IEEE MASS 2020
2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)

Fine-grained Urban Prediction via Sparse Mobile CrowdSensing

Wenbin Liu[1], Yongjian Yang[1], En Wang[1], and Jie Wu[2]

[1]*College of Computer Science and Technology, Jilin University, China*

[2]*Department of Computer and Information Sciences, Temple University, Philadelphia, USA*

Outline

Outline

I. **Background, Motivation, and Challenges**

II. Problem Formulation

III. Matrix Completion

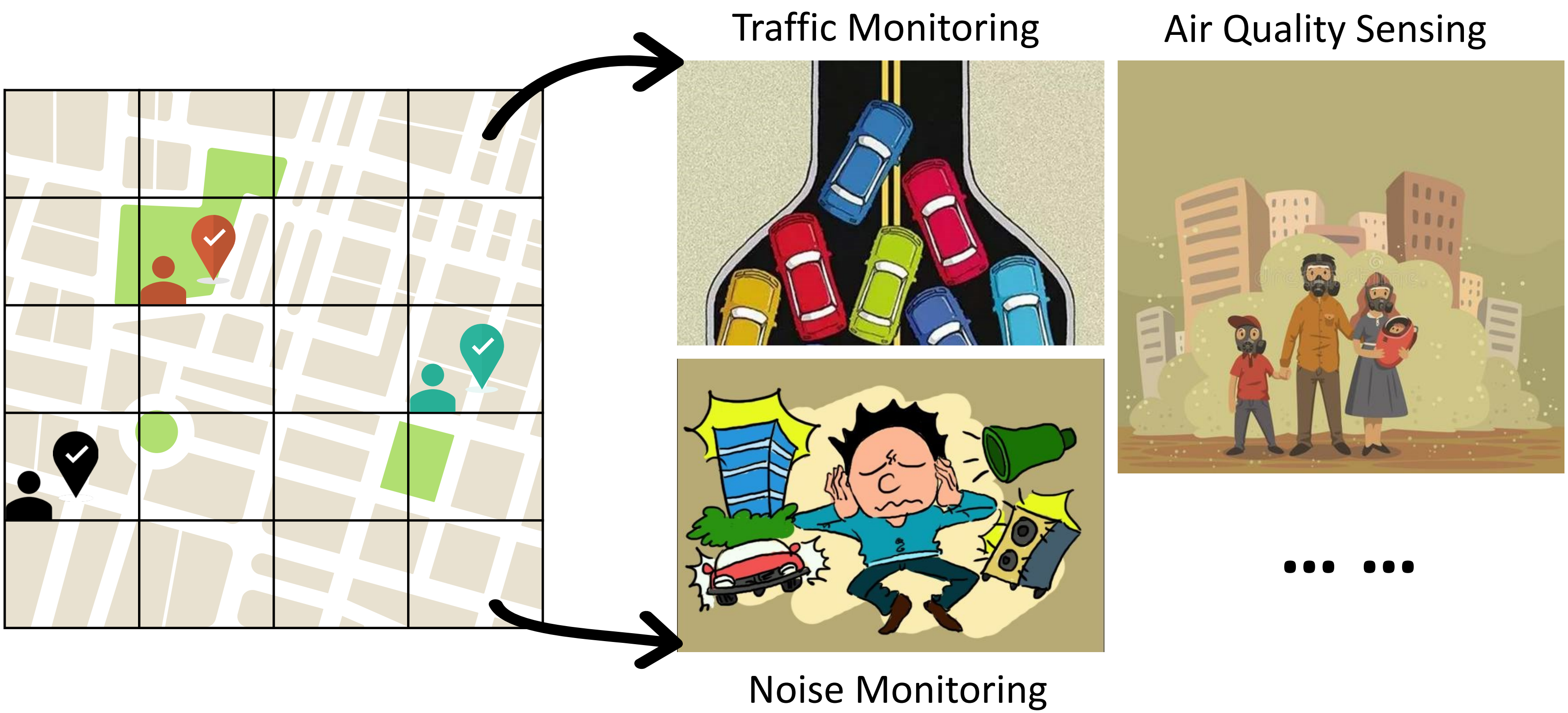
IV. Urban Prediction

V. Performance Evaluation

