CA 1340/7-1 Phase II: From Massive MIMO to Massive Wireless Networks

Co-PIs: G. Caire, G. Kutyniok, G. Wunder
Co-workers: R. Levie, C. Yapar, M. Barzegar Khalilsarai

Technische Universität Berlin, Freie Universität Berlin

CoSIP Retreat
Aachen, February 2020

Outline

Outline

WP1: The (approximate) Common Eigenvector Problem

- Consider an M-dimensional antenna array and let $\mathbf{h} \in \mathbb{C}^M$ be the corresponding (random) channel vector with covariance $\boldsymbol{\Sigma}$. Consider a family of beamforming vectors $\mathbf{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_M\}$. The covariance matrix of the channel in the beam-space domain is $\widetilde{\boldsymbol{\Sigma}} = \mathbf{U}^{\mathsf{H}} \boldsymbol{\Sigma} \mathbf{U}$.
- If $\mathbf{U}$ is the matrix of eigenvectors of $\boldsymbol{\Sigma}$, then $\widetilde{\boldsymbol{\Sigma}}$ is diagonal. For typical propagation scenarios, the diagonal elements are quite “sparse”.
- Given a set of $M \times M$ covariance matrices $\{\boldsymbol{\Sigma}_k : k = 1, \ldots, K\}$, find a set of beamforming vectors $\mathbf{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_M\}$ such that for all k the beam-space beam-space domain covariances $\{\mathbf{U}^{\mathsf{H}} \boldsymbol{\Sigma}_k \mathbf{U} : k = 1, \ldots, K\}$ are “close to diagonal” and the diagonals are “as sparse as possible”.
- Main application: active channe sparsification (ACS) precoding in FDD massive MIMO:

  Khalilsarai, M.B., Haghighatshoar, S., Yi, X. and Caire, G., 2018. FDD massive MIMO via UL/DL channel covariance extrapolation and active channel sparsification. IEEE Transactions on Wireless Communications, 18(1), pp.121-135.

WP2: Learning the Network Soft-Topology

- In problems such as link scheduling in D2D communications (e.g., a high density V2V scenario), or user-cell association in dense small cell deployments, it is very helpful to know, for any two points x_1 and x_2 on the network region (e.g., on the plane), the propagation loss function $g(x_1, x_2)$ in the case of isotropic antennas.
- The scope of this WP2 is to develop Deep Learning techniques to predict the pathloss function for general complicated topologies with blocking, scattering, and diffraction.

WP3: Efficient Multiband Splicing

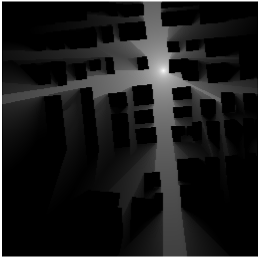
- In several applications (e.g., channel sounding, indoor localization) we wish to obtain a high resolution estimation of the CIR.
- Normally, the resolution of the estimation is limited by $1/W$, where W is the measurement bandwidth. E.g., high resolution channel sounders have front-end bandwidth of 1GHz and more, and are therefore expensive.
- Most communication systems (e.g., WiFi) have limited channel bandwidth, but use multiple channel bands. We wish to pooling together multiple narrowband measurements (spectrum "splicing") to obtain the equivalent of a very large measurement bandwidth and therefore high timing resolution CIR estimation, with cheap commercial devices.

Outline

Setting of Pathloss Function Prediction

A set of transmitter–receiver links in an urban environment.

Pathloss = loss of signal strength
between a transmitter and receiver.



Freie Universität Berlin

CoSIP

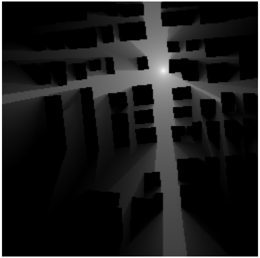
Setting of Pathloss Function Prediction

A set of transmitter–receiver links in an urban environment.

Pathloss = loss of signal strength between a transmitter and receiver.

Radio map

- Fixed transmitter location
- Pathloss at all locations is $R : \mathbb{R}^2 \to \mathbb{R}$.



Freie Universität Berlin

CoSIP

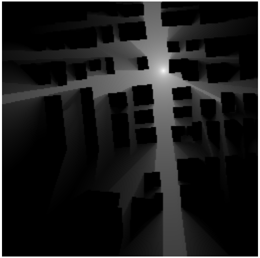
Setting of Pathloss Function Prediction

A set of transmitter–receiver links in an urban environment.

Pathloss = loss of signal strength
between a transmitter and receiver.

Radio map

- Fixed transmitter location
- Pathloss at all locations is $R : \mathbb{R}^2 \to \mathbb{R}$.

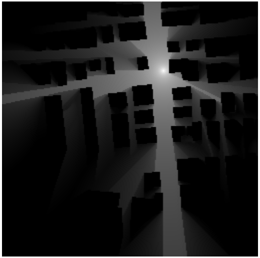


Examples

Setting of Pathloss Function Prediction

A set of transmitter–receiver links in an urban environment.

Pathloss = loss of signal strength
between a transmitter and receiver.



Radio map

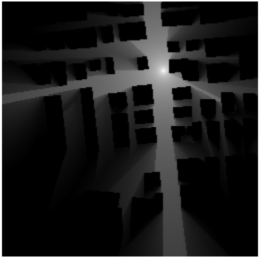
- Fixed transmitter location
- Pathloss at all locations is $R : \mathbb{R}^2 \to \mathbb{R}$.

Examples
device to device

Setting of Pathloss Function Prediction

A set of transmitter–receiver links in an urban environment.

Pathloss = loss of signal strength
between a transmitter and receiver.



