AutoEncoder based Active Signal Map Crowdsourcing

Chengyong Liu, Yanchao Zhao, Kun Zhu

College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics

Sheng Zhang

State Key laboratory for Novel Software Technology, Nanjing University

Jie Wu

Center for Networked Computing, Department of Computer and Information Sciences, Temple University

Abstract

Signal maps, which consist of signal strengths at different locations, play an important role in site spectrum monitoring[1], LBS[2] and network construction[3]. But constructing the signal maps through on-site surveys, is time-consuming and laborious. And the signal may change with time and surroundings, which leads to a dilemma (Figure 1) that the acquisition and updating of the signal map could hardly meet the requirements of the applications, no matter in quality or timing.

Crowdsourcing schemes are proposed to address this issue, but collecting signals through crowdsourcing suffers from random and insufficient participants, which leads to incomplete or low-quality measurements.

In this work, we study how to effectively reconstruct and update the signal map in the case of partially measured signal maps and propose an auto-encoder-based active signal map reconstruction method (AER).

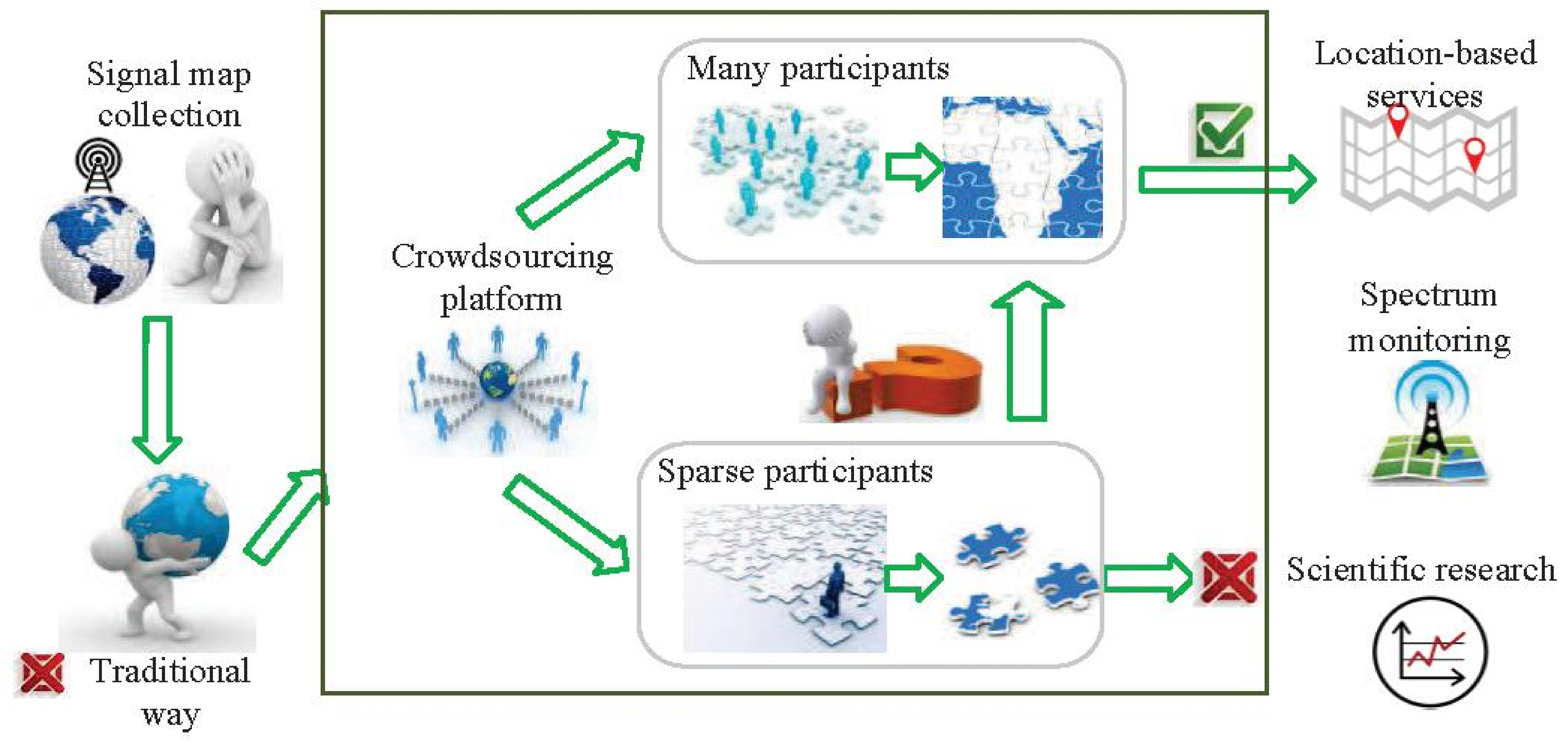


Figure 1: Application scenario and the dilemma

Methodology

In order to solve these problems in a comprehensive way, we propose a novel signal map reconstruction algorithm based on auto-encoder based active matrix completion.

In this work, we adopt the auto-encoder to learn the signal correlation existing in the historical signal and effectively reconstruct the signal map. The algorithm is divided into the offline training phase and the online update phase (as shown in Fig. 2).