VI. Conclusion

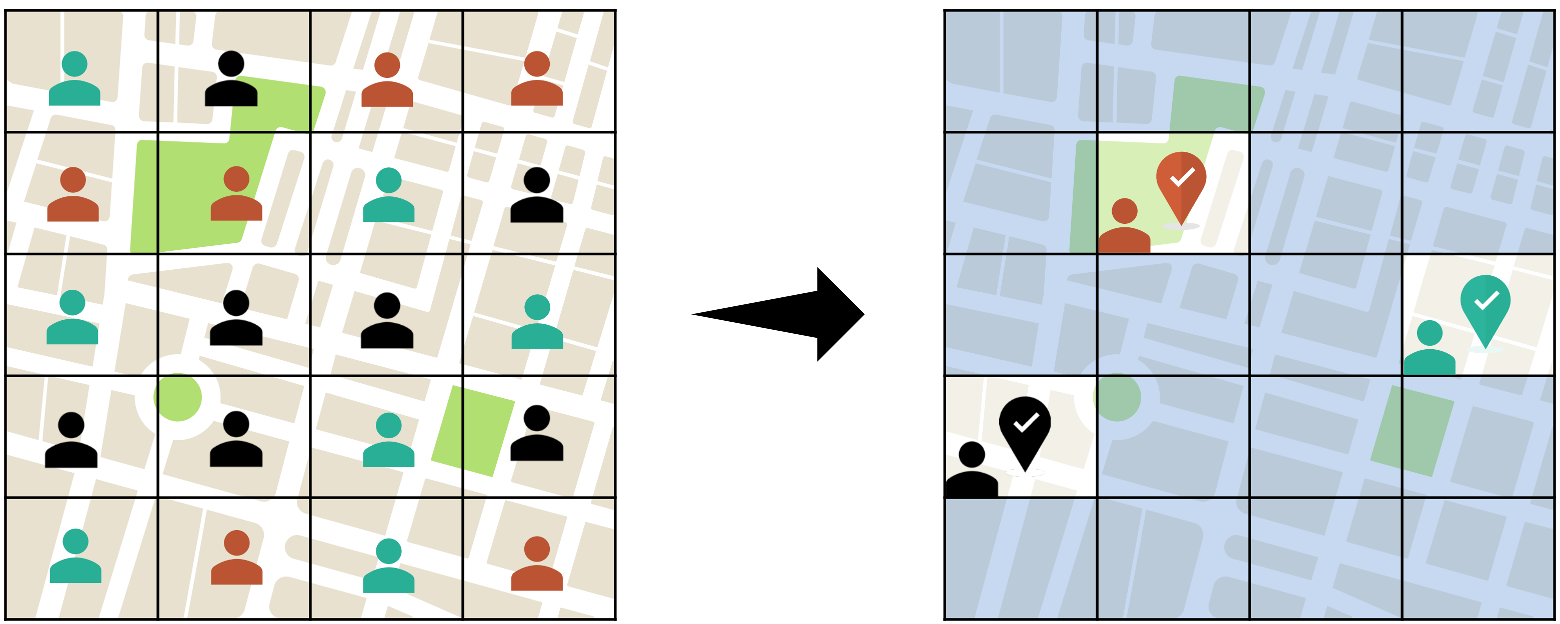
Mobile CrowdSensing (MCS)

- Recruit users to collect various urban data



Sparse Mobile CrowdSensing

- MCS: a large number of users
- Sparse MCS: sense a few subareas and infer the rest ones

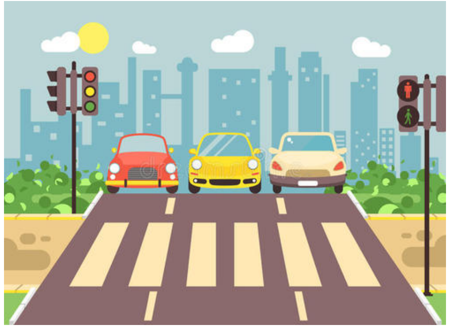


Sparse Mobile CrowdSensing

- Sparse MCS provides an effective way for urban sensing
  - ✓ Infer full map from sparse data
  - Data inference
    - Compressive Sensing
    - Matrix Completion
  - Subarea selection
    - Active Learning
    - Reinforcement Learning

