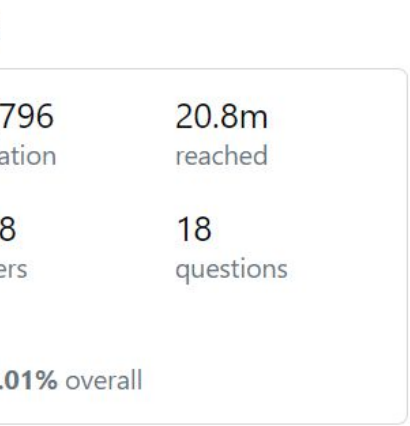
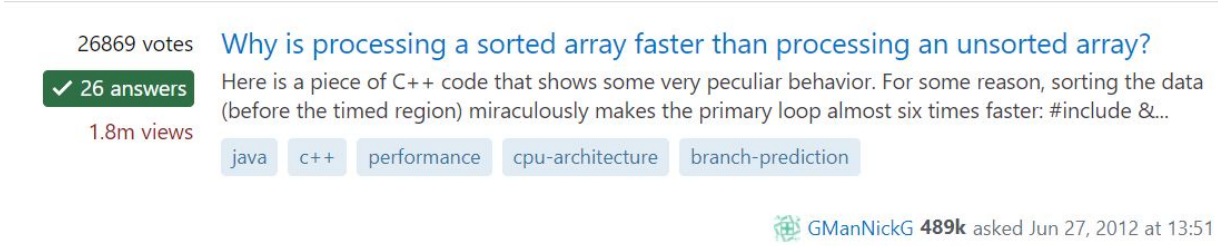
Simulating Stackoverflow Interactions using ABM

Stackoverflow

- Open community and Q&A platform
- Covers many computer programming topics
- Ask/answer questions, upvote
- Getting upvotes increases reputation
- Ranking based on reputation



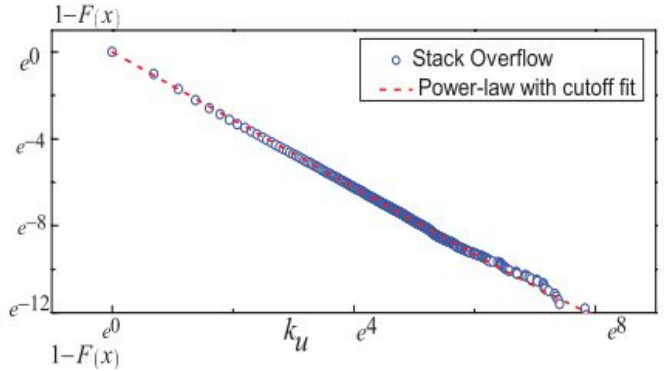


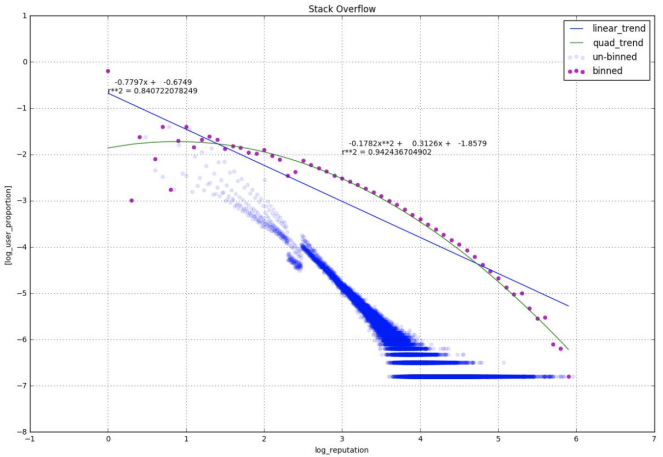
Emerging behaviour

[1][2]

- **75%** of users only ask one question, **65%** only answer one question and only **8%** of users answer more than 5 questions
- 0.46% with a reputation greater than 5000
- Most questions are answered by few people

1) The distribution of the number of upvotes for both questions and answers
2) The distribution of the reputation





ABM to Model Stackoverflow

Why is ABM appropriate?

