



Understanding Localness of Knowledge Sharing: A Study of *Naver KiN* “Here”

Sangkeun Park, Yongsung Kim, Uichin Lee (KAIST, Knowledge Service Engineering)
Mark Ackerman
(University of Michigan, School of Information and Department of Computer Science
and Engineering)

KAIST



Social Q&A

Conventional Q&A

NAVER 지식iN

YAHOO!
ANSWERS



Quora

Social Q&A has been widely used for information seeking

Social Q&A

Location-based Q&A

“Please recommend me a good Korean restaurant near Hyatt Regency Toronto.”

Seeking location-based information via local experts

Related Work

Conventional social Q&A analysis

- Question topic/type
- User behavior patterns
- Motivations

Location-based Q&A services analysis

- Research prototypes for design exploration
- Small scale user studies

Research Motivation

Previous studies focused on

- Conventional Q&A analysis
- Small-scale feasibility tests on location-based Q&A services

Our research goal

- **Understanding localness of knowledge sharing with**
 - Real-world dataset
 - Geographical analysis

Research Questions

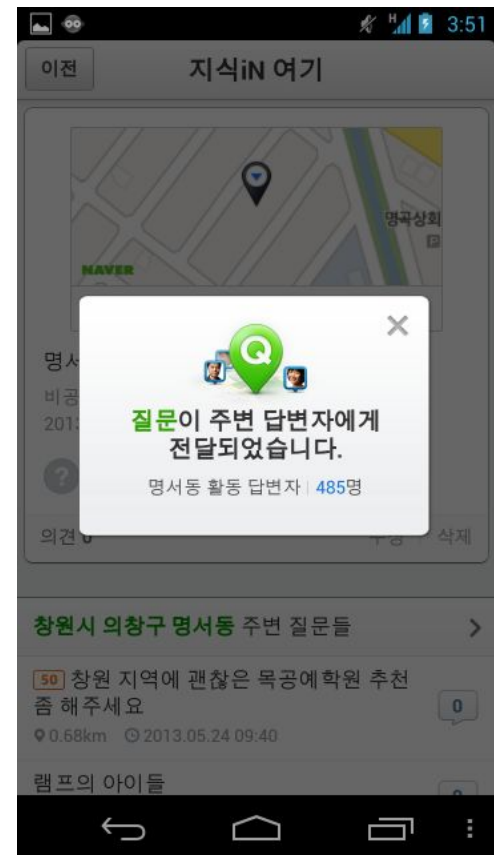
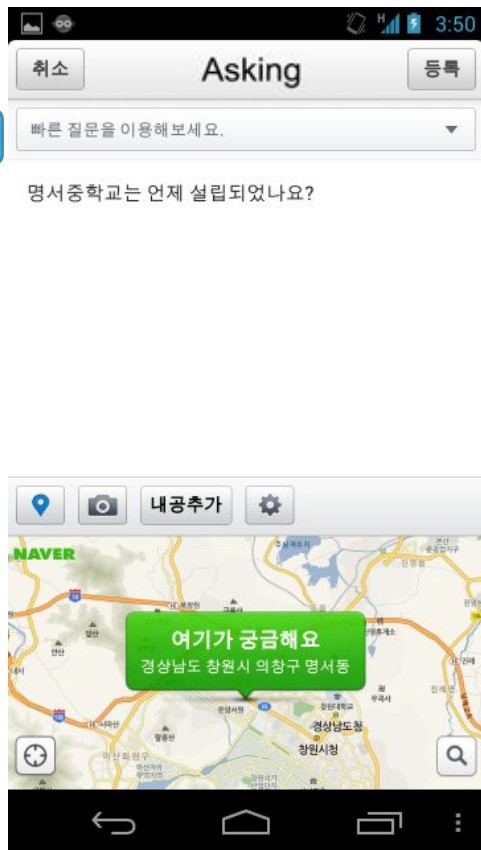
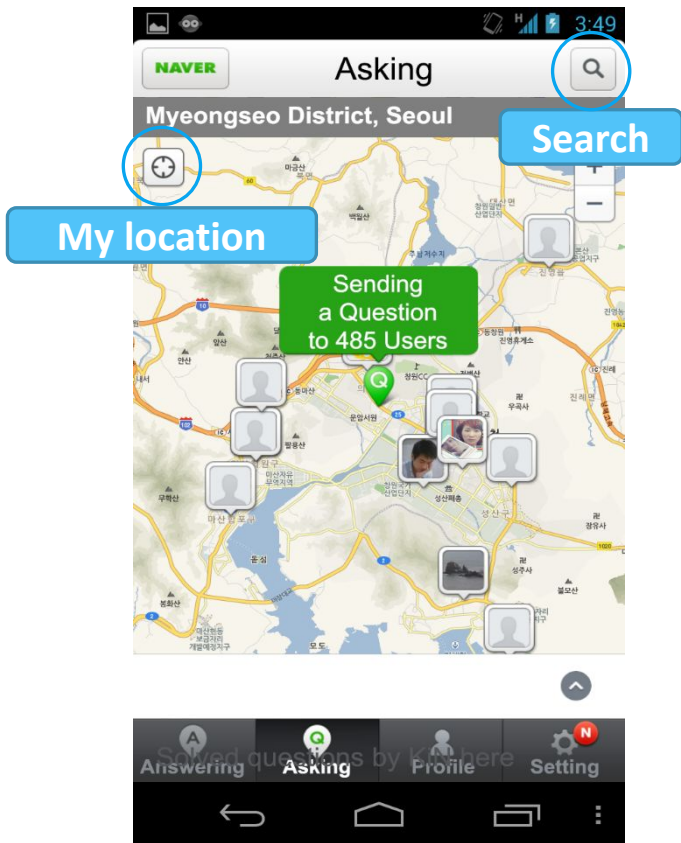
RQ1) Are topical/typological patterns related to geographic characteristics, and how?

