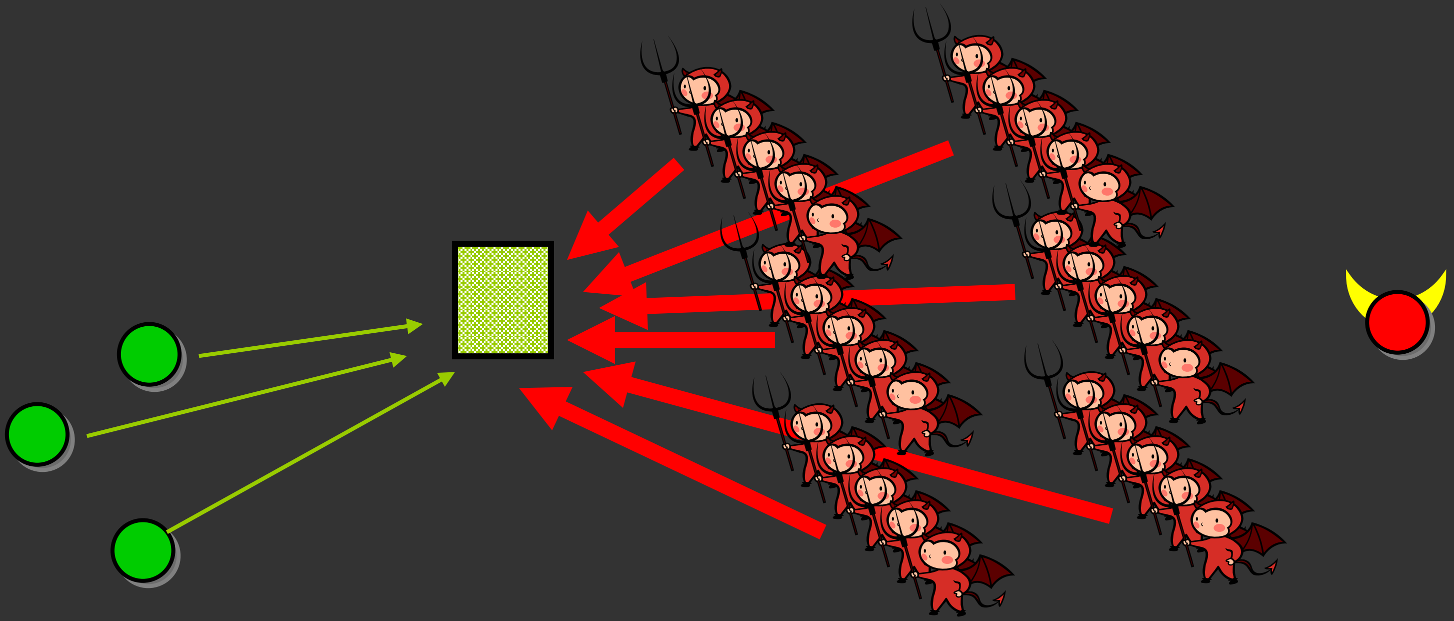


However, after consumption of product 4, Bob finds this product is related to other things, e.g., soccer.

We say the votes are misleading.

# Background: A Story … …

The ultimate form of the misleading attack is to launch SYBIL ATTACK … …



# Background: A Story … …

How to defend against sybil attacks?

Sybil defense is considered challenging. There are many many papers that aim to defend against sybil attack, but most without having a good solution … …

# Background: A Story … …

How to defend against sybil attacks?

Social-network-based defense:

- SybilGuard [SIGCOMM'06]
- SybilLimit [Oakland'08]
- SybilInfer [NDSS'09]
- SumUp [NSDI'09]
- Whanau [NSDI'11]
- Ostra [NSDI'08]
- Gatekeeper [PODC'10, NetEcon'10] … …

# Background: A Story … …

However, recommendation systems are more vulnerable … …

- Byzantine Tolerance: # = n/3;
- DHT: # = n/4;
- Rec Systems: # = n/500.

# sybil identities we can tolerate (n identities total)

# Background: A Story … …

Therefore, social-network-based defenses are not sufficiently strong for Rec systems … …

E.g., to create n/500 sybil identities: compromise only 1 node out of every 5000 honest nodes is sufficient.

**How to defend against sybil attacks in recommendation systems?**

# Background: A Story … …

There is an ancient idea: adjust “trust” to an identity based on its historical behavior … …

# Background: A Story … …

Could trust sufficiently diminish the influence of sybil identities in recommendation systems ?

Aim for provable guarantees under all attack strategies (including worst-case attack from intelligent attackers)

# Background: A Story … …

DSybil answers the question:

- Based on feedback and trust
- Loss (# of bad recommendations) is provable

  $O(D \log M)$ even under worst-case attack

  D: Dimension of the objects (less than 10 in Digg)

  M: Max # of sybil identities voting on each obj
- The authors prove DSybil's loss is optimal

# Roadmap

I Background and existing efforts ?