Radio map

- Fixed transmitter location
- Pathloss at all locations is $R : \mathbb{R}^2 \to \mathbb{R}$.

Examples

device to device **cellular network**

Freie Universität Berlin CoSIP TU berlin

Applications relying on Radio Maps

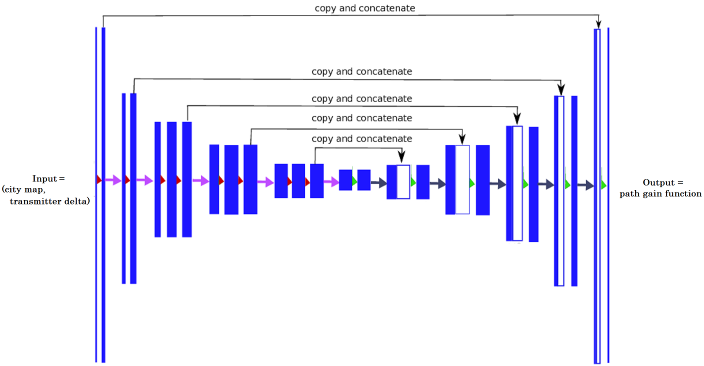
- **Device to device**: link scheduling
- **Cellular network**: cellular base station assignment

Additional applications:

- fingerprint based localization
- physical-layer security
- power control in emerging systems
- activity detection

Goal: Estimate the Radio Map

Physical simulation is too slow $\longrightarrow$ Use UNet instead.



Supervised learning:
Dataset = city maps with simulated radio maps.

RadioMapSeer Dataset

- 700 maps from **OpenStreetMap**, converted to **morphological images**.
- 80 devices per map.
- **Simulated radio-maps** (radio network planning software *WinProp*):
  Dominant Path Model (DPM)
  Intelligent Ray Tracing (IRT)
- The obtained results are converted to **gray level**.

Different Settings of RadioUNet

Deep learning radio map estimator: **RadioUNet.**

Different Settings of RadioUNet

Deep learning radio map estimator: **RadioUNet.**

Two **network input** scenarios

- Only the city map and transmitter location is given
- The city map and transmitter are given + some measurements.

Different Settings of RadioUNet

Deep learning radio map estimator: **RadioUNet.**

Two **network input** scenarios

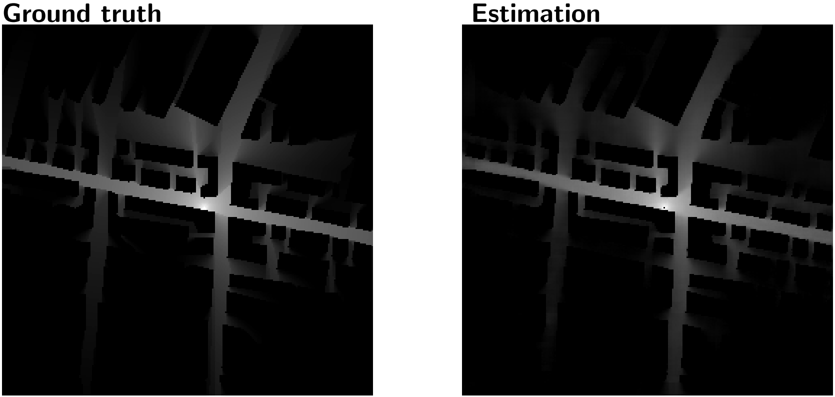
- Only the city map and transmitter location is given
- The city map and transmitter are given + some measurements.

Two **map** scenarios

- The accurate map is given
- A perturbed map is given

Example: Accurate Map, No Measurements

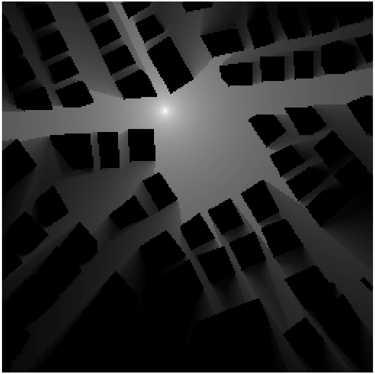
(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)



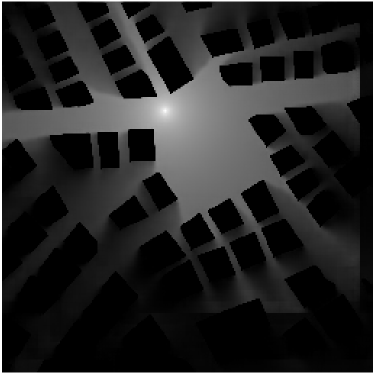
Example: Accurate Map, No Measurements

(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)

Ground truth



Estimation



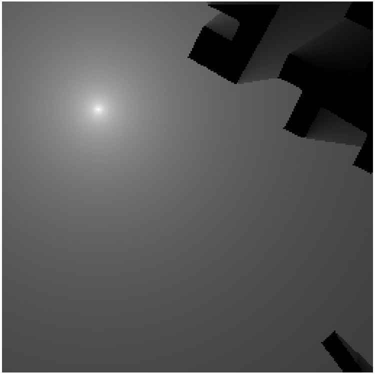
Freie Universität Berlin

CoSIP

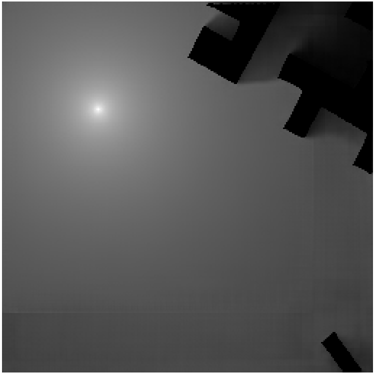
Example: Accurate Map, No Measurements

(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)

