An effective approach for accurate estimation of trust of distant information sources in the Semantic Web

T. G. Papaioannou, V. Bintzios, and G. D. Stamoulis

Networks Economics & Services Lab, Dept.of Informatics, Athens Univ. of Economics & Business (AUEB)
http://nes.aueb.gr

SecPerU'06, *June 2006 - Lyon, France*

Work partly supported by EuroNGI Network of Excellence (IST-2003-507613)

Overview

- Introduction and Problem Definition
- Related Work: End-to-End Approaches
- Our Approach: FACiLE
- Experimental Results
- Conclusions and Future Work

Introduction and Problem Definition

Introduction

- Trust metrics to access the trustworthiness of information sources' referrals are often employed
  - E.g. in the World Wide Web, mobile ad-hoc networks
- Direct own experience rarely suffices for sites visited only occasionally
- Accuracy of inferred trust for "distant" sources may considerably deteriorate due to
  - "noise"
  - the intervention of malicious nodes

  This gets worse with distance

In the Semantic Web…

- Transaction → exchange of information
- All referrals and query responses are information
- Objective: Access the trustworthiness of distant nodes
  - Our approach is based on path algebra …
  - yet, in a more effective way

Related Work

- Simple aggregation functions
  - E.g. sum +1, -1 votes
- Linear algebra
  - Matrix multiplication of direct trust values, probabilistic interpretation
- Path algebra:
  - Directed weighted graph, algebra on the path
- Multi-dimensional trust metrics, e.g. context factors, interests etc.

FACiLE vs. End-to-End Approaches

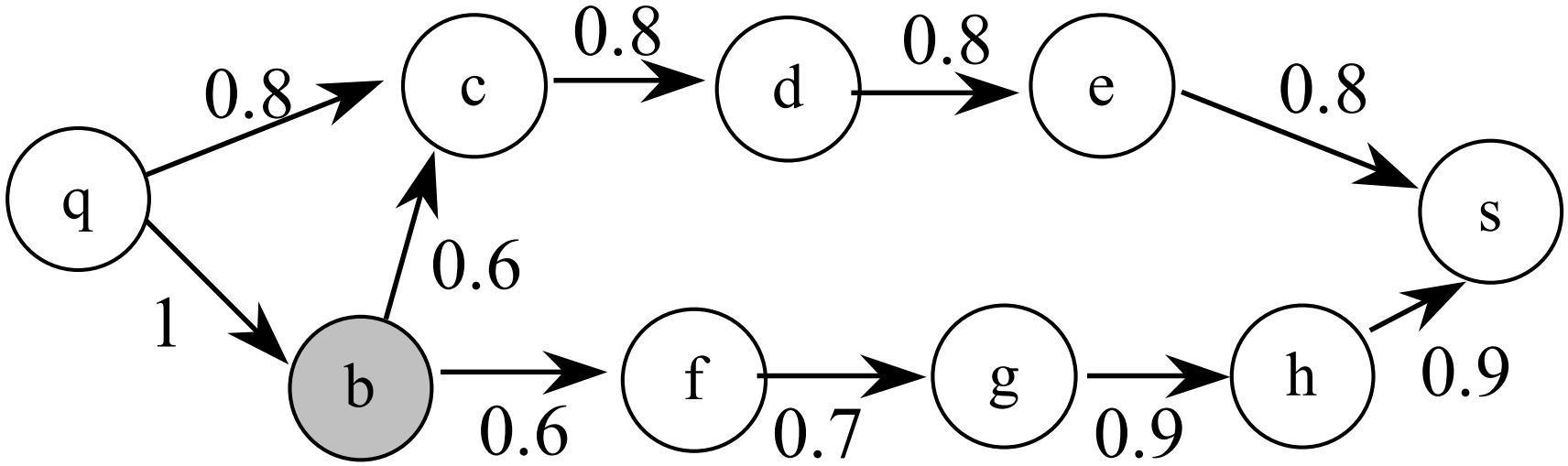
End-to-End Approaches

- Path algebra treats trust networks as directed weighted graphs
- Trust → a link's weight in range [0, 1]
  - Results from direct experience with the node
  - No link ⇔ unawareness of trust
- End-to-end trust inference
  - Find alternative paths terminating to information source
  - Concatenate trust values along path
    - max, min, harmonic mean
  - Aggregate calculated trust values along different paths
    - sum, average, max