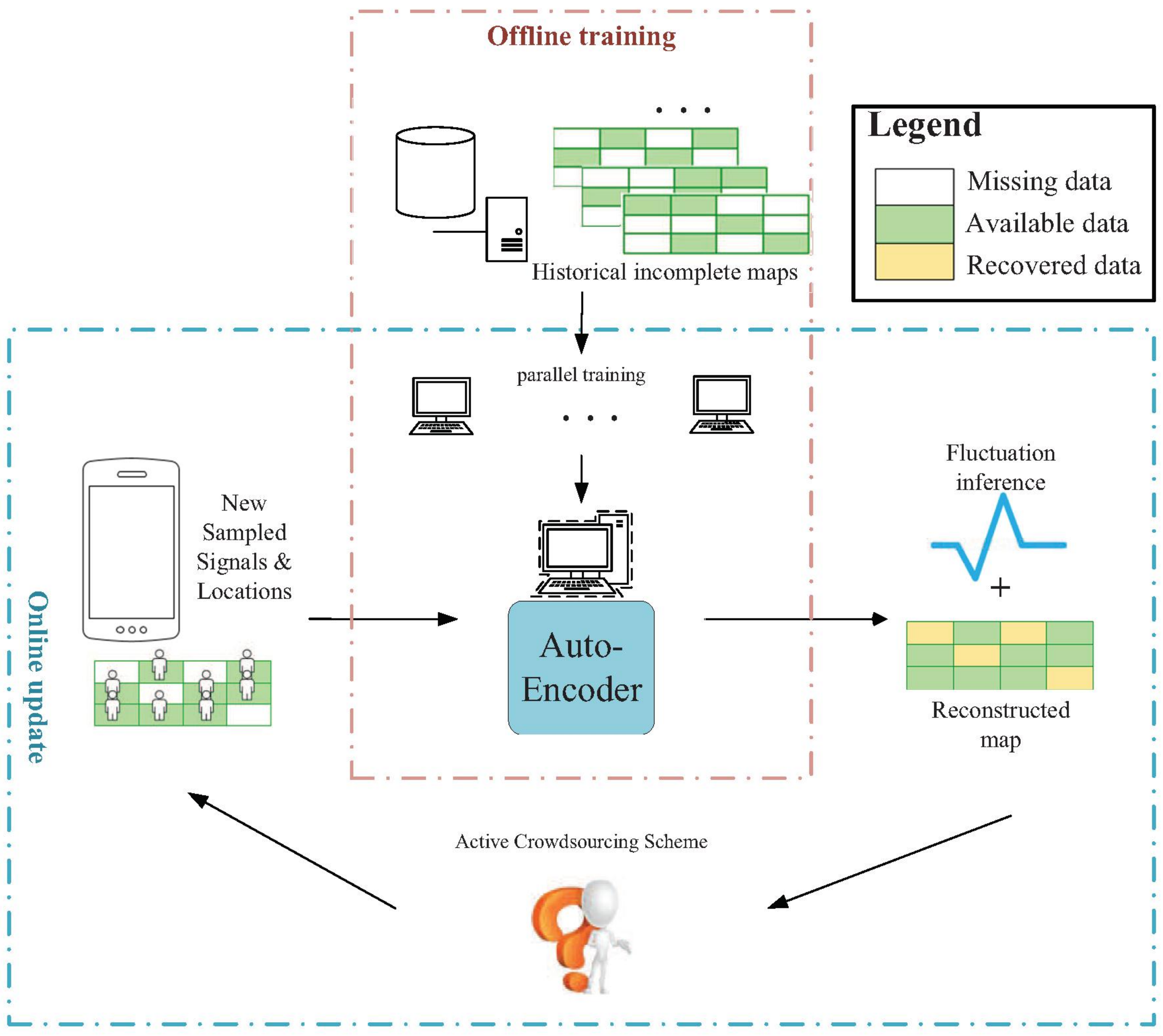


Figure 2 :The workflow of the AER.

Offline training phase

Faced with a large amount of incomplete historical data, we extend the training of auto-encoder with parallel training method.

- We divide the massive crowdsourced historical signals and then use each compute node to train part of the fragmentation data.
- We combine the gradient parameters generated by multiple compute nodes and update them through the parameter server to train the auto-encoder.

Online update phase

We use the uploaded signal to reconstruct the signal map in time.

- The AER takes the partial signal collected at time t as input and reconstructs the signal map using the previously learned features.
