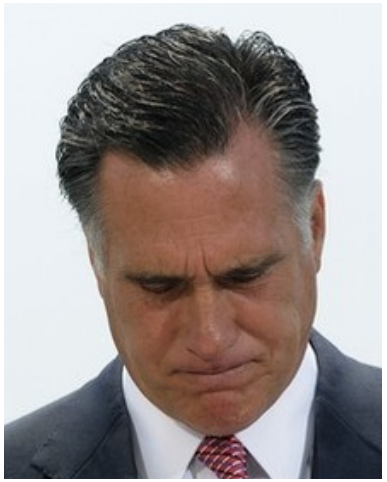
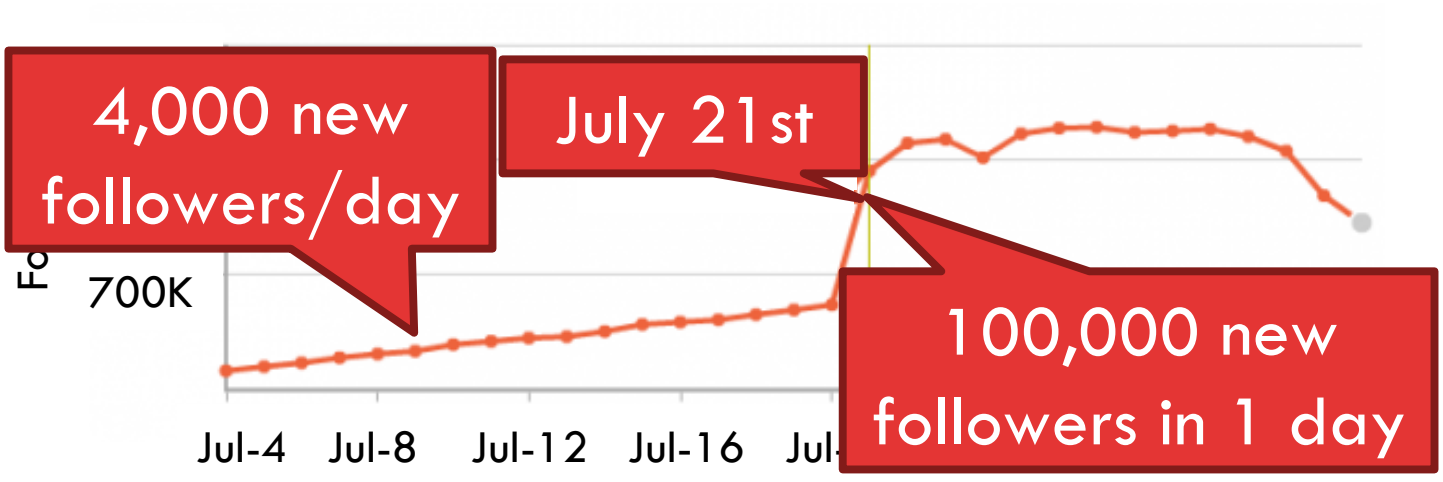
Sybil In Online Social Networks (OSNs)

- Sybil (*sɪbəl*): fake identities controlled by attackers
  - Friendship is a pre-cursor to other malicious activities
  - Does not include benign fakes (secondary accounts)

- Research has identified malicious Sybils on OSNs
  - Twitter [CCS 2010]
  - Facebook [IMC 2010]
  - Renren [IMC 2011], Tuenti [NSDI 2012]

Real-world Impact of Sybil (Twitter)





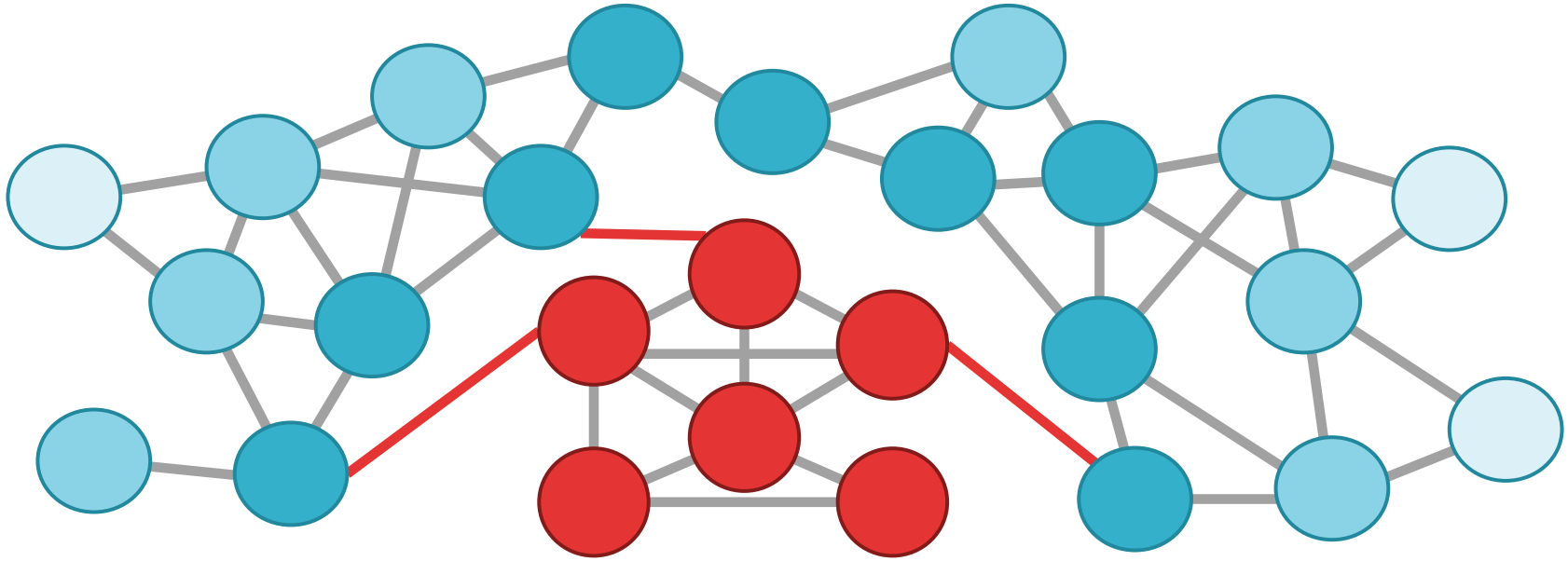
- Russian political protests on Twitter (2011)
  - 25,000 Sybils sent 440,000 tweets
  - Drown out the genuine tweets from protesters

Security Threats of Sybil (Facebook)

- Large Sybil population on Facebook
  - August 2012: 83 million (8.7%)
- Sybils are used to:
  - Share or Send Spam