RQ2) How are users geographically focused in their asking/answering activities?

RQ3) What are the answer motivations that are unique in location-based social Q&A?

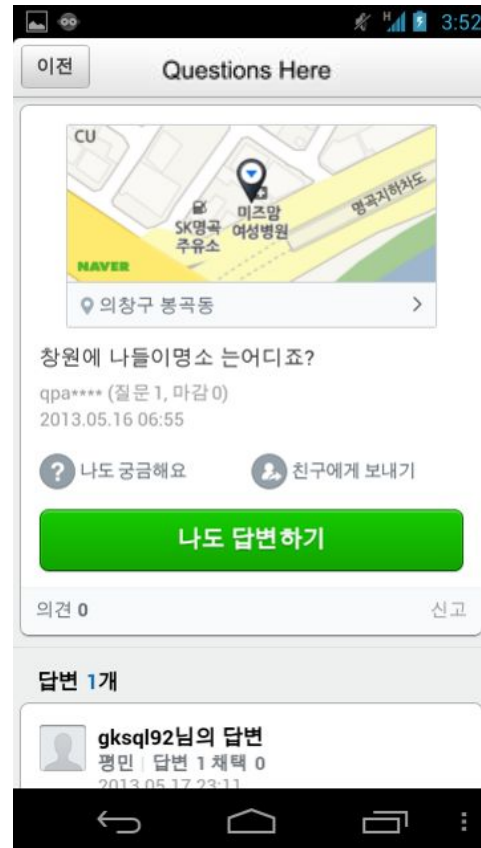
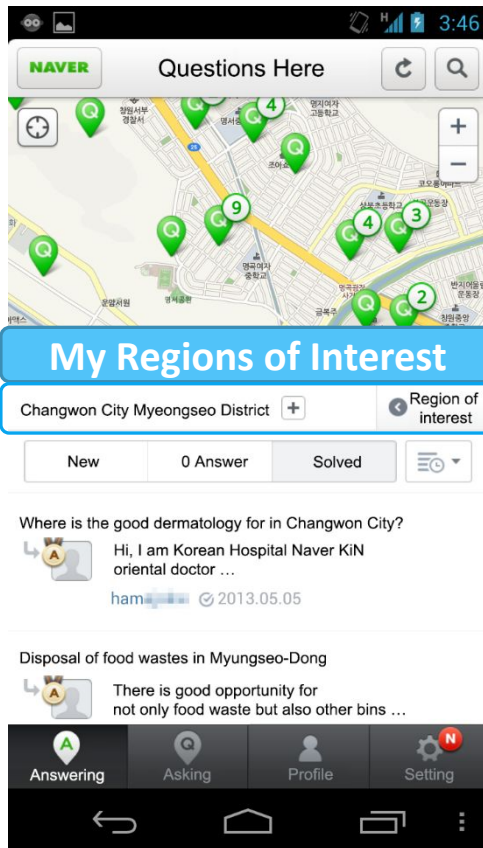
Interaction