II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# Roadmap

I Background and existing efforts ?

II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# The Challenges in Design

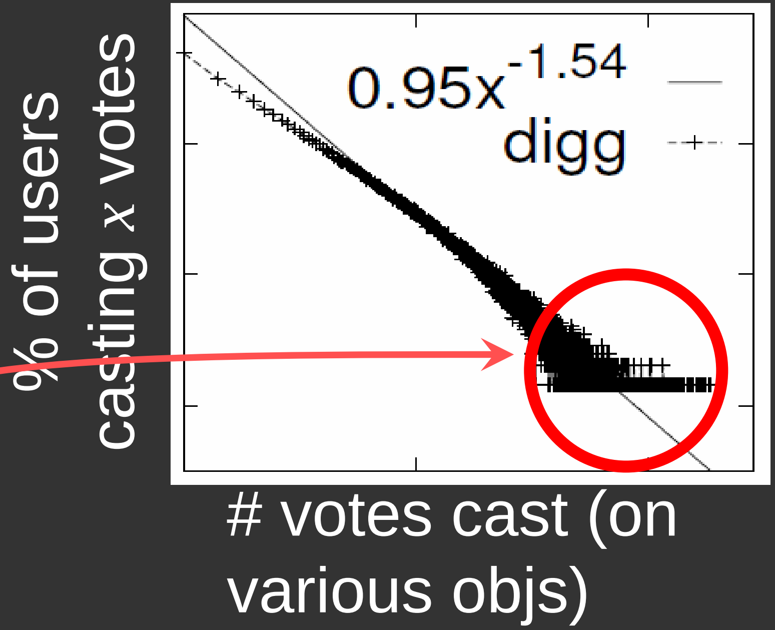
There are some subtle aspects of using trust:

1. How to identify “correct” but “non-helpful” votes?
2. How to assign initial trust to new identities?
3. How exactly to grow trust?
4. How exactly to make recommendations?

# Two Key Insights

Key #1: Leveraging typical voting behavior of honest users:

- Heavy-tail distribution
- Exist very active users who cast many votes



# Two Key Insights

Key #2: If user is already getting “enough help”, then do not give out more trust:

This insight enables us to avoid giving trust to some sybil identities;

The insight can make us strike an optimal balance.

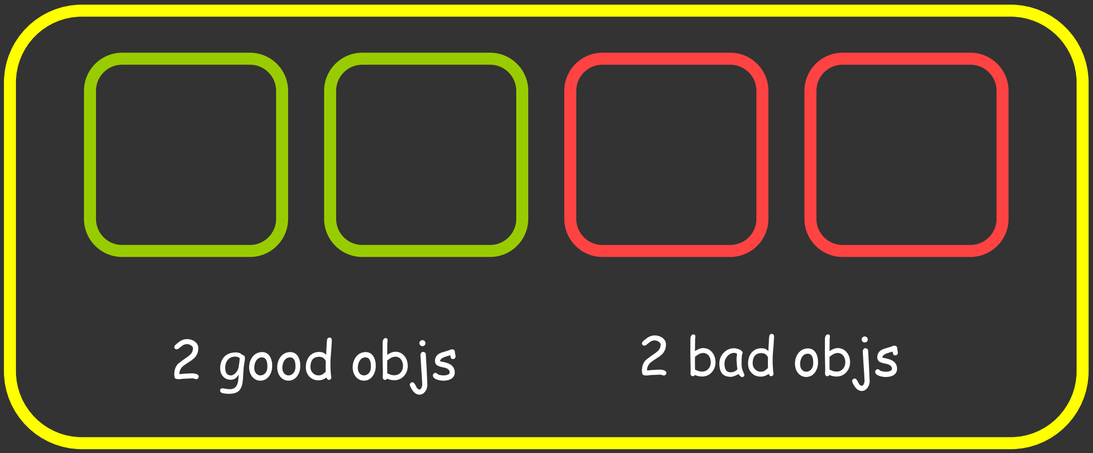
# System Model

- Objects (or products) to be recommended are either good or bad (e.g., Digg);

- Votes are positive. Namely, DSybil only has positive votes.

- DSybil is personalized:
  - ✓ Each user may have different subjective opinions;
  - ✓ Different users may get different recommendations;
  - ✓ Run by either Alice or a central server (simple).