1) The interaction between agents are **complex** and active:
   - Asking
   - Answering
   - Upvoting

2) Agents are **heterogeneous** in their activeness and their interests, and their interaction with other agents

3) Complex behaviour (feedback mechanism based on **local interaction**)

Entities and Attributes

User

- P_{ask}
- P_{answer}
- P_{upvote}
- P_{active}
- Tag
- Reputation
- Upvote bias

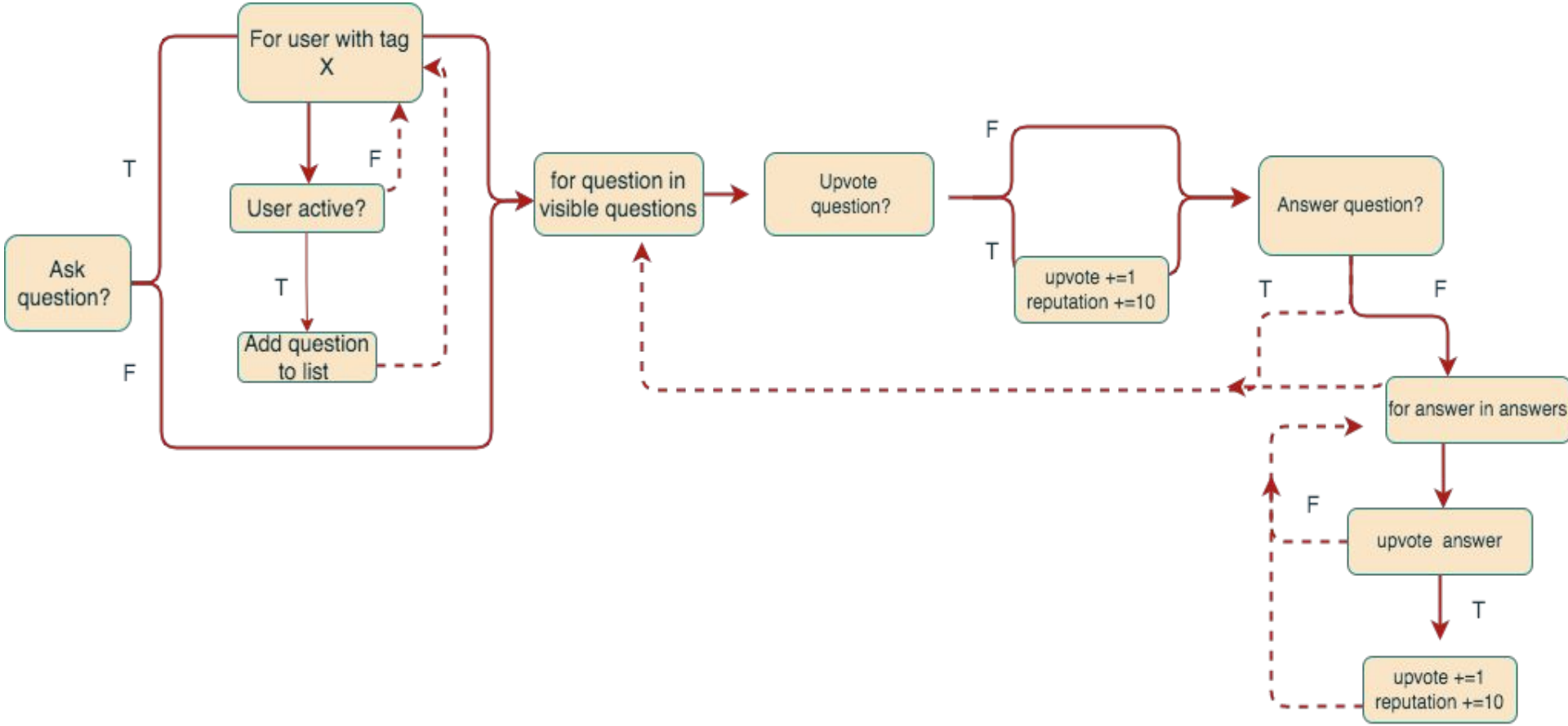
Question

- Asker
- Tag
- List of answers
- # upvotes

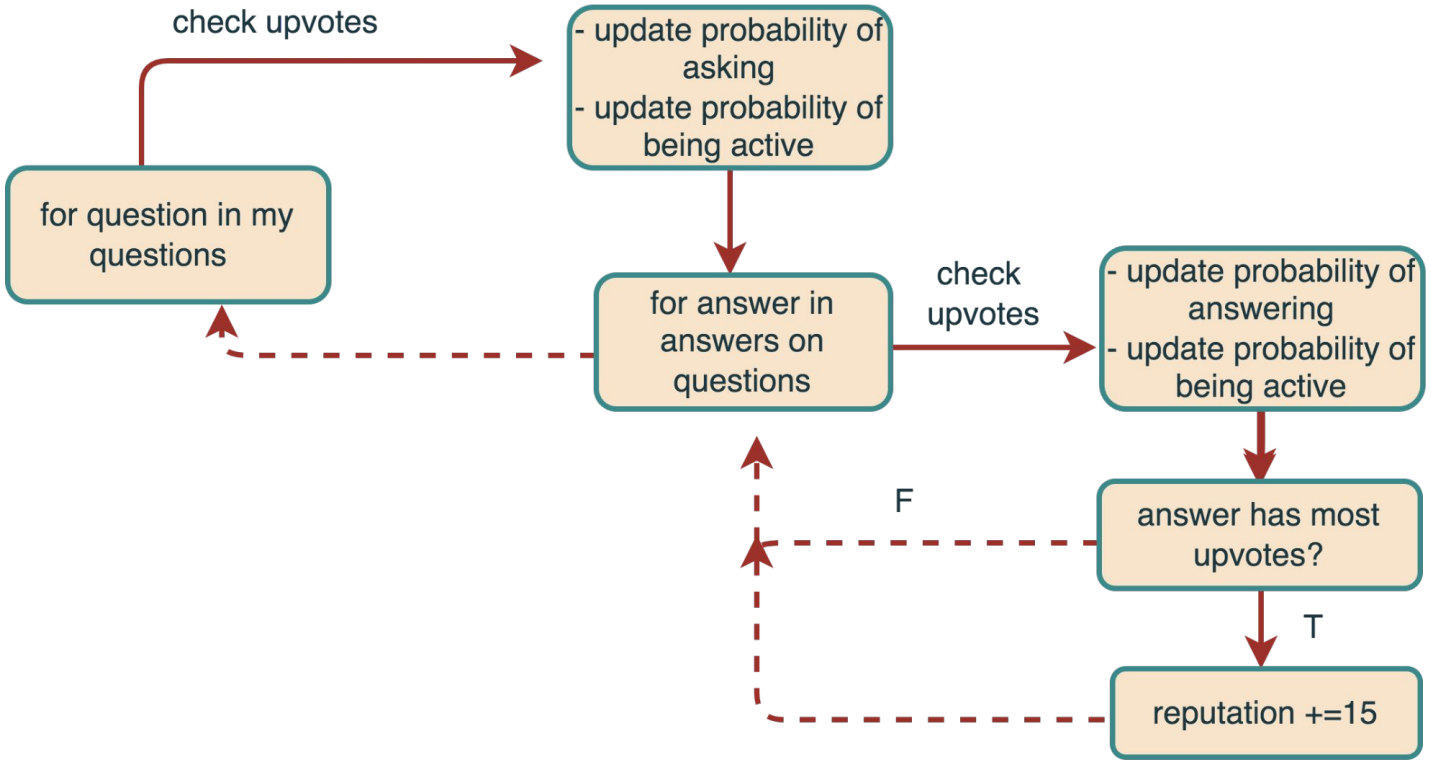
Answer

- Responder
- # upvotes

Proposed Model



Feedback Mechanism



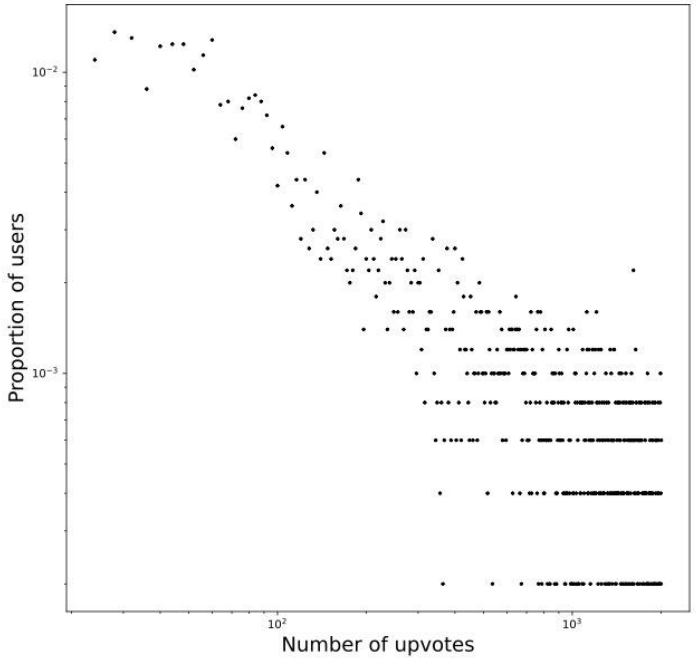
Experiment

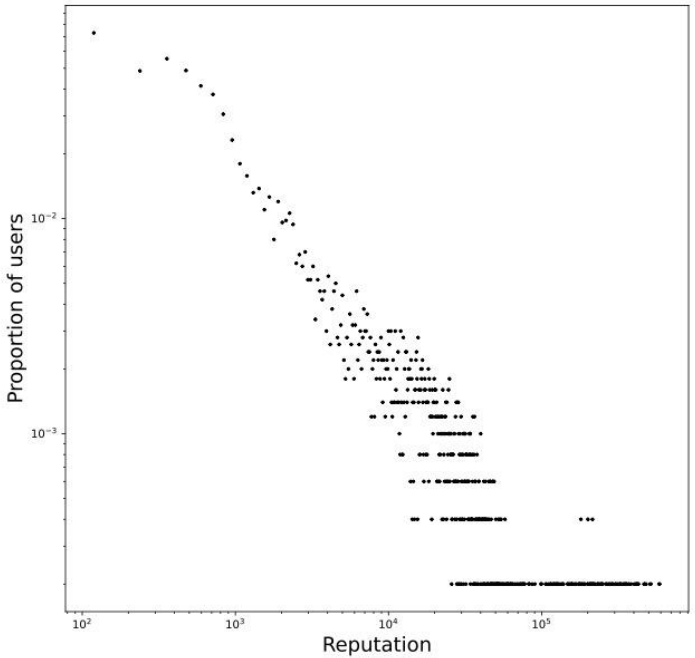
1. 3 Probabilities + basic stackoverflow dynamics

