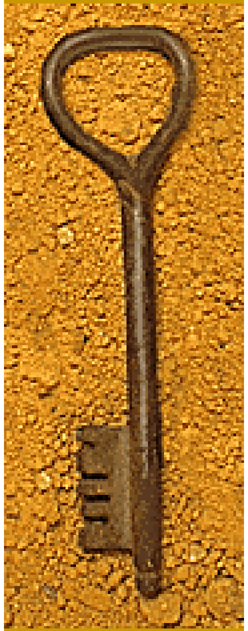
Trust Management for Host-based Collaborative Intrusion Detection

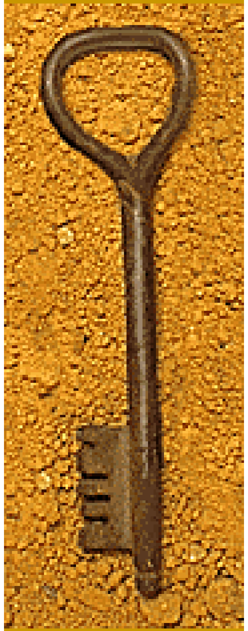
Carol Fung, Olga Baysal, Jie Zhang,
Issam Aib and Raouf Boutaba

David R. Cheriton School of Computer Science
University of Waterloo



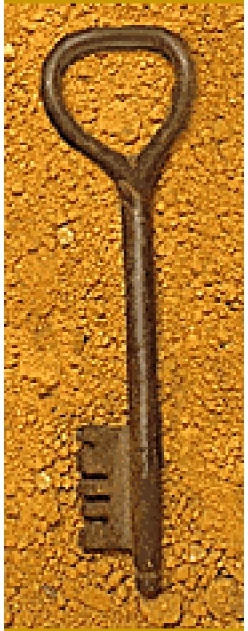
Outline

- Introduction
- Motivation
- Trust Model
- Robustness
- Attacks Prevention
- Conclusions



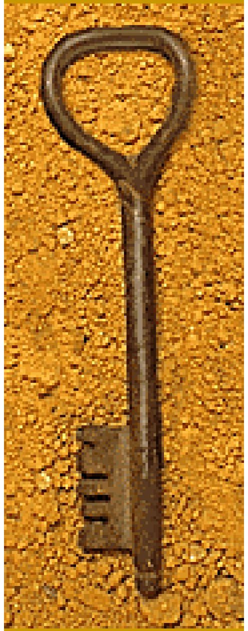
Introduction

- The worldwide impact of malicious intrusions is estimated to be over $10 Billion annually
- Intrusions include viruses/worms, spyware, spam, DoS, unauthorized login.
- Traditional isolated IDSes are inefficient in detection unknown intrusions.



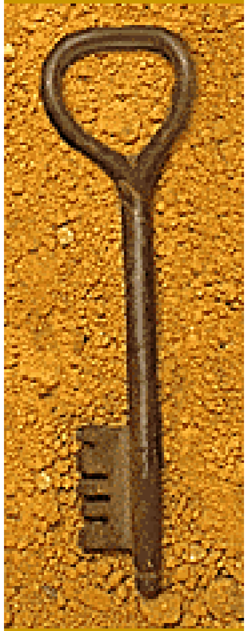
Benefits of IDS Collaboration

- Accurate alert ranking
- Effective intrusion detection and prevention
- Worm spreading warning



Related Work

- Current collaborative architectures have strong assumptions on the trustworthiness (all IDSes are trusted and faithfully report intrusion events) [1,2,3,4]
- Used trust models are naïve [5, 6]
- Many efficient trust management models in other areas such as e-market and P2P networks [7,8,9]