Sparse Mobile CrowdSensing

- In some cases,
  - More interested in predicting **the future full map**
    - ✓ Rather than inferring the current data
- For example,
  - Traffic congestion or parking capacity monitoring
    - ✓ Users still need some time to drive there
    - ✓ Current data are not very important



Urban Prediction via Sparse MCS

- Infer the current → **predict the future**
  - ✓ Based on the sparse sensed data
- Two challenges:
  - ➢ How to utilize the sparse sensed data
    - Complete the matrix
    - Preserve the temporal-spatial correlations
  - ➢ How to capture the temporal-spatial correlations
    - Non-linear temporal relationships (among different cycles)
    - Pairwise spatial correlations (between two subareas)

Outline

I. Background, Motivation, and Challenges

II. Problem Formulation

III. Matrix Completion

IV. Urban Prediction

V. Performance Evaluation

VI. Conclusion

Problem Formulation

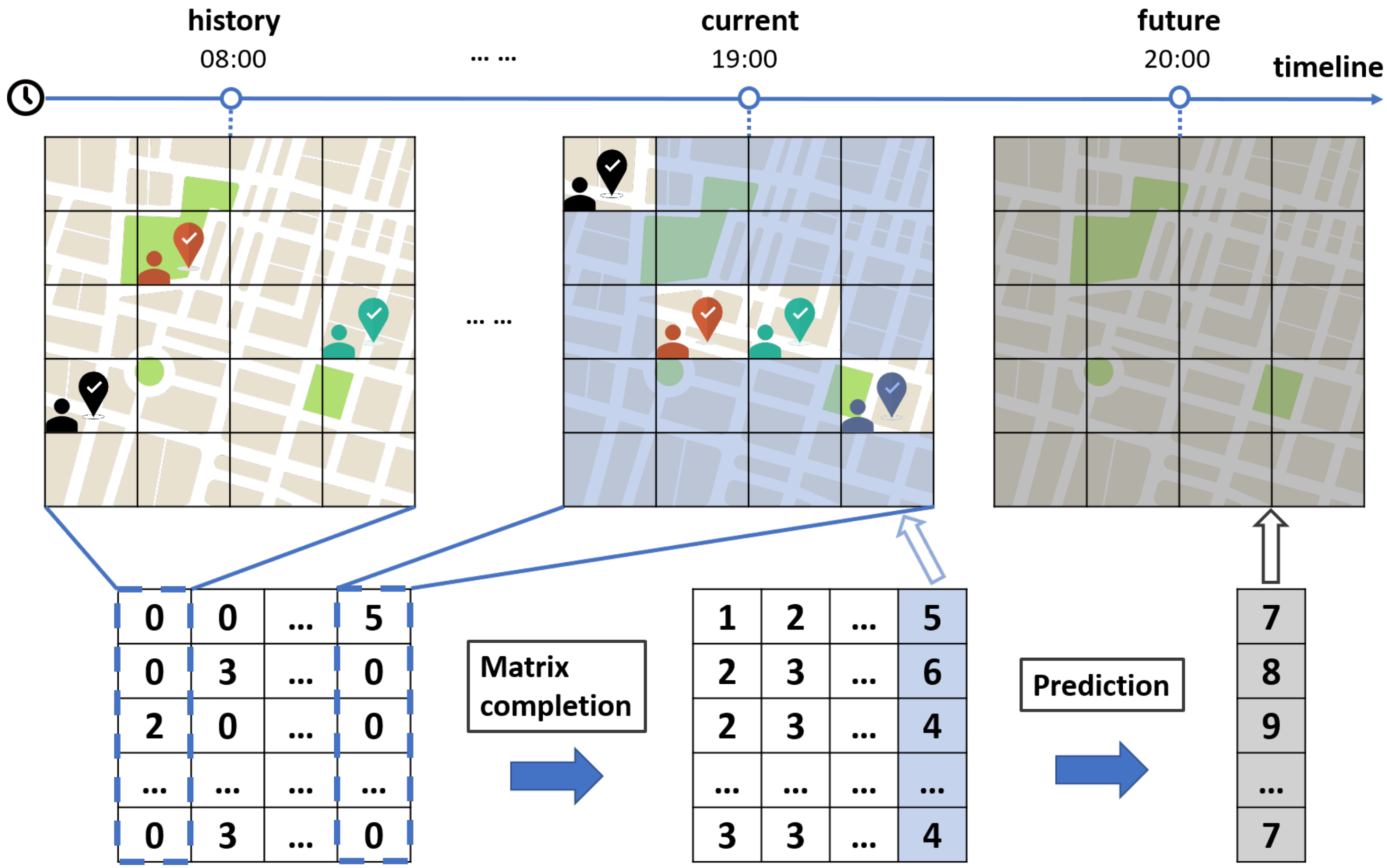
- **Problem** [Urban Prediction via Sparse MCS]:
- Given a MCS task with m subareas and n sensing cycles
  - for each cycle, sense data from a few subareas
  - predict the full maps of k future cycles
- Goal: minimizing the prediction errors