2. Add Feedback mechanism

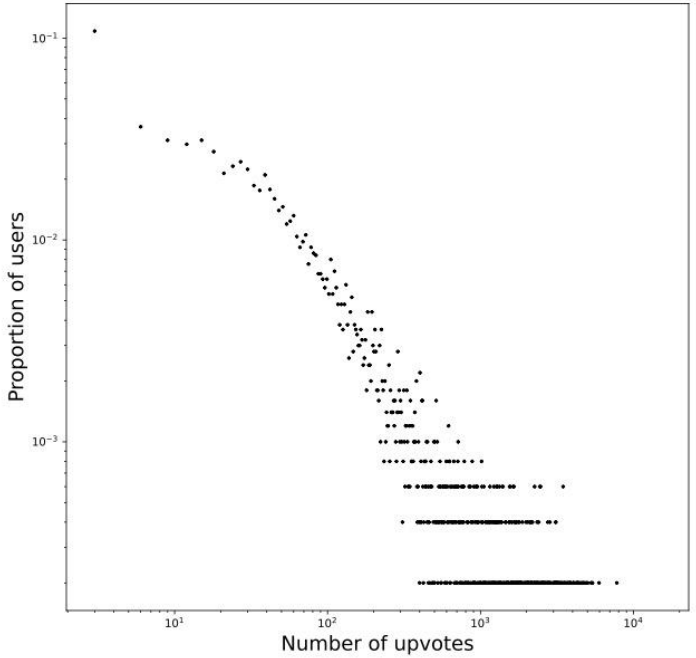
3. Add the probability of being active

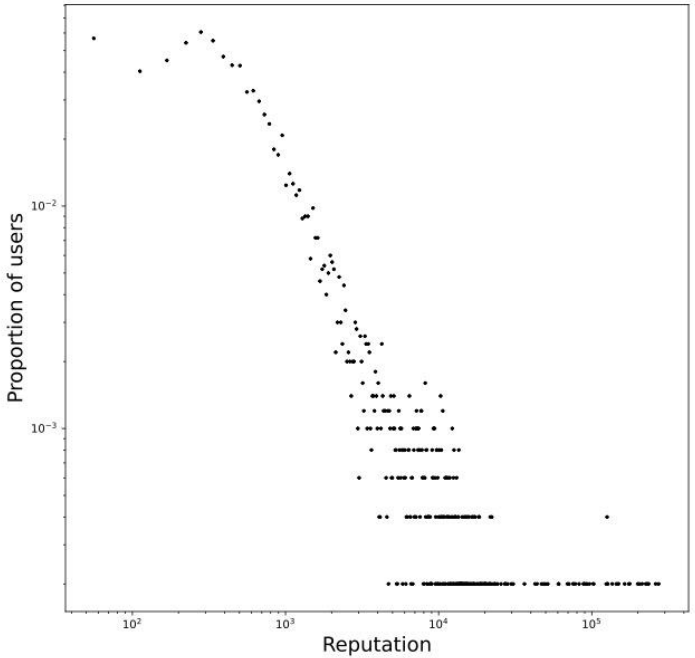
Result 1: Using normal distributions





Result 1: Exponential distribution for P_{upvote}





Reference

Lu, X. Y., Lin, J. H., Guo, Q., & Liu, J. G. (2015). Empirical Analysis of the Online Rating Systems. *arXiv preprint arXiv*:1510.08142.