Ground truth



Estimation



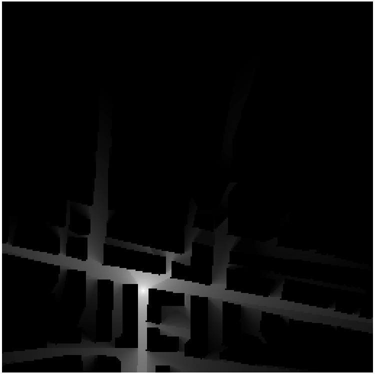
Freie Universität Berlin

CoSIP

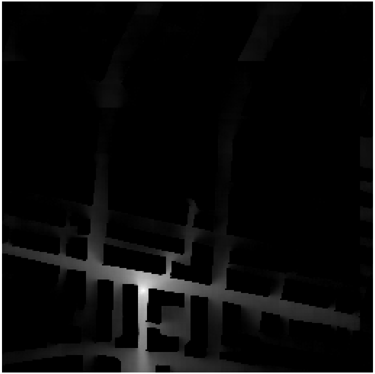
Example: Accurate Map, No Measurements

(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)

Ground truth



Estimation



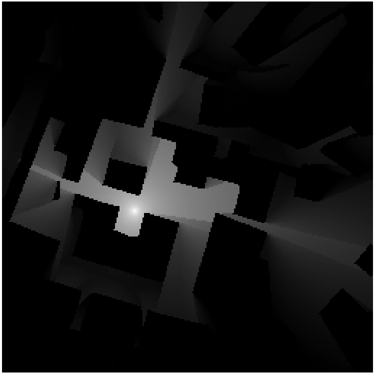
Freie Universität Berlin

CoSIP

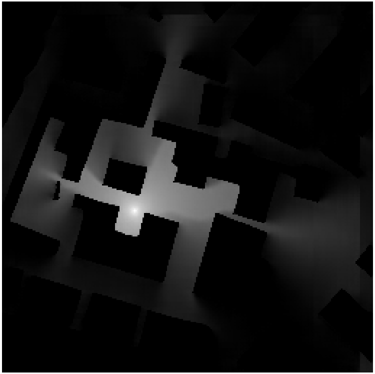
Example: Accurate Map, No Measurements

(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)

Ground truth



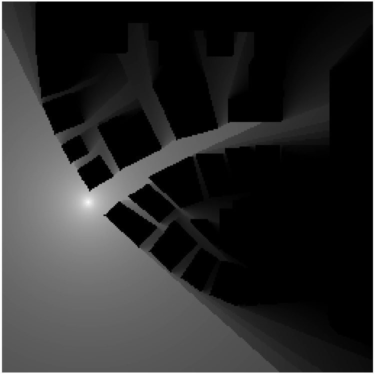
Estimation



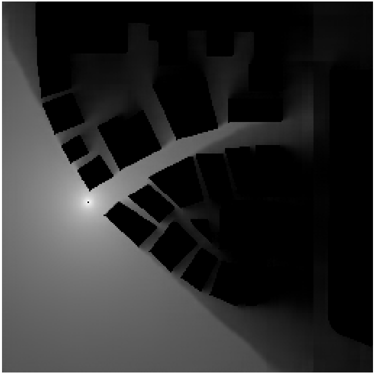
Example: Accurate Map, No Measurements

(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)

Ground truth



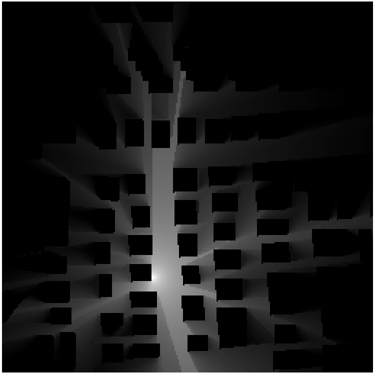
Estimation



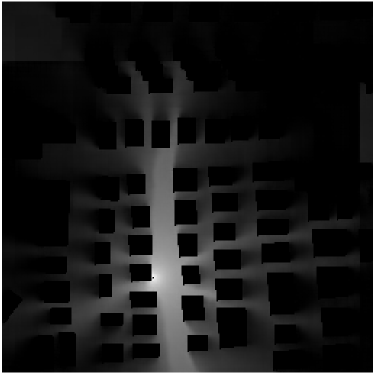
Example: Accurate Map, No Measurements

(pathloss $\in (-147dB, -47dB)$, l_2 error$= 2.18dB$)