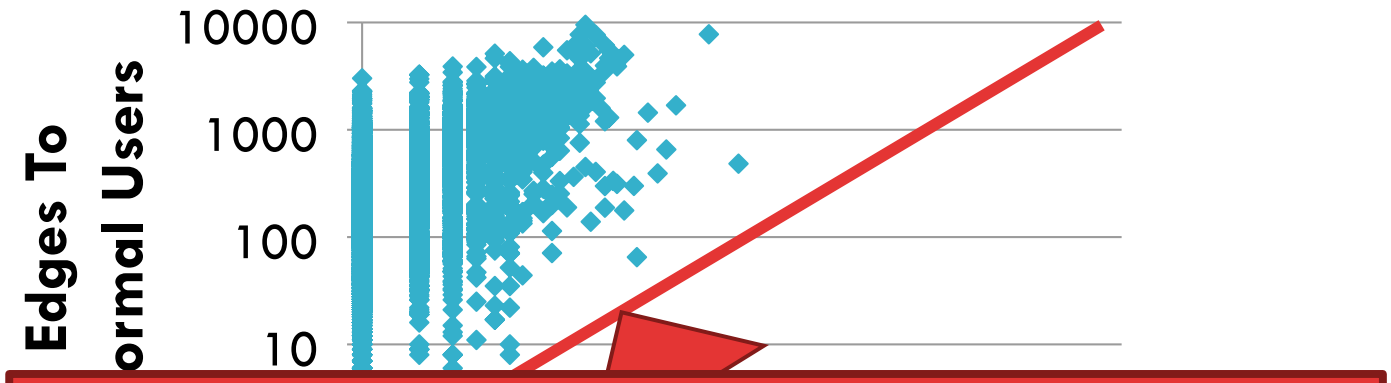
Community-based Sybil Detectors

- Prior work on Sybil detectors
  - SybilGuard [SIGCOMM'06], SybilLimit [Oakland '08], SybilInfer [NDSS'09]
  - Key assumption: **Sybils form tight-knit communities**
    - Sybils have difficulty "friending" normal users?



Do Sybils Form Sybil Communities?

- Measurement study on Sybils in the wild [IMC'11]
  - Study Sybils in Renren (Chinese Facebook)
  - Ground-truth data on 560K Sybils collected over 3 years
- Sybil components: sub-graphs of connected Sybils



- Sybil components are internally sparse
- Not amenable to community detection
- New Sybil detection system is needed

Detect Sybils without Graphs

- Anecdotal evidence that people can spot Sybil profiles
  - 75% of friend requests from Sybils are rejected
  - Human intuition detects even slight inconsistencies in Sybil profiles
- Idea: build a crowdsourced Sybil detector
  - Focus on user profiles
  - Leverage human intelligence and intuition
- Open Questions
  - How **accurate** are users?
    - What **factors** affect detection accuracy?
  - How can we make crowdsourced Sybil detection **cost effective**?

Outline

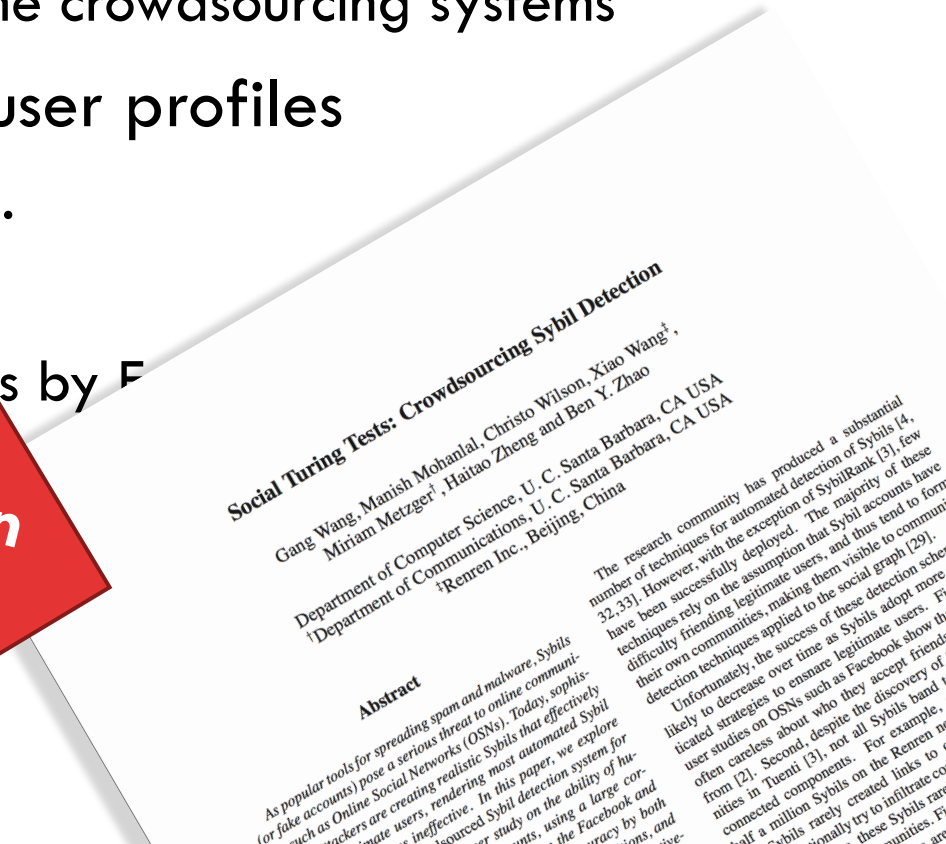
- Introduction
- User Study
  - Feasibility Experiment
  - Accuracy Analysis
  - Factors Impacting User Accuracy
- Scalable Sybil Detection System
- Conclusion

Details in Paper

User Study Setup*

- User study with 2 groups of testers on 3 datasets
- 2 groups of users
  - Experts – Our friends (CS professors and graduate students)
  - Turkers – Crowdworkers from online crowdsourcing systems
- 3 ground-truth datasets of full user profiles
  - Renren – given to us by Renren Inc.
  - Facebook US and India – crawled
    - Sybils profile
    - Legitimate p

