भारतीय प्रौद्योगिकी संस्थान दिल्ली
Indian Institute of Technology Delhi

iitd CSE

A Distributed Trust-based Recommendation System on Social Networks

Karan Sarda

Priya Gupta

Debdoot Mukherjee

Smruti Padhy

With Prof. Huzur Saran

Motivation

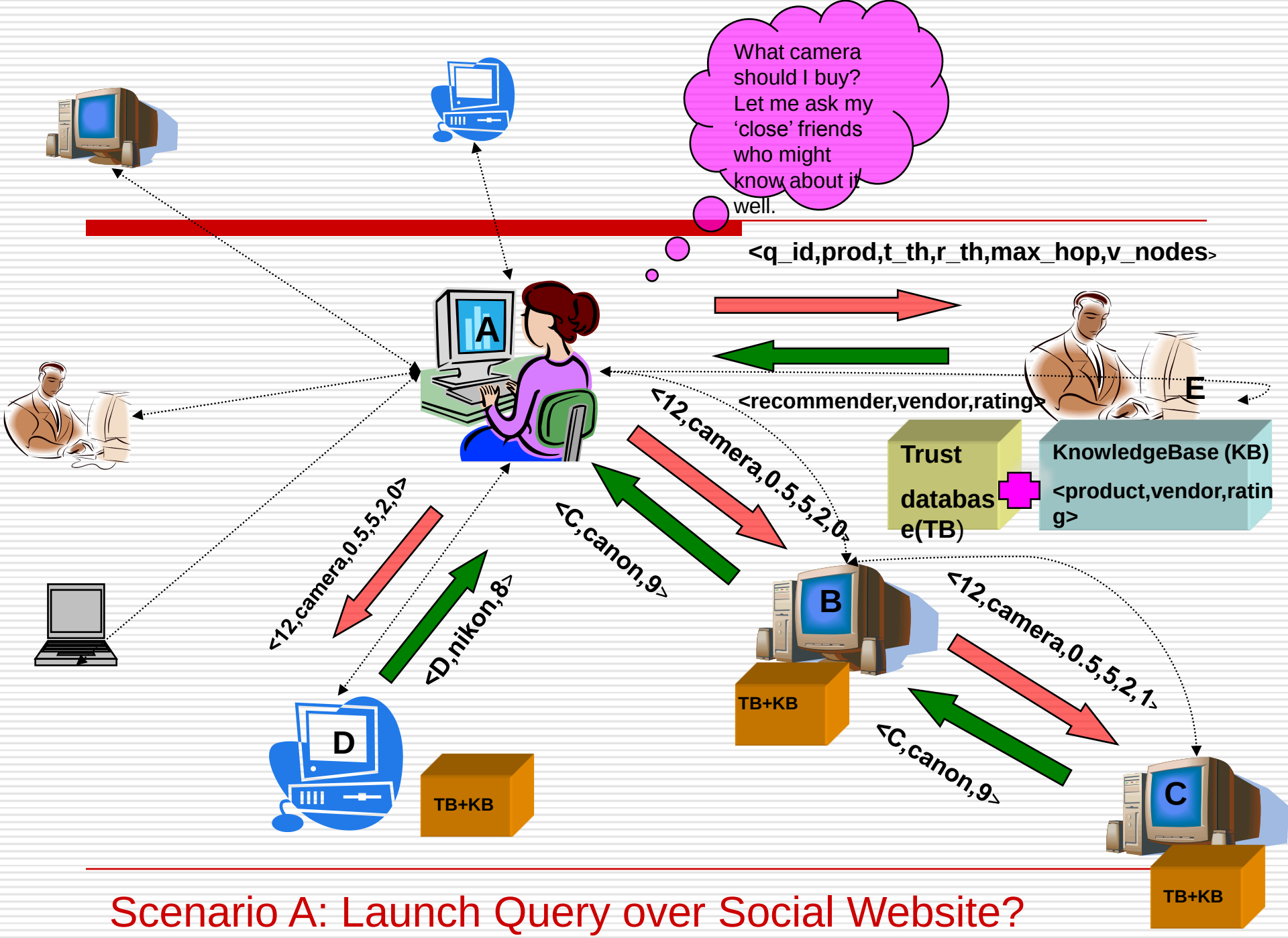
- Need recommendations from others in everyday life for buying goods and services
- For widely used commodity articles movies, books, phones – taking a vote of global opinion as in many review sites is sufficient
- Social network crucial for more specific articles – eg good car mechanic in the area, rely on recommendations coming from 'relevant' friends
  - *Trust* we have in our recommender friend
  - Relevance or domain expertise of the recommender