The FACiLE Approach

- For trust inference on a distant node:
  1. Ask neighbors for their trust assessment
  2. and adopt them based on their own relative inferred trust
- Neighbor's trust is inferred based on
  - Concatenation and aggregation
- Innovative last step: combination
  - Direct trust values of to the distant node are combined based on their own inferred trust

Example

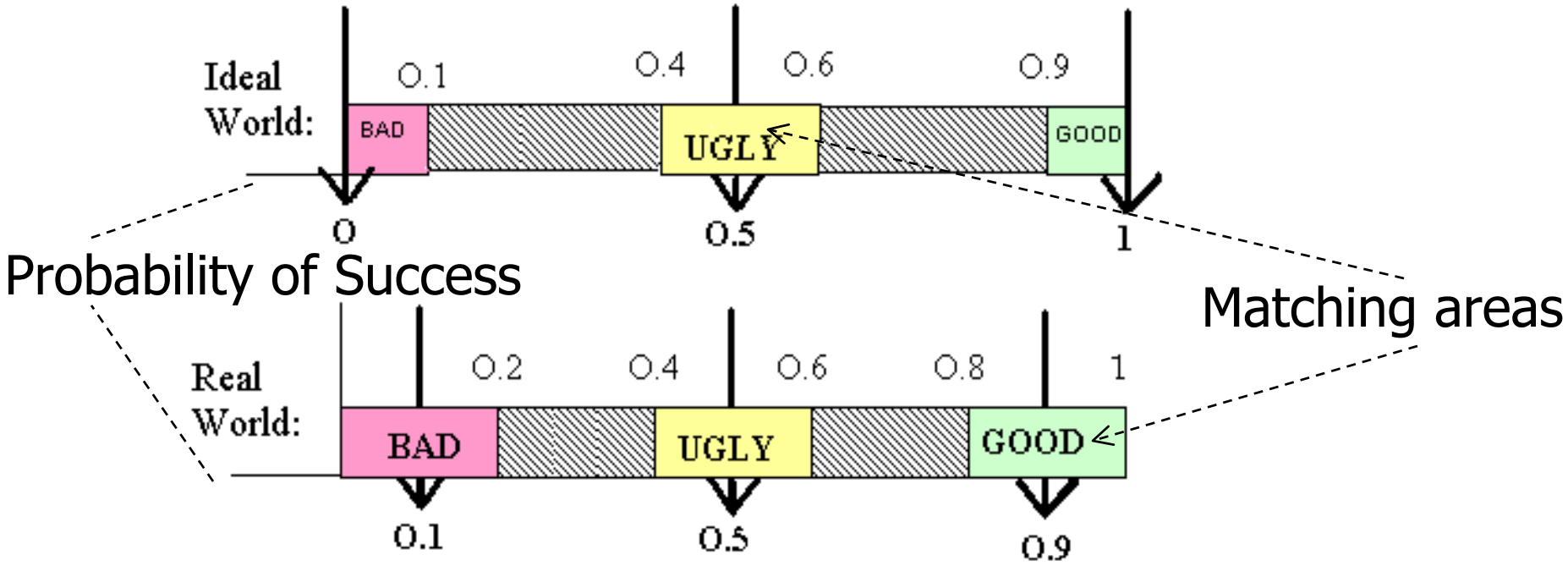


- End-to-end inferred trust from q to s
  - Maximum: 0.4096, Minimum: 0.3072
- FACiLE's inferred trust from q to s
  - 0.8 or above

Experimental Results

The Model (I)

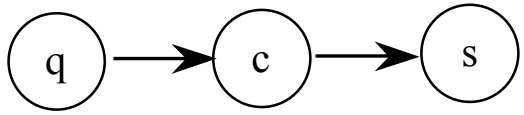
- 100-node power-law graph with some short-cuts ensuring small-word properties
- Three node types: “Good”, “Bad”, “Ugly”