Ground truth

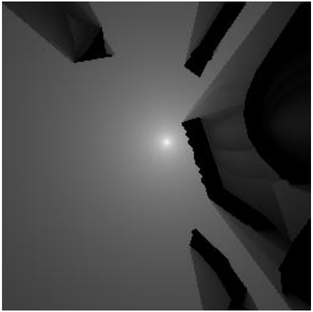


Estimation

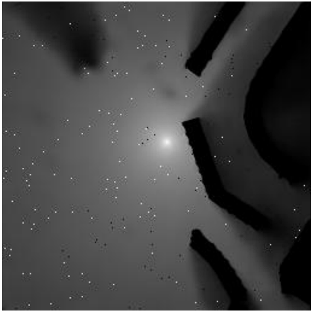


Example: Missing Building, Measurements

Ground truth

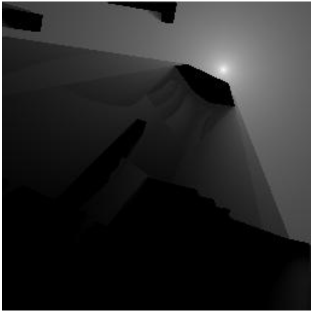


Estimation

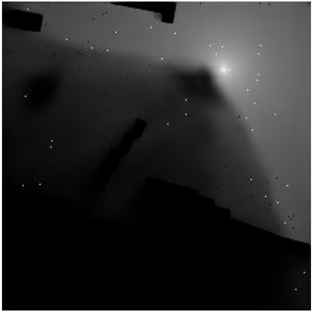


Example: Missing Building, Measurements

Ground truth



Estimation

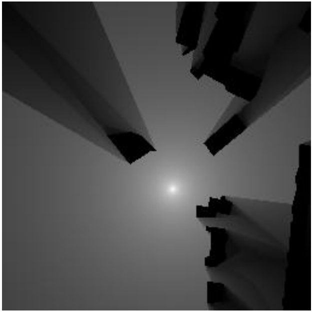


Freie Universität Berlin

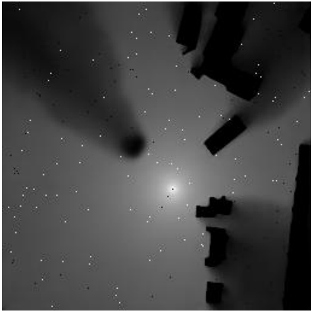
CoSIP

Example: Missing Building, Measurements

Ground truth



Estimation



Classification of Radio Map Estimation Methods

- **Data driven interpolation methods**
  (radial basis function interpolation, tensor completion, support vector regression, matrix completion)

Classification of Radio Map Estimation Methods

- **Data driven interpolation methods**
  (radial basis function interpolation, tensor completion, support vector regression, matrix completion)
- **Model based predictions/simulations**
  (ray-tracing, dominant path model, and empirical model)

RadioUNet vs Model Based Simulations

Run-time

- Dominant path method $\approx$ 1sec

RadioUNet vs Model Based Simulations

Run-time

- Dominant path method $\approx$ 1sec
- Intelligent ray tracing $\approx$ 10sec

RadioUNet vs Model Based Simulations

Run-time

- Dominant path method $\approx$ 1sec
- Intelligent ray tracing $\approx$ 10sec
- RadioUNet $\approx 10^{-3}$sec
  with accuracy

$$\frac{\|\mathrm{RadioUnet} - \mathrm{Simulation}\|^2}{\|\mathrm{Simulation}\|^2} \approx 10^{-2}$$

Freie Universität Berlin CoSIP

RadioUNet vs Data Driven Interpolation Methods

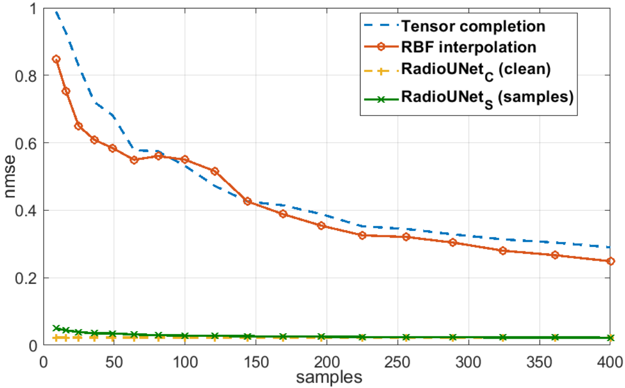
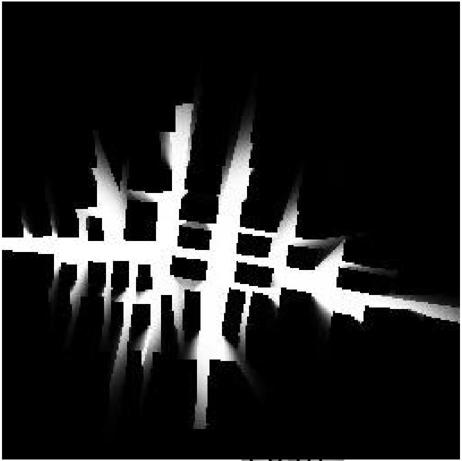


Figure: Estimation error of the radio map reconstruction methods as a function of the number of measurements. RadioUNet$_C$ has zero samples, and is given as a baseline.

Freie Universität Berlin

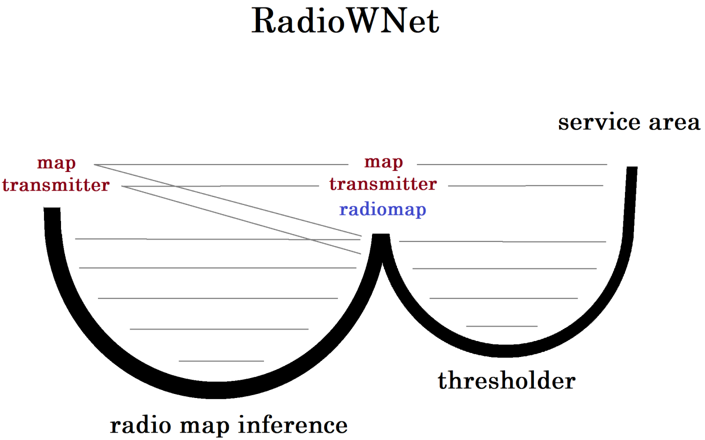
Service Area Classification

- Classify if devices can **receive wanted signal**.
- Classify if devices **receive unwanted signal**.