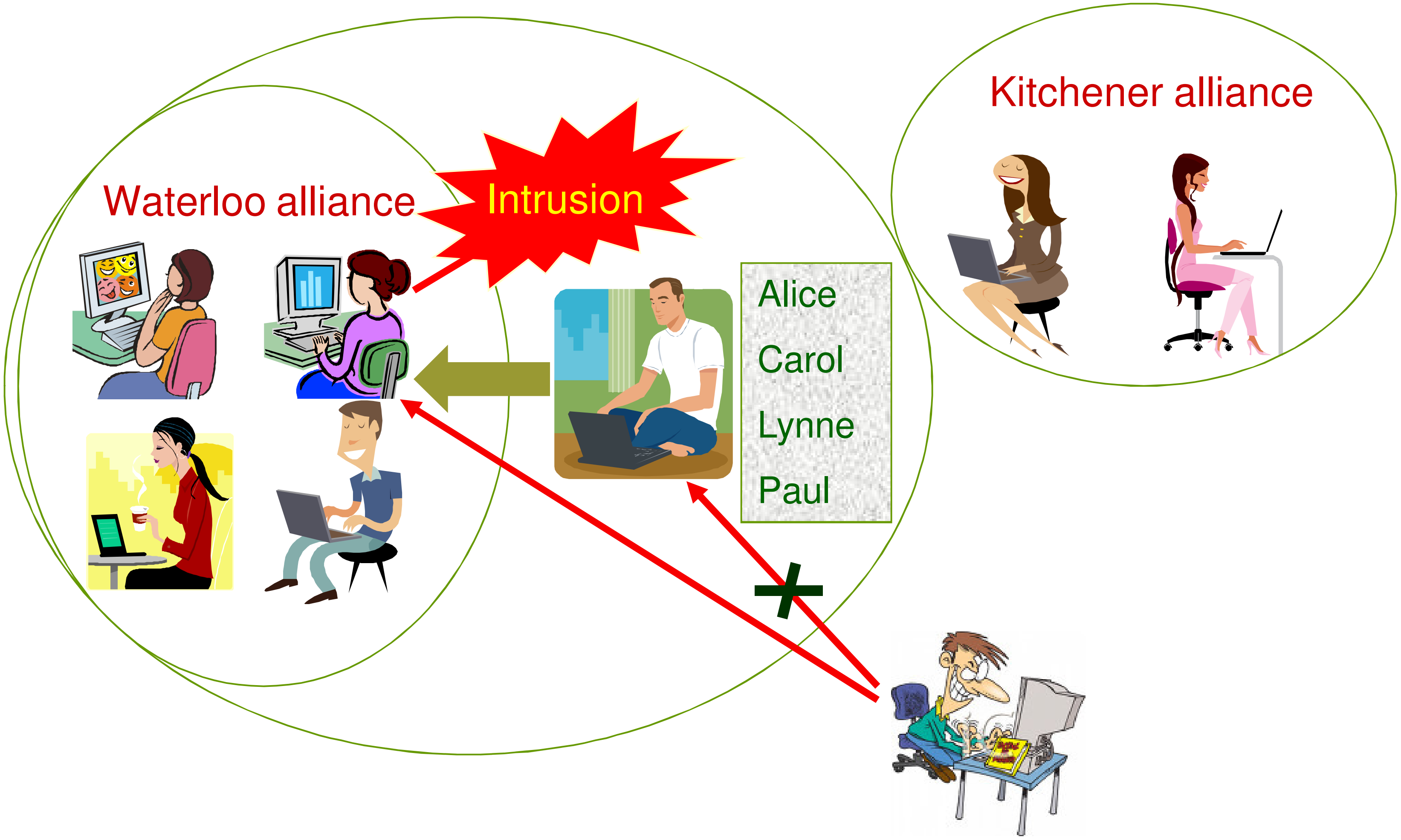


Contribution

1. Build a trust management model for IDS collaboration

2. Propose a framework for efficient collaboration of IDSes

Scenario



Waterloo alliance

Intrusion

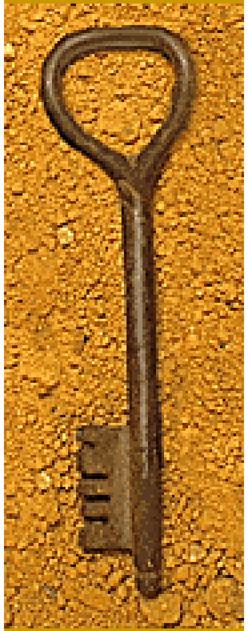
Kitchener alliance

Alice

Carol

Lynne

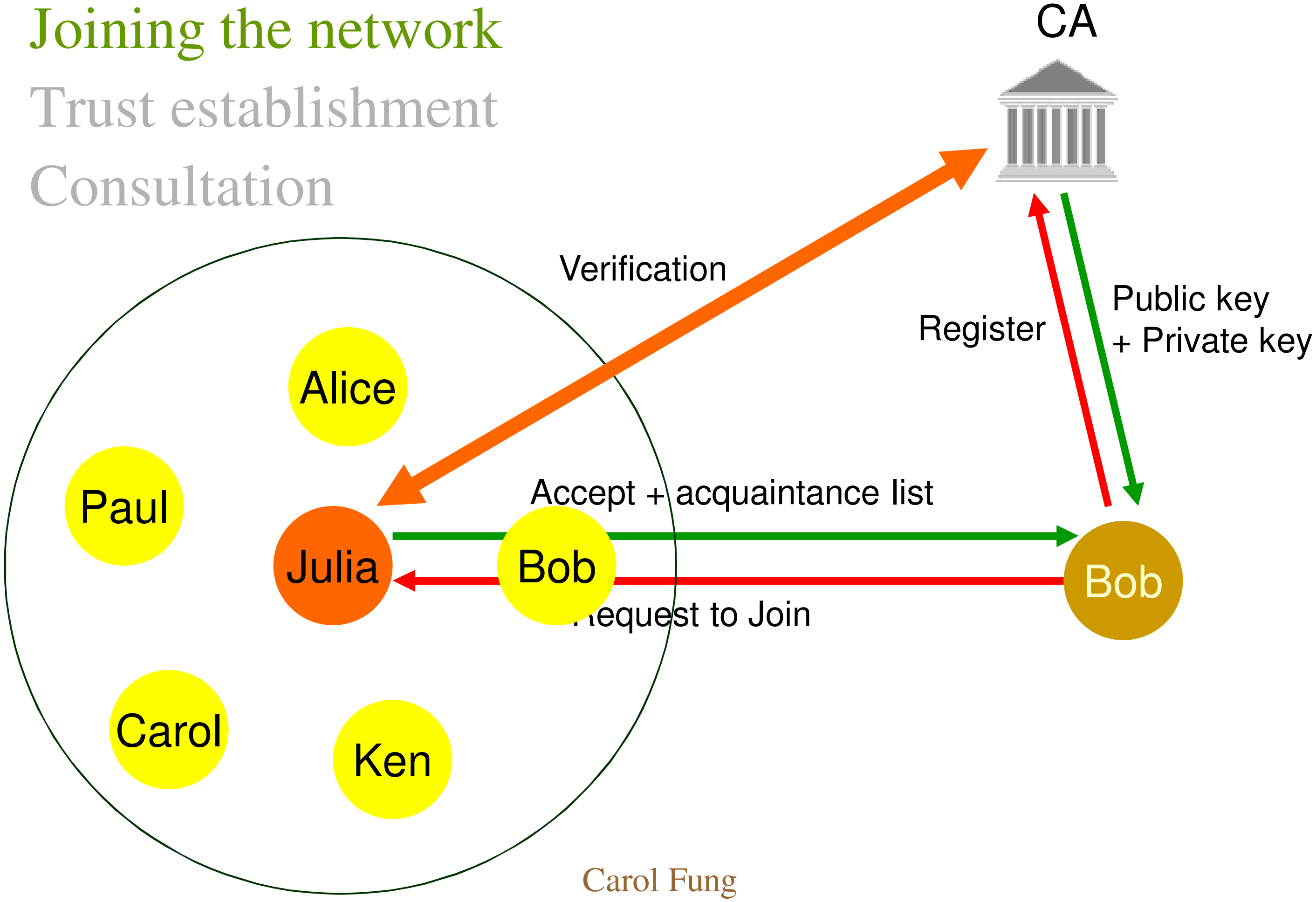
Paul

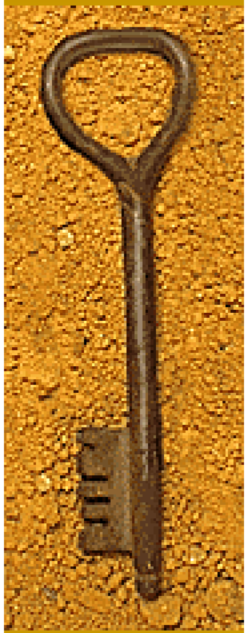


Framework

Three components:

- Joining the network
- Trust establishment
- Consultation

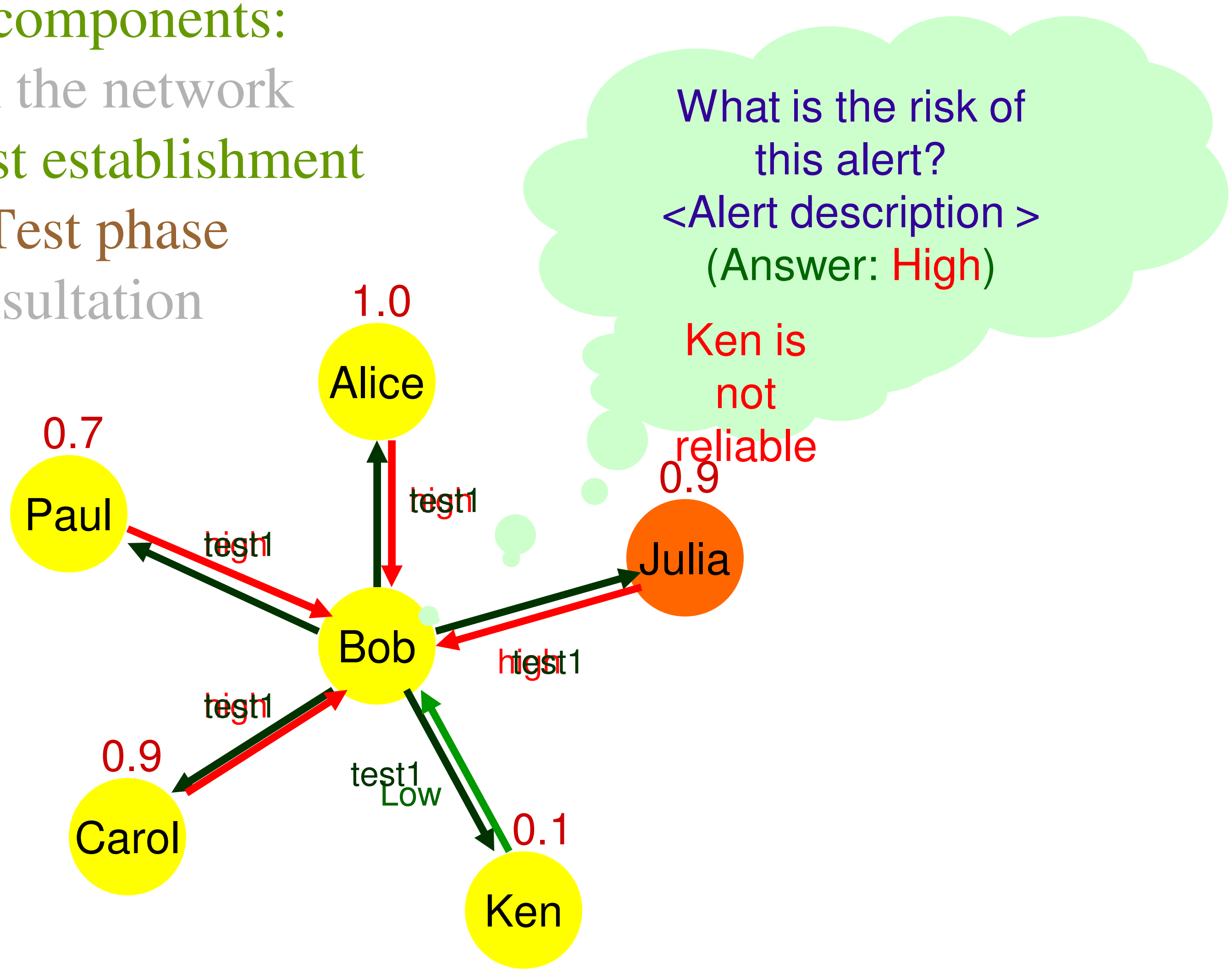


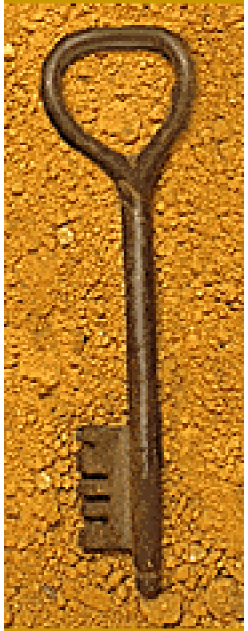


Framework

Three components:

- Join the network
- Trust establishment
  - Test phase
- Consultation

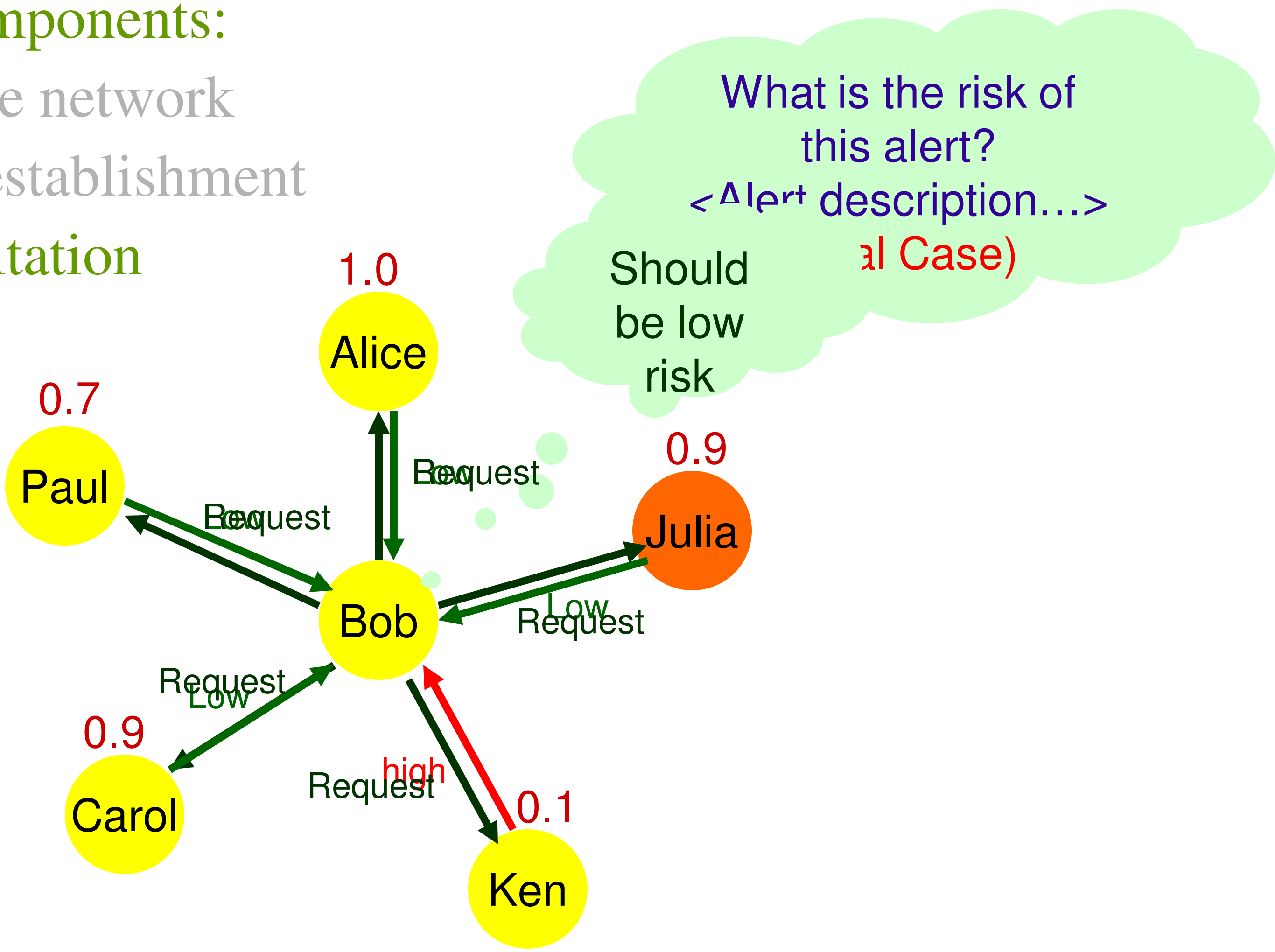


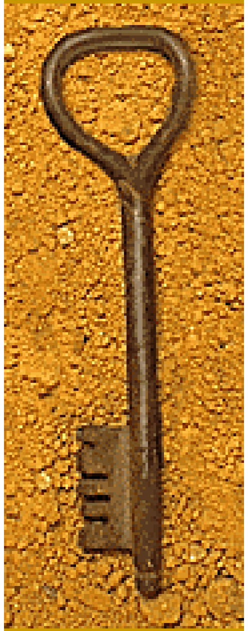


Framework

Three components:

- Join the network
- Trust establishment
- Consultation





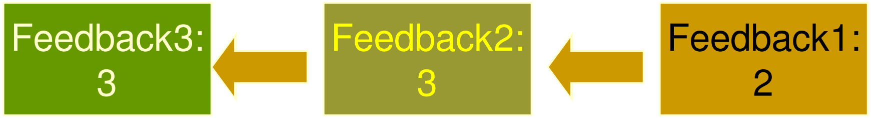
Trust Establishment

Trust level is built on history

- Satisfactory level of past feedbacks
- Helpfulness

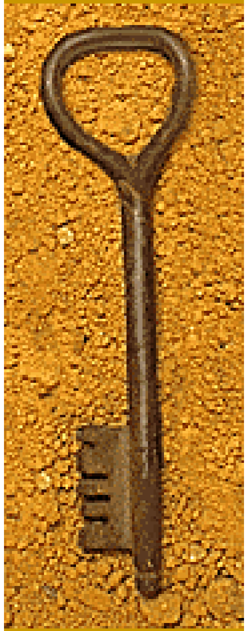
Feedback	Expected Answer	Satisfaction Level
High	High	3
Medium	High	1
Low	High	0

Paul’s History


Feedback3: 3 ← Feedback2: 3 ← Feedback1: 2

$$\xcancel{TrustLevel = \frac{3+3+2}{3+3+3}} \Rightarrow TrustLevel = \frac{3+3\lambda+2\lambda^2}{3+3\lambda+3\lambda^2}$$

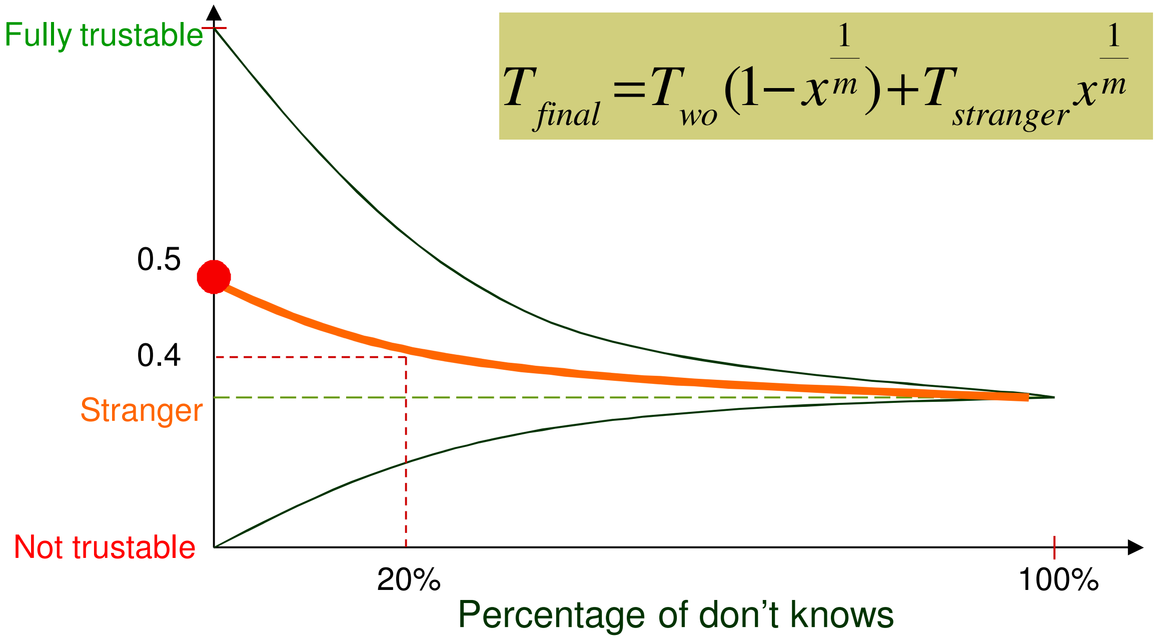
Naive!