Data collection details

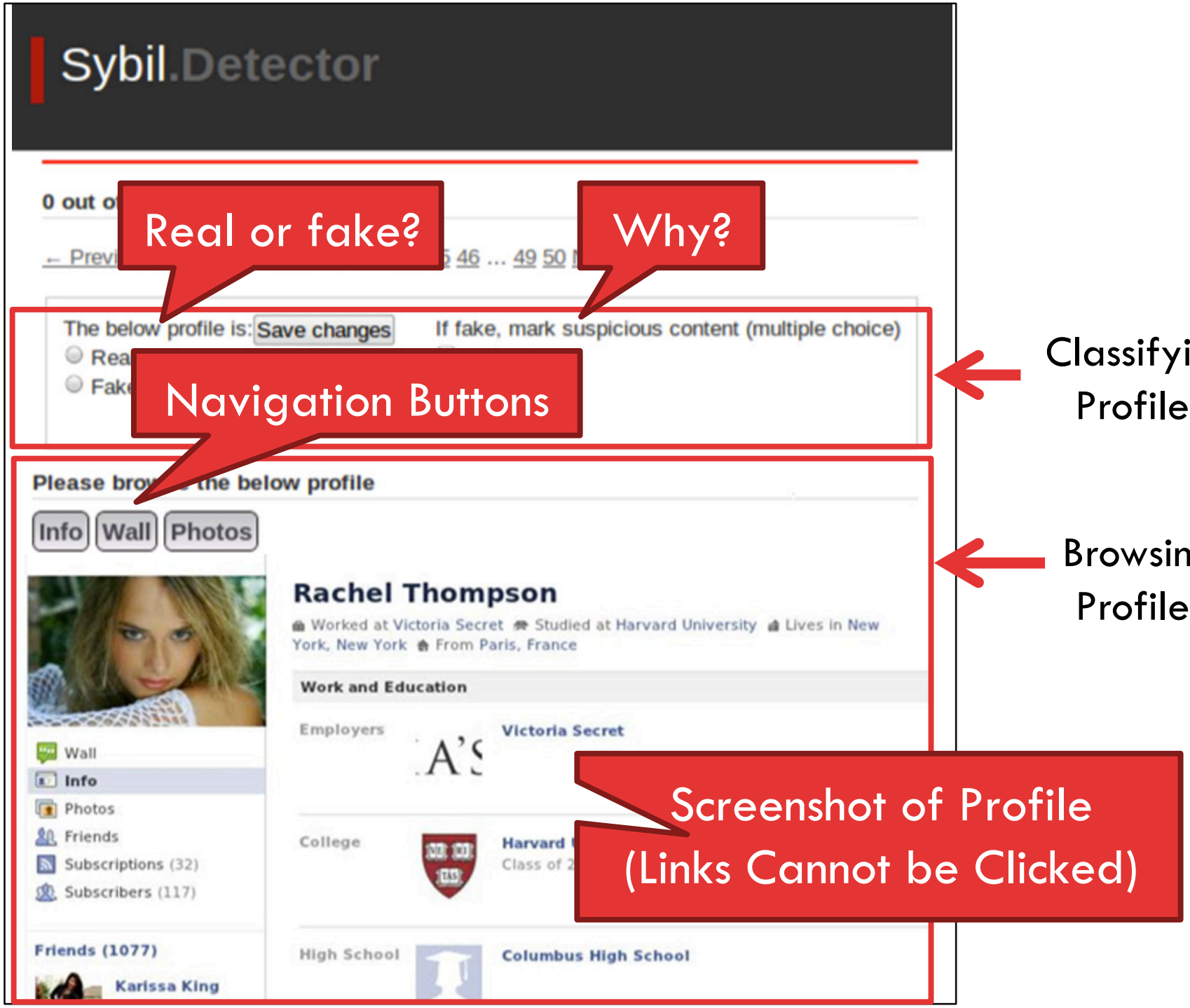

Social Turing Tests: Crowdsourcing Sybil Detection

Gang Wang, Manish Mohanlal, Christo Wilson, Xiao Wang‡, Miriam Metzger†, Haitao Zheng and Ben Y. Zhao

Department of Computer Science, U. C. Santa Barbara, CA USA
†Department of Communications, U. C. Santa Barbara, CA USA
‡Renren Inc., Beijing, China

Abstract

*IRB Approved

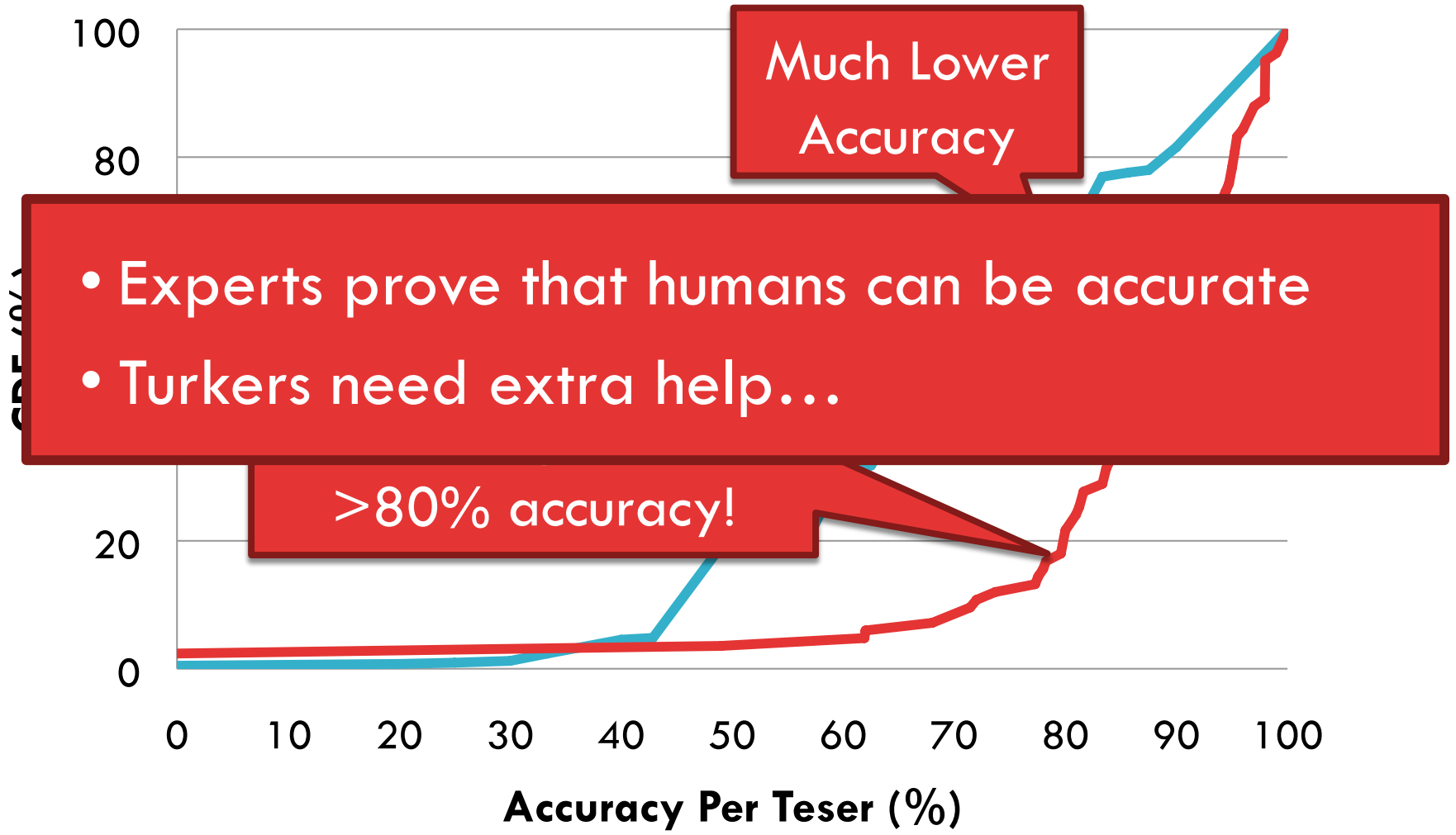

Sybil.Detector
Real or fake?
Why?
Navigation Buttons
The below profile is:
Save changes
If fake, mark suspicious content (multiple choice)
Info
Wall
Photos
Rachel Thompson
Worked at Victoria Secret
Studied at Harvard University
Lives in New York, New York
From Paris, France
Work and Education
Employers
Victoria Secret
College
Harvard
High School
Columbus High School
Friends (1077)
Karissa King
Classifying Profiles
Browsing Profiles
Screenshot of Profile
(Links Cannot be Clicked)