The Model (II)

- Ideal- and Real-World models
- Ideal World: Good always honest, Bad always dishonest
- Real World: Inverted response with probability 0.1, or "Noisy" observation
  - Ugly give random response with probability 0.5
- Efficiency criteria: hit ratio
  - Count a "hit" if inferred trust matches true type

Operators Considered for Each Function (I)

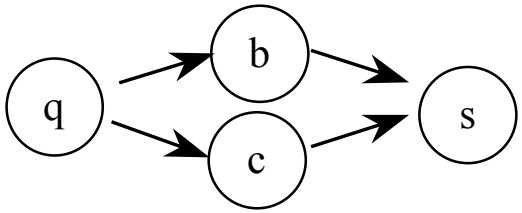


- Concatenation
  - Multiplication (MULTI): $t_{qs} = t_{qc} \cdot t_{cs}$
  - Harmonic Mean (HARM): $t_{qs} = \frac{t_{qc} \cdot t_{qc}}{t_{qc} + t_{cs}}$
  - Hybrid Mean (HYBRID): If $t_{qc} + t_{cs} < 1$ then HARM else MULTI

Operators Considered for Each Function (II)

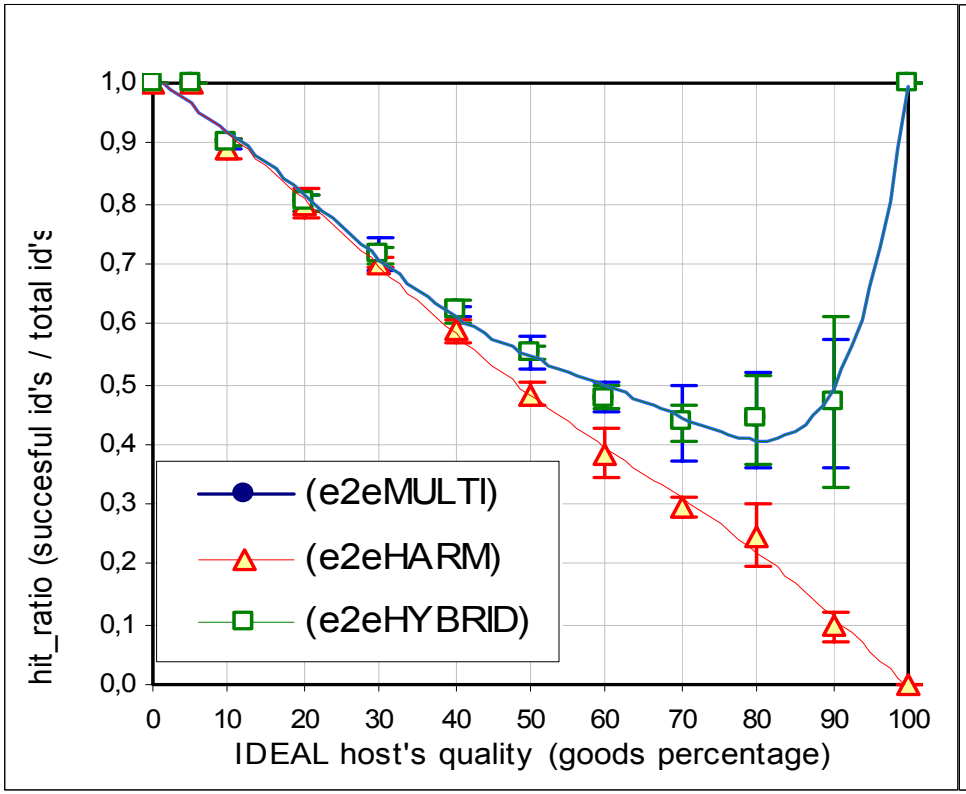
- Aggregation
  - Maximum: Path with max inferred trust