# System Model

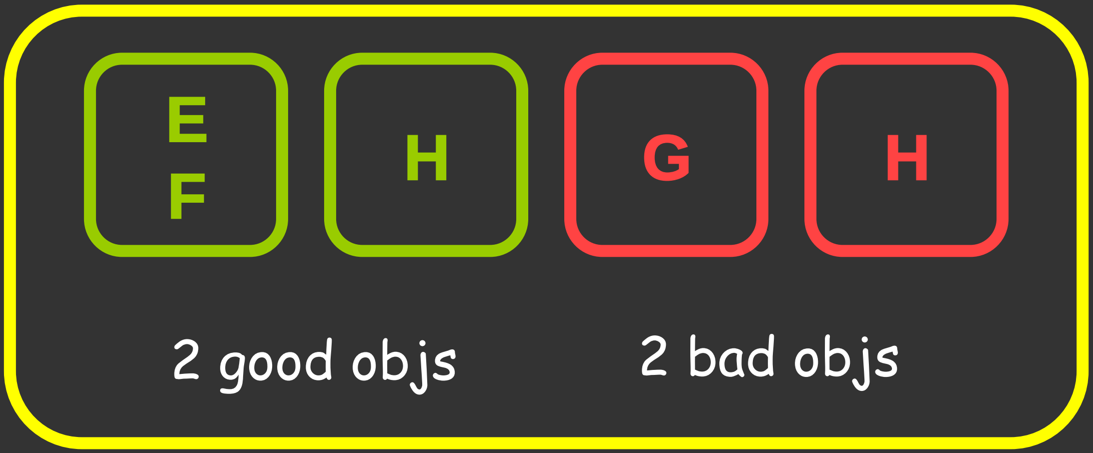


DSybil does not know which are good … …

Each round has a pool of objects:

- DSybil recommends one object for Alice to consume;
- Alice provides feedbacks after consumption;
- DSybil adjusts trust based on the feedbacks.

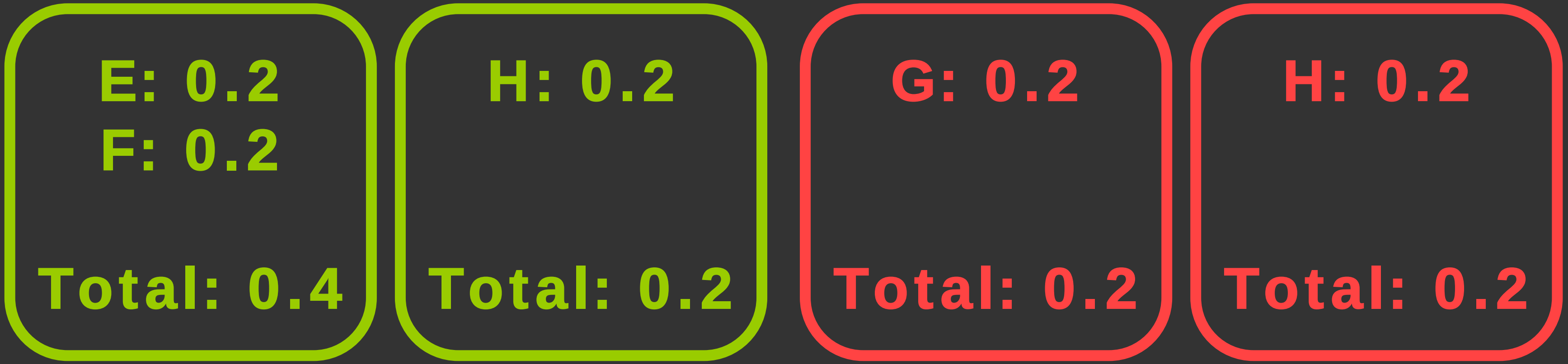
# System Model



Each identity is able to cast at most one vote/object.

At most *M* (e.g., 10^10) sybil identities voting on each product.

# Initial Round: Classifying Objs



E: 0.2
F: 0.2
Total: 0.4

H: 0.2
Total: 0.2

G: 0.2
Total: 0.2

H: 0.2
Total: 0.2

Each identity starts with initial trust 0.2 … …

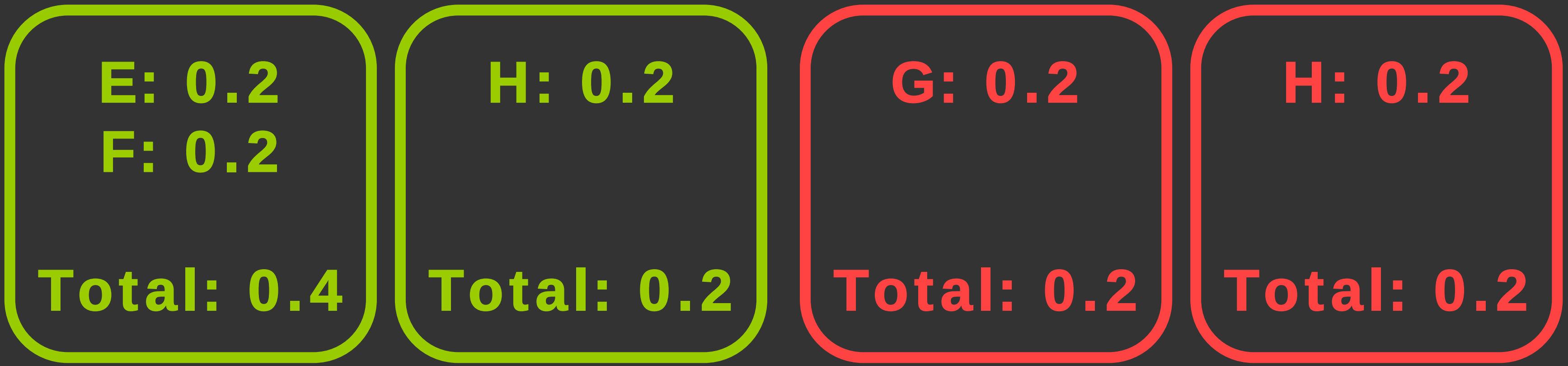
An object is overwhelming if total trust >=*C* (*C* = 1.0)

