

ソーシャルメディアにおける発言位置特定とその匿名化

磯 颯, 若宮 翔子, 荒牧 英治 (NAIST)



I. Introduction

- * SNS 上で共有した発言が発言者の意図しない聴衆の目まで届いてしまう可能性
- * 個人のプライバシーが損なわれる恐れ

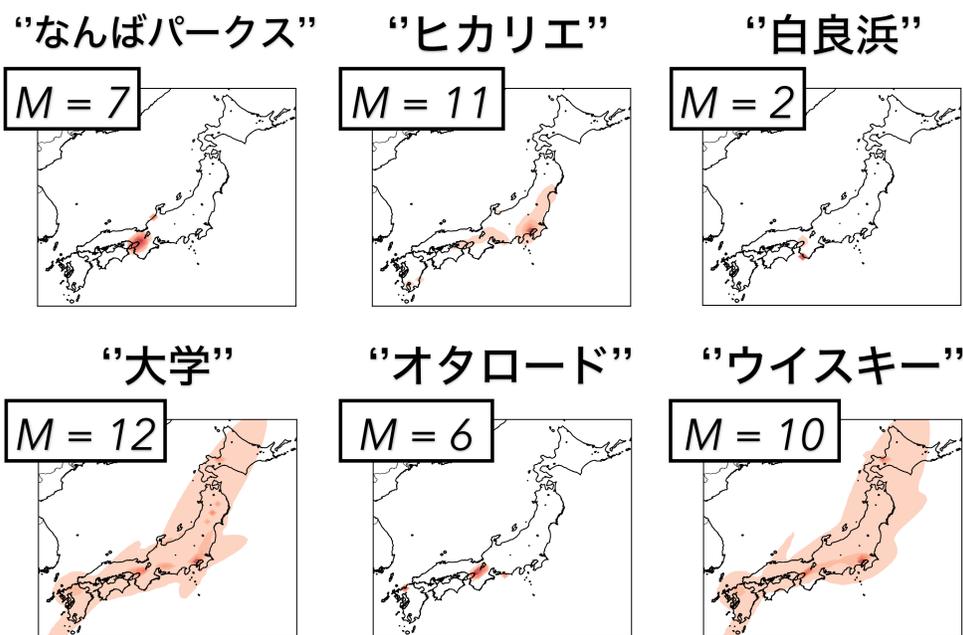


- * 発言位置の秘匿化に焦点を当て、位置特定能力を評価関数とした匿名化手法を導入

II. Gaussian Mixture Model

- * 形態素ごとの推定発言位置分布: GMM で推定
- * 混合数 M の推定法: AIC (H. Akaike, 1974)

$$p_v(x) = \sum_{m=1}^M \alpha_{v,m} \phi(x; \mu_{v,m}, \Sigma_{v,m}) \quad \text{s.t.} \quad \sum_{m=1}^M \alpha_{v,m} = 1$$



III. Weighted Sum of GMM

- * 位置特定能力に応じた固有の重みを形態素 v ごとに定義

$$w_v = \frac{1}{\text{error}_v^{\alpha}} \left(\begin{array}{l} \text{error}_v = \frac{1}{N_v} \sum_{N_v} x_i - \lambda_v \\ \lambda_v = \arg \max_x p_v(x) \end{array} \right)$$

- * 任意の発言位置の推定分布を形態素ごと推定分布に対する重みづけ和で表現

$$p_{\text{tweet}}(x) = \frac{\sum_{v \in \text{tweet}} w_v p_v(x)}{\sum_{v \in \text{tweet}} w_v}$$

