Data Clustering using Matrix Factorization Techniques for Wireless Propagation Map Reconstruction

Junting Chen and Urbashi Mitra

Department of EE, University of Southern California, Los Angeles, CA, USA, {juntingc, ubli}@usc.edu

USC Viterbi School of Engineering

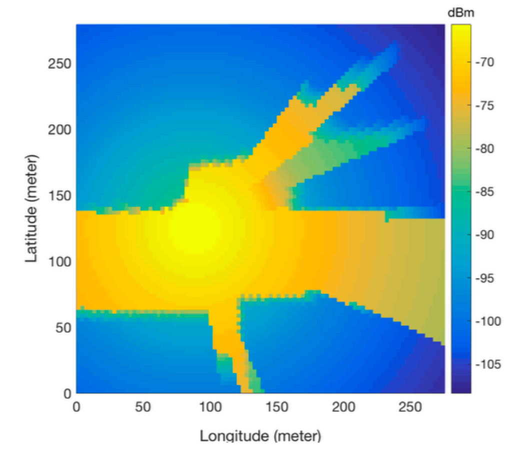
Ming Hsieh Department of Electrical Engineering

Background: Learn a Radio Map for UAV Placement

Motivational Applications:

- Place a UAV to build a good relay channel for ground terminals
- Learn the shadowing environment to exploit it

Learning the radio map is the key!



Radio Map: Geographically-indexed dataset of the channel quality

Existing Learning Model

- Probabilistic channel / Stochastic terrain model (Hourani *et. al*'14 TCOML, Mozaffari *et. al*'16 ICC)
- "Fingerprint" model (Romero *et. al*'17 TSP)
- 3D city map based (Monserrat *et. al.*'15 IJAP)

A good model would simplify the learning

Classical propagation model $y = \beta + \alpha d_{\mathrm{u}}(\mathbf{x}) + \xi$

Proposed: Segmented propagation model

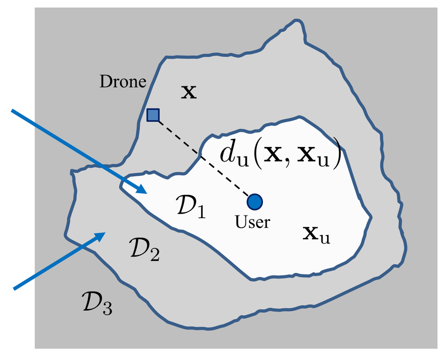
$$y(\mathbf{x}) = \sum_{k=1}^{K} \Big(g_{\mathrm{u}}^{(k)}(\mathbf{x}) = \beta_k + \alpha_k d_{\mathrm{u}}(\mathbf{x}) + \xi_k \Big) \mathbb{I}\Big\{ \mathbf{x} \in \mathcal{D}_k \Big\}$$

(Model parameters to learn)

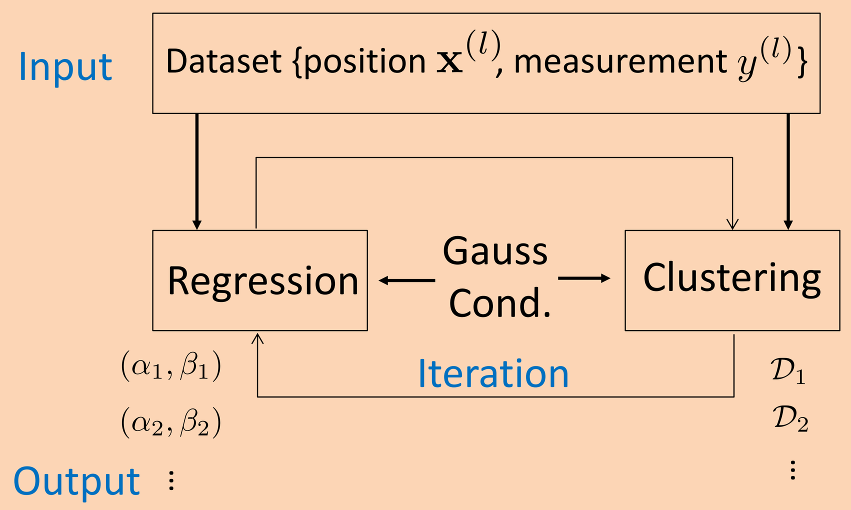
Challenges: (1) nonparametric (D_k), (2) coupling