Trust

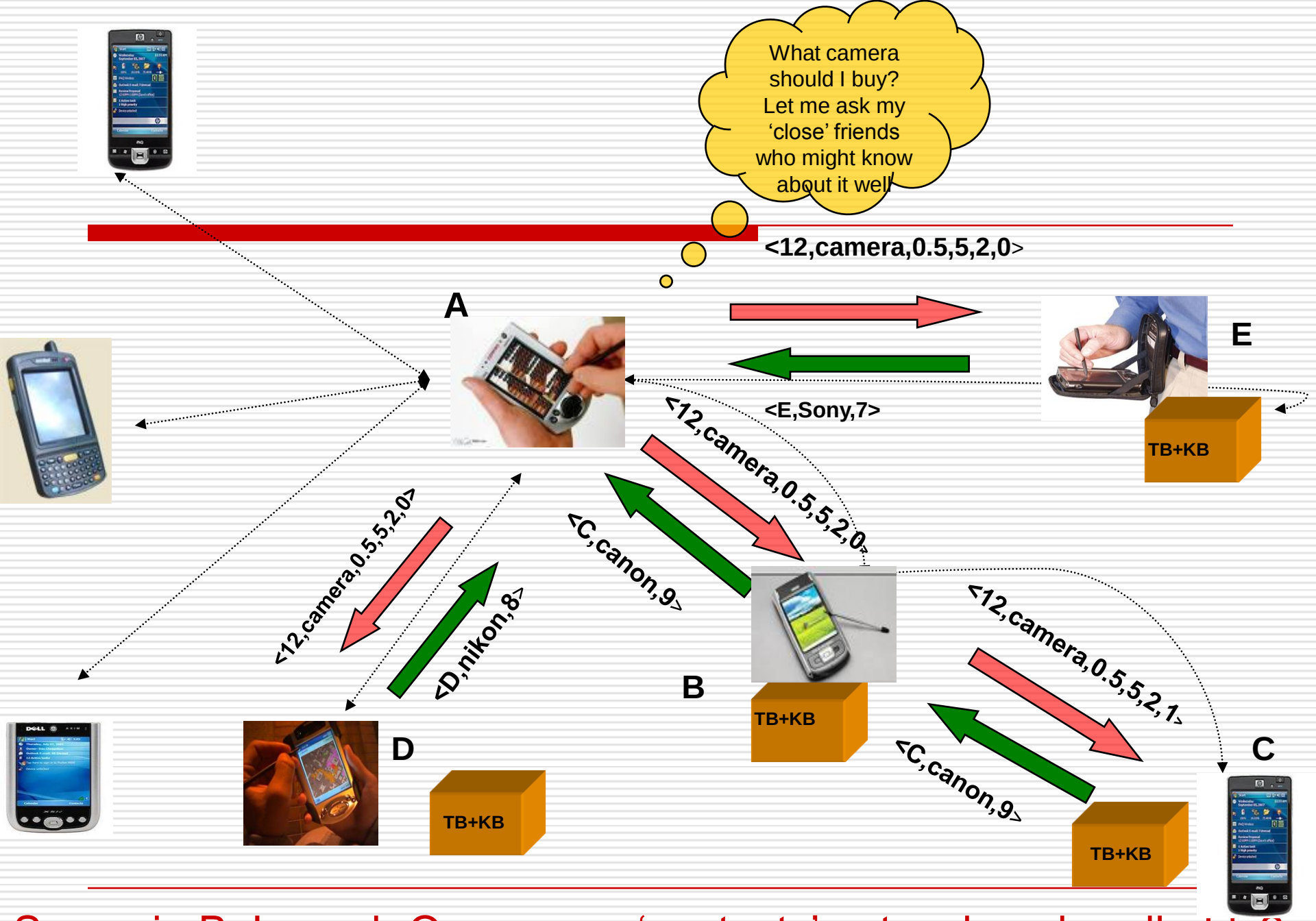
- Friendship Trust
- Domain-expertise Trust

Problem Statement

HOW TRUST MAY BE USED TO BUILD AN EFFECTIVE RECOMMENDATION SYSTEM OVER SOCIAL NETWORK?