# Rounds Without Overwhelming Objs

| | | | |
|---|---|---|---|
| E: 0.2<br>F: 0.2 | H: 0.2 | G: 0.2 | H: 0.2 |
| Total: 0.4 | Total: 0.2 | Total: 0.2 | Total: 0.2 |

1. Recommend uniformly random object

# Rounds Without Overwhelming Objs



E: 0.2
F: 0.2
Total: 0.4

H: 0.2
Total: 0.2

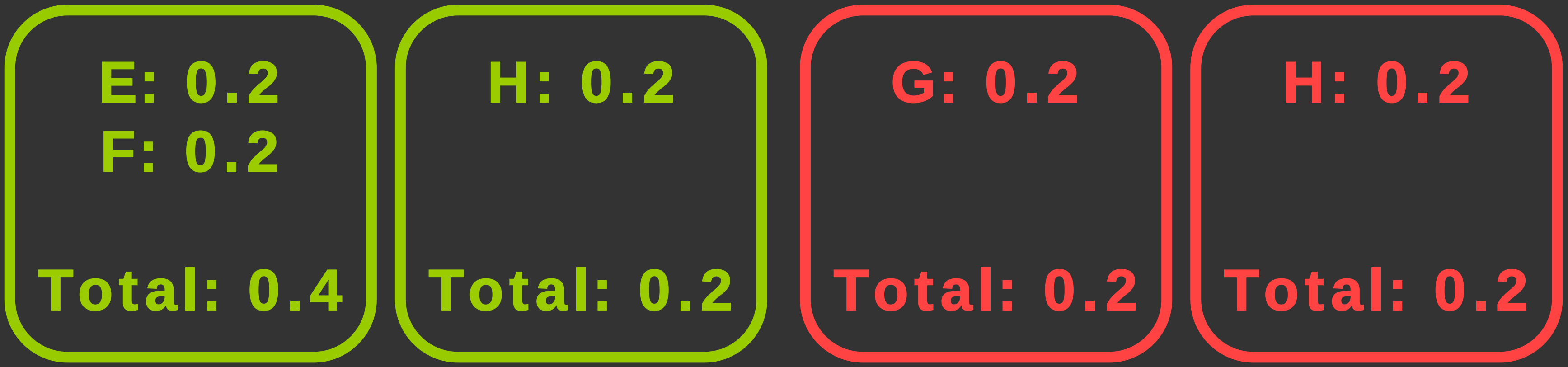
G: 0.2
Total: 0.2

H: 0.2
Total: 0.2

1. Recommend uniformly random object

Notice that recommending obj with the largest total trust would result in linear loss … …

# Rounds Without Overwhelming Objs



E: 0.2
F: 0.2
Total: 0.4

H: 0.2
Total: 0.2

G: 0.2
Total: 0.2

H: 0.2
Total: 0.2

1. Recommend uniformly random object
2. Adjust trust after feedback
   - if obj is bad, multiply trust of voters by $\beta = 0.5$
   - if obj is good, multiply trust of voters by $\alpha = 2$

# Rounds With Overwhelming Objs

E: 1.0
H: 0.2

Total: 1.2

G: 0.2
H: 0.2

Total: 0.4

F: 1.0

Total: 1.0

1. Recommend arbitrary overwhelming object
   - Will confiscate sufficient trust if obj is bad … …
2. Adjust trust after feedback
   - if obj is bad, multiply trust of voters by $\beta = 0.5$
   - if obj is good, no additional trust given out (#2 key)

