

SezHoo: Using Reputation to Increase the Trustworthiness of Information

Mark A. Kramer

781-271-3296 • mkramer@mitre.org

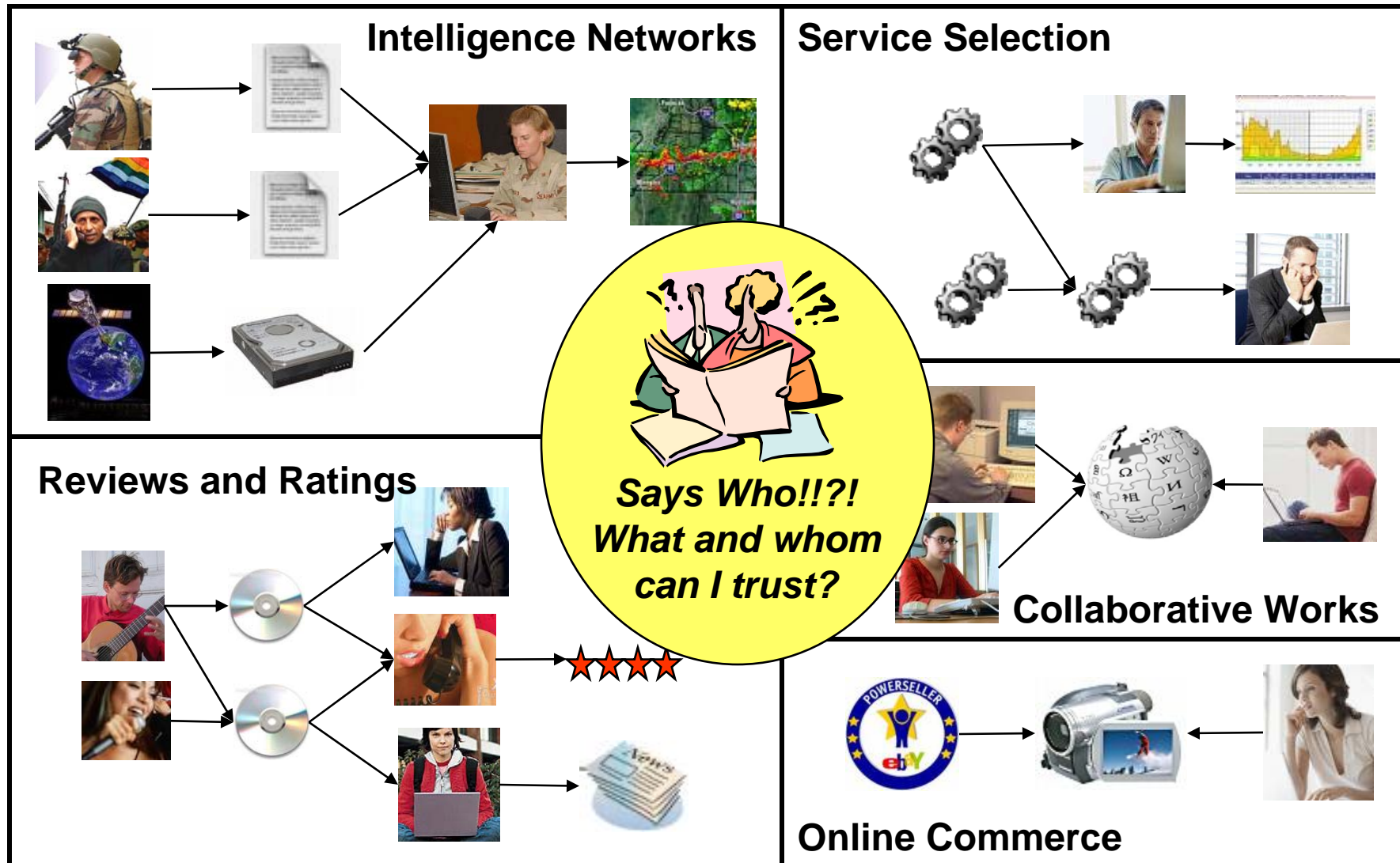
MITRE Sponsored Research



Problem

- **Cooperation transcending traditional hierarchical boundaries is a hallmark of future operations**
- **Many participants will create, share, and judge information**
- **Notwithstanding information assurance, NOT all sources, data, and analysis are equally trustworthy**
- **It is difficult for the consumer to know what to believe or who is a reliable source of information**

