Cost-effective Signal Map Crowdsourcing with Auto-Encoder based Active Matrix Completion

Reporter: Chengyong Liu
Outline

- Background
- System Model
- Algorithms
- Experiments
- Conclusions
Signal map consists of signal strength at different locations.


Location-Based Services (LBS) Spectrum Monitoring
Traditional signal map construction

Full site survey

Professionals
Professional equipments

Interpolation reconstruction based on a small number of signals

Related Works

- KNN
- Gaussian Processes
- Compressive Sensing
- Matrix completion

Low accuracy
Complex model
Sparse property
Prior knowledge
• Smartphones are programmable and equipped with a set of cheap but powerful embedded sensors.
• Mobile phones are quite pervasive.

Crowdsourcing method to collect signals

Sensorly: https://www.sensorly.com/
OpenSignal: https://www.opensignal.com/
Outline

- Background
- System Model
- Algorithms
- Experiments
- Conclusions
System Model

\[
\begin{align*}
\min \quad & \varepsilon = error(\tilde{M}, M) \\
\text{s.t.} \quad & \tilde{M} = \Psi_{\text{old}}(O_{m \times n}) \\
& O = \Omega \cdot M
\end{align*}
\]
Signals fluctuate significantly during different times of day, and this fluctuation is non-linear.

Auto-encoder can learn nonlinear features in matrices.

Signals fluctuate significantly during different times of day, and this fluctuation is non-linear.

Auto-encoder can learn nonlinear features in matrices.

Auto-Encoder

Legend
- Missing data
- Available data
- Recovered data

Signals fluctuate significantly during different times of day, and this fluctuation is non-linear.

Auto-encoder can learn nonlinear features in matrices.

Signals fluctuate significantly during different times of day, and this fluctuation is non-linear.

Auto-encoder can learn nonlinear features in matrices.

---

For reconstruction algorithms, the signals at different locations have different effects on the reconstruction accuracy.

- The original image
- The 30% sampling rate
- The 40%(30%+10%) sampling rate
- The reconstructed image
- The reconstructed image
Active Crowdsourcing Scheme

Difference in reconstructed signal maps in different batches

The degree of changes in signals

\[ I^{i,j} = \text{abs}(x_{t-1}^{i,j} - x_t^{i,j}) \]

\[ I_{\text{initial}} = \text{abs}(\tilde{M}_0 - \text{mean}(M_{\text{his}})) \]
The Signal Dynamics

\[ s_x^2 = \frac{1}{k-1} \sum_{i=1}^{k} \frac{\omega_i \cdot ||x_i - \bar{x}||^2}{|\omega_i|} \]

Sample relationship with the population

Auto-Encoder

New Sampled Signals & Locations

Advanced
Experiment Setup

- The simulated WiFi indoor positioning dataset
  - The ray tracing technology generates 5000 signal maps with random changes of channel as historical signal maps
  - 50% missing rate
  - Signal maps from the same channel random variation as test data

- Baseline algorithms
  - BCS Model signal map reconstruction as a compressive sensing model
  - LmaFit A popular alternating least-squares method for matrix completion
**Experiment Results**

AER can achieve lower errors with higher probability

<table>
<thead>
<tr>
<th></th>
<th>5%</th>
<th>6%</th>
<th>7%</th>
<th>8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AER</td>
<td>0.0530±0.0011</td>
<td>0.0523±0.0000</td>
<td>0.0517±0.0011</td>
<td>0.0510±0.0013</td>
</tr>
<tr>
<td>BCS</td>
<td>0.0716±0.0000</td>
<td>0.0716±0.0000</td>
<td>0.0715±0.0000</td>
<td>0.0716±0.0000</td>
</tr>
<tr>
<td>MC</td>
<td>0.0908±0.0024</td>
<td>0.0859±0.0017</td>
<td>0.0831±0.0015</td>
<td>0.0810±0.0013</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>9%</th>
<th>10%</th>
<th>15%</th>
<th>20%</th>
<th>25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AER</td>
<td>0.0507±0.0015</td>
<td>0.0502±0.0015</td>
<td>0.0489±0.0020</td>
<td>0.0486±0.0022</td>
<td>0.0484±0.0020</td>
</tr>
<tr>
<td>BCS</td>
<td>0.0716±0.0000</td>
<td>0.0716±0.0000</td>
<td>0.0715±0.0000</td>
<td>0.0715±0.0000</td>
<td>0.0715±0.0000</td>
</tr>
<tr>
<td>MC</td>
<td>0.0797±0.0000</td>
<td>0.0786±0.0000</td>
<td>0.0752±0.0000</td>
<td>0.0714±0.0124</td>
<td>0.0611±0.0000</td>
</tr>
</tbody>
</table>

The relative error of AER is at least 2% lower than the other two algorithms
Experiment Results

Under the same number of acquisitions, the active method is far superior to other methods.

The proposed method can achieve more than 90% coverage.

![Graph showing experimental results](image_url)
Conclusions

A comprehensive solution for signal map construction

- The offline training phase
- The online reconstruct phase

An active crowdsourcing scheme for better performance

A more realistic signal map model with the description of the signal dynamics
Future Works

- Impact of different types of collection equipment on signal collection
- How to accurately determine the signal collection location of historical signals
- How to design an active mechanism more reasonably
Thanks for coming

Have a nice day!