Experiment Overview

Dataset	# of Profiles		Test Group	# of Testers	Profile per Tester
	Sybil	Legit.			
Renren	100	100	Chinese Expert	24	100
			Chinese Turker	418	10
Facebook US	32	50	US Expert	40	50
			US Turker	299	12
Facebook India	50	49	India Expert	20	100
			India Turker	342	2

More Profiles per Experts

Individual Tester Accuracy



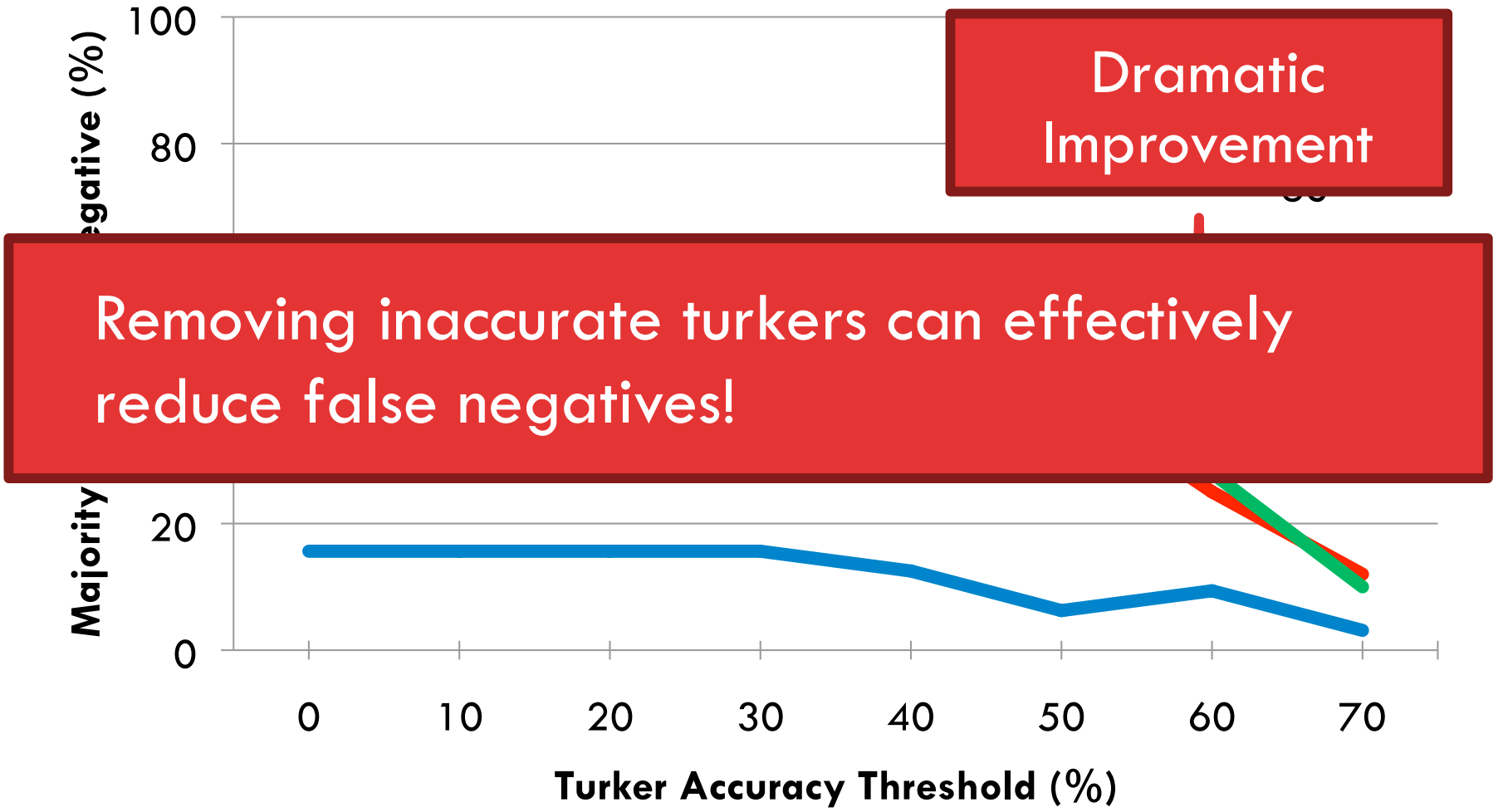
Wisdom of the Crowd

- Is wisdom of the crowd enough?

- Majority voting
  - Treat each classification by each tester as a vote
  - Majority vote determines final decision of the crowd

- False positive rates are excellent
- What can be done to improve turker accuracy?

Eliminating Inaccurate Turkers



Outline

- Introduction
- User Study
- Scalable Sybil Detection System
  - System Design
  - Trace-driven Simulation
- Conclusion

A Practical Sybil Detection System

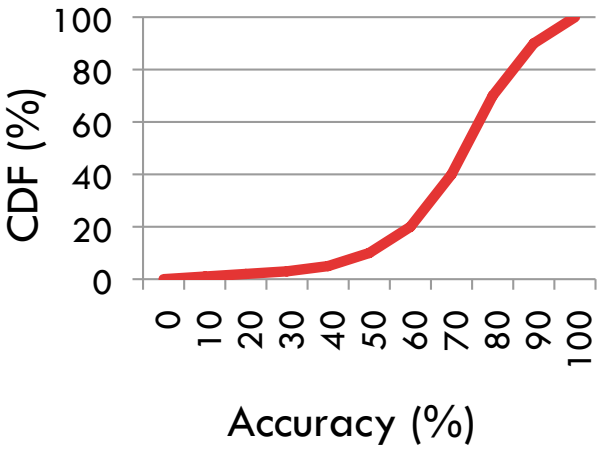
1. Scalability
   - Must scale to millions of users
   - High accuracy with low costs