Background



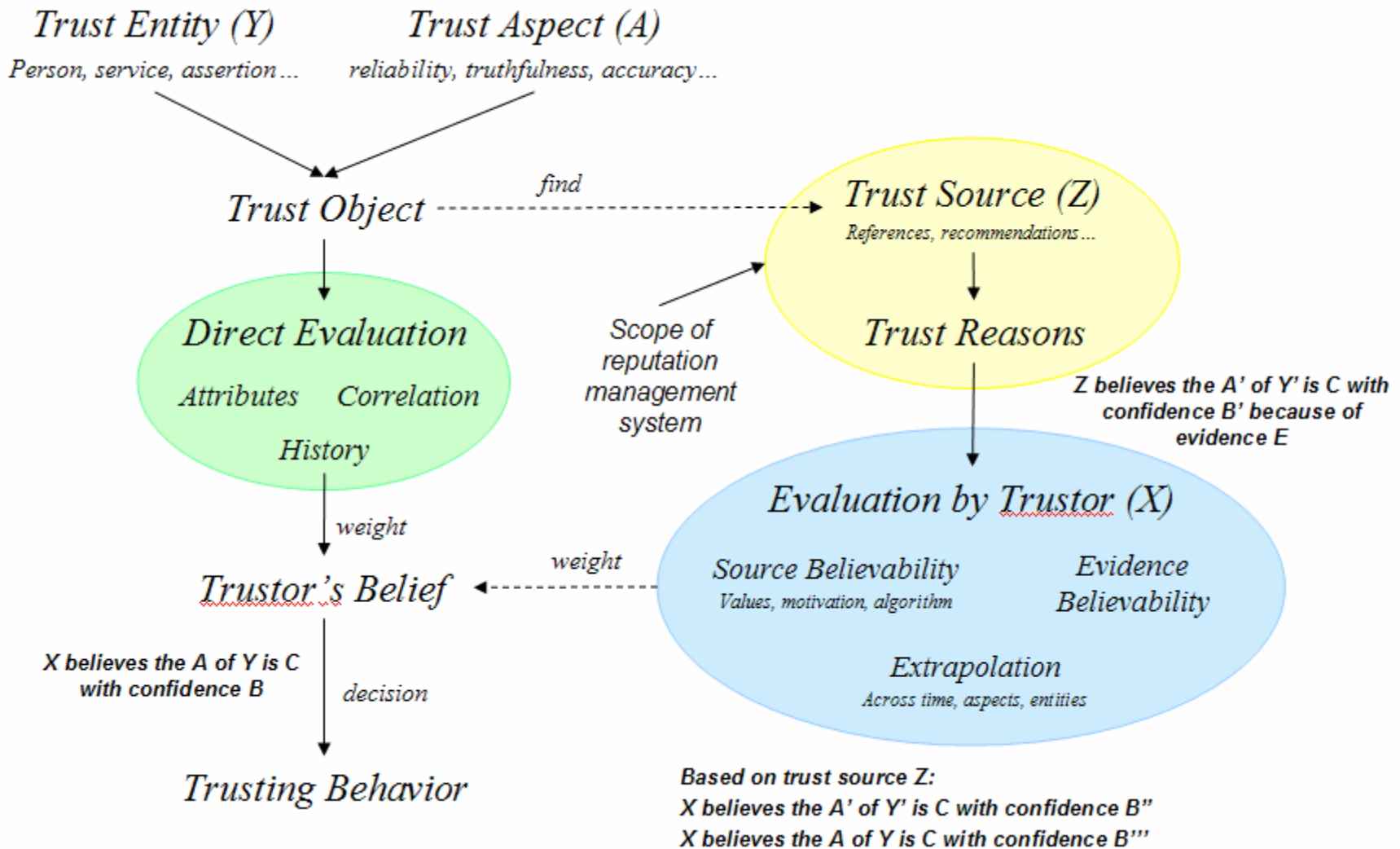
Objective

- Develop models of **trust and reputation** for information systems involving human judgment and expertise
- Evaluate **incentives and sanctions** that increase the overall trustworthiness of information, without driving out potentially useful information
- Identify and **counteract vulnerabilities** of reputation systems to strategic behavior by non-cooperative information providers

Activities

- Apply modeling approach to determine how best to **estimate information reliability**, participant trustworthiness, and the value of analyst and source contributions
- Create computational environment to **simulate and test reputation algorithms**, using test problem or competition framework
- Create **case studies** applying reputation systems to collaborative works, Web service selection, and intelligence networks