Naver KiN "Here" (Asking)



Interaction

Naver KiN "Here" (Answering)



Methodology

Crawled dataset from Naver KiN “Here”

- Dec 17, 2012 ~ Dec 31, 2013 (13 months)
- 508,334 questions / 567,156 answers



Online survey

- 285 users participated (Aug, 2013)
- Questionnaire is composed of two parts
 - The number of selected regions of interests,
A list of those names as well as the reason for each choice
 - Motivation answered questions (open-ended)



Research Questions

RQ1) Are topical/typological patterns related to geographic characteristics, and how?

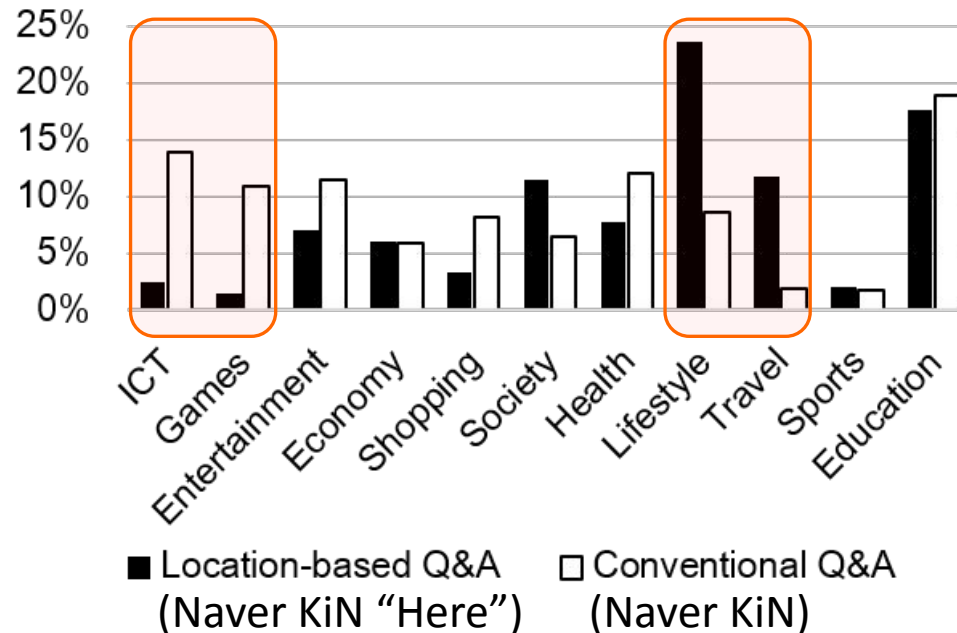
- **What kinds of questions topics?**
- **What kinds of questions types?**

RQ2) How are users geographically focused in their asking/answering activities?

RQ3) What are the answer motivations that are unique in location-based social Q&A?