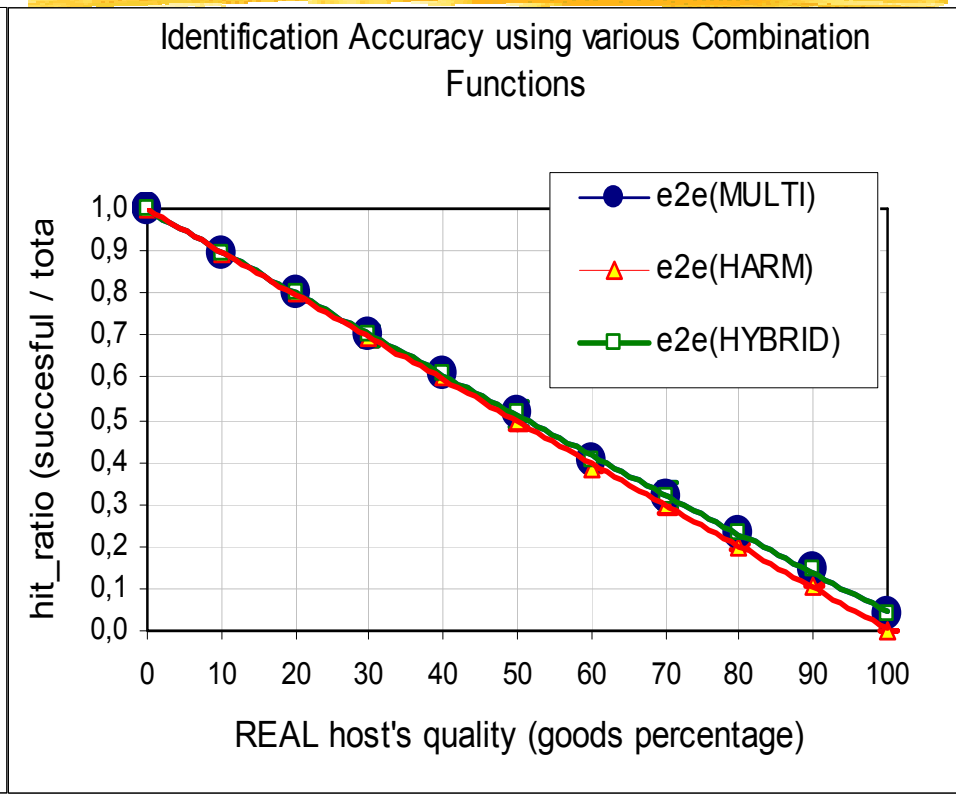
- Combination



  - Maximum(Max): if $t_{qb} > t_{qc}$ then $t_{qs} = t_{bs}$ else $t_{qs} = t_{cs}$
  - Weighted Average (WeiAvg): $$t_{qs} = \frac{t_{qb} \cdot t_{bs} + t_{qc} \cdot t_{cs}}{t_{qb} + t_{qc}}$$