$$\min \quad \sum_{j=1}^{n-k} \varepsilon(y_{j+k}, \hat{y}_{j+k})$$

$$\text{s.t.} \quad p(m(Y_j'), k) = \hat{y}_{j+k}, \forall j \in \{1, 2, \ldots, n-k\}$$

Running example

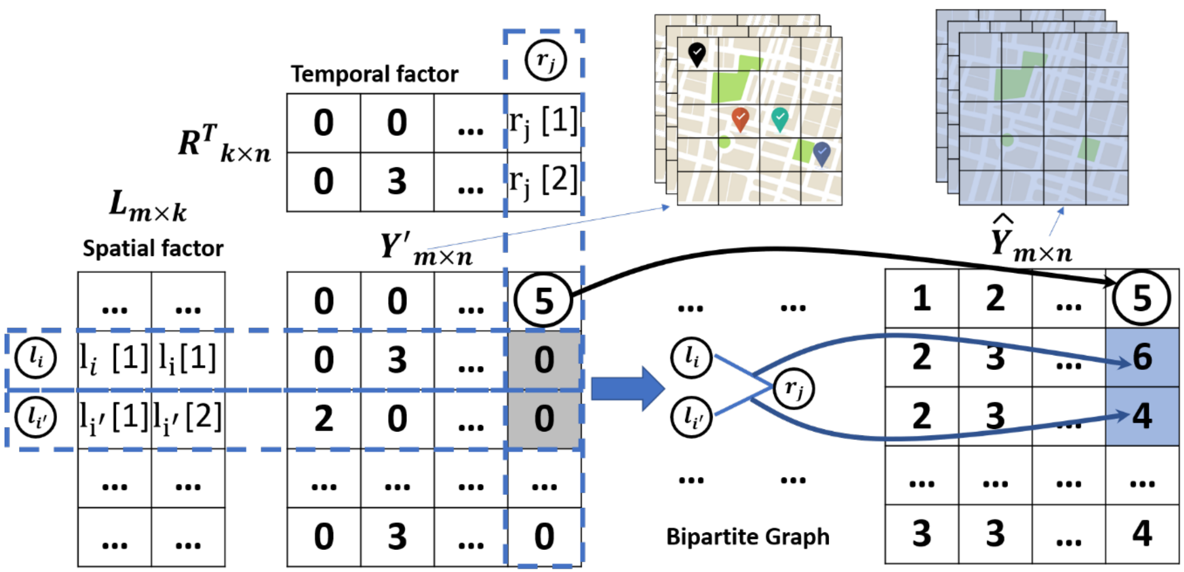
- An example of urban prediction via Sparse MCS



Outline

Matrix Factorization

- Sensing data exhibit strong temporal-spatial correlations
  - Sensing matrix Y usually has **the low-rank property**
- Given $rank(Y) = k$, we factor the inferred matrix $\hat{Y}$ into
  - a spatial factor matrix $L_{m\times k}$ and a temporal factor matrix $R_{n\times k}$