$g_{\mathrm{u}}^{(1)}(\mathbf{x}) = \beta_1 + \alpha_1 d_{\mathrm{u}}(\mathbf{x})$

$g_{\mathrm{u}}^{(2)}(\mathbf{x}) = \beta_2 + \alpha_2 d_{\mathrm{u}}(\mathbf{x})$



The Problem: Data Clustering



Classical Learning Diagram (MLE)

MLE Formulation (our prior work [1])

$$\underset{\{\alpha_1^{(k)}, \alpha_2^{(k)}, \sigma_k, \pi_k, \bar{z}_k^{(l)}\}}{\text{maximize}} \quad \sum_{l=1}^{N} \sum_{k=1}^{K} \bar{z}_k^{(l)} \Big[\log \pi_k + \log p_k(\mathbf{x}^{(l)}, y^{(l)}) \Big]$$

where $p_k(\mathbf{x}, y) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\Big\{ -\frac{(y - \alpha_1^{(k)} d_{\mathrm{u}}(\mathbf{x}) - \alpha_2^{(k)})^2}{2\sigma_k^2} \Big\}$

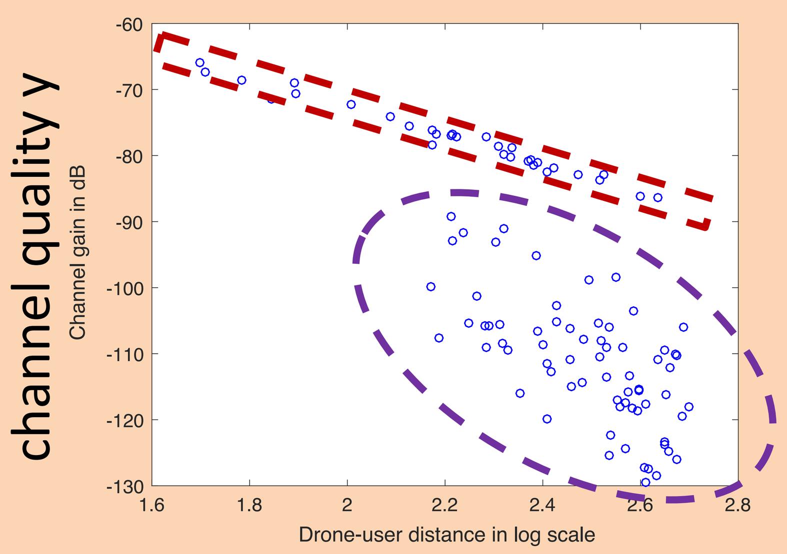
Drawback:

- Assumptions too ideal: Gaussian assumption, knowing perfect user location, *etc.*
- Very high complexity (large dataset)

Question: Can we obtain the regression parameters *without* relying on the temporary clustering result and Gaussian assumption?

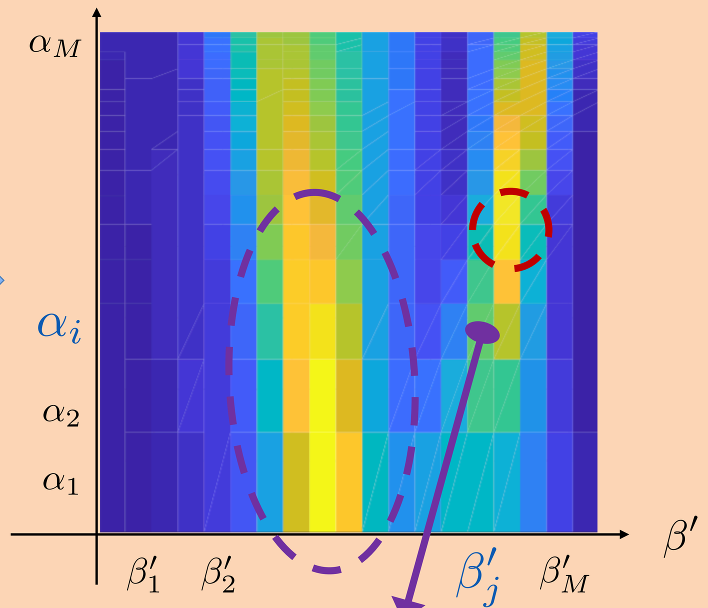
Compress the Data to a Feature Matrix

Data clustering problem (scatter plot of the dataset)