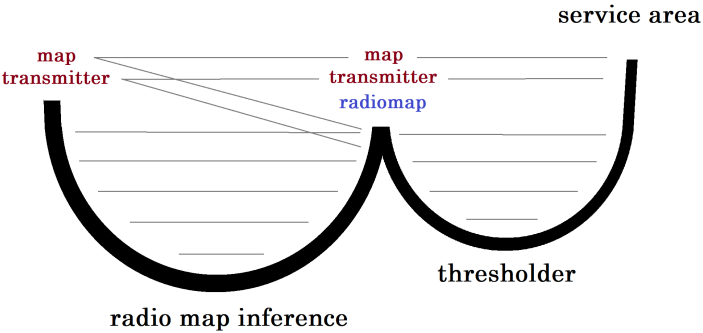


Freie Universität Berlin

RadioWNet

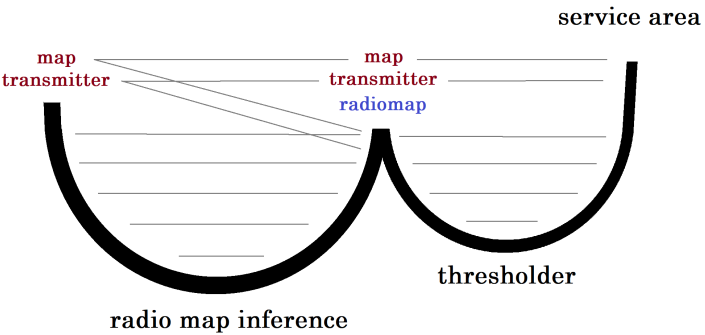

RadioWNet
service area
map
transmitter
map
transmitter
radiomap
thresholder
radio map inference

RadioWNet



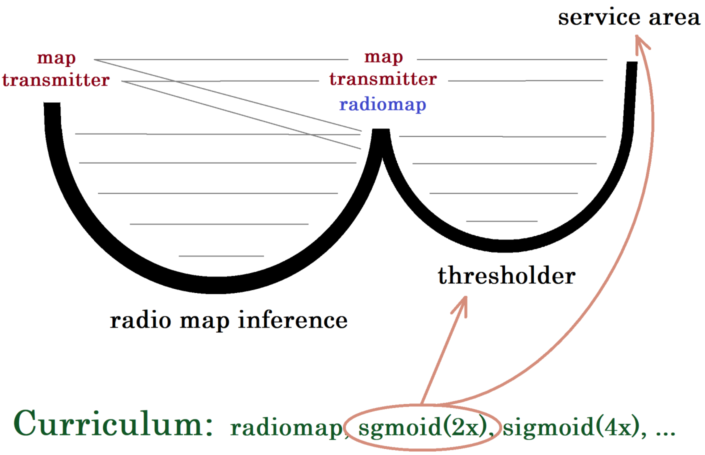
Curriculum: radiomap, sgmoid(2x), sigmoid(4x), ...

RadioWNet



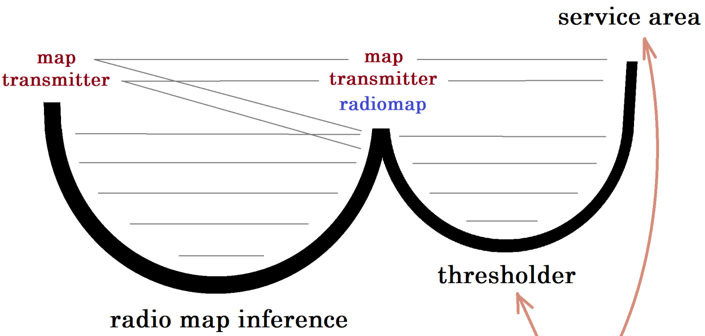
Curriculum: radiomap, sgmoid(2x), sigmoid(4x), ...

RadioWNet



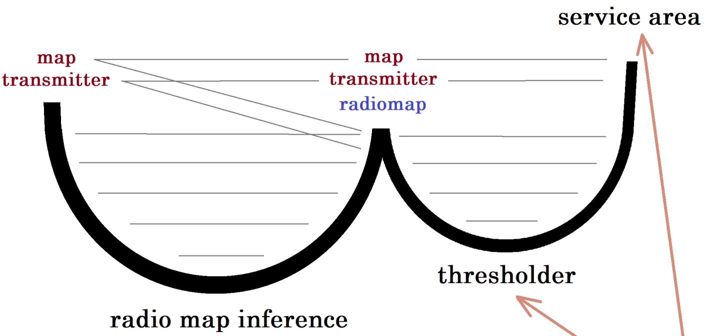
Curriculum: radiomap, sgmoid(2x), sigmoid(4x), ...

RadioWNet



Curriculum: radiomap, sgmoid(2x), sigmoid(4x), ...

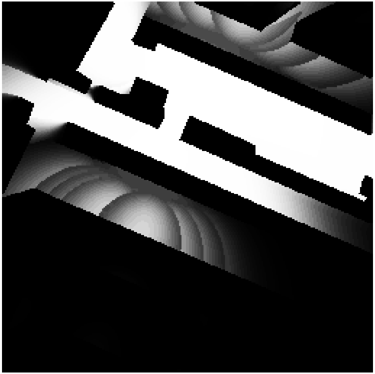
RadioWNet



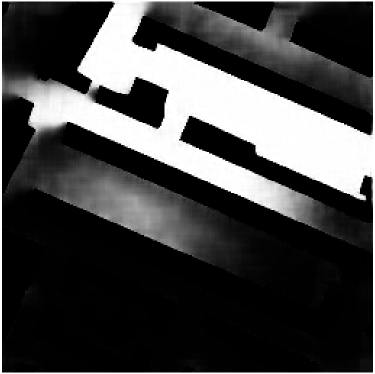
Curriculum: radiomap, sgmoid(2x), sigmoid(4x), ...

