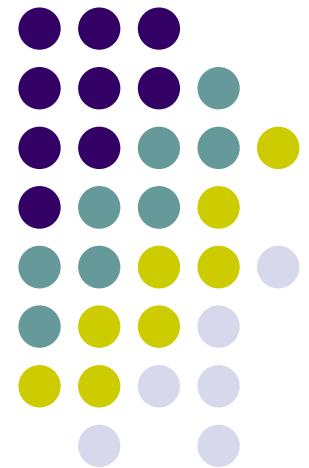**CS 525M S13**

When Recommendation Meets Mobile: Contextual and Personalized Recommendation On The Go

Zijian Liu

Electrical and Computer Engineering Dept.

Worcester Polytechnic Institute (WPI)

Introduction: motivation

- Mobile phone has become a recommendation terminal customized for the individuals

  --what is around? What to do?

- Existing research focused on recommendation relying on text input. However, it is a tedious job for phone users.

- Voice-to-search (like Siri) has limitations: quiet environment, short expressions, not context-aware(time, locations, etc.)

Introduction: motivation

- Based on analysis result on a real-world large-scale click-through-data collected from a commercial mobile search engine.

Introduction: solution

- Their Idea: Mobile recommendation without requiring input, rich context (implicit input), to rank both entity types[1] and entities[2].

- A probabilistic recommendation approach:
  - To rank both entity types and entities
  - Relevant to user and sensory context.

[1] Entity Types: Coffee Shop, Shopping

[2] Entities: Starbucks, Wal-Mart

Introduction: solution

- An application based on Windows Phone 7 for evaluation—Easylife.

a) user and sensory context

b) Rank of entity types

C) Rank of entities

(a) user and sensory context   (b) rank of entity types   (c) rank of restaurant entities

Related Work

- Query suggestion and auto-completion
  - User does not need to type the whole query
  - However, User intent on mobile is quite different
    Actually user's input on mobile is short.
- Object-level vertical/local search
  - Vertical search engine focuses on specific segment of online content.(Local business, sites)

Related Work

- Recommendation System
  - Requires long query history and heavy computation

Analysis on click-through-data

- Analysis on a real-world large-scale click-through-data collected from a commercial mobile search engine.

    - Collect a large-scale query log data from2009-09-30 to 2010-03-28.

| #raw query	#user	#entity query	#entity user
75,221,037	13,711,497	11,492,382	4,012,030

Analysis on click-through-data

- Distribution of mobile queries in US
- shows mobile search is becoming pervasive, especially in big cities.