log-distance $d_{\mathrm{u}}(\mathbf{x})$

"Source" localization problem from a feature matrix



$$V(\alpha_i, \beta_j'; \lambda) = \sum_{l=1}^{N} \exp\big\{ -\lambda \big[y^{(l)} - \alpha \hat{d}_{\mathrm{u}}(\mathbf{x}^{(l)}) - \beta' \big]^2 \big\}$$

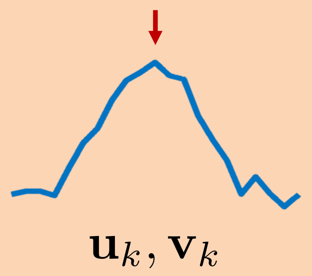
Value function to evaluate the similarity to the parametric model for the whole dataset

Peak localization of the *k*th unimodal component yields the model parameter of the *k*th data group

Multi-source Localization using unimodal matrix factorization (UMF)

$$\underset{\{\alpha_k, \mathbf{u}_k, \mathbf{v}_k\}}{\text{minimize}} \quad \Big\| \mathcal{P}_{\Omega}(\hat{\mathbf{H}} - \sum_{k=1}^{K} \alpha_k \mathbf{u}_k \mathbf{v}_k^{\mathrm{T}}) \Big\|_F^2$$

subject to $\quad \mathbf{u}_k, \mathbf{v}_k$ are unimodal



+ nonparametric hypothesis testing to learn D_k

Why it works

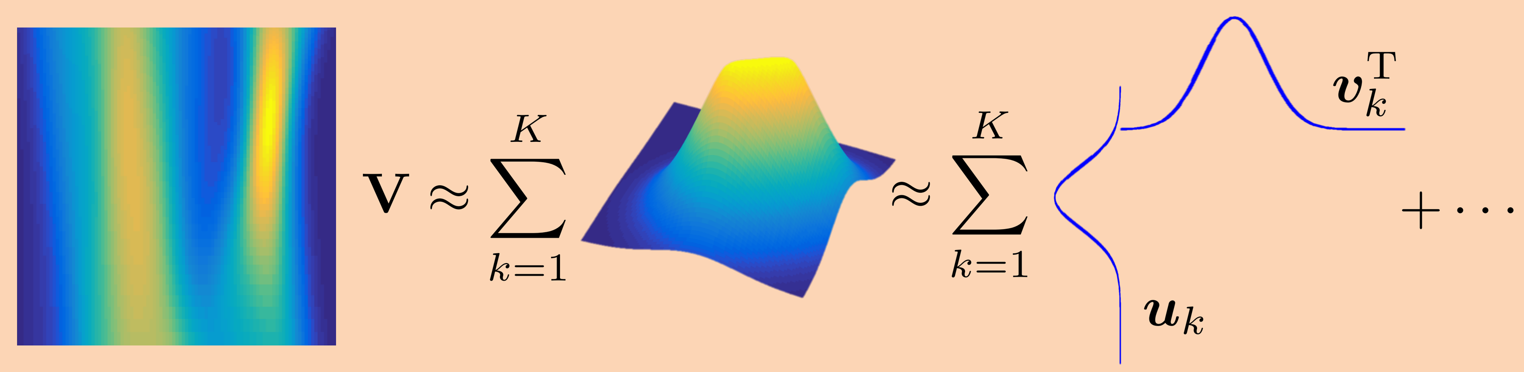
Structured model of the feature matrix:

$$\mathbf{V} = \mathbf{V}^{(1)} + \mathbf{V}^{(2)} + \cdots + \mathbf{V}^{(K)} + \mathbf{N}$$

where $\mathbf{V}^{(k)}$ is a component matrix of a particular data subset with

$$V_{ij}^{(k)} = \sum_{l=1}^{N} \exp\big\{ -\lambda \big[y^{(l)} - \alpha_i \hat{d}_{\mathrm{u}}(\mathbf{x}^{(l)}) - \beta_j \big]^2 \big\} \mathbb{I}\{\mathbf{x}^{(l)} \in \mathcal{D}_k\}$$

Main result: Under small enough parameter λ and some mild regularity conditions, $\mathbf{V}^{(k)}$ is a unimodal matrix.