IV. Kullback Leibler divergence

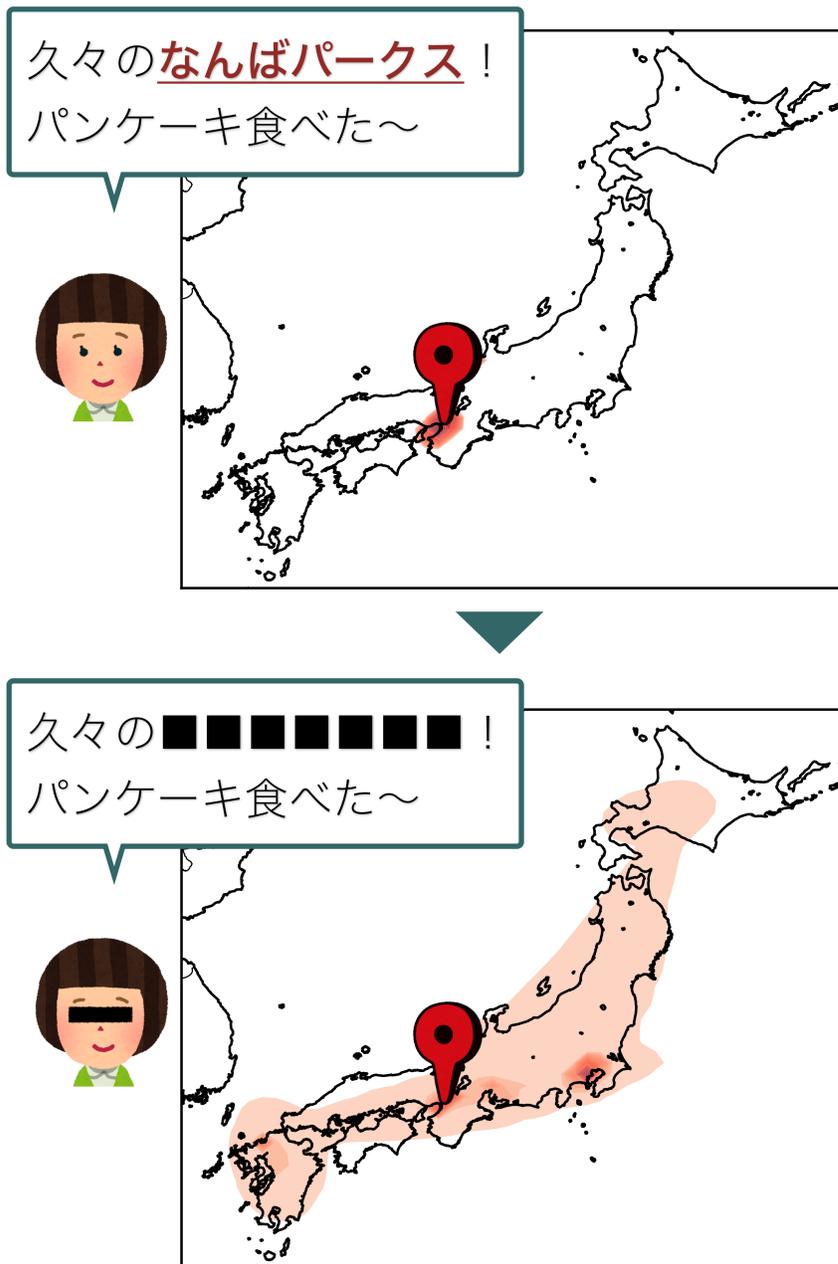
位置特定が容易 \leftrightarrow 分布のエントロピー 小

- * Kullback Leibler divergence に基づき、最も分布を変化させる形態素 v の削除

$$\arg \max_v D_{KL}(P||Q) \left(= \int_{\mathbb{R}^2} p(x) \log \frac{p(x)}{q(x)} dx \right)$$

$$q_{\text{tweet}}(x) = \frac{\sum_{v \in \text{tweet} \setminus \{v_{del}\}} w_v p_v(x)}{\sum_{v \in \text{tweet} \setminus \{v_{del}\}} w_v}$$

V. Result



: 実際の発言位置

VI. Future Work

- * 複数形態素の組合せによる位置特定能力と削除後の分布の変化の検証