Topical/Typological patterns

Topic distribution



Lifestyle and Travel were dominant

ICT and Game were minimal

Topical/Typological Patterns

Topic distribution of top 10 cities/districts



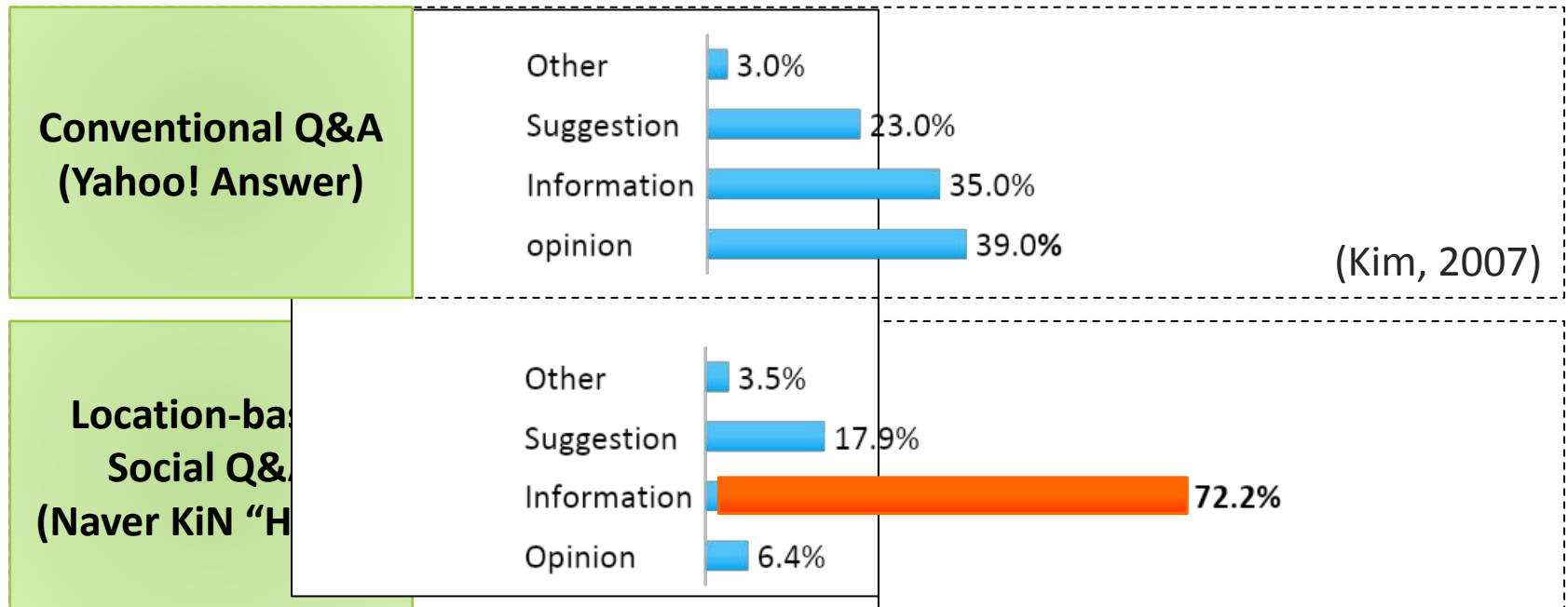
Top 10 Cities

The geographic characteristics were well reflected

The topical distributions were largely dependent on the size and functional complexity of the region

Topical/Typological Patterns

Typological distribution



Information is dominant in location-based social Q&A

Research Questions

RQ1) Are topical/typological patterns related to geographic characteristics, and how?

RQ2) How are users geographically focused in their asking/answering activities?

- **The degree of geographic focus**
- **Granularity of spatial locality**
- **Categories of regions of users' interest?**

RQ3) What are the answer motivations that are unique in location-based social Q&A?