Examples: Service Area Classification (PDF l_2 error = 0.12)

Ground truth



Estimation

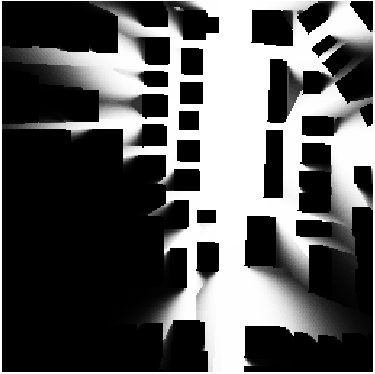


Freie Universität Berlin

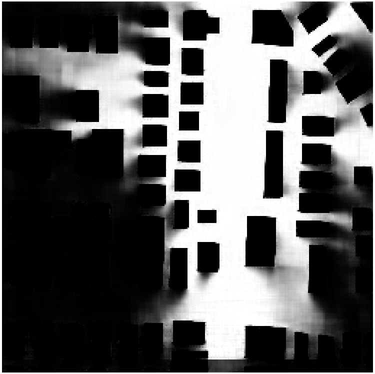
CoSIP

Examples: Service Area Classification (PDF l_2 error $= 0.12$)

Ground truth



Estimation



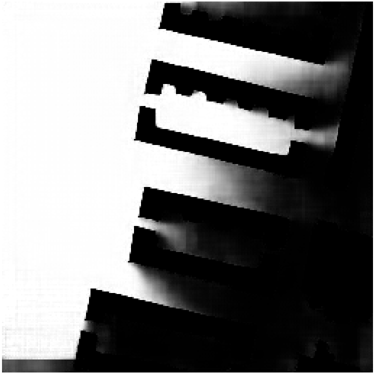
Freie Universität Berlin

CoSIP

Examples: Service Area Classification (PDF l_2 error $= 0.12$)

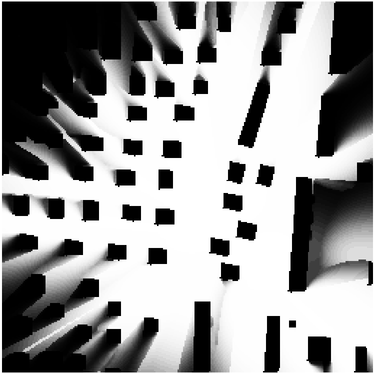
Ground truth

Estimation

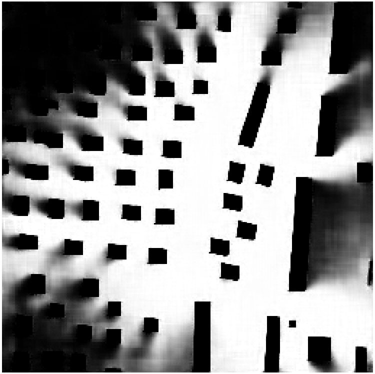


Examples: Service Area Classification (PDF l_2 error = 0.12)

Ground truth



Estimation



Freie Universität Berlin

CoSIP

Future work

- Dataset of sparse measurements.

Future work

- Dataset of sparse measurements.
- Supervised data driven link scheduling.

Future work

- Dataset of sparse measurements.
- Supervised data driven link scheduling.
- Fingerprint based localization.

Future work

- Dataset of sparse measurements.
- Supervised data driven link scheduling.
- Fingerprint based localization.
- More...

Outline

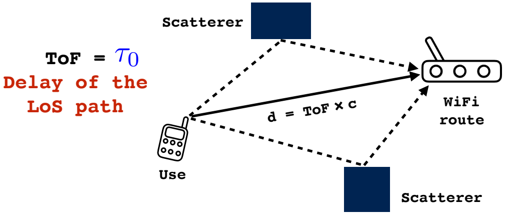
System Setup: Channel Model

- OFDM pilot transmission from a user to a Base Station over M frequency bands, each with N subcarriers.
- The **sparse** channel impulse response (CIR): $h(\tau) = \sum_{k=0}^{K-1} c_k \delta(\tau - \tau_k)$
- Channel frequency response (CFR) over band m: $\tilde{\mathbf{h}}^{(m)} \in \mathbb{C}^N$ where

$$\tilde{\mathbf{h}}_n^{(m)} = \sum_{k=0}^{K-1} c_k e^{-j2\pi f_{m,n}\tau_k}, \;\; n = -\frac{N-1}{2}, \ldots, \frac{N-1}{2},$$

where $f_{m,n}$ is the frequency of the n-th subcarrier of the m-th band.

- An example: indoor localization using raw WiFi pilot data



Freie Universität Berlin

System Setup: Distorted Measurements

- Carrier frequency offset, sampling frequency offset and packet detection delay distort the received signal.
- Phase-distorted and noisy received signal over band m:

$$\mathbf{y}_n^{(m)} = e^{-j\phi_{m,n}}\tilde{\mathbf{h}}_n^{(m)} + \mathbf{z}_n^{(m)},$$

where $\phi_{m,n} := 2\pi n f_s \delta_m + \psi_m$ is the affine phase distortion, where f_s is the subcarrier spacing, and $\mathbf{z}_n^{(m)}$ is the AWGN.

- The expression of receiver pilot measurements over all bands:

$$\mathbf{y} = \mathbf{\Phi}\tilde{\mathbf{h}} + \mathbf{z}$$

$\mathbf{\Phi} \in \mathbb{C}^{MN \times MN}$ is an unknown diagonal matrix, with unit-modulus entries.

Main Problem

Given the **distorted measurements** vector $\mathbf{y}$, estimate the **sparse** CIR $h(\tau) = \sum_{k=0}^{K-1} c_k \delta(\tau - \tau_k)$ and therewith the ToF plus the ranging distance.