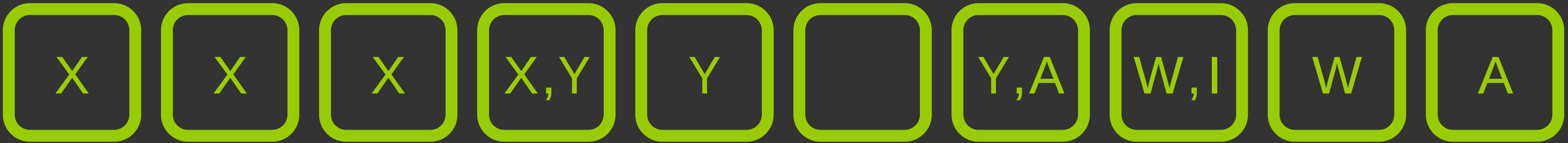
# Definition of Guides and Dimension

Guides: Honest users with same/similar "opinion" with Alice. Namely the guy never/seldom votes for bad objects … …

Dimension: the minimal # of guides needed to "cover" large fraction (e.g., 60%) of the good objs --- called critical guides

# Definition of Guides and Dimension

Dimension: the minimal # of guides needed to "cover" large fraction (e.g., 60%) of the good objs --- called critical guides.



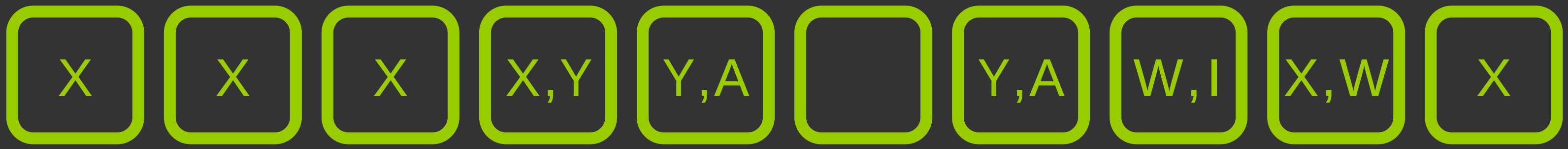
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| X | X | X | X,Y | Y | | Y,A | W,I | W | A |

Dimension = 2; Critical guides = {X,Y} or {X,W} or {X,A}

Notice that DSybil does not know who are the guides or what the dimension is

# Definition of Guides and Dimension

Dimension: the minimal # of guides needed to "cover" large fraction (e.g., 60%) of the good objs --- called critical guides.