Scenario A: Launch Query over Social Website?



Scenario B: Launch Query over 'contacts' network on handhelds?

Major Components of Our System

- Social Network with Friendship trust
  - Realized as a website or as a distributed network
- Trust computation between Non-adjacent nodes
- Knowledge Base (KB)
  - Stores prior history of service experiences
  - Can be leveraged to respond to queries automatically with minimal user interaction
- Query Propagation
  - Propagate to ‘selected’ friends
- Query Response Accumulation
  - **Possible to add feedback accumulation after service is utilized**

Distributed Local Trust Model

- Every user stores the trust values for every other user on the network whom it trusts above a certain threshold
- With a million nodes, every node requires a few MBs of storage space.
- A model to dynamically compute trust information is not encouraged
- A node polls its neighbors either periodically or on a change event

Computational Trust Model

We use the model Proposed by Golbeck (2005)

$$t_{is} = \frac{\sum_{j \in N(i)} \begin{bmatrix} (t_{js} * t_{ij}) & \text{if } t_{ij} \geq t_{js} \\ (t_{ij})^2 & \text{if } t_{ij} < t_{js} \end{bmatrix}}{\sum_{j=0}^{n} t_{ij}}$$

where,

t_{is} : Trust i has in non-adjacent node s

N(i): Neighbhors of i

Trail Levels - [To model domain specificity]

- History: links which have given good responses for a certain product are likely to do the same for `similar' products.
- Classify products into different product categories
- High trail level for a particular product-category on a link signifies that high quality responses on products of that class have been received through that link.

Query Propagation

- At each node when a query arrives:
  - If an entry is in its knowledge base for this service, with a satisfactory rating (>Rating threshold), query not propagated further
  - Response created for this entry and sent back.
  - Else if the number of hops done till now less than Maximum Hops, query propagated to friends based on
    - trust threshold
    - attractiveness of a link, $\eta = \beta * trust + (1-\beta) * trail$
    - β is 0.5 for evaluation
    - few random links from top N attractive links

Response Accumulation

- At each node at which a response arrives:
  - find out the trust between this node and the recommender
  - calculate the transient score of this response using the following formula, $score = trust^{\gamma} rating^{\delta}$
  - Update the trail levels using $new\_trail = \rho * score + (1-\rho) * old\_trail$
  - γ and δ decide the relative importance of trust and rating. Currently, we have set both the values to 1.0, so that score is just the product of trust and rating. ρ is weight given to the present response against past response and we set its value to be 0.7.
  - Send only one response – the one with max score

Initial Evaluation of Approach

Limitations:

* Lack of adequate real world data
* Anecdotal evidence
* Build a model using available data
    Orkut
    A survey to get an initial trust distribution

Hi Rose, Could you please tell me good shop nearby for buying a camera and also the brand?

Hi Alice, I do not have much idea about it. Let me ask my friend Bob. He is a professional photographer .

Hi Anny, Could you please tell me good shop nearby for buying a camera and also the brand?

Hello Alice, my friend , Bob has suggested Nikon from Priceless Showroom.

Hi Bob, could you suggest me a good camera and the showroom to buy?

Alice

Rose

Hi Alice, there is Best Buy showroom, you can try that one and I have very good experience with Canon camera.

Hello Rose, from my experience, I would recommend Nikon and from Priceless showroom.