Analysis on click-through-data

- number of queries (#word) with different lengths. 62.3% less than three words
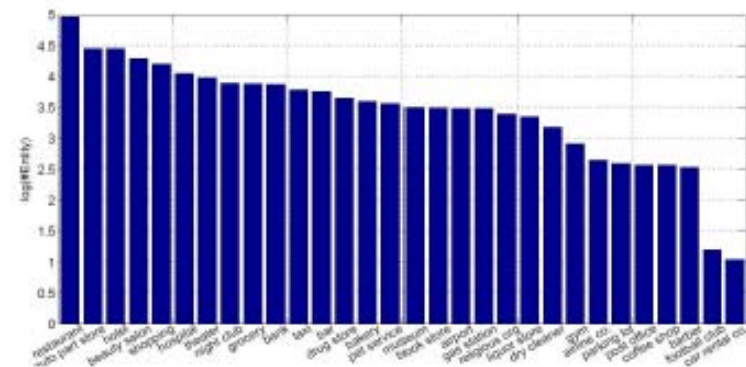
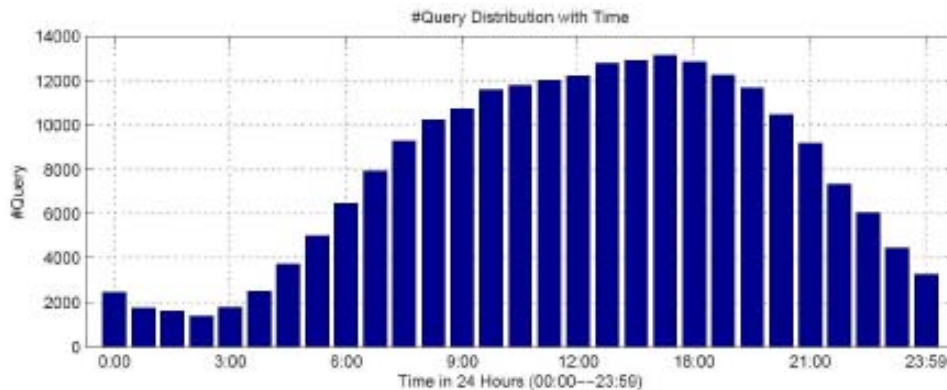- shows queries on mobile platform are usually short.
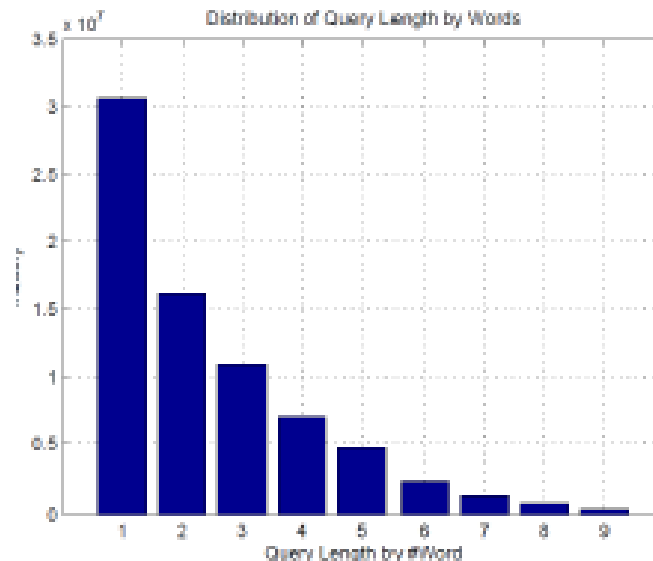
(b) distribution of query length.

Analysis on click-through-data

- Search Is Local and Context-Sensitive
  - very sensitive to location. (Commercial area)
  - The highest peak occurs near 5–6 pm, lowest point occurs at about 2–3 am.
  - 15.8% is entity query (target on search entities)

Analysis on click-through-data

- characteristics of mobile query motivates the design of a recommendation system that is context-aware, personalized, and without requiring any typing of queries.

Approach

1. Entity extraction which detects and recognizes entities from a textual query log

2. Entity ranking which ranks a candidate set of entities and the corresponding entity types to the user.

Approach: Entity Extraction

- Use the algorithms of previous Entity Extraction "Know it all"
  - Extractor automatically create a collection of extraction rules for each kind of entity types and attributes
  - Pass initial extracted ones to retrieve more entities
  - Pattern learner filter out high-quality entities for expansion

Approach: Entity Extraction

- The common attributes of extracted entities from three examples. Each entity type has its unique attributes.

Name	Country	Region	Locality	Address	EntityType	Latitude	Longtitude	Phone
Starbucks	U.S.	CA	Irivne	115 Fortune Dr	Coffee Shop	33.6504	-117.7461	(617)277-0087
FrenchDress	U.S.	MA	Boston	49 River St	shopping	42.3572	-71.0703	(617)723-4968
Walmart	U.S.	UT	HeberCity	435 Airport Rd	shopping	40.4827	-111.4207	(435)654-6436

Probabilistic Entity Ranking

- Key notions used in this paper

Table 3. List of key notations.