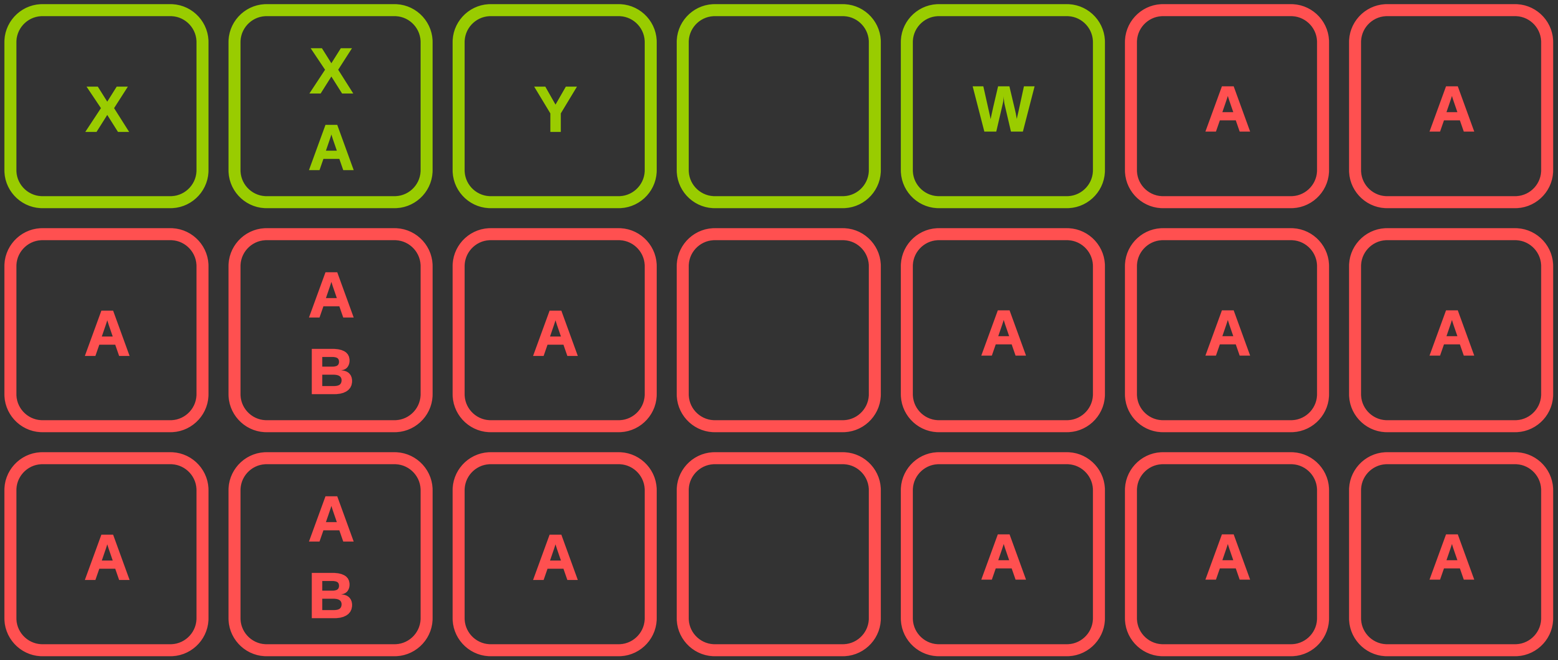


X X X X,Y Y,A Y,A W,I X,W X

Dimension = 1; Critical guide = {X}

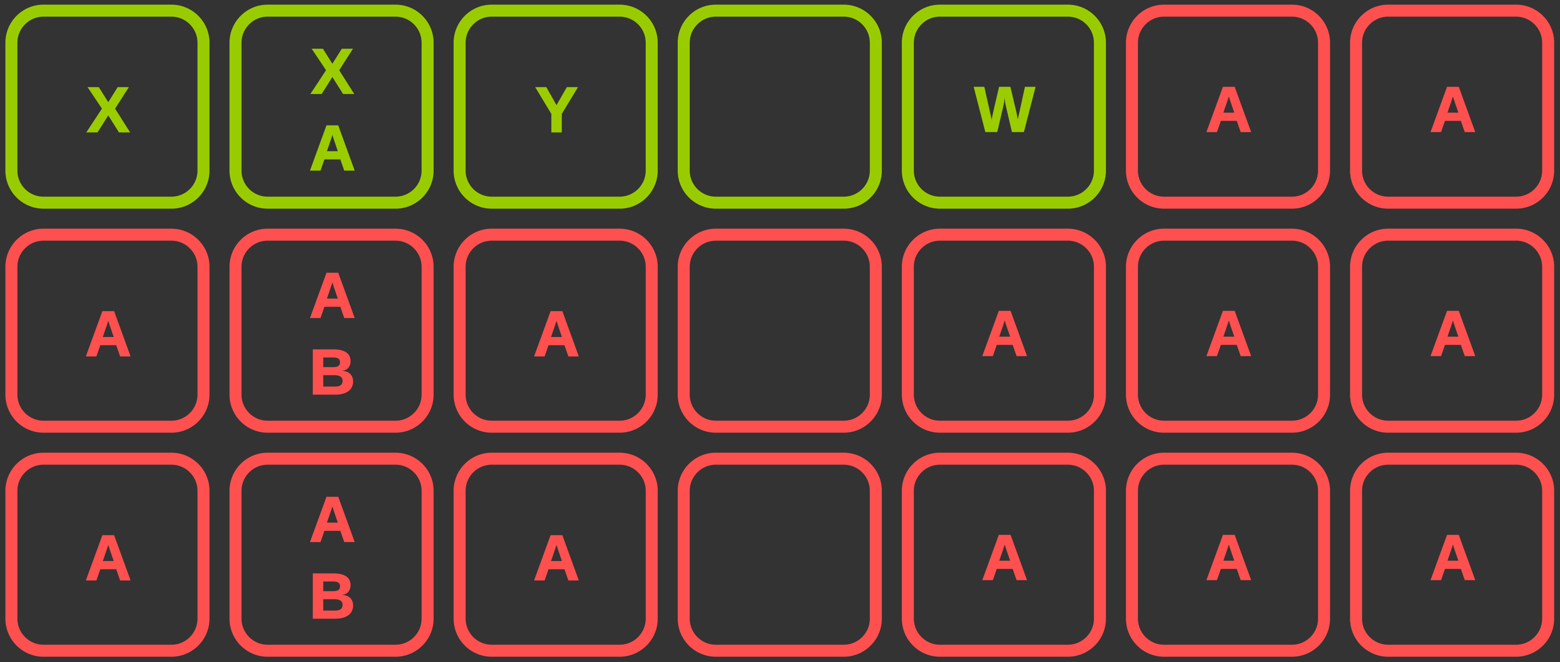
Notice that DSybil does not know who are the guides or what the dimension is

# Definition of Guides and Dimension



| | | | | | | |
|---|---|---|---|---|---|---|
| X | X<br>A | Y | | W | A | A |
| A | A<br>B | A | | A | A | A |
| A | A<br>B | A | | A | A | A |