Anny

Bob

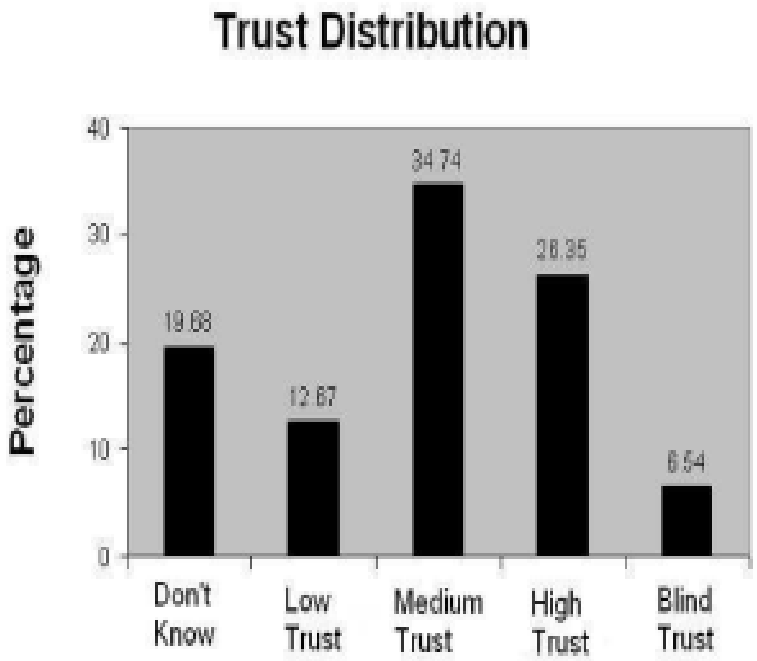
Anecdotal Evidence for our proposed recommendation system framework

Social Network Structure

- Crawled about 200,000 users ~ 0.3% of Orkut population
- Snowball sampling: BFS from random seed
- Removed leaf nodes which hadn't been crawled
- Network of ~10,000 nodes
- Overestimates node degree, but preserves other characteristics

Initial Trust Distribution

- Distribution from trust survey:
  - Blind Trust: 6.54%
  - High Trust: 26.35%
  - Medium Trust: 34.74%
  - Low Trust: 12.67%
  - Don't Know: 19.68%



Knowledge Base

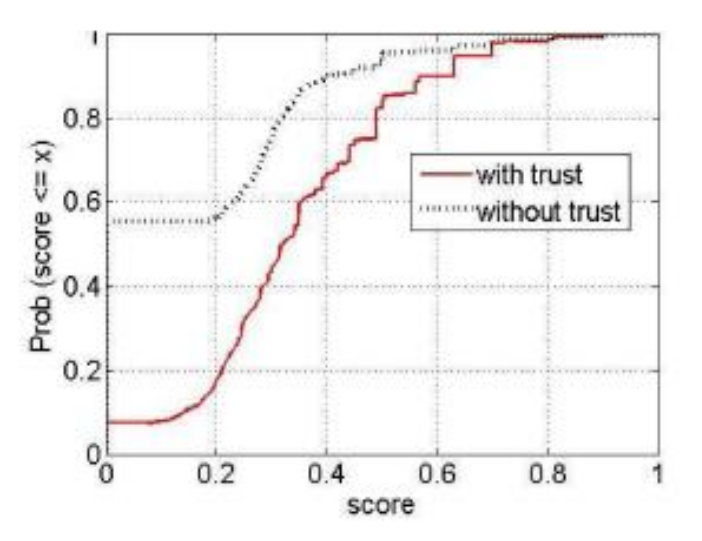
- A service ontology tree with 8000 products having 4 levels
- Tree pruned to include only 10 *product categories* at the highest level
- A service ontology tree with 10 categories at level 1, 60 product-classes at level 2, and about 240 products at the 3rd level