Temporal-Spatial Matrix Factorization

- Temporal and spatial constraint matrices, $\mathbb{T}$ and $\mathbb{S}$
  - Important and naturally occurring correlations
  - Help data inference and preserve the correlations
- $\mathbb{T}$ constraints that two continuously sensed data from the same subarea are usually similar
- $\mathbb{S}$ constraints that the data sensed from the closer subareas usually have the similar values.

$$\min \ \|Y' - \hat{Y} \bullet C\|_F^2 + \lambda_t \|\hat{Y}\mathbb{T}^T\|_F^2 + \lambda_s \|\mathbb{S}\hat{Y}\|_F^2,$$

Graph-based Matrix Completion

- 1. With $L_{m\times k}$ and $R_{(n-1)\times k}$, calculate the new r_n
- 2. Iteratively train and update the factor L and R

Algorithm 1 Graph-based Matrix Completion

Input: $Y'_{m\times n} = \{Y'_{m\times(n-1)}, y_n'^T\}$, $L_{m\times k} = \{l_1, l_2, ..., l_m\}$, $R_{(n-1)\times k} = \{r_1, r_2, ..., r_{n-1}\}$

Output: $\hat{Y}_{m\times n}$

Init r_n, $R_{n\times k} = \{R_{(n-1)\times k}, r_n\}$, $count = 0$;

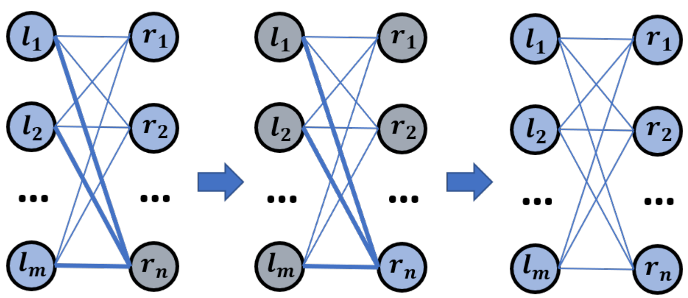
Build the linear system by using y_n', $L_{m\times k}$, and $R_{(n-1)\times k}$, and then calculate r_n;

while not convergent **and** $count <$ MAX_ITER **do**

    Fix $R_{n\times k}$ and treat $L_{m\times k}$ as unknown, build the linear system by using $Y'_{m\times n}$, $L_{m\times k}$ and $R_{n\times k}$, and then calculate and update $L_{m\times k}$;

    Fix $L_{m\times k}$ and treat $R_{n\times k}$ as unknown, build the linear system by using $Y'_{m\times n}$, $L_{m\times k}$ and $R_{n\times k}$, and then calculate and update $R_{n\times k}$, $\hat{Y} = LR^T$, and $count$++;

return $\hat{Y}$.