- A fluctuation inference method infers the fluctuation range of the reconstructed signal map.
- An active crowdsourcing method finds the signal with the highest informativeness.

References

[1] Yang Bo, Suining He, and S. H. Gary Chan. Updating wireless signal map with bayesian compressive sensing. In Proc. of Acm MSWiN, 2016.

[2] Simon Yiu, Marzieh Dashti, Holger Claussen, and Fernando PerezCruz. Wireless rssi fingerprinting localization. Signal Processing,131:235–244, 2017.

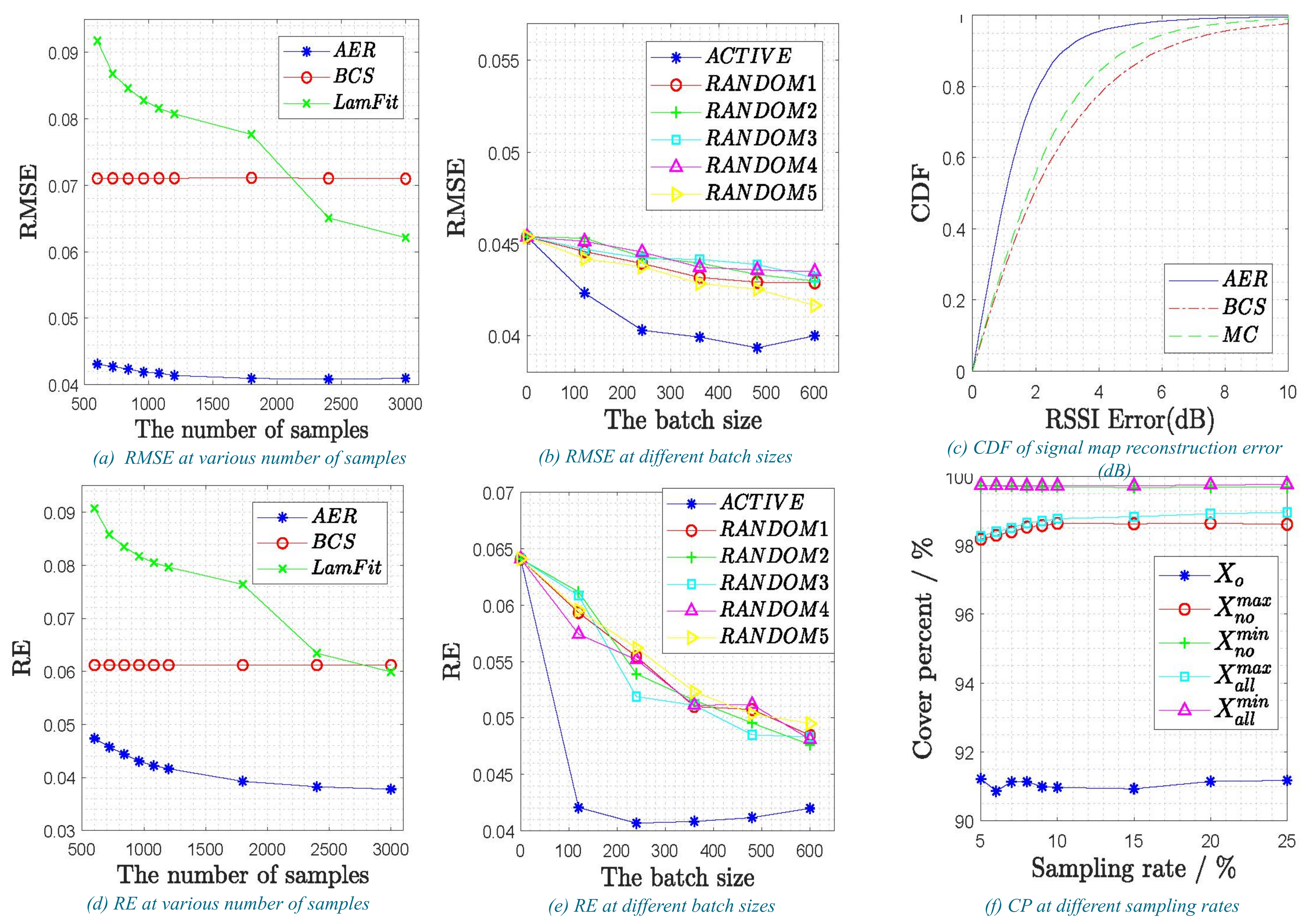
[3] Zhijing Li, Ana Nika, Xinyi Zhang, Yanzi Zhu, Yuanshun Yao, Ben Y. Zhao, and Haitao Zheng. Identifying value in crowdsourced wireless signal measurements. In Proc. of Springer WWW, 2017.