CIR Estimation via Phase Retrieval

Solution Idea

Estimate the CIR by applying a **phase retrieval (PR)** algorithm to the **magnitude** of the measurements, i.e.

$$\mathbf{u}_i := |\mathbf{y}_i|^2 = |\tilde{\mathbf{h}}_i|^2 + \tilde{\mathbf{z}}_i, \quad i = 0, \ldots, MN-1.$$

In particular:

1. recover the CIR autocorrelation from the magnitude measurements.
2. recover the CIR from its estimated autocorrelation.
3. resolve ambiguities in the estimated CIR via handshaking.

- CFR magnitude corresponds to the Fourier transform of the autocorrelation:

$$|\tilde{\mathbf{h}}_i|^2 = \mathcal{F}\left\{R(\xi)\right\}|_{f=f_i}, i = 0, \ldots, MN-1$$

$$R(\xi) := h(\tau) \star h^*(-\tau)|_\xi = \sum_{k=0}^{K-1} \sum_{\ell=0}^{K-1} c_k c_\ell^* \delta(\xi - (\tau_k - \tau_\ell)), \;\; \xi \in [-\tau_{\max}, \tau_{\max}].$$

CIR Estimation via Phase Retrieval Cont'd

- **Step 1: Sparse Recovery of** $R(\xi)$
- The autocorrelation is sparse for $K \sim \mathcal{O}(\sqrt{MN})$.
- Approximate $R(\xi)$ on a dense, discrete grid

$$\mathcal{G} = \{\xi_0, \ldots, \xi_{G-1}\} \subset [-\tau_{\max}, \tau_{\max}], \quad G \gg MN.$$

- Estimate the sparse discretized autocorrelation vector $\mathbf{x}$ using LASSO:

$$\begin{aligned} \mathbf{r}^\star = &\underset{\mathbf{r} \in \mathbb{C}^G}{\text{minimize}} \;\; \tfrac{1}{2}\|\mathbf{A}\mathbf{r} - \mathbf{u}\|_2^2 + \lambda \|\mathbf{r}\|_1 \\ &\text{subject to} \;\; \mathbf{r} = \text{flip}(\mathbf{r})^*, \end{aligned} \tag{1}$$

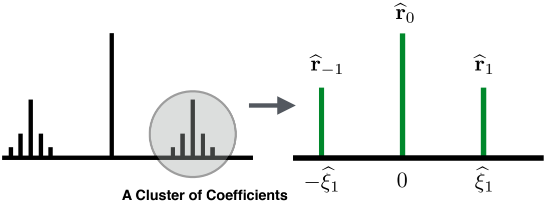
where $\lambda > 0$ is a regularization scalar, flip($\mathbf{r}$) is the flipped version of $\mathbf{r}$, and

$$[\mathbf{A}]_{k,\ell} = \frac{1}{\sqrt{MN}} e^{-j2\pi f_k \xi_\ell}, \quad k = 0, \ldots, MN-1, \; \ell = 0, \ldots, G-1.$$

- The constraint ensures conjugate symmetry of the solution, as expected for $R(\xi)$.

CIR Estimation via Phase Retrieval Cont'd

- In order to obtain exactly $\frac{K(K+1)}{2}$-sparse solutions, we run a k-means Alg. on the support of $\mathbf{r}^{\star}$.



- Alternative methods such as superresolution can be used to obtain the sparse autocorrelation, but are complex for large MN.
- The corresponding coefficients are obtained using simple Least-Squares.
- Eventually, the autocorrelation estimate is given as: $\widehat{R}(\xi) = \sum \widehat{\mathbf{r}}_i \delta(\xi - \widehat{\xi}_i)$

CIR Estimation via Phase Retrieval Cont'd

- **Step 2: Recovering the CIR $h(\tau)$ from the Autocorrelation Estimate**
- The autocorrelation support is the **difference set** of the support of the CIR.

$$\mathcal{D} = \boldsymbol{\Delta}\mathcal{T} = \mathcal{T} - \mathcal{T} \quad \text{(Minkowski subtraction)}$$

$$\mathcal{D} = \{\xi_i\}_{i=1}^{\frac{K(K+1)}{2}}, \quad \mathcal{T} = \{\tau_k\}_{k=0}^{K-1}$$

- We adopt a successive support and magnitude recovery to estimate the CIR support $\{\tau_k\}_{k=0}^{K-1}$ up to a **shift** and **conjugate reflection** and the path gains $\{c_k\}_{k=0}^{K-1}$ up to a **global phase shift** [1].

[1] G. Baechler, M. Kreković, J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Super resolution phase retrieval for sparse signals," arXiv preprint arXiv:1808.01961, 2018.

CIR Estimation via Phase Retrieval Cont'd

Algorithm 1 CIR Support Estimation

Input: Estimated autocorrelation support $\widehat{\mathcal{D}} = \{\widehat{\xi}_i\}_{i=1}^{\frac{K(K+1)}{2}}$
Initialize the sets $\mathcal{X}_2 = \{0, \widehat{\xi}_1\}$ and $\mathcal{P}_2 = \widehat{\mathcal{D}} \backslash \mathcal{X}_2$.
for $k = 2$ to $K-1$ **do**
    Select $\widehat{\xi}_i \in \mathcal{P}_k$ such that $\{\mathcal{X}_k \cup \widehat{\xi}_i\} - \{\mathcal{X}_k \cup \widehat{\xi}_i\} \subseteq \widehat{\mathcal{D}}$
    $\mathcal{X}_{k+1} = \mathcal{X}_k \cup \widehat{\xi}_i$
    $\mathcal{P}_{k+1} = \widehat{\mathcal{D}} \backslash \mathcal{X}_{k+1}$
end
$\mathcal{X}_K \leftarrow \mathcal{X}_K - \min\{\mathcal{X}_K\}$ ▷ Shift the solution set such that $\min\{\mathcal{X}_K\} = 0$.
Output: $\widehat{\mathcal{T}} = \mathcal{X}_K$