(1) New cycle (2) Caculate new r_n (3) Update all

Outline

I. Background, Motivation, and Challenges

II. Problem Formulation

III. Matrix Completion

IV. Urban Prediction

V. Performance Evaluation

VI. Conclusion

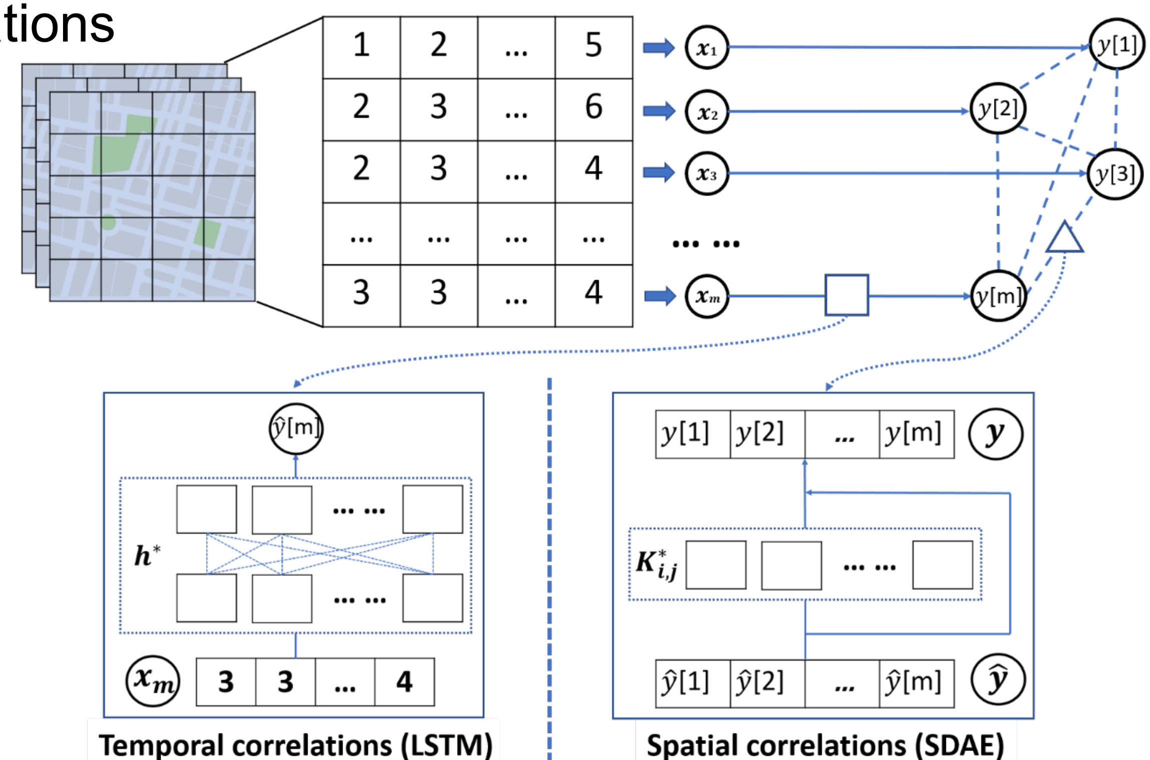
Urban Prediction

- Continuous Conditional Random Field (CCRF)
  1. relationships between the input and output data
     - ➢ Temporal relationships among different sensing cycles
     - ✓ Long Short-Term Memory (LSTM)
  2. correlations among the output data
     - ➢ Spatial correlations between different subareas
     - ✓ Stacked Denoising Auto-Encoder (SDAE)

Urban Prediction

- CCRF

1. LSTM for Temporal Relationships
2. SDAE for Spatial Correlations



1. Preliminary estimations by LSTM
2. Constrain and smooth by SDAE

Outline

I. Background, Motivation, and Challenges

II. Problem Formulation

III. Matrix Completion

IV. Urban Prediction

V. Performance Evaluation

VI. Conclusion

Three real-world data sets

- Sensor-Scope[1], U-Air [2], and TaxiSpeed [3]
  - Five typical urban sensing tasks:
    - ✓ Temperature, Humidity, PM2.5, PM10, and Traffic speed
  - Collected by static sensors
    - Can use mobile devices to collect the same data

[1] F. Ingelrest, G. Barrenetxea, G. Schaefer, M. Vetterli, O. Couach, and M. Parlange, "Sensorscope:application-specific sensor network for environmental monitoring," ACM Transactions on Sensor Networks, vol. 6, no. 2, pp. 1–32, 2010.

