SocialQ&A: An Online Social Network Based Question and Answer System

Haiying Shen *, Guoxin Liu*, Nikhil Vithlani†
*Department of Electrical and Computer Engineering, Clemson University, Clemson, USA
†Amazon.com, Seattle, USA

Presenter: Guoxin Liu
PhD candidate
Outline

• Introduction
• Background
• Social Q&A
• Results and Analysis
• Conclusion and Future work
Introduction

- Important source of information
  - Yahoo! Answers – 10 million users in first 2 years
  - Currently 200 million users
  - 15 million visits everyday
Introduction

- **Scope for improvement**

Red: Answered  Green: Unanswered  Black: Unsatisfied

**Yahoo! Answers**

**Baidu Zhidao**
Introduction

Quality of answers

Wait Time
Introduction

Contributions of this paper

- Design of SocialQ&A
- Implementation of SocialQ&A
- Trace-driven experiments and real usage of SocialQ&A
Outline

- Introduction
- Background
- Social Q&A
- Results and Analysis
- Conclusion and Future Work
Background

History of search engines

- Manual Indexing
- Automatic Indexing
- Google
Background

Information retrieval paradigms

Library paradigm

Village Paradigm
Background

Q&A systems: Evolution

- Community based Question & Answer systems
- Online social based Question & Answer systems
Outline

• Introduction
• Background
• SocialQ&A
• Results and Analysis
• Conclusion and Future work
SocialQ&A

Components of SocialQ&A

- The user interface
- User Interest Analyzer
- Question Categorizer
- Question-User Mapper
SocialQ&A

Core algorithms

- User Interest Analyzer
- Question Categorizer
- Question-User Mapper
User Interest Algorithm

Analyzes the user profile and associates each user with the user interest vector
User Interest Algorithm

Tokenization

I love country music
User Interest Algorithm

Matching and association

Tokens
- love
- music
- country

Synset
- rock
- metal
- country
- reggae

User X’s interest vector
# User Interest Algorithm

<table>
<thead>
<tr>
<th></th>
<th>Rock</th>
<th>Metal</th>
<th>Country</th>
<th>Reggae</th>
</tr>
</thead>
<tbody>
<tr>
<td>User x</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
SocialQ&A

Core algorithms

- User Interest Analyzer
- Question Categorizer
- Question-User Mapper
Question Categorizer

Analyzes the question and associates the question with tags.
Updates the user interest vector
what is the best country band?
Question Categorizer

Matching and association

Tokens
- what
- best
- country
- band

Synset
- rock
- metal
- country
- reggae

Question Tags
SocialQ&A

Core algorithms

- User Interest Analyzer
- Question Categorizer
- Question-User Mapper
Question-User Mapper

Generates a list of potential answer providers by leveraging the asker’s and answers’ interest vectors and the tags associated with the question.
Question-User Mapper

Identify answer providers

Interest Similarity

Social closeness
Question-User Mapper

Factors considered for calculating the interest similarity

- User interest vector
  - User’s profile data
  - Type of questions asked by the user
  - Answers provided by the user

- Question interest vector

\[ \psi_{I,U_k} = \text{Sum}(W_{U_k}^{I_j}) \]
Question-User Mapper

- Interest Similarity
- Social closeness

Identify answer providers
Factors considered for calculating the social closeness factor

- Interest similarity between two users ($P_S$)
- Past interactions between the users ($P_A$)
- Number of common friends to both the users ($P_C$)

$$\psi_{C,U_k} = \gamma_S \cdot P_S + \gamma_A \cdot P_A + \gamma_C \cdot P_C$$
Question-User Mapper

- Interest Similarity
  \[ \beta \]
- Social Closeness
  \[ (1 - \beta) \]

Identify answer providers
Outline

- Introduction
- Background
- Social Q&A
- Results and Analysis
- Conclusion and Future work
Results and Analysis

Statistics
- 1000 Facebook users
- 9419 Yahoo! Answer questions with best answers

Metrics
- Precision
- Recall
- Wait time
Results and Analysis

High precision + high recall = high answer quality
Results and Analysis

Short wait time

- SocialQ&A
- Interest
- Social
- Random

Wait time (min)

Num. of selected answerers
Prototype usage of SocialQ&A

Number of users - 124
Number of questions - 163
Number of answers - 282
Number of users from India - 70
Number of users from US - 35
Number of users from UK - 1
Results and Analysis

Questioning and answering activity
Results and Analysis

Answer quality: high ranked

- **Average rating**
  - % of questions
  - 1 2 3 4 5 6 7 8 9 10

- **Maximum rating**
  - % of questions
  - 1 2 3 4 5 6 7 8 9 10
Results and Analysis

User questioning and answering activity

- Percentage of unanswered questions (SocialQ&A) – Approximately 15%
- Percentage of unanswered questions (Yahoo! Answers) – Approximately 16%
- Percentage of unanswered questions (Baidu Zhidao) – Approximately 40%
Outline

- Introduction
- Background
- Social Q&A
- Results and Analysis
- Conclusion and Future work
Conclusion and future work

SocialQ&A
- Social network based answerer selection
- High quality potential answerer selection

Achievement
- High Quality
- Short wait time

Future work
- Questions outside the social community
Outline

- Introduction
- Background
- Social Q&A
- Results and Analysis
- Conclusion and Future work
If you have any questions and comments please contact guoxinl@g.clemson.edu.