Geographical Focus

Measuring geographic focus

$$Entropy = -\sum_{k=1}^K p_k \log_2 p_k$$



City C



City A



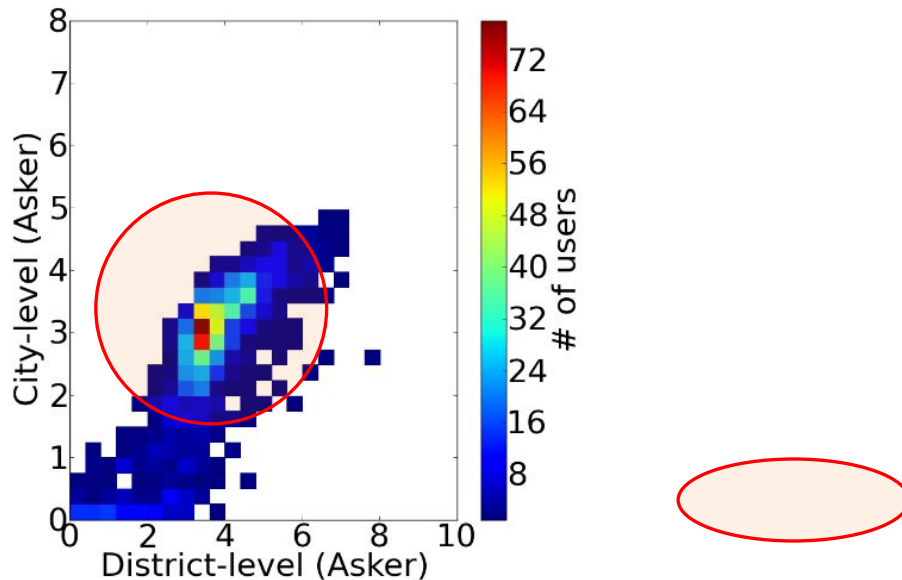
City B



City D

Geographical Focus

Entropy comparison btw asking and answering

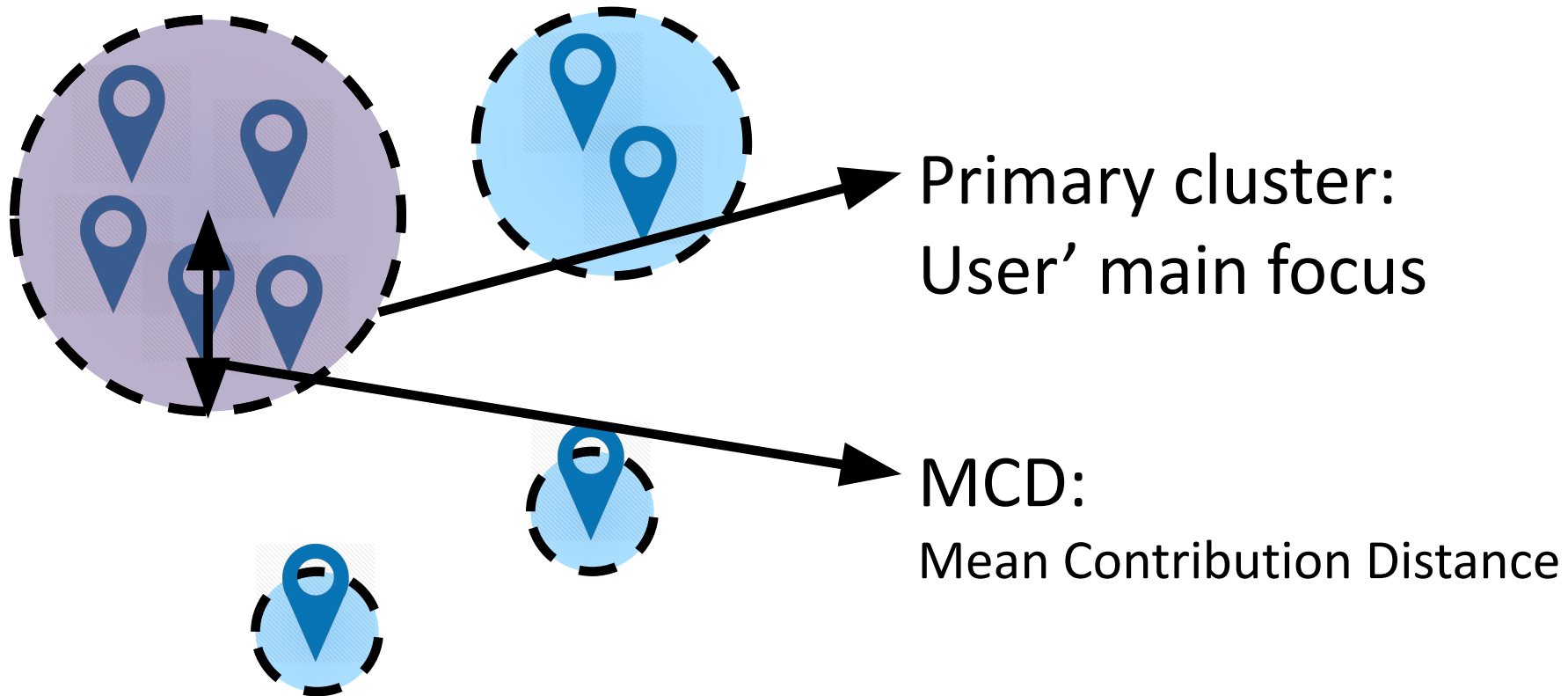


Askers' activities typically spanned multiple districts/cities
(Mean=3.5/2.5)

Answerers' activities focused on multiple districts in a few cities
(Mean=2.6/0.8)