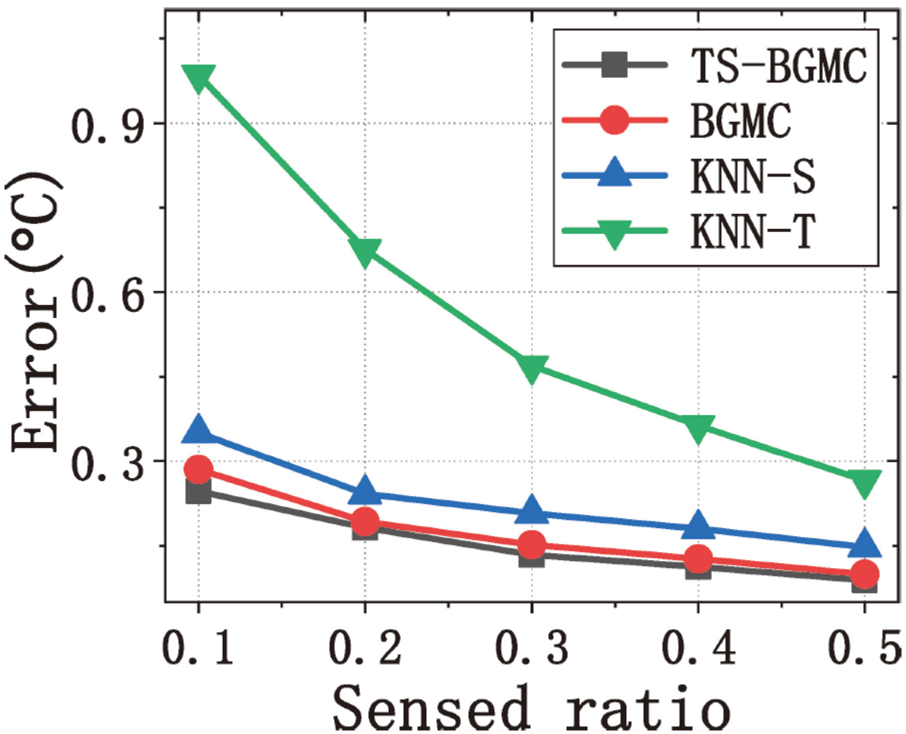
[2] Y. Zheng, F. Liu, and H. P. Hsieh, "U-air:when urban air quality inference meets big data," in ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2013, pp. 1436–1444.

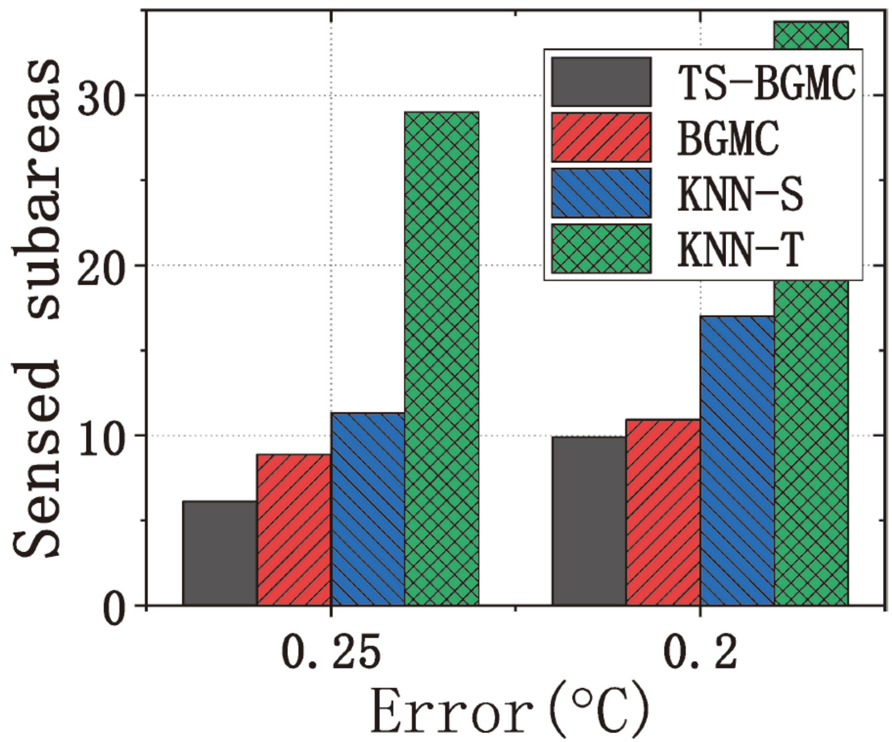
[3] J. Shang, Y. Zheng, W. Tong, E. Chang, and Y. Yu, "Inferring gas consumption and pollution emission of vehicles throughout a city," in ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, pp. 1027–1036.