Symbol	Description
$\mathbb{N}_u$	$\{1, \ldots, u\}$, a set of integers up to u
N_u	$|\mathbb{N}_u|$, the cardinality of $\mathbb{N}_u$
$\mathcal{U}$	$\{\mathbf{u}_i : i \in \mathbb{N}_u\}$, a collection of N_u users
Q	the collection of all the queries
U, E, Z, T, L	represents user, entity, entity type, time, location, respectively
$a.B$	the attribute B of object a

Probabilistic Entity Ranking

- Framework of building a personal user based on click-through data

Probabilistic Entity Ranking
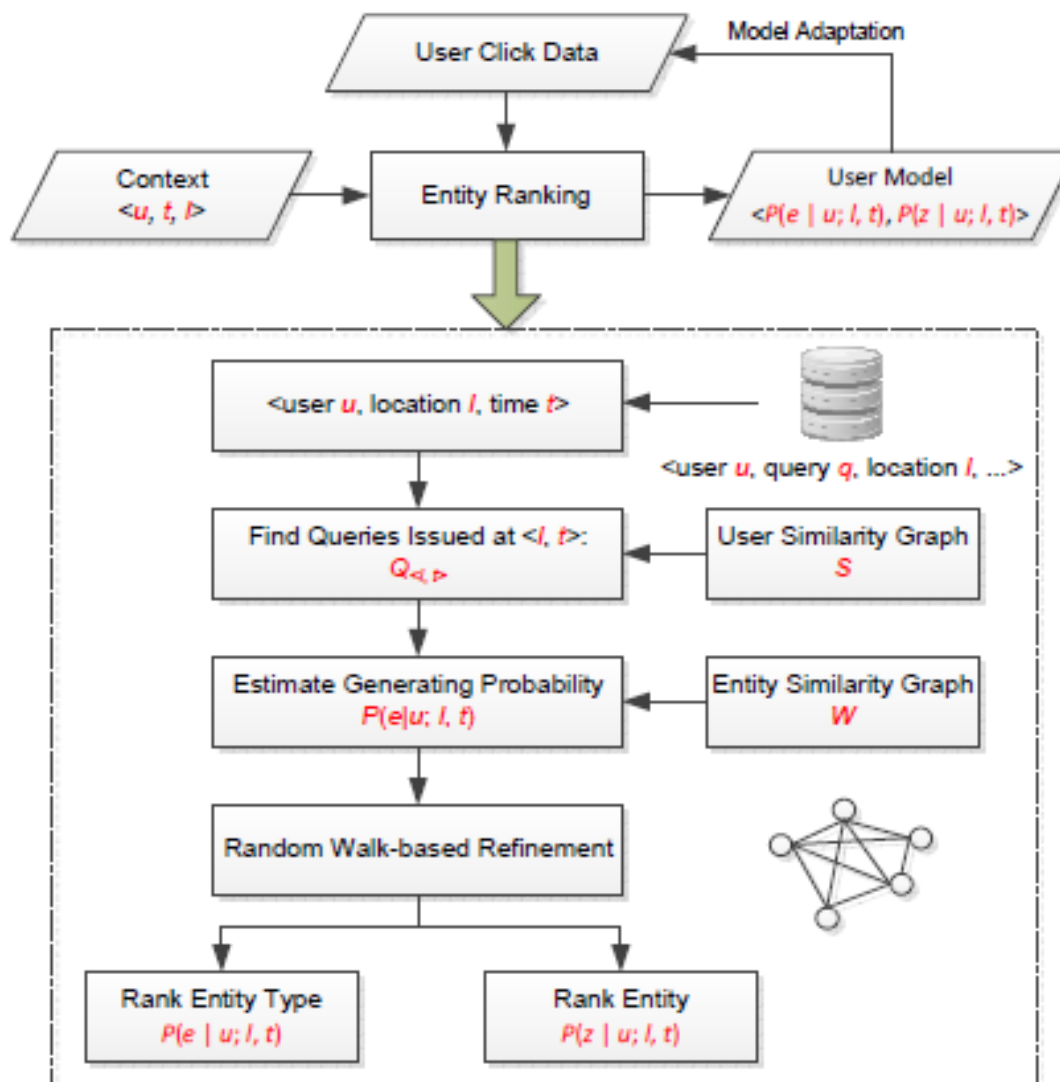
Probabilistic Entity Ranking

- entity ranker can estimate the conditional probabilities of entities and entity types for a given user under certain context.