Theorem [2]: The dominant singular vectors of a unimodal matrix are unimodal.

Theorems justify the low-rank unimodal-structured model for the feature matrix!

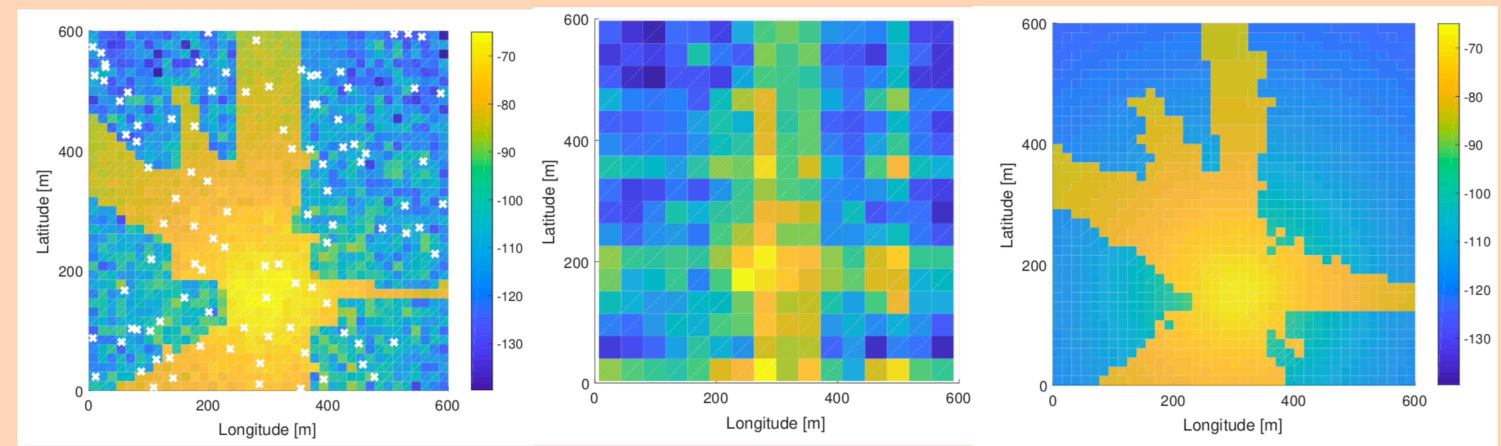
Numerical Evaluation

Model for general training data: $y = \alpha \log_{10} d(\mathbf{x}) + \beta + n$

where $(\alpha^{(1)}, \beta^{(1)}) = (-22, -28), n \sim \mathcal{N}(0, 1)$ for LOS region

$(\alpha^{(2)}, \beta^{(2)}) = (-36, -22), n \sim \mathcal{N}(0, 8^2)$ for NLOS region

$\lambda = 1/(2\sigma^2), \sigma = 8$ [dB]



Sample 100 random locations | Direct matrix completion from original data | Proposed method recovers both region pattern & value

Main contributions

- A nonparametric **learning tool** to compress data to a feature matrix to discover unknown (clustered) data structure
- Robust & low complexity
- Use of UMF as a useful tool for component analysis

References

[1] Chen, Esrafilian, Gesbert, Mitra, Globecom, 2017.

[2] Chen & Mitra, ArXiv: 1711.07457, 2018.

ACKNOLEDGEMENT: This research has been funded in part by one or more of the following grants: ONR N00014-15-1-2550, NSF CNS-1213128, NSF CCF-1718560, NSF CCF-1410009 , NSF CPS-1446901, and AFOSR FA9550-12-1-0215