Details in Paper

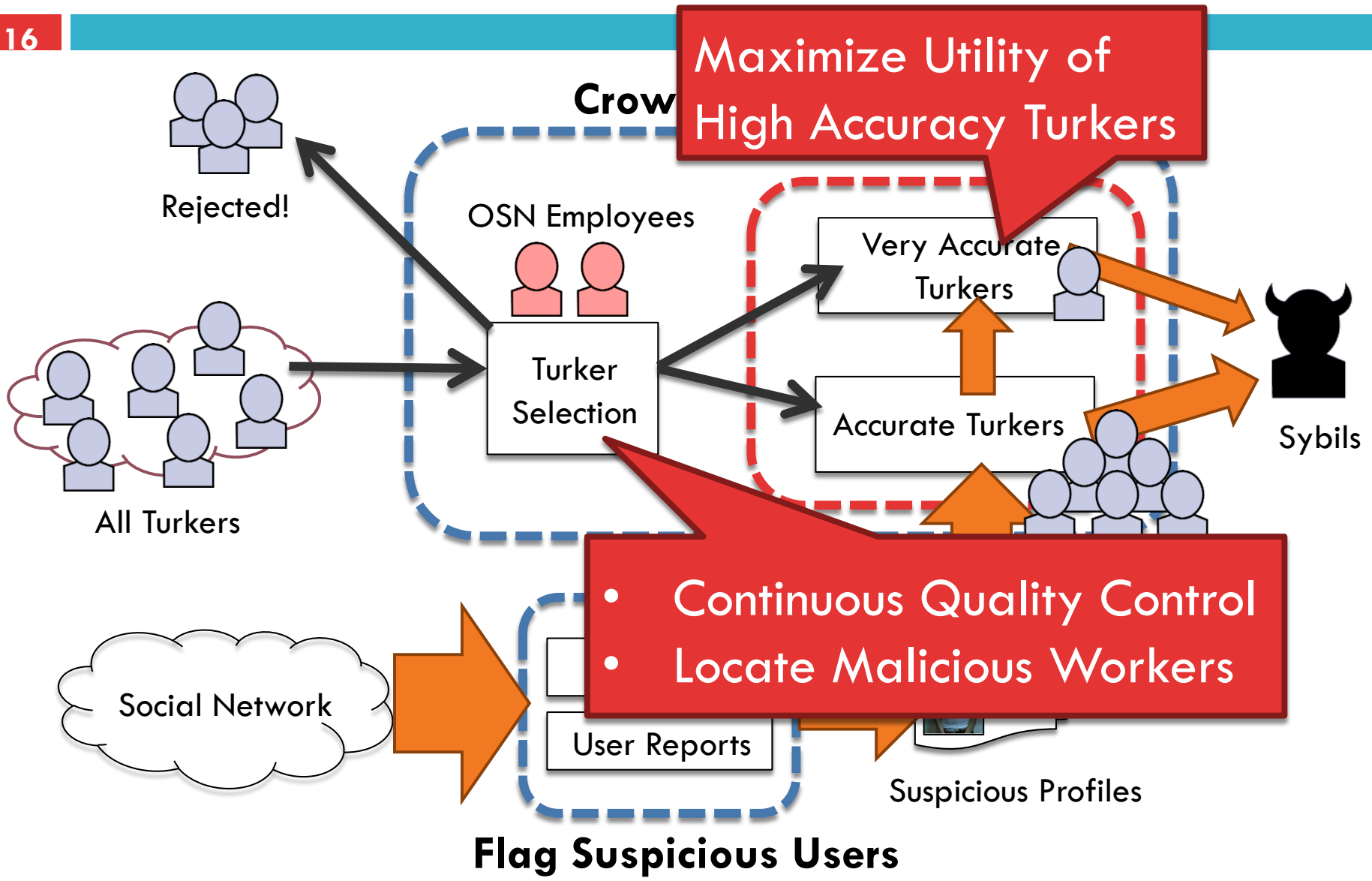
2. Preserve user privacy when giving data to turkers

Key insight to designing our system

- Accuracy in turker population highly skewed
- Only 10% turkers > 90% accurate

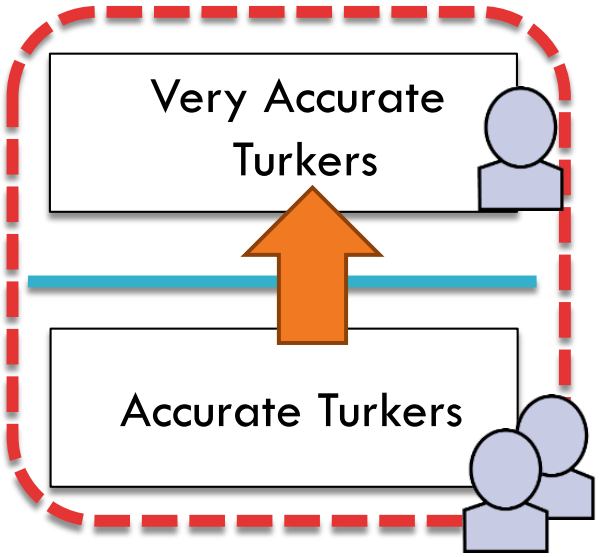


System Architecture



Trace Driven Simulations

- Simulation on 2000 profiles
- Error rates drawn from survey data
- Calibrate 4 parameters to:
  - Minimize false positives & false negatives
  - Minimize votes per profile (minimize cost)



Results++

- Average 8 votes per profile
- <0.1% false positives
- <0.1% false negatives

Estimating Cost

- Estimated cost in a real-world social networks: Tuenti
  - 12,000 profiles to verify daily
  - 14 full-time employees
  - er hour) → $2240 per day
- Cost with malicious turkers
  - 25% of turkers are malicous
  - $504 per day
  - 20sec/profile, 8 hour day turkers
  - Facebook wage ($1 per hour) → $400 per day

Augment existing automated systems

*http://www.glassdoor.com/Salary/Tuenti-Salaries-E245751.htm

Conclusion

- Designed a crowdsourced Sybil detection system
  - False positives and negatives <1%
  - Resistant to infiltration by malicious workers
  - Low cost