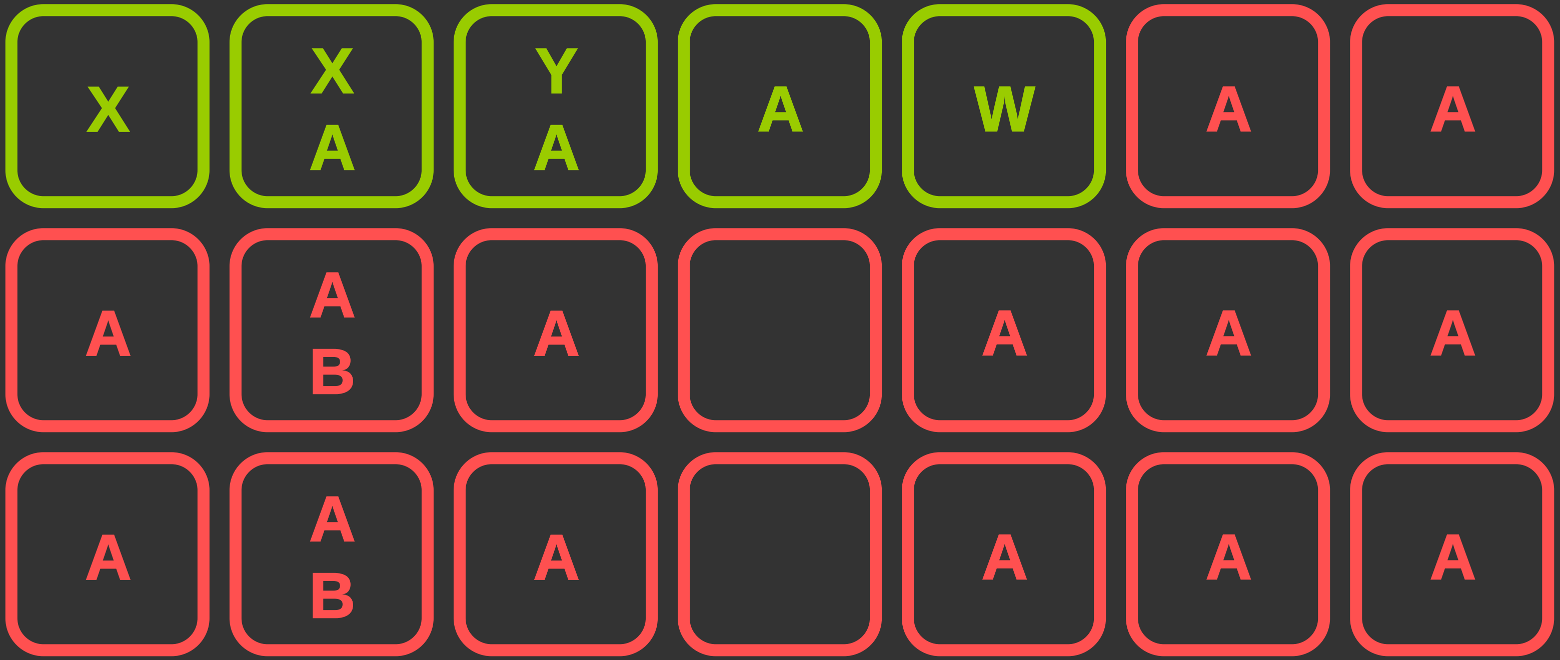
# Definition of Guides and Dimension



| | | | | | | |
|---|---|---|---|---|---|---|
| X | X<br>A | Y | | W | A | A |
| A | A<br>B | A | | A | A | A |
| A | A<br>B | A | | A | A | A |