Integration of Don’t Knows

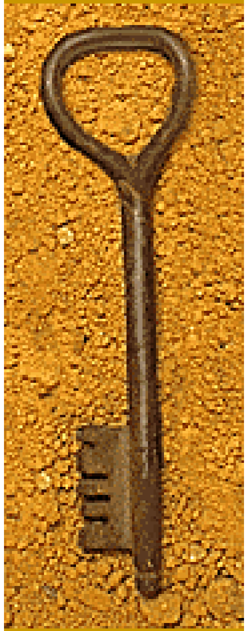
Reply “don’t know” is allowed

- Trust value will approach to the level of stranger



$$T_{final} = T_{wo}(1 - x^{\frac{1}{m}}) + T_{stranger}\, x^{\frac{1}{m}}$$

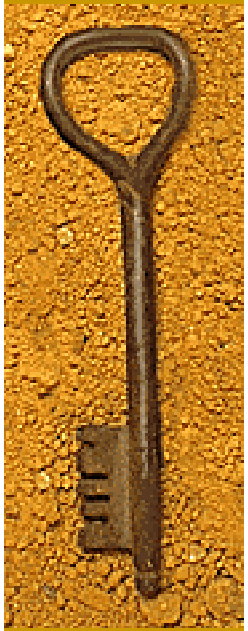
Carol Fung



Feedback Aggregation

Depends on:

- Peers' trust values
  - Trust weight
- Peer's location
  - Proximity weight



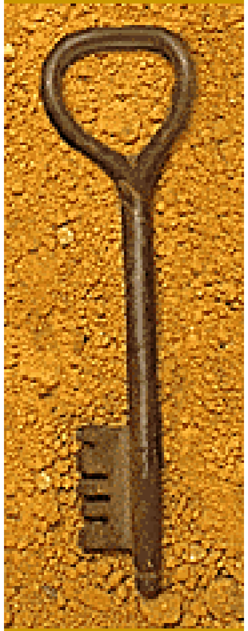
Feedback Aggregation

- Weighted average
- Threshold

Name	Trust	Proximity	Ranking
Alice	1	1 (Waterloo)	High(3)
Carol	0.9	1 (Waterloo)	High(3)
Julia	0.9	0.9 (Toronto)	High(3)
[redacted]	[redacted]	[redacted]	[redacted]
Paul	0.7	0.7 (US)	Medium(2)

High Risk!

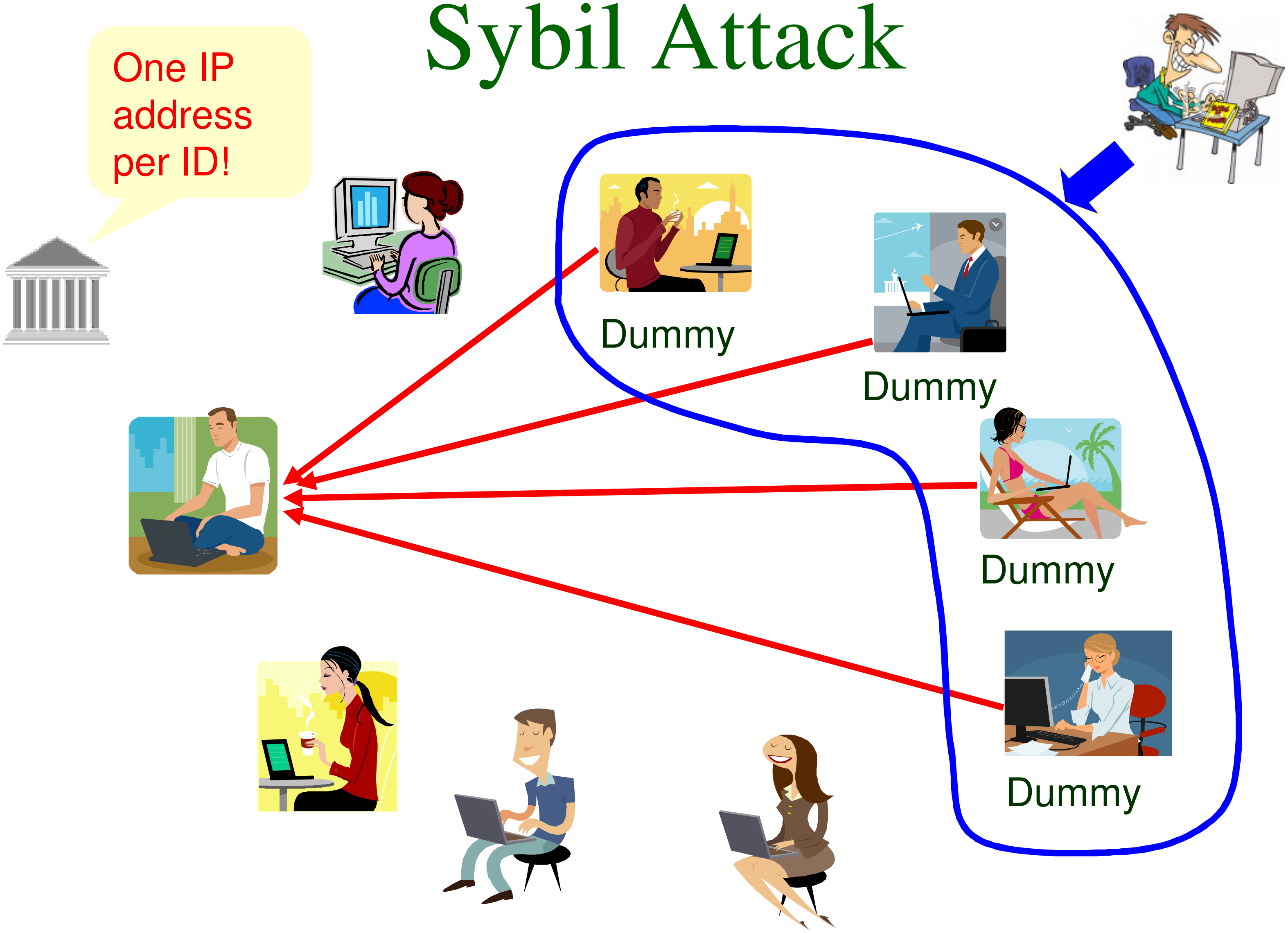
$$finalRanking = \frac{3\cdot 1\cdot 1 + 3\cdot 0.9\cdot 1 + 3\cdot 0.9\cdot 0.9 + 2\cdot 0.7\cdot 0.7}{1\cdot 1 + 0.9\cdot 1 + 0.9\cdot 0.9 + 0.7\cdot 0.7}$$
$$= 2.85$$