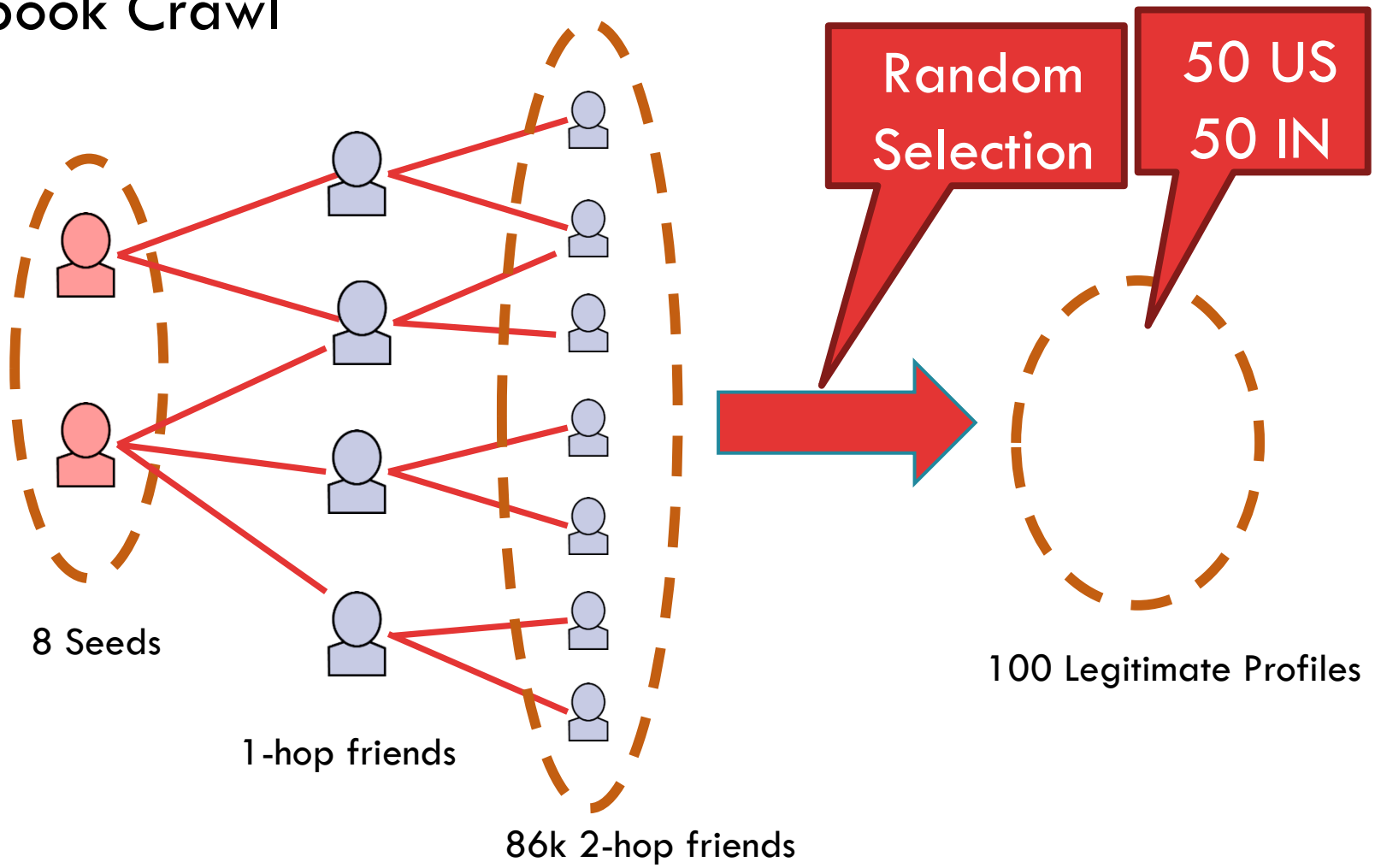
- Currently exploring prototypes in real-world OSNs

Questions?

Thank you!

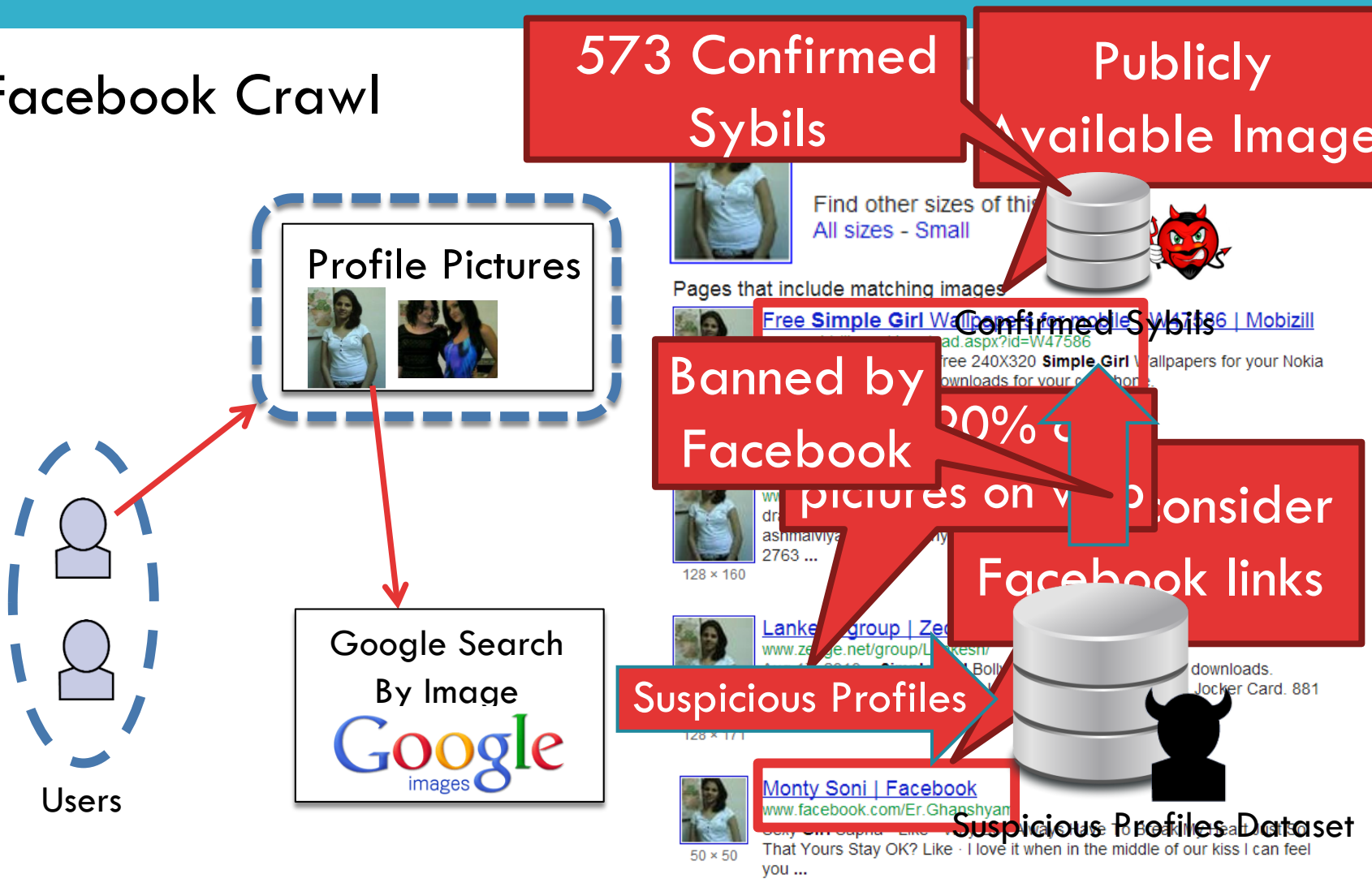
Ground-truth Data Collection (Legit.)