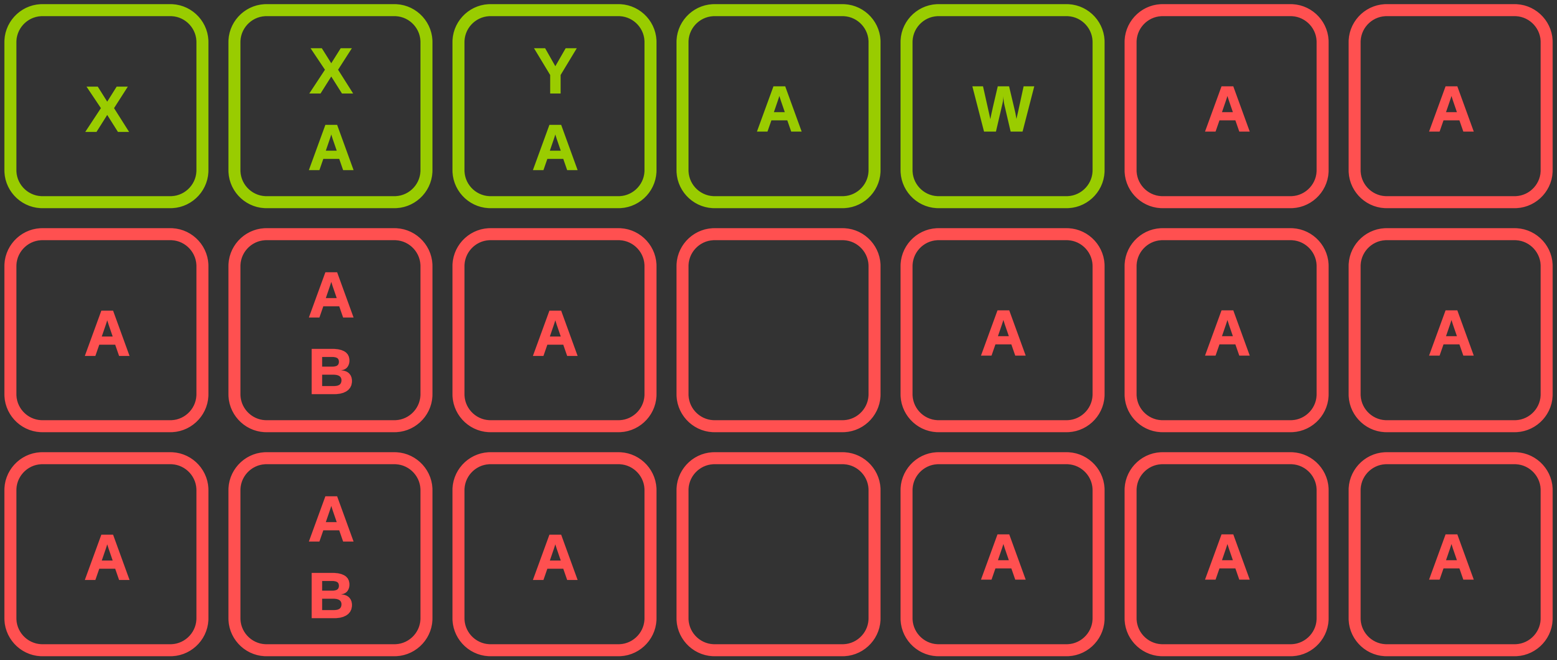
Dimension = 2; Critical guides = {X,Y} or {X,W}

# Definition of Guides and Dimension



| | | | | | | |
|---|---|---|---|---|---|---|
| X | X<br>A | Y<br>A | A | W | A | A |
| A | A<br>B | A | | A | A | A |
| A | A<br>B | A | | A | A | A |

# Definition of Guides and Dimension



| | | | | | | |
|---|---|---|---|---|---|---|
| X | X<br>A | Y<br>A | A | W | A | A |
| A | A<br>B | A | | A | A | A |
| A | A<br>B | A | | A | A | A |

Dimension = 1; Critical guide = {A}

# Key #1: Leveraging Small Dimension

Dimension is typically small in practice … …

Small dimension -> will encounter critical guides frequently when picking random objs:

* Trust to critical guides quickly grow to C
* This will result in overwhelming objs … …

# Key #2: Help is Sufficient

Consuming good overwhelming obj = Alice already has sufficient help

Thus do not give out additional trust:

- * Prevent sybil identities from getting trust "for free"
- * May hurt honest identities (but remember this is optimal … …)

# Strong Guarantee

Proof:

… … … ^^^ … ++==………

… … ~~~ … ……

,,l. ……………………

End of Proof

Proof for $O(D \log M)$ loss even under worst-case attack

# Roadmap

I Background and existing efforts ?

II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# Roadmap

I Background and existing efforts ?

II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# Results

* One-year Digg dataset with half-million users

- Pessimistically assuming guides are only 2% of the honest users (see paper for more details … )
- To cover 60% of good objs, need only 3 guides

* Robustness: Remove guides --- 5 new guides to cover 60%

* The experiments mainly prove the heavy-tail distribution of votes cast by individual users (the only assumption of DSybil) in real world.

# Roadmap

- I Background and existing efforts ?
- II The design of DSybil ?
- III Experimental results ?
- IV My evaluations on DSybil and future ideas ?

# Roadmap

I Background and existing efforts ?

II The design of DSybil ?

III Experimental results ?

IV My evaluations on DSybil and future ideas ?

# My Evaluation

1. DSybil reveals an interesting phenomenon in voting-based systems (i.e., heavy-tail distribution) … …  We can use it …

2. DSybil reveals some interesting ideas to design reputation systems or other scoring systems … …

3. After reading DSybil, I know how to prove the optimality of this kind of system.

# Future Ideas

1. The authors fail to provide guarantee to the convergency of DSybil. We can show the guarantee through introducing social networks.

2. By improving DSybil's algorithm (introducing some reasonable assumptions), actually we can obtain a lower upper bound.

3. We can develop a spam-proof tagging system through designing an approach like DSybil.

# Future Ideas

4. Recent years, the development of P2P reputation systems has been not very good. There is little paper on this topic … …

5. We are able to propose some attacks to overwhelm DSybil. Although the optimality of DSybil was proved, we can generate some attacks to overwhelm DSybil in real world. (implementing a DSybil and finish it … …)

Thank you !!!