Highlight



Highlight

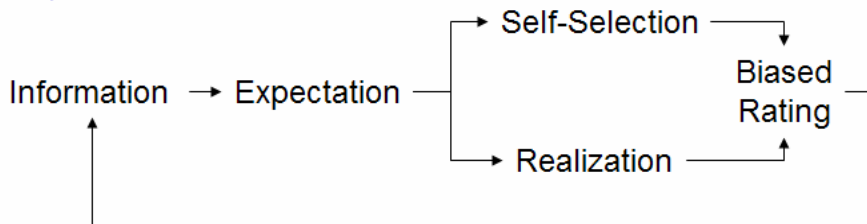
Self-Selection Bias in Reputation Systems

Observation:

Reputation systems are inherently biased toward better-than-average ratings

- Amazon average rating = 3.9 stars out of five
- Netflix average rating = 3.6 stars out of five

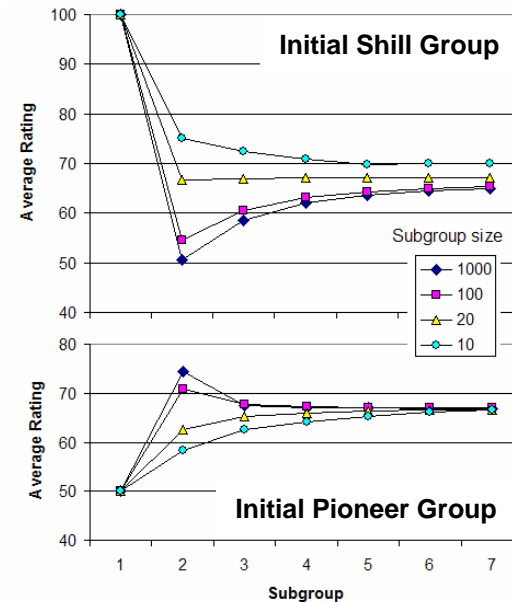
Hypothesis:



Analysis:

- Self-Selection: $P(\text{Selection}|\text{Expectation}) > P(\text{Selection}|\sim\text{Expectation})$
 Realization: $P(\text{Satisfaction}|\text{Expectation}) > P(\text{Satisfaction}|\sim\text{Expectation})$
 Independence: $P(\text{Satisfaction}|\text{Expectation}, \text{Selection}) = P(\text{Satisfaction}|\text{Expectation})$
Implies
 Bias: $P(\text{Satisfaction}|\text{Selection}) > P(\text{Satisfaction})$

Results:



Fair rating = 0.5, final rating = 0.67

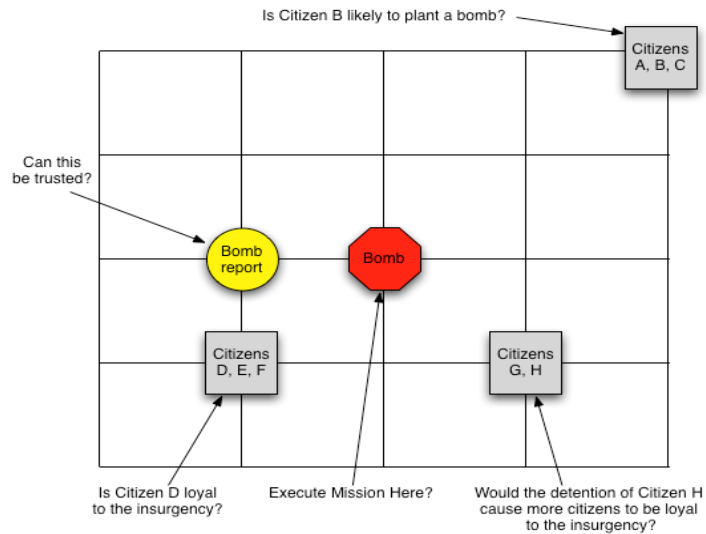
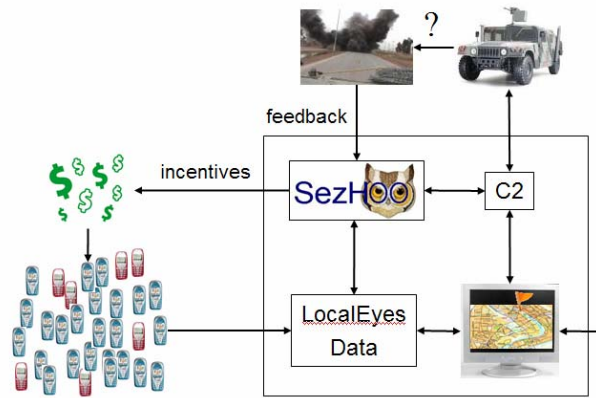


Impacts

- **Wider information sharing** with reduced need for “blind trust” via consumer knowledge of source reputation
- **More effective intelligence analysis** by evaluation of the reliability of information, analysts, and sources
- Additional **resistance to misinformation**, intentional or otherwise

Future Plans

Believability of HUMINT



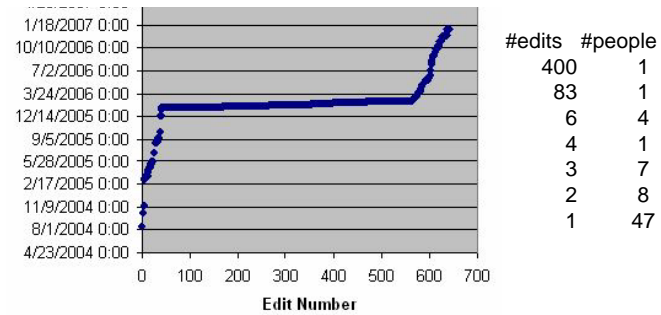
Believability of Wiki articles

Military Record

1000 visitors → In an SBVT television ad, Dr. Lewis Letson asserted "I know John Kerry is lying about his first Purple Heart because I treated him for that injury," but did not specify what the alleged lie was. Kerry's medical records list a medic, J. C. Garrean, as the "person administering treatment" for this wound. Dr. Letson's name does not appear on the record, but he claims that it was common for medics to sign the paperwork even though Letson would treat the patient. However, the claim cannot be verified as Garrean died in 1967. Kerry's citation for bravery under fire is false because neither Kerry's boat nor any of the others was under hostile fire. In a sworn affidavit about the incident, Thurlow testified, "I never heard a shot." [10] Of the three boat commanders present besides Kerry and Thurlow, two are SBVT members who now claim that there was no hostile fire during the incident. But one of them was seriously wounded with a concussion and the

1000 visitors →

20 changes



Consumer/Resource Variance Assignment