- Facebook Crawl



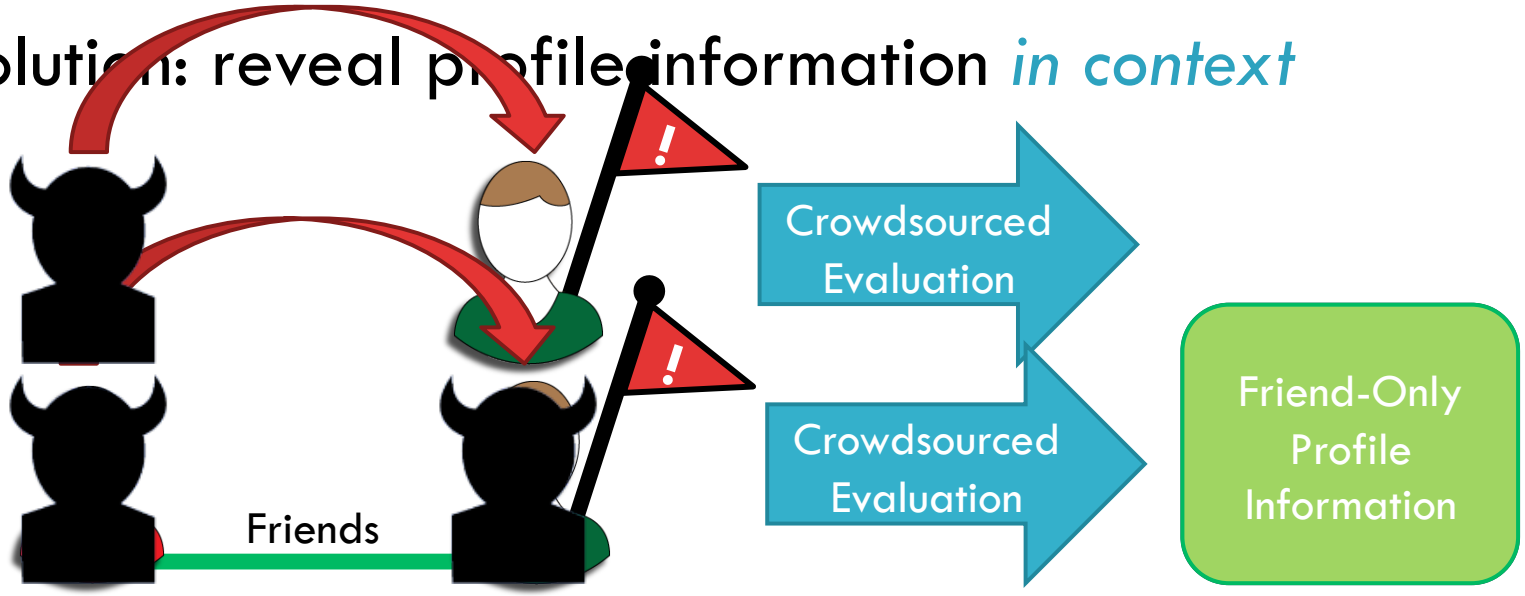
Ground-truth Data Collection (Sybil)

- Facebook Crawl

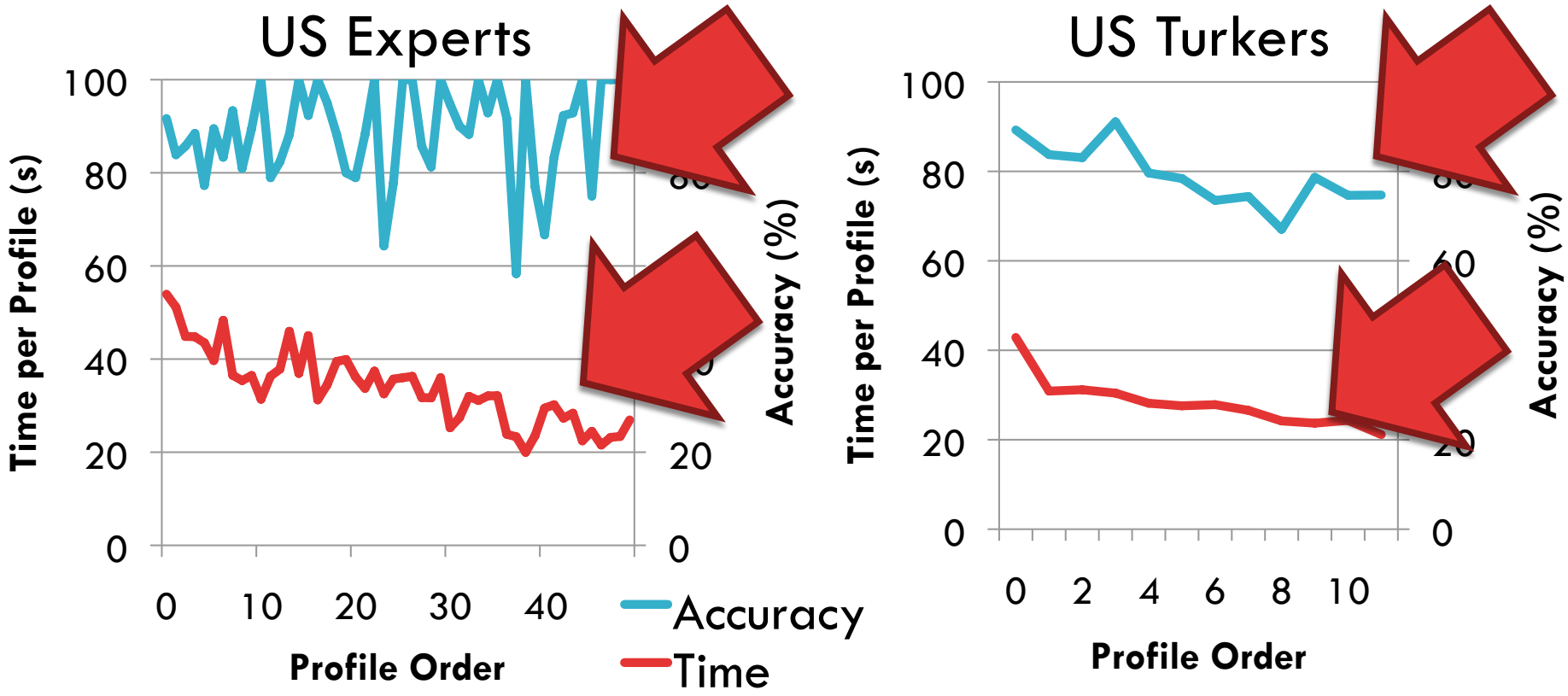


Preserving User Privacy

- Showing profiles to crowdworkers raises privacy issues
- Solution: reveal profile information *in context*



Survey Fatigue



Wisdom of the Crowd

- Treat each classification by each tester as a vote
- Majority vote determines final decision