Performance of End-to-End



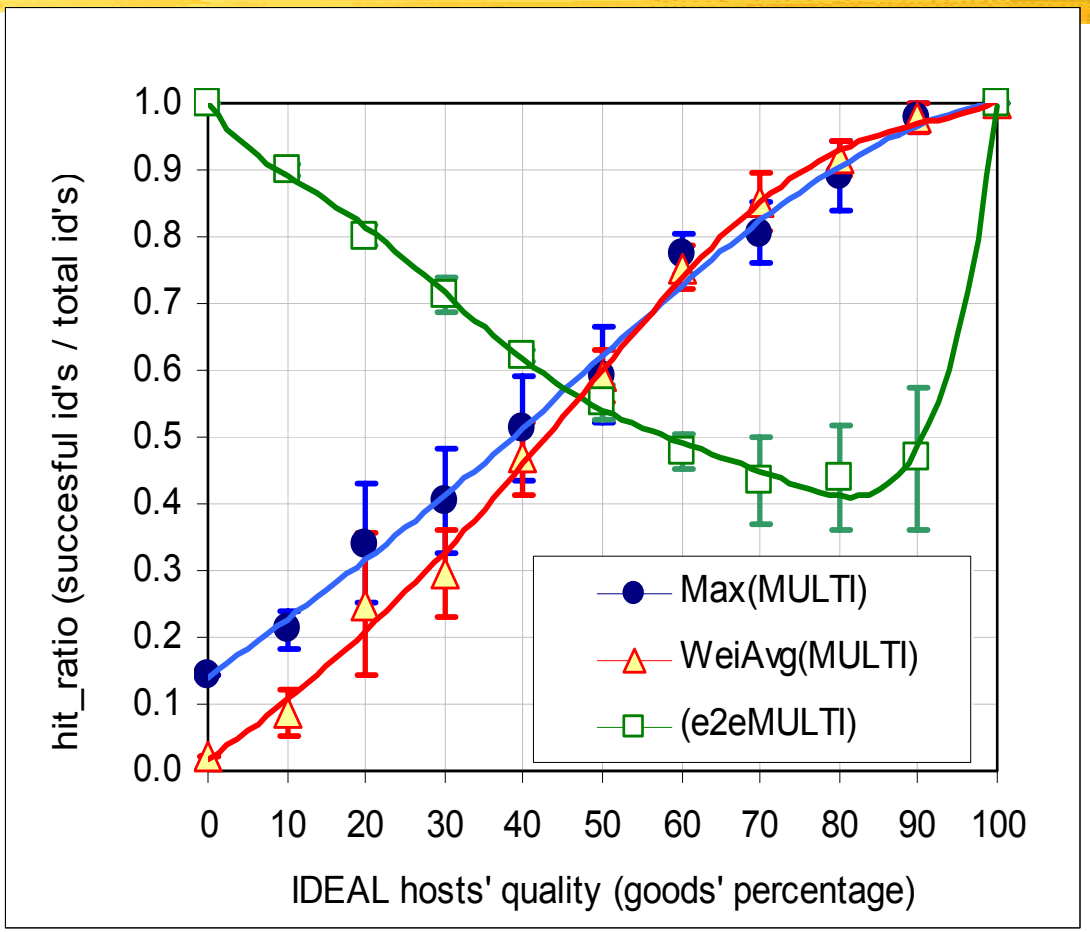


Ideal World

Real World

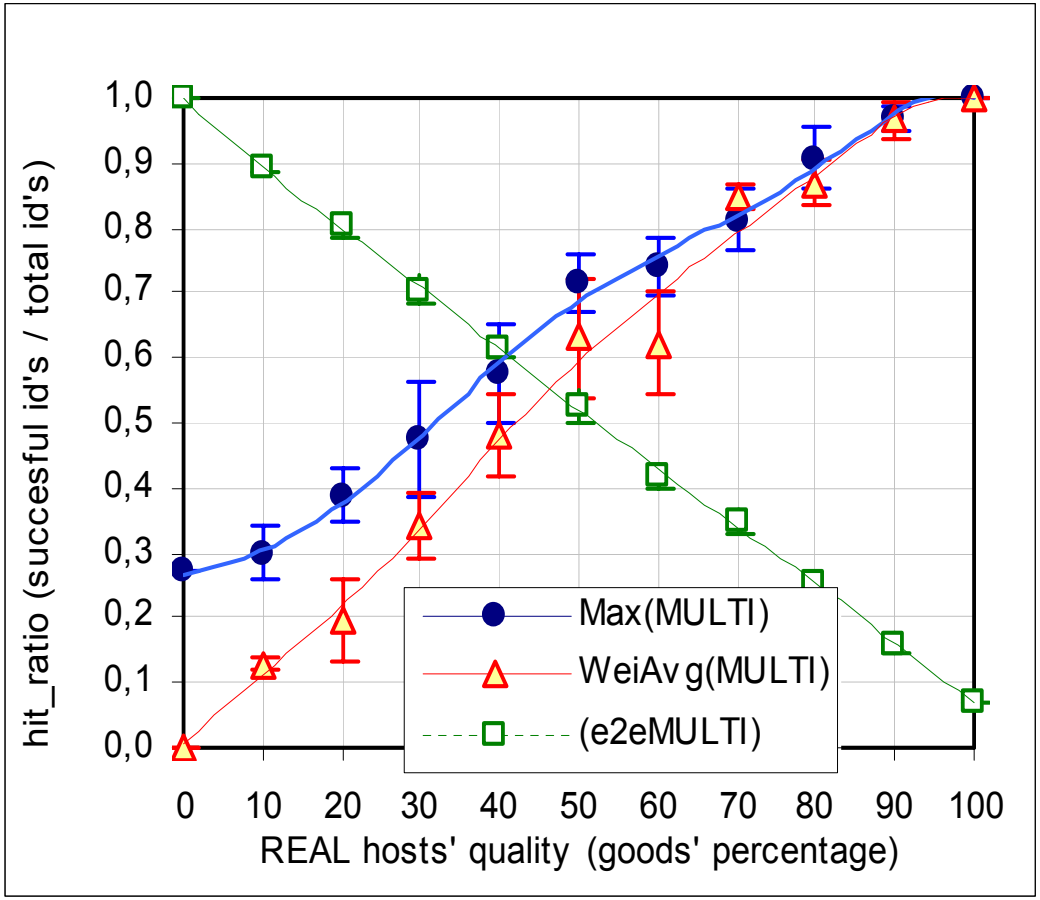
- Low performance in all interesting cases, i.e. "Good" nodes over 50%