- Each mobile query is a 5-dimensional tuple:

$$Q = <E, Z, U, L, T>$$

Probabilistic Entity Ranking

$$P(e_j|u_i;l,t) = \sum_{k=1}^{N_z} P(e_j|z_k, u_n; l,t) \sum_{n=1}^{N_u} P(z_k|u_n; l, t) P(u_n|u_i)$$

$$\propto \sum_{k=1}^{N_z} P(e_j|z_k; l,t) \sum_{n=1}^{N_u} P(z_k|u_n; l,t) s(u_n, u_i), \quad (1)$$

$$P(e_j|z_k; l,t) := \frac{|Q_e : \forall q \in Q_e, q.E = e_j, q.Z = z_k, q.L = l, q.T = t|}{|Q_z : \forall q \in Q_z, q.Z = z_k, q.L = l, q.T = t|}$$

$$P(z_k|u_n; l,t) := \frac{|Q_z : \forall q \in Q_z, q.Z = z_k, q.U = u_n, q.L = l, q.T = t|}{|Q_u : \forall q \in Q_u, q.U = u_n, q.L = l, q.T = t|}$$

Probabilistic Entity Ranking

- User similarity S(*,*): each user can be represented by a query history record.

- Three level similarity function : entity-based, entity-type-based, and entity-attribute-based similarity.

- For example, two users, interested in McDonalds, KFC. may like Burger King since they both need fast food.

Probabilistic Entity Ranking

- Ranking Refinement by Random Walk

Restaurant for dinner → Bar for night life

- use the number of users that exhibit temporal patterns to measure the transition probability between two entities

$$w(e_i, e_j) = |\mathcal{U} : \forall u \in \mathcal{U}, \exists q, q' \in \mathcal{Q}_u, q.E = e_i \text{ and}$$
$$q'.E = e_j \text{ and } q'.T - q.T < \epsilon|,$$

Probabilistic Entity Ranking

- Ranking Entity Types

$$P(z|u; l, t) \propto \sum_{n=1}^{N_u} P(z|u_n; l, t) s(u_n, u).$$

EXPERIMENTS

- ## Data and settings
  - First five months: build the user similarity graph and entity similarity graph based on the mobile click-through data
  - last one month: randomly selected 2,000 users use their queries in the March of 2010 as test data

$$q \in Q_{t.}$$

58,111 query records   test set

EXPERIMENTS

- Data and settings
  - Location and time <l, t>: split the time into 7 intervals
  - Extract a set of queries from the query database with the context <l, t>
  - Then we sort the entities contained in these queries by ranker.
  - extract the queries conducted within five kilometers to user l and same t

EXPERIMENTS

- Data and settings
  - Accuracy of three kinds of recommendation: entity, intent (entity types), entity in each of entity types

$$Accuracy(Q_t, k) = \frac{\sum_{q \in Q_t} \mathbb{I}(\pi(q.E) \leq k)}{|Q_t|},$$

where $\mathbb{I}(c)$ returns 1 if c is true and 0 otherwise.

EXPERIMENTS

- ## Data and settings

  - Examined schemes: Baseline 1 to Baseline 6, PCAR-T, PCAR-E

  Figures: Page 160

EXPERIMENTS

- Experiment 1: Top-k Recommendation Accuracy -- Figure 5

- Experiment 2: Sensitivity to Context – Figure 6

- Experiment 3: Top-k Recommendation Accuracy – Table 5

Conclusion

- conduct an analysis on a large-scale mobile click-through data collected from a commercial mobile search engine.

- a query-free entity recommendation approach to understand implicit user intent on the mobile devices.

- a recommendation application based on Windows Phone 7 and evaluation

Future Works

- Exploring other machine learning techniques for a better recommendation

- Leveraging social signals for improving user similarity (such as Facebook)

- collecting real-world click-through data through the developed mobile application and evaluation.

Thank you!