Evaluation Results

- Temperature:

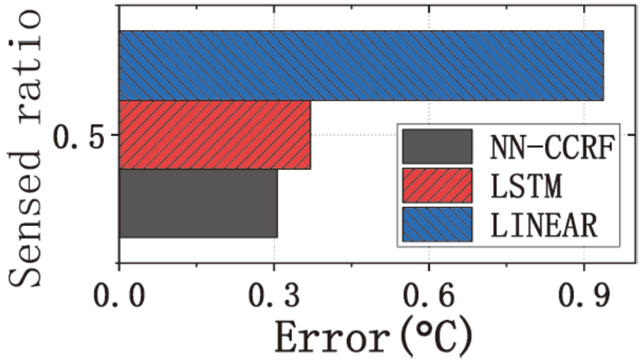
1) Inference accuracy; 2) Number of sensed subareas

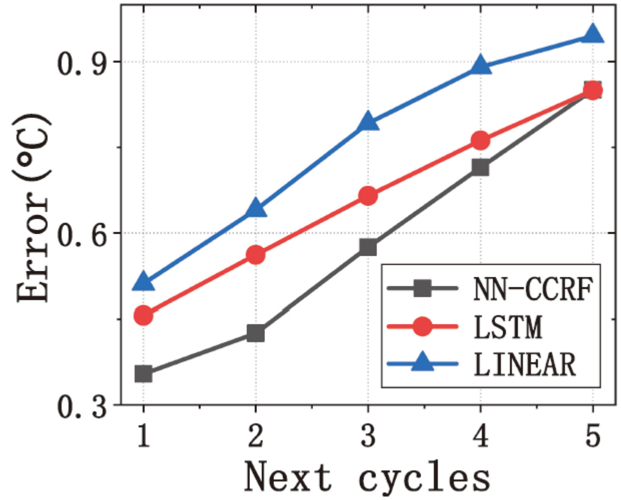




Evaluation Results

- 3) Sensed ratio; 4) Next cycles; 5) Running time





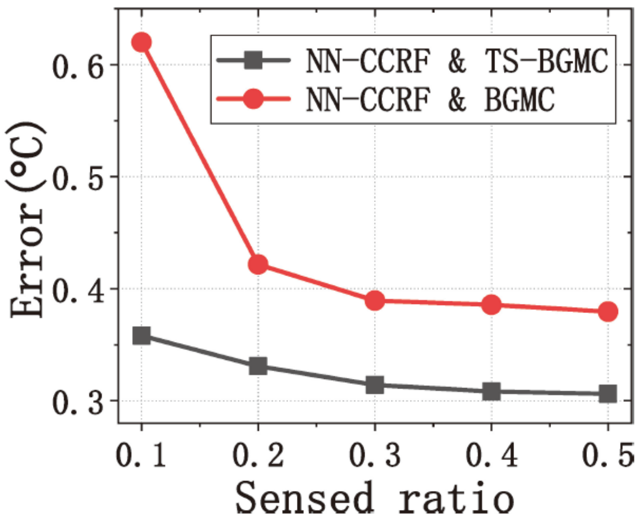


TABLE II: Running time for main methods

	Tem.	Hum.	PM2.5	PM10	Tra.
TS-BGMC	0.45s	0.45s	0.33s	0.34s	0.62s
LSTM	2.10ms	2.13ms	1.52ms	1.52ms	1.00ms
NN-CCRF	0.12s	0.12s	0.06s	0.06s	0.10s

Outline

I. Background, Motivation, and Challenges

II. Problem Formulation

III. Matrix Completion

IV. Urban Prediction

V. Performance Evaluation

VI. Conclusion

Conclusion

- Urban Prediction via Sparse Moblie CrowdSensing
  - Predict the future full map from sparse sensed data
- Matrix Completion with Temporal-Spatial constraints
  - Preserve temporal-spatial correlations
- Urban Prediction by Continuous Conditional Random Field
  - LSTM and SDAE for temporal and spatial correlations
- Extensive Evaluation
  - Three real-world data sets with five typical urban sensing tasks

Thank you!

Q&A