

R24-3: Effective Location Identification from Microblogs

Guoliang Li, Jun Hu, Jianhua Feng (Tsinghua Univ.), Kian-lee Tan (NUS)

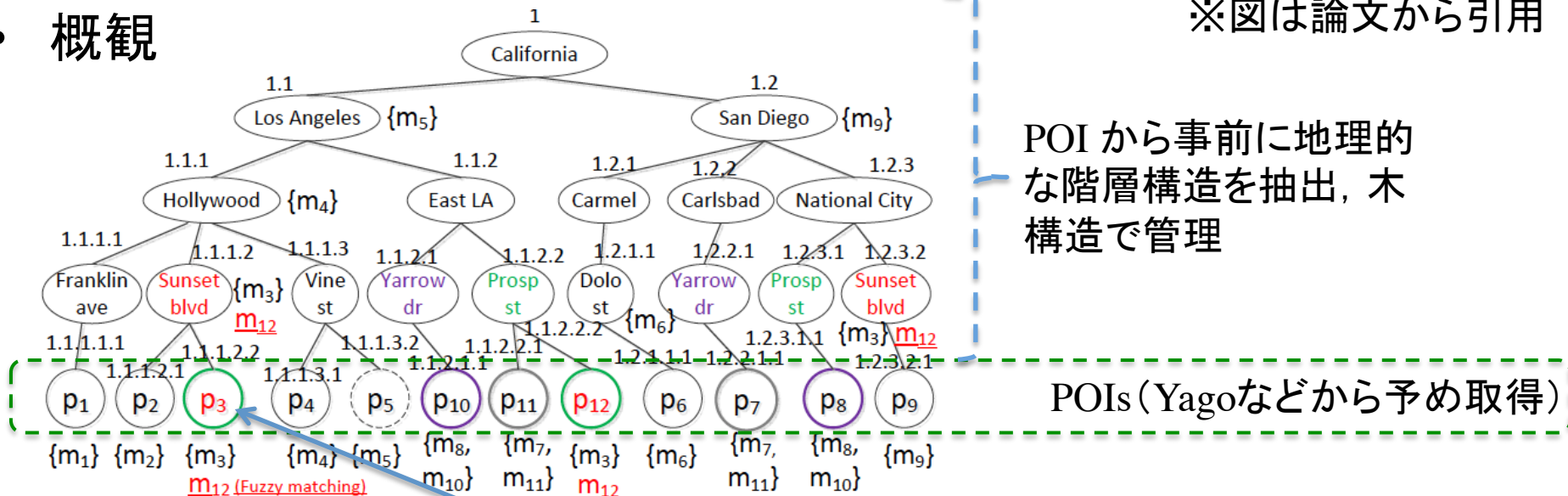
- マイクロブログからマイクロブログとユーザの位置推定
 - マイクロブログサービスに正確な位置を登録するユーザが少ない(16% [17])
 - 位置がわかると位置に基づく情報推薦などが可能になる
- 提案手法 (GLITTER) : ユーザの位置推定手法
 - ユーザの位置 ← interested locations
 - (おそらく)よく行く場所, 行きたい場所
 - 1. 個々のマイクロブログの位置 (POI) を推定 (fuzzy manner, top-k)
 - POI = Point of Interest, e.g., 映画館, ショッピングモール
 - 2. 複数のマイクロブログの位置からユーザの位置を推定 (top-k)
 - マイクロブログの位置はPOIであるのに対し, ユーザの位置はPOIでない
 - 推定したマイクロブログの POI からユーザの位置を推定する
 - 3. 他のマイクロブログからマイクロブログの位置推定を改良
 - “Olympia theater is so nice.” + “Manhattan is my favorite place.”
 - Olympia theater in Manhattan, New York

R24-3: Effective Location Identification from Microblogs

Guoliang Li, Jun Hu, Jianhua Feng (Tsinghua Univ.), Kian-lee Tan (NUS)

概観

※図は論文から引用



POI から事前に地理的な階層構造を抽出, 木構造で管理

POIs (Yagoなどから予め取得)

Fig. 1. Tree-based Location Structure and Indexes.