- **Coefficient estimation:** construct the matrix $\mathbf{C} \in \mathbb{R}^{K \times K}$ such that

$$\mathbf{C}_{k,\ell} \triangleq \begin{cases} 0 & k = \ell, \\ \log|\widehat{r}_{s(k,\ell)}| = \log|\widehat{c}_k| + \log|\widehat{c}_\ell| & k \neq \ell. \end{cases}$$

Once we have ordered delays, the index $s(k,\ell)$ is such that $\xi_{s(k,\ell)} = \tau_k - \tau_\ell$. Note that $\sum_{\ell=0}^{K-1} \mathbf{C}_{k,\ell} = (K-2)\log|\widehat{c}_k| + \sum_{\ell=0}^{K-1} \log|\widehat{c}_\ell|$ for $k = 0, \ldots, K-1$. Define $\beta \triangleq \sum_k \sum_\ell \mathbf{C}_{k,\ell} = 2(K-1)\sum_{\ell=0}^{K-1} \log|\widehat{c}_\ell|$. For $K > 2$, using these equations we can obtain $|\widehat{c}_k|$, $k = 0, \ldots, K-1$ as

$$\log|\widehat{c}_k| = \frac{1}{(K-2)} \left(\sum_{\ell=0}^{K-1} \mathbf{C}_{k,\ell} - \frac{\beta}{2(K-1)} \right).$$

Freie Universität Berlin   CoSIP   TU Berlin

CIR Estimation via Phase Retrieval Cont'd

- **Step 3: Resolving Ambiguities via Handshaking**
- The zero subcarrier in band $m \in [M]$ only contains the constant phase error term ψ_m with different signs at the transmitter and the receiver:

$$\mathbf{y}_{0,tx}^{(m)} = \tilde{\mathbf{h}}_0^{(m)} e^{j\psi_m} + \mathbf{z}_{0,tx}^{(m)}, \quad \mathbf{y}_{0,rx}^{(m)} = \tilde{\mathbf{h}}_0^{(m)} e^{-j\psi_m} + \mathbf{z}_{0,rx}^{(m)}$$

- Exchanging these measurements $\{\mathbf{y}_{0,tx}^{(m)}, \mathbf{y}_{0,rx}^{(m)}\}_{m=0}^{M-1}$ we have

$$\mathbf{y}'_m := \mathbf{y}_{0,tx}^{(m)} \mathbf{y}_{0,rx}^{(m)} = (\tilde{\mathbf{h}}_0^{(m)})^2 + \mathbf{z}'_m,$$

which is an **information used to resolve the ambiguities.**

CIR Estimation via Phase Retrieval Cont'd

- Consider two hypotheses about the CIR corresponding to two possible ambiguities:
  1. Time Shift: $H_1 : f_+(\tau;\tau_\epsilon) = \sum_{k=0}^{K-1} \widehat{c}_k \delta(\tau - \widehat{\tau}_k - \tau_\epsilon)$
  2. Time Shift + Time Reflection: $H_2 : f_-(\tau;\tau_\epsilon) = \sum_{k=0}^{K-1} \widehat{c}_k^* \delta(\tau + \widehat{\tau}_k - \widehat{\tau}_{K-1} - \tau_\epsilon)$
- Let $\mathbf{p}_+(\tau_\epsilon) = [F_+^2(f_{0,0};\tau_\epsilon), \ldots, F_+^2(f_{M-1,0};\tau_\epsilon)]^\mathsf{T}$ and $\mathbf{p}_-(\tau_\epsilon) = [F_-^2(f_{0,0};\tau_\epsilon), \ldots, F_-^2(f_{M-1,0};\tau_\epsilon)]^\mathsf{T}$ denote squared frequency samples of the tentative solutions for a shift τ_ϵ and define:

$$g(\tau_\epsilon|H_1) = \|\mathbf{p}_+(\tau_\epsilon) - \mathbf{y}'\|_2, \quad g(\tau_\epsilon|H_2) = \|\mathbf{p}_-(\tau_\epsilon) - \mathbf{y}'\|_2$$

- We select H_1 over H_2 if $\min_{\tau_\epsilon \in [0,\bar{\tau}]} g(\tau_\epsilon|H_1) < \min_{\tau_\epsilon \in [0,\bar{\tau}]} g(\tau_\epsilon|H_2)$, and H_2 over H_1 otherwise.
- In addition, the optimal value of the shift parameter τ_ϵ is given by

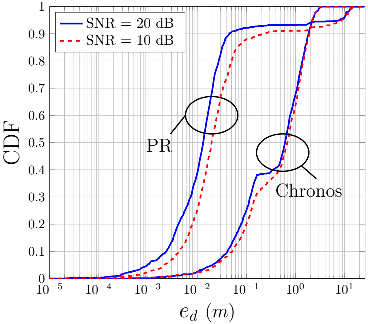
$$\tau_\epsilon^\star = \arg\min_{\tau_\epsilon} g(\tau_\epsilon|H_i),$$

  where H_i is the winning hypothesis.
- The time of flight is estimated as: $\boxed{\widehat{\text{ToF}} = \widehat{\tau}_0 = \tau_\epsilon^\star.}$

Freie Universität Berlin CoSIP

Simulation Results

- We compare our method to Chronos [2].
  - $M = 32$ adjacent bands, each with $N = 33$ subcarriers
  - $K = 3$ propagation paths, with complex Gaussian gains and uniformly random delays
- The ranging error is given by $e_d = |\tau_0 - \widehat{\tau}_0|c$, with c being the speed of light.



[2] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point." in NSDI, vol. 16, 2016, pp. 165-178.

Questions?