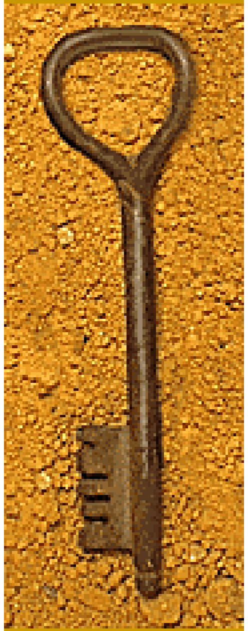


Threat Model

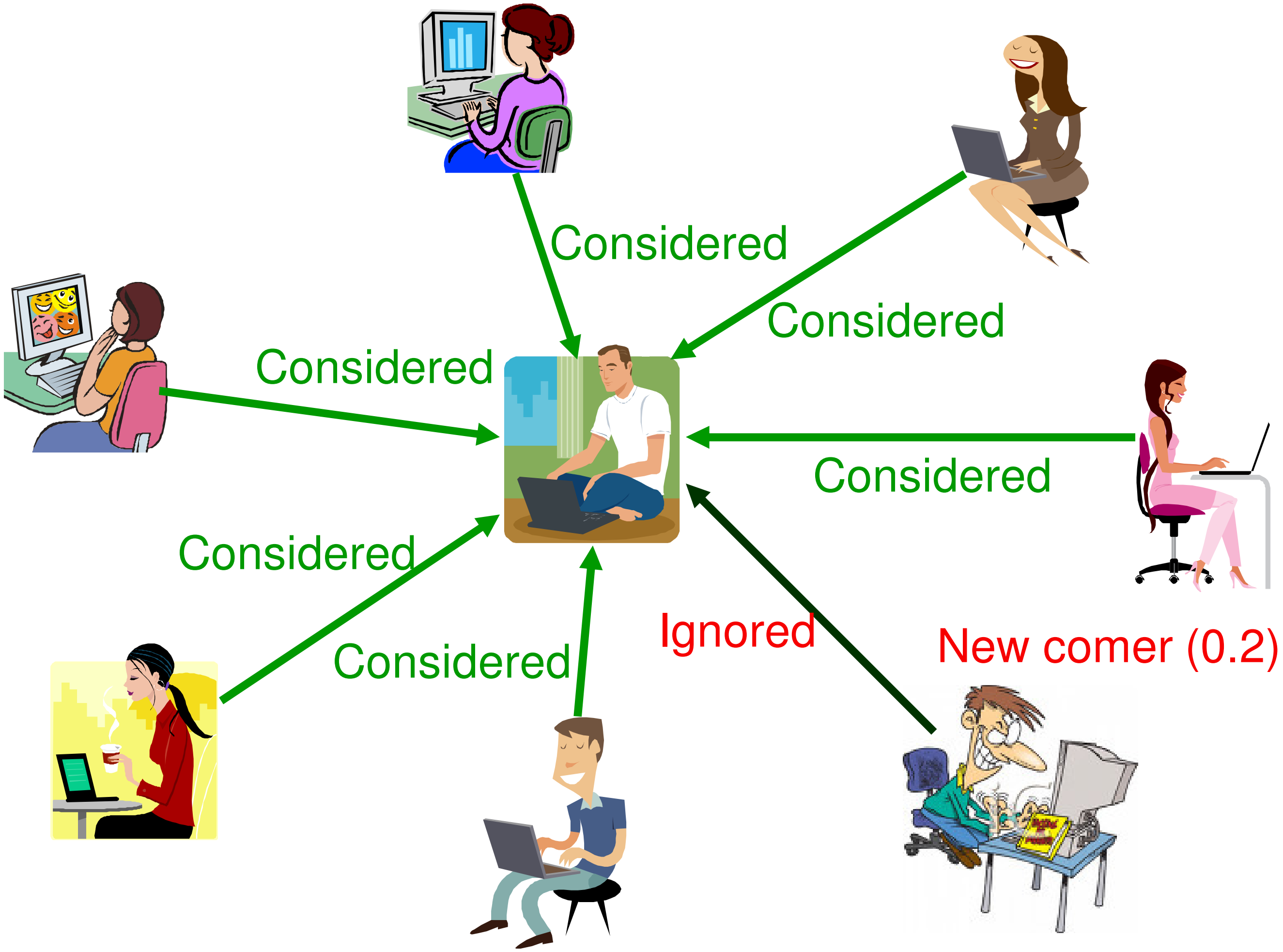
- Sybil Attack
- New Comer Attack
- Identity Cloning Attack
- Betrayal Attack
- Collusion Attack