Best End-to-End Combination vs. FACiLE in Ideal World



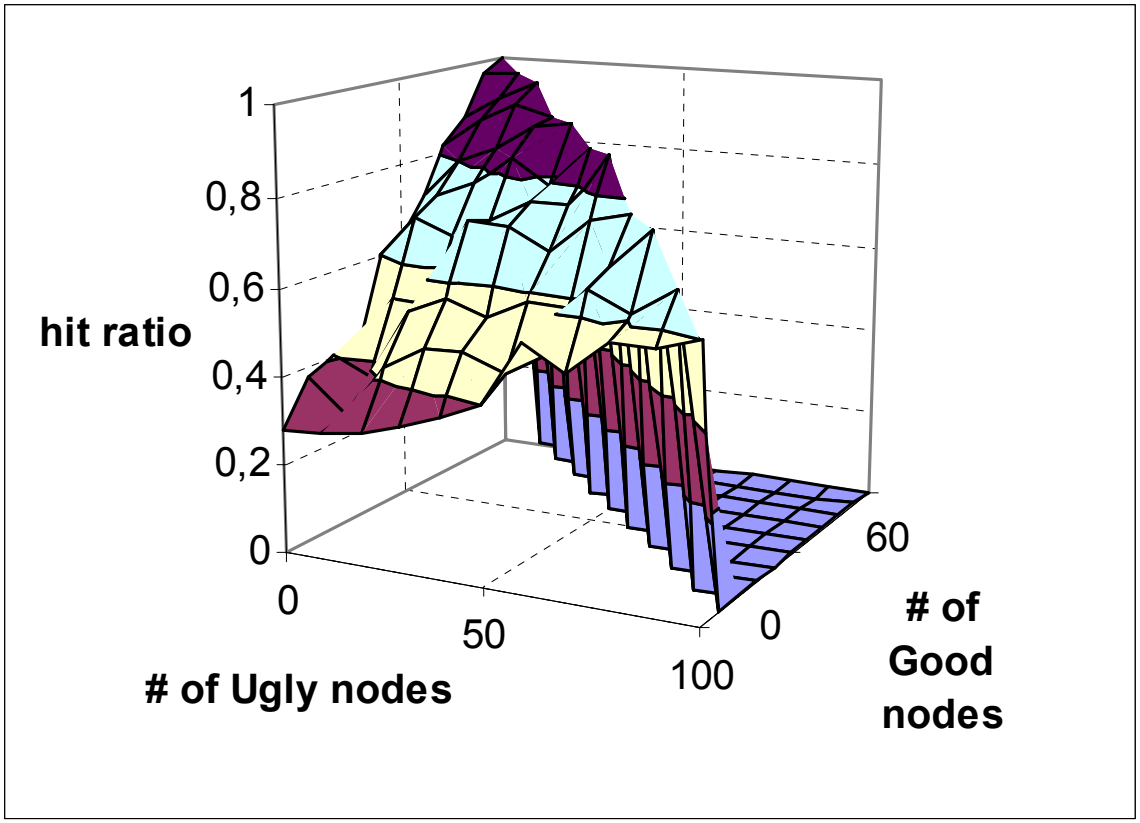
- WeiAvg and Max perform better than end-to-end approaches for all interesting cases