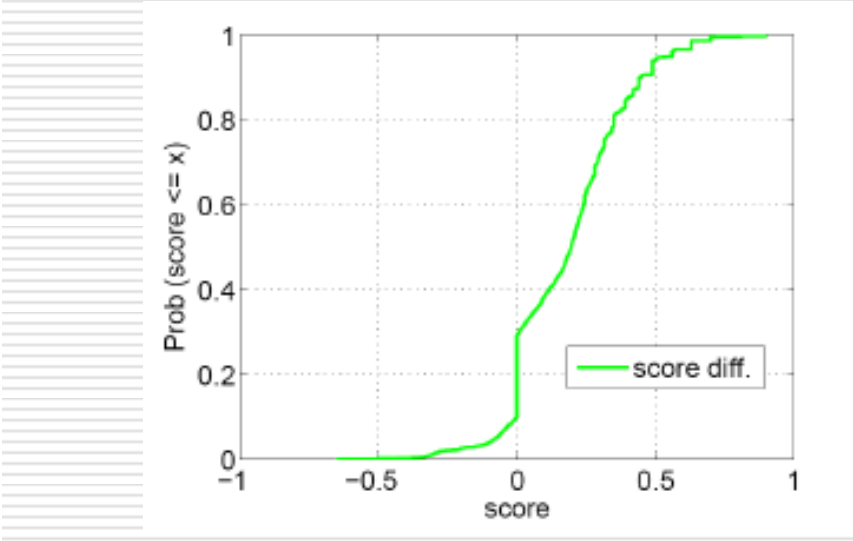
www.ksl.stanford.edu/projects/DAML/UNSPSC.daml

Vendors and Ratings

- 100 vendors assigned to each of the 10 service classes
- Each service assigned a random number of vendors (between 5 and 15)
- This completes the database of possible services and their vendors
- A random number of <service, vendor, rating> tuples (between 20 and 30) at each node, ratings chosen randomly

Comparison with no-trust system





Parameter values used in Simulation

Parameter	Value
Trust Threshold	**0.27**
Rating Threshold	**0.3**
Maximum Hops	**4**
Maximum propagation	**20**

Observations

- Original System
  - 75% scores lie in [0.2, 0.6]
  - Mean score = 0.355 , Std Dev. = 0.175
- Without trust system
  - 50% of scores are 0
  - 45% lie in [0.2, 0.6]
  - Mean score = 0.162, Std Dev. = 0.198

Modeling/Implementation Issues

- Quantifying Interpersonal trust and obtaining initial trust values
- Estimation of transitive trust between two individuals who are not directly connected on a social network
- Efficient and effective query propagation mechanism
- System needed to work even when some nodes on the network are down.
- Bootstrapping of the system.
- Needed more queries for system to learn and stabilize

Discussions

- Our proposed framework can be implemented on Mobile Ad hoc network, P2P network with the formation of social networks on these networks.
- The notion of distrust and its impact in our proposed framework may give interesting insight as it may affect the transitive trust computation. However, it does not impact on the response received since queries will not be routed on 'dis-trusted' path.
- The notion of trail levels provides simplistic and elegant way of query propagation.
- Observation of trail levels may give useful insight into the communities formation in social network.

Limitations and Future Work

- Validation of proposed framework needs set up of online social network with friendship trust requiring large user participation over time.
- In current work, we have used a simplistic and static Knowledge base to validate our model. Exploring dynamic knowledge base which is based on feedback mechanism would be interesting.
- As a future work, we aim to incorporate the dynamism in friendship trust value.
- Require Sensitivity analysis of system with respect to trust thresholds and other simulation parameters. Search for better performance metric to analyze the system.

Conclusion

- Empirically validated that trust can significantly enhance the utility of a recommendation system,
- Pave way for personalized search
- Used a dual notion of trust for query propagation:
  - *friendship-trust:* general belief in a friend's recommendations
  - Trail levels:- context-specific domain-expertise based on history of queries.
- Simulation of a large social network and knowledge base helps in bringing the model closer to real life

Thank You

References

- Walter Frank E., Battiston Stefano, Schweitzer Frank, A Model of a Trust-based Recommendation System on a Social Network, Nov. 2006.
- Marsh, S. Formalising Trust as a Computational Concept. PhD thesis, University of Stirling, 1994.
- Jennifer Golbeck, Computing and Applying Trust in Web-Based Social Networks, Doctoral Thesis, 2005
- Abdul-Rahman Alfarez, Hailes Stephen, Supporting trust in virtual communities, Proceedings of the 33rd Hawaii International Conference on System Sciences – 2000
- Gambetta D. Can We Trust Trust?. In, Trust: Making and Breaking Cooperative Relations, Basil Blackwell. Oxford, 1990
- Massa Paolo and Bhattacharjee Bobby, Using Trust in Recommender Systems: An Experimental Analysis, In Proceedings of the Second International Conference on Trust Management (ITRUST 2004)
- http://genie.iitd.ernet.in/trust/welcome.html
- http://www.ksl.stanford.edu/projects/DAML/UNSPSC.daml