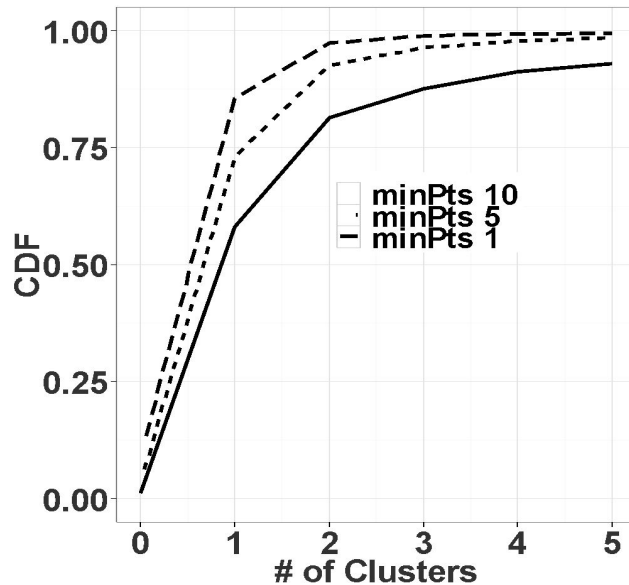
Geographical Focus

Measuring granularity of spatial locality



Geographical Focus

Spatial cluster analysis



More than 75% of users geographically focused on 1~2 clusters

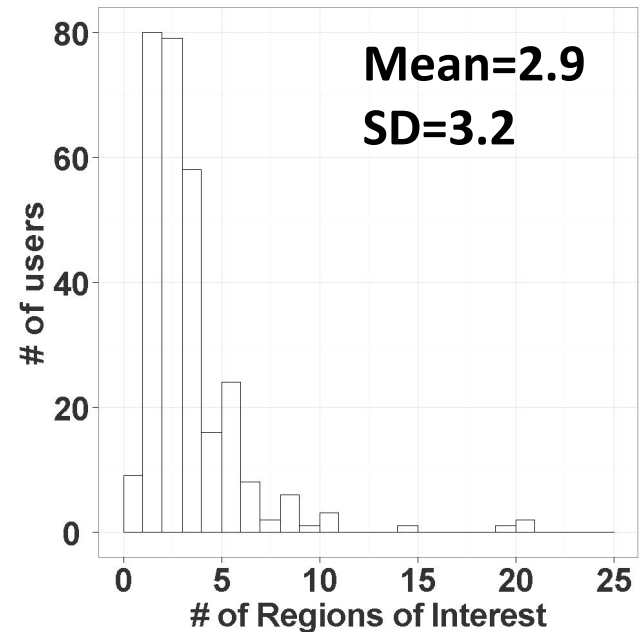
The primary cluster covered a few nearby districts (Mean=2.3Km, minPts=1)

Geographical Focus

Regions of interest

Users' regions of interests were:

- Home (93.7%)
- Work/School (23.9%/28.9%)
- Downtown (24.6%)



Regions of interest is closely related to life experience

Research Questions

RQ1) Are topical/typological patterns related to geographic characteristics, and how?

RQ2) How are users geographically focused in their asking/answering activities?

RQ3) What are the answer motivations that are unique in location-based social Q&A?

- **Are there unique motivations related to localness?**

Answer Motivations

1. Knowledge exchange (24.9%)
2. Altruism (18.2%)
3. **Ownership of local knowledge (10.1%)**
4. Points (9.8%)
5. Pastime (9.2%)
6. **Sense of community (7.0%)**
7. Business promotion, learning, etc.

Answer Motivations:

Ownership of local knowledge

“

Because I know everything about my town as I have lived in my town for a long time

”

- Ownership of local knowledge -

“

I think kind and sincere answering is one of the representative images of the area, and I want to build a good image of my area.

”

- Sense of community -

Discussion

Design implications

Leveraging topical/typological patterns

- Filtering location-based questions by topics
- Local search by archived factual information Q&A dataset

Leveraging the geographical activity analyses

- Extending the radius of questions notification
- Recommending neighboring districts for subscribing

Motivating user contributions based on localness

- Community-level symbols such as badges
- Regional competition such as ranking

Conclusion

Topical/typological patterns

- In general, lifestyle and travel were dominant
- The geographic characteristics were well reflected
- Information is dominant

Geographical focus

- Askers' activities typically spanned multiple districts/cities
- Answerers' activities focused a very few cities
- A primary clustered region covered a few nearby districts

Unique motivators

- Ownership of local knowledge
- A sense of community

Understanding Localness of Knowledge Sharing: A Study of *Naver KiN “Here”*



Sangkeun Park, Yongsung Kim, Uichin Lee
(KAIST, Knowledge Service Engineering)

Ackerman Mark S.

(University of Michigan, School of Information and Department of Computer Science and Engineering)

Type Manual Coding

Randomly selected 1000 questions

Coded by two external raters.

- 200 questions together
- separately coded the remaining 800 questions (i.e. 400 questions each).
- Cohen's Kappa: 0.84 - substantial agreement.