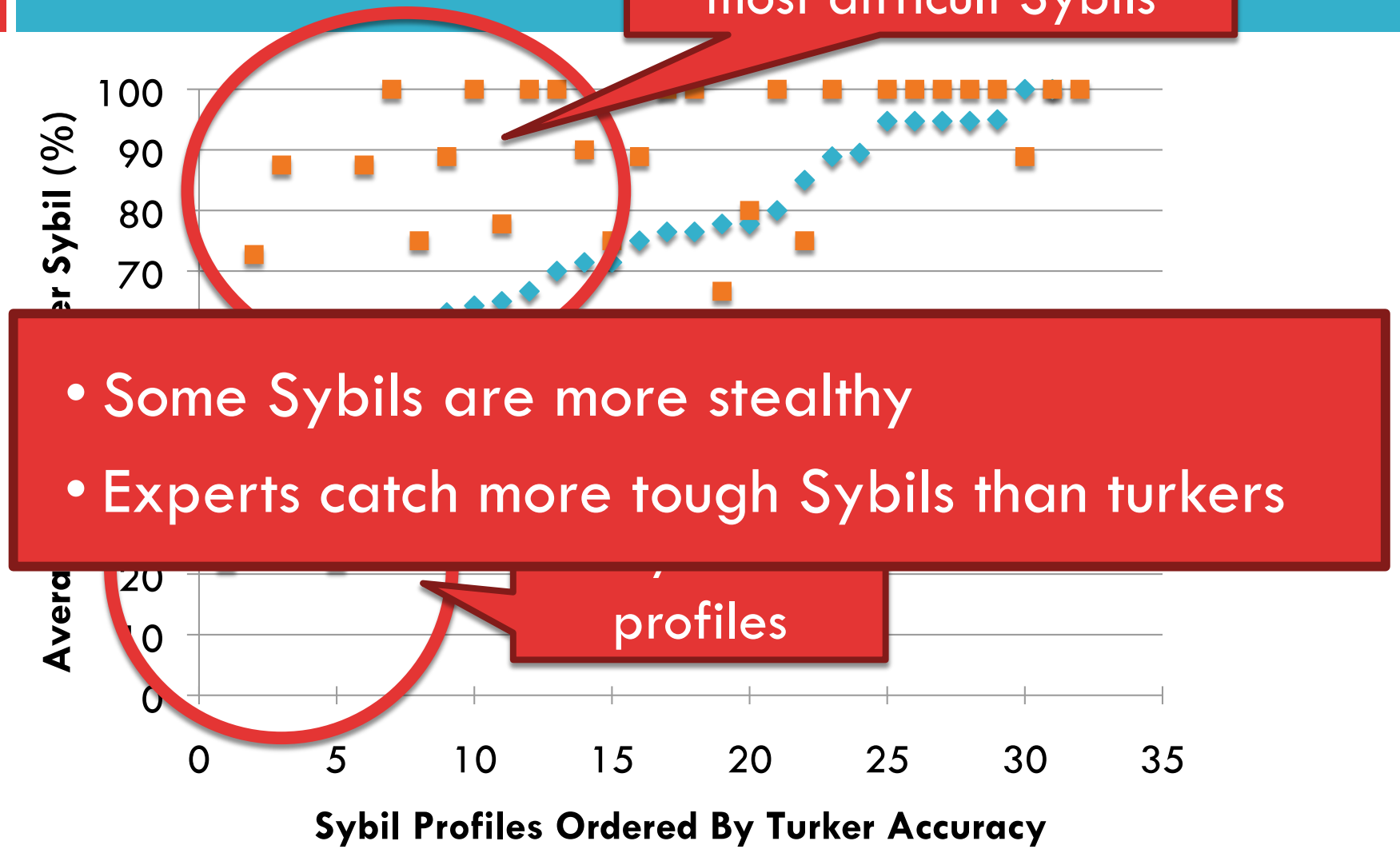
Almost Zero

Experts

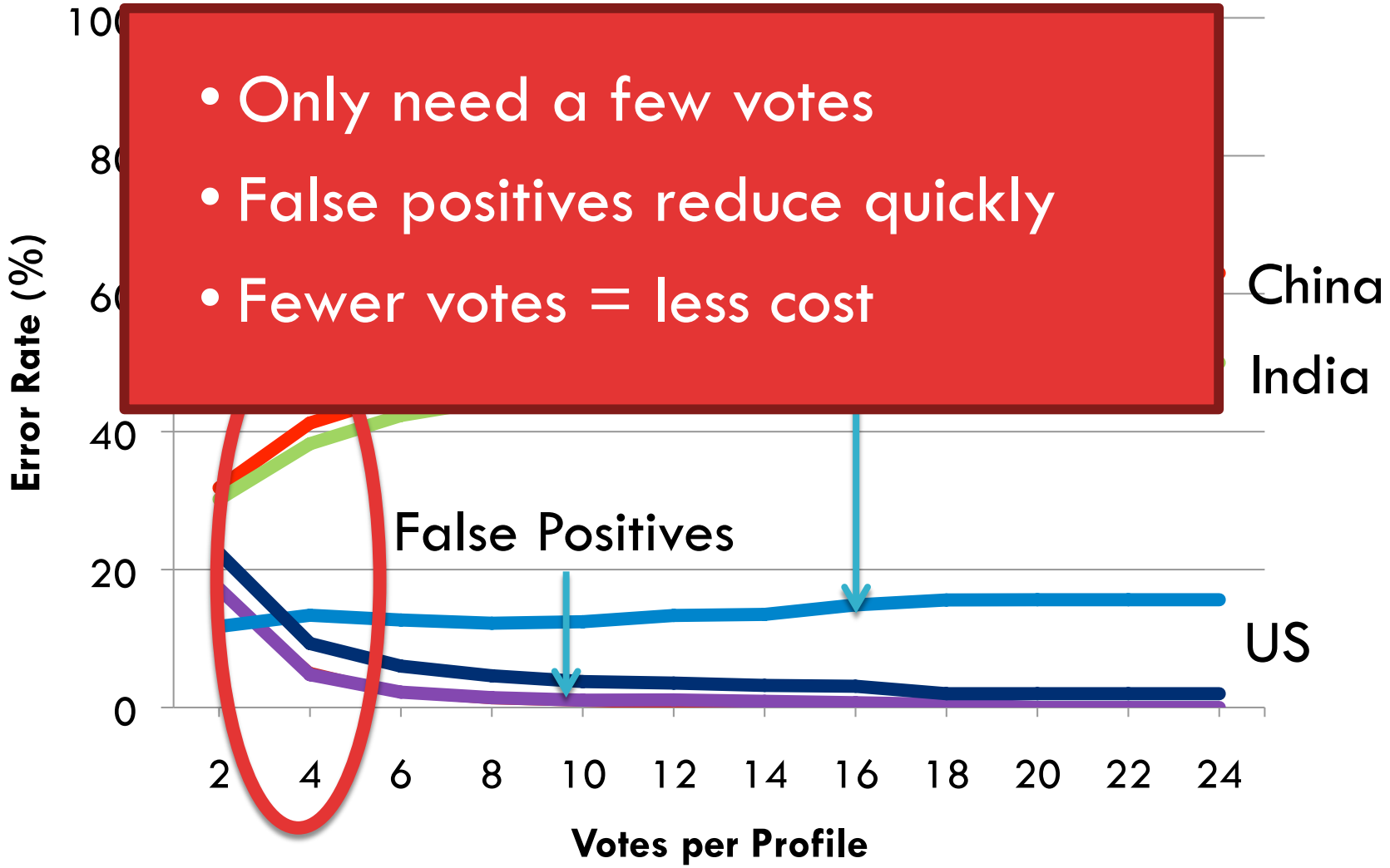
Dataset		False Positives	False Negatives
Facebook India	India Expert	0%	16%
	India Turker	0%	50%

- False positive rates are excellent
- Turkers need extra help against false negatives
- What can be done to improve accuracy?

Sybil Profile Diffic



How Many Votes Do You Need?



Individual Tester Accuracy