Related Work

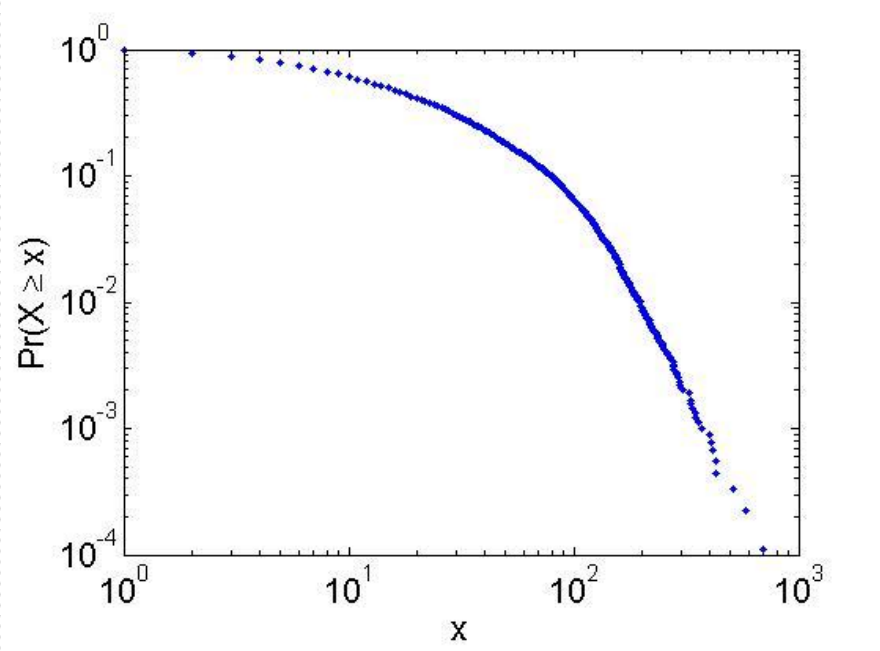
- Trust in social science (Morton Deutsch 1958)
- Marsh's formalization (1994)
- Advagato's maximum flow trust metric
- Appleseed Algorithm by Ziegler and Lausen (2005), trust/distrust propagation model by Guha et al.(2004), simple multiplication trust model by Walter et al. (2006)
- Goldbeck model (2005)

Related Work: Trust and Reputation

- Reputation is the global trust acquired by a user on the basis of her past behaviour and independent of interpersonal faith or belief
- Recommendation Systems– by O'Donovon(2005), Montaner et al(2002), ebay.com uses reputation scheme.
- Walter et al (2006) uses social network information in recommendation system.
  - Trust used in their system is domain-expertise in our system
  - Their work considered limited number of agents(or nodes) and considered random directed graph as underlying graph whereas we work on real online social network, Orkut.

Analysis of Network Structure

- **Power-law degree distribution**
- $P(x) \sim x^{\gamma}$
- $\gamma \sim 3.5$ for degree < 100
- $\gamma \sim 1.8$ for degree > 100
- Mean degree ~ 30
- Agree with other literature [Ahn 2007]



x = node degree

Trust

Trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or a group will perform a particular action, both before he can monitor such action (or independently of his capacity of ever to be able to monitor it) and in a context in which it affects his own action.

(Gambetta 1988)

Query Propagation

Algorithm 1 Query Propagation

```
if product ∈ KB_C with vendor V and rating r (> rating_thresh) then
    Send response as ⟨C, V, r⟩ to B
else if length(visited_nodes) < max_hops then
    for all friends X of C s.t. trust_CX > trust_thresh do
        Compute η_X = β * trust_CX + (1 − β) * trail_CX
    end for
    Add C to visited_nodes
    Send query to few random links from top N in order of η
end if
```

Query Response Accumulation

Algorithm 2 Response Retrace

1: Compute $score = trust_{BY}{}^{\gamma} \ * \ rating^{\delta}$
2: Compute $new\_trail = \rho * score \ + \ (1-\rho) * old\_trail$
3: Update trail level for link B→C to new_trail
4: Send response to A, who had sent the query to B

Variation with β and N