Categorization

Extracted keywords from title

Searched the extracted keywords on Naver KiN which is a topic-based Q&A service

In the top 100 results, most frequent topic was selected as a topic of the questions

Manually coded the topics of 100 randomly selected questions, then, inter-rater agreement: $k = 0.87$

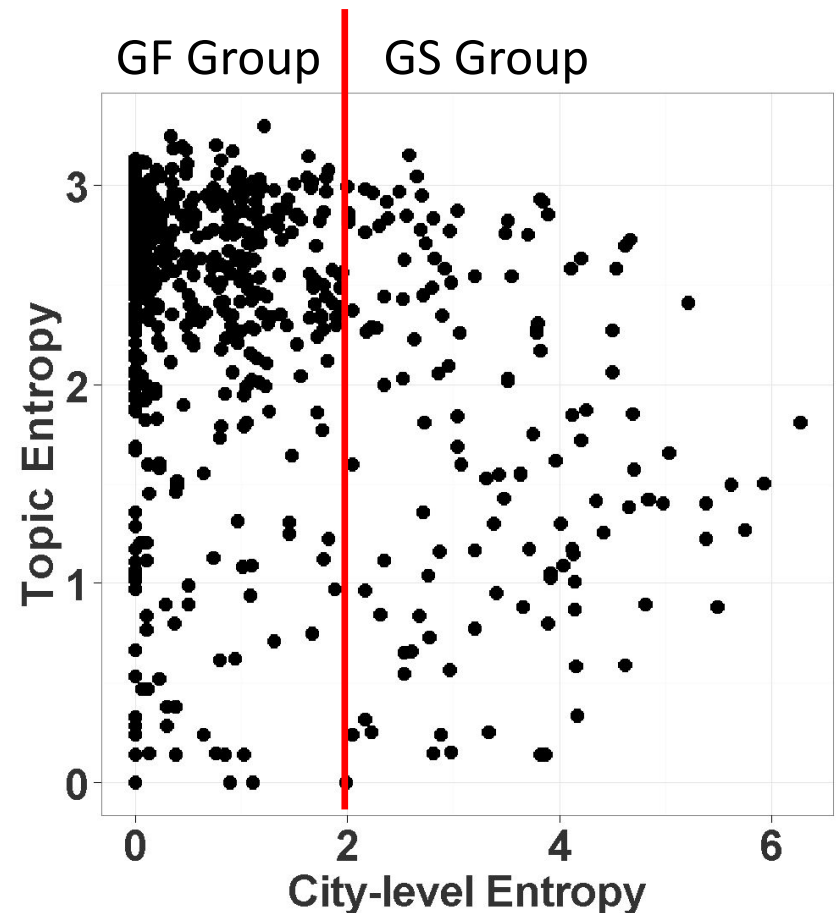
Geographical Focus

behavioral difference by geographical focus

GF: Geographically Focused (81%)

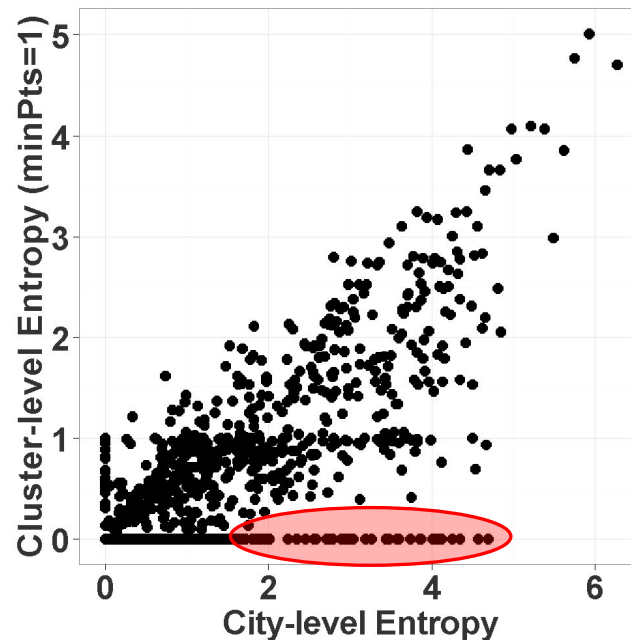
GS: Geographically Scattered (19%)

- The active web searchers
- The province-level experts



Geographical Focus

Spatial cluster analysis



A majority of users focused on one cluster