Film School, Subset blvd, Hollywood, Los Angeles, California

1. マイクロブログの位置を推定 (m_i に対応する p_j を見つける)
 - 地理的なエンティティを含むマイクロブログを木のノードに関連付ける
 - m_3 = “I was able to get a tour at Film School, Subset blvd.”
2. マイクロブログと地理エンティティの類似度を基に “適当な” top-k 祖先地理エンティティをユーザ位置として推定 (詳しくは論文で)
3. 推定したユーザの位置を参考にマイクロブログの位置を修正

R24-3: Effective Location Identification from Microblogs

Guoliang Li, Jun Hu, Jianhua Feng (Tsinghua Univ.), Kian-lee Tan (NUS)

※図は論文から引用

実験

データセット

- Twitter (3.5k users, 2.2m microblogs, 635 microblogs/user)
- Foursquare (21k users, 1m microblogs, 48 microblogs/user)

評価方法: Precision, Recall, F-measure

- マイクロブログ: 緯度経度が付いているものに対して, 推定位置が100m以内なら正解とする
- ユーザ: ランダムに選んだユーザに top-k interested locations を聞いて正解とする

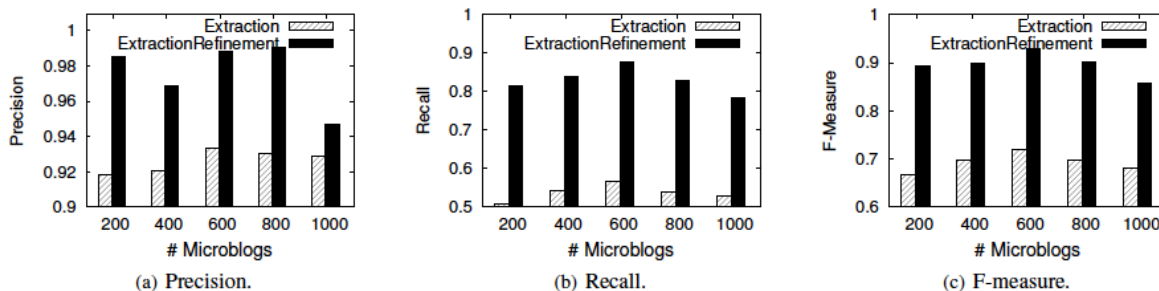


Fig. 6. Evaluating quality of *top-k* locations of microblogs on Twitter dataset (exact match, $k = 3$).

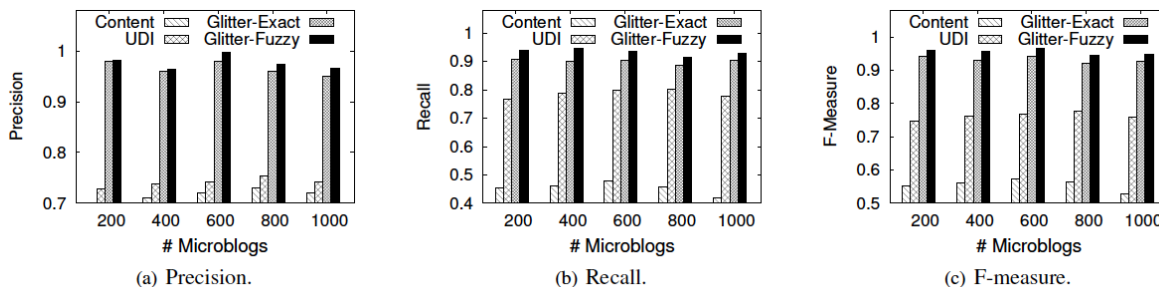


Fig. 9. Comparing quality of *top-k* locations of users on Twitter dataset ($k = 3$).