Routing in Question & Answer Networks

Simon Fleming

Foundations of Software Systems School of Informatics University of Sussex

Multi-Service Networks 2010 Abingdon, Oxfordshire, England 08th July 2010





Introduction: Q&A

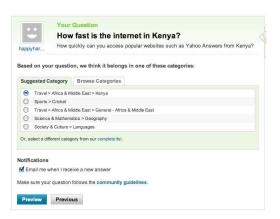
- ▶ We all need help from time to time...work, life and play
- ▶ What is electronic question and answering (Q&A)?
- Exploit 'the wisdom of the crowds'
- ▶ Contextual, subjective, opinions and advice √





Existing Q&A Technologies: issues?

- Identity and accounts privacy?
- ► Knowledge Markets public search
- ► Human Attention!! resource to optimise
- Centralized bottleneck, failure, control and ownership







My research: overview

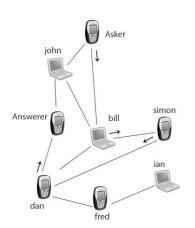
- Distributed question and answer service
- ▶ Q&A over ad-hoc networks: mobiles, laptop, access points...
- Decentralized lower requirements on single nodes
- Investigation through simulation





My research: aims

- Routing strategies to reduce the attention required from network users to get satisfactory answers.
- Improve privacy though plausible deniability
- ► A robust/usable and effective Q&A service







My research: tactics

- ► Swarm intelligence: stigmergic approach dynamic networks
- Strengthening/reinforce links to desirable network members (experts)
- Reward good behaviour, punish bad behaviour, prevent 'bombardment'
- Experiments comparing stigmergic against flooding and random approaches



Questions/ Answers act as simulated ants.

Network nodes act as path choices

Simulated pheromones aid path selection.

Feedback?



simulation: simple user model

- ► Yahoo! answers database: distributions, facts & figures
 - ➤ Yahoo! Webscope Datasets Catalog (L6) Yahoo! Answers Comprehensive Questions & Answers version 1.0 http://www.stanford.edu/class/cs345a/YahooData.pdf
- Range of interest categories
- Users answer questions which match interests
- Priority based question queues
- Markov model (attention / idle)
- Question answering monkeys!!





Simulation: specifics

- ▶ **Answer quality** : how good our the users?
 - Best answer counts distribution
 - example: "Computers & Internet" (0.78...) 37 out of 47

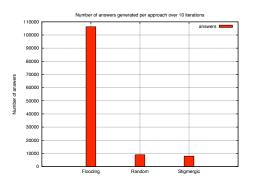
 - Dynamic ranking naive
- Fixed question asking probability
- ▶ Feedback messages are used with the stigmergic approach.





Preliminary Results

- Comparison show stigmergic/random approaches will dramatically reduce required attention in comparison to flooding (number of answers)
- ► Attention consumed by: reading, thinking(*) and writing(*) responses per user. f(number of answers)







Number of answers per question over 10 iterations 6000 Flooding Random Stigmergic 5000 4000 Number of observations 3000 2000 1000

Number of answers

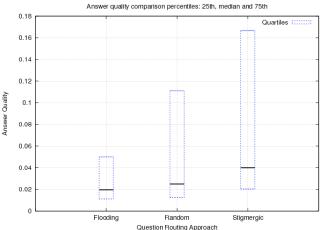
3



10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

Preliminary Results

▶ Answer quality is improved with a stigmergic approach.





Conclusion & Future Work

- Stigmergy helps to locate network *experts*.
- ...while reducing user attention
- ... while improving privacy through plausible deniability
- ► → Improve user model, answer quality assessment, network realism and fine tune stigmergic routing approaches





Fin.

Thank you kindly for listening! =) questions?



References

- PlanetSim: Object Oriented Simulation Framework for Overlay Networks: http://projects-deim.urv.cat/trac/planetsim/
- ► Yahoo! Webscope Datasets Catalog (L6): http://www.stanford.edu/class/cs345a/YahooData.pdf