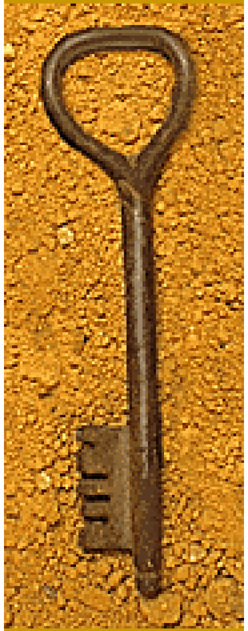
Sybil Attack



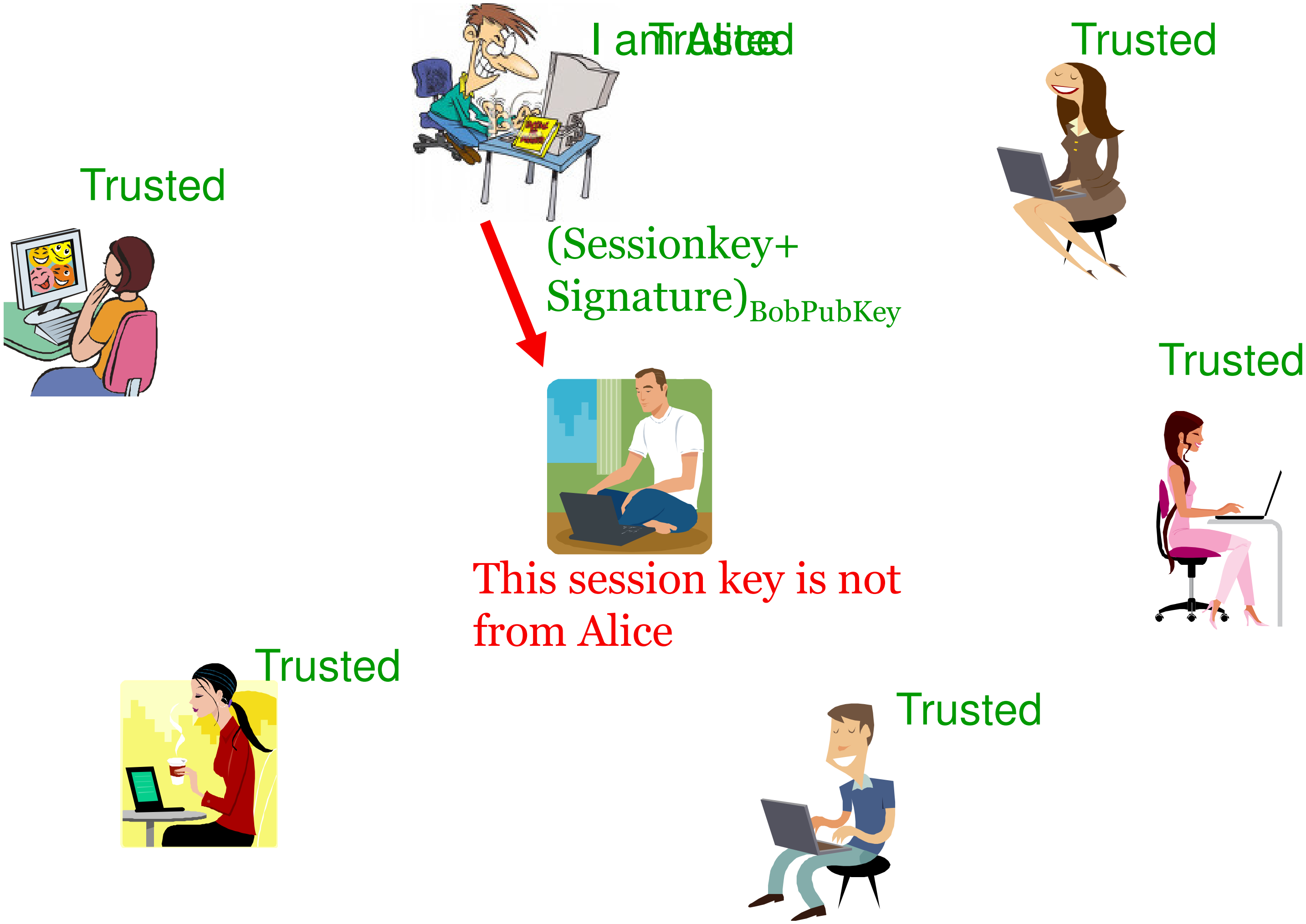


New Comer Attack





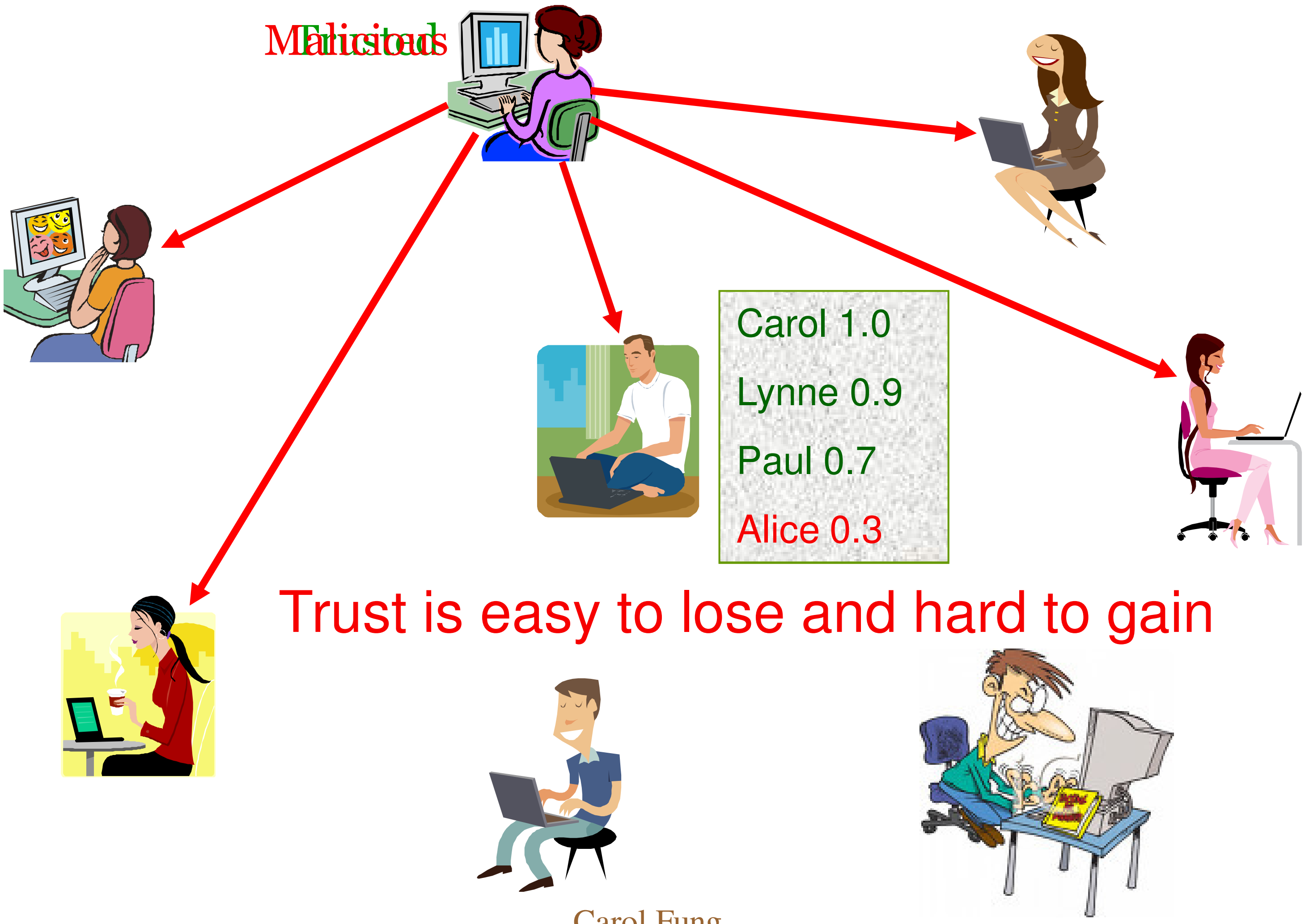
Identity Cloning Attack



$(\text{Sessionkey+}\ \text{Signature})_{\text{BobPubKey}}$

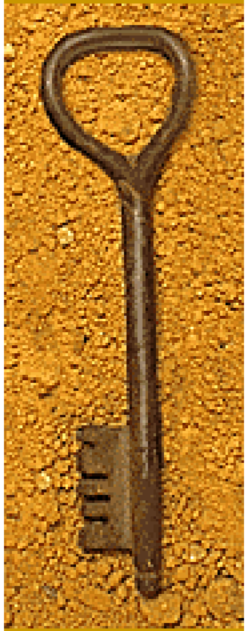
This session key is not from Alice

Betrayal Attack

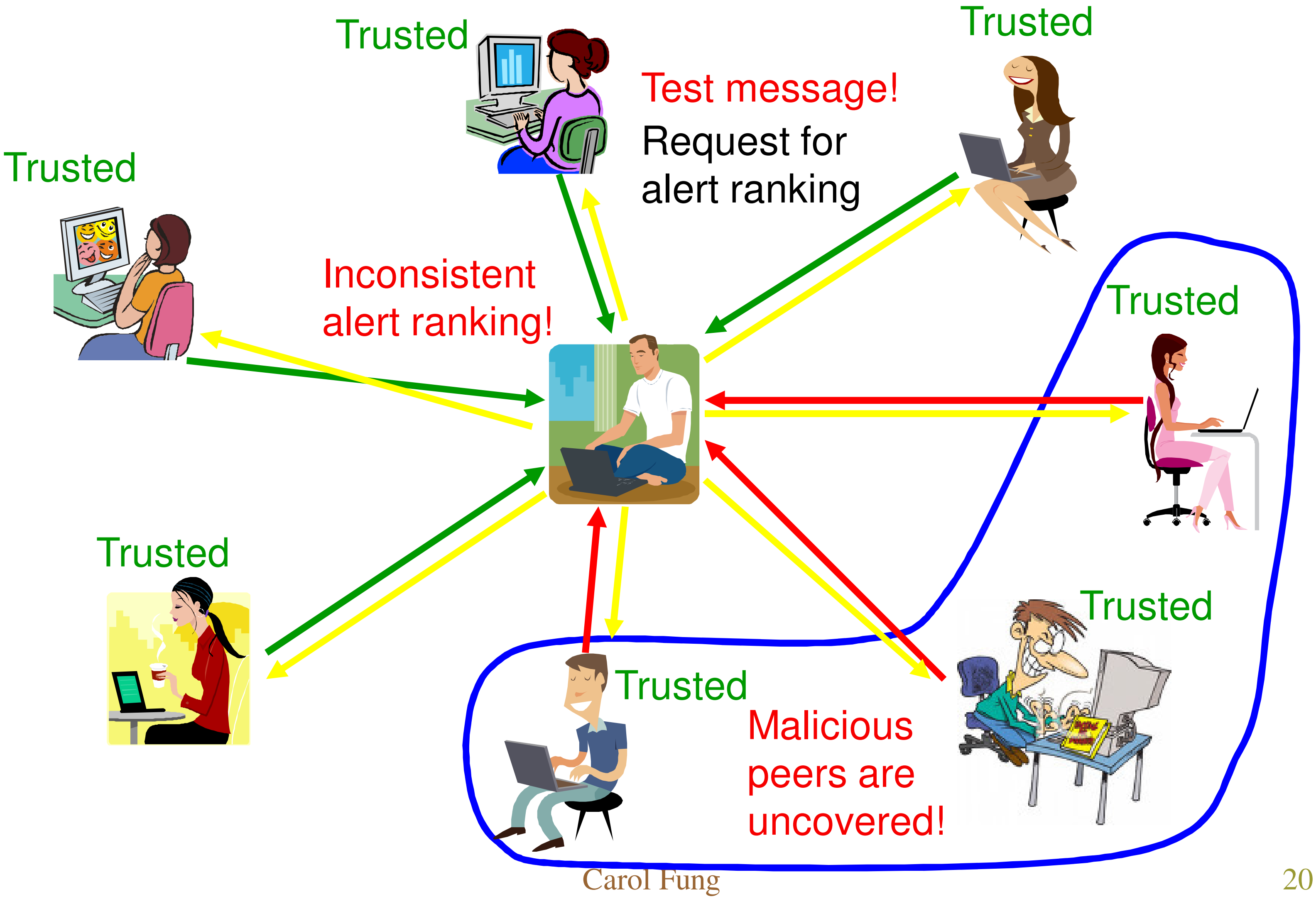


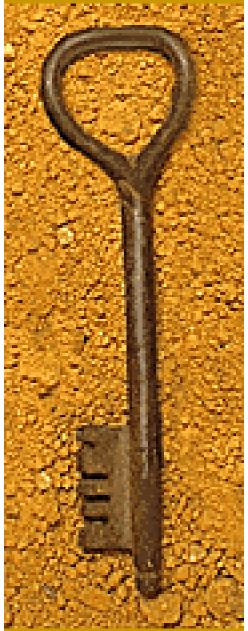
Carol	1.0
Lynne	0.9
Paul	0.7
Alice	0.3

Trust is easy to lose and hard to gain



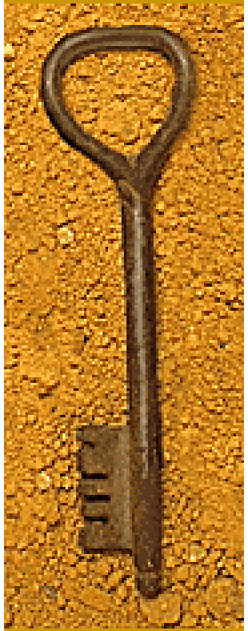
Collusion Attack


Trusted
Trusted
Test message!
Request for alert ranking
Trusted
Inconsistent alert ranking!
Trusted
Trusted
Trusted
Trusted
Malicious peers are uncovered!



What’s Next?

- Simulation design and implementation
- Design more sophisticated trust management model
  - Alert categorization
  - Expertise in intrusion detection



Conclusion

- Proposed a trust-based IDS collaboration model
  - More accurate intrusion detection
  - Robust to several attacks
- Novel ideas
  - Use of test messages in trust establishment
  - Integration of “don’t knows” into trust value
  - Introduction of proximity
  - Aggregation threshold