Best End-to-End Combination vs. FACiLE in Real World



- WeiAvg and Max perform better than end-to-end approaches for all interesting cases

FACiLE with “Ugly” Nodes too



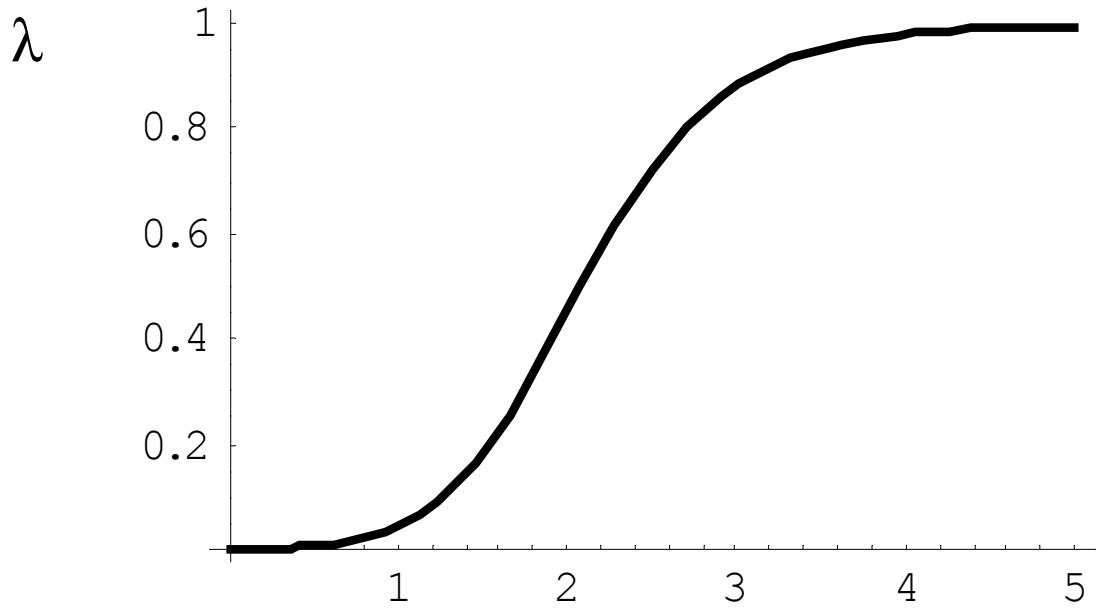
- MULTI concatenation, Max combination
- Real-world model

- FACiLE achieves high hit ratios, provided that Bad nodes are fewer than 50% of the system

Incorporate Direct Trust

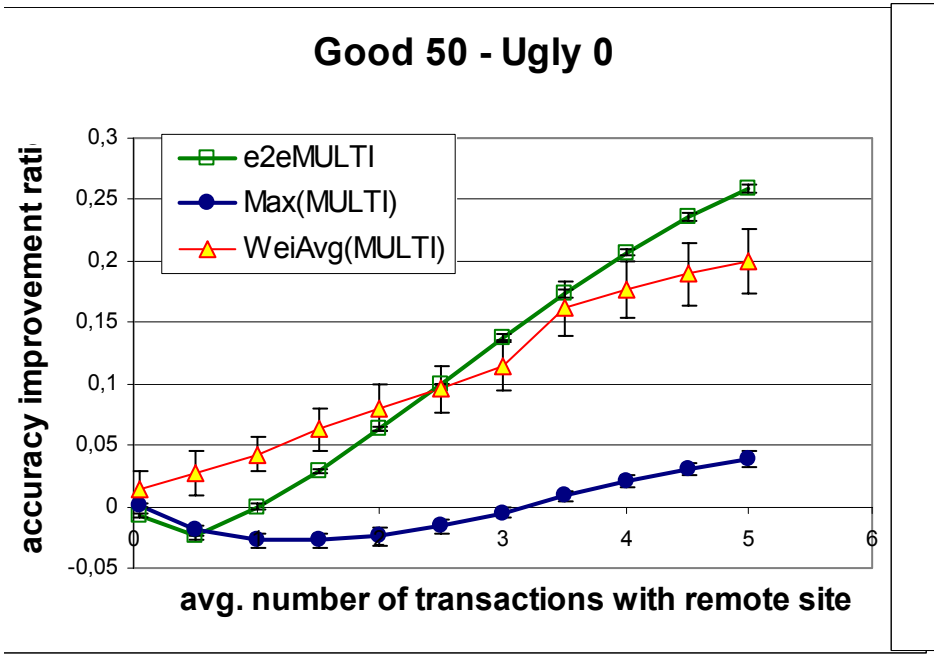
- Trust to distant nodes is given as weighted sum:

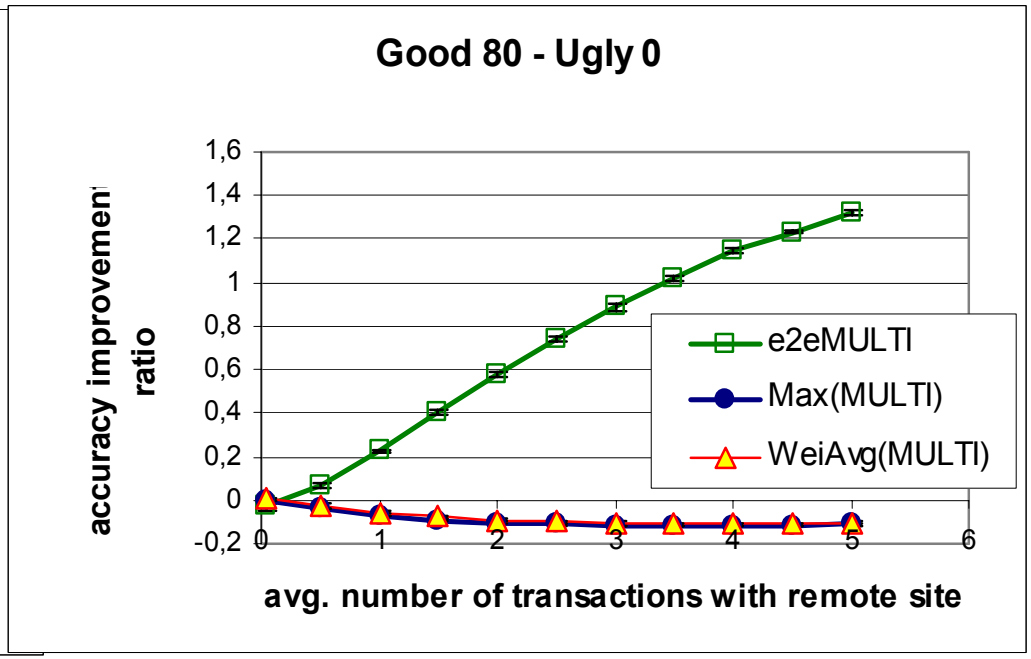
$$\lambda \cdot \text{direct\_trust} + (1-\lambda) \cdot \text{inferred\_trust}$$



Number of transactions
with the same distant node

Effect of Direct Trust to FACiLE





- Direct trust is beneficial for FACiLE only if the system has few "Good" nodes

Conclusions

Conclusions

- Developed a new approach for trust inference over occasionally visited nodes in the Semantic Web
- FACiLE reveals that referrals from "trustworthy" nodes "near" the target-node for trust inference are:
  - more informative, and
  - more resistant to "noise" and malicious collectives
- FACiLE has high hit ratios and performs better than end-to-end approaches in all interesting cases
  - I.e. systems with more trustworthy nodes
- FACiLE is as effective as direct trust for trust graphs with more than half of the nodes being trustworthy

Future Work

- Apply FACiLE to other specific contexts
  - e.g. mobile ad-hoc networks, grid
- Use different concatenation and aggregation operators
  - e.g. max-flow