[4] Feng Yin and Fredrik Gunnarsson. Distributed recursive gaussian processes for rss map applied to target tracking. IEEE Journal of Selected Topics in Signal Processing, 11(3):492–503, 2017.

[5] Zaiwen Wen. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. Mathematical Programming Computation, 4(4):333–361, 2012.

[6] S. He and K. G. Shin. Steering crowdsourced signal map construction via bayesian compressive sensing. In Proc. of IEEE INFOCOM, pages 1016–1024, April 2018.

Results



(a) RMSE at various number of samples *(b) RMSE at different batch sizes* *(c) CDF of signal map reconstruction error (dB)* *(d) RE at various number of samples* *(e) RE at different batch sizes* *(f) CP at different sampling rates*

Figure 3 :Experiment Result (The upper part is the result of no noise and the lower part is the result of noise.)

Relative Error

From Fig.3(a) and Fig.3(d), we can clearly see that our algorithm is far superior to the other two algorithms. And we find that the Relative Error of our algorithm is 1% lower or even 2% lower than the error of BCS, and AER is much better than MC. Then we compare the cumulative distribution of errors in Fig.3(c). We can see that the performance of our proposed algorithm far exceeds the other two algorithms. AER's RSSI error is about 90% less than 3.5dB, MC's RSSI error is about 90% less than 5dB, and BCS's RSSI error is about 90% less than 5.5dB.

Active Crowdsouring Scheme

For the active crowdsourcing problem, we found that the signal obtained by the active method is much better than the random sampling method, and under the same errors, the active AER only needs less than half of the number of random sampling in Fig.3(b) and Fig.3(e).

Fluctuation Estimation

We use the fluctuation estimation algorithm to estimate the signal fluctuation range, making the reconstructed signal map more practical. From Fig.3(f), we can see that even at the lowest sampling rate, we can still guarantee that more than 90% of the signal range contains the ground-truth of the signal.

Conclusion

In this paper, we propose a comprehensive solution for signal map acquisition, where auto-encoder is used to learn the nonlinear features of and compose an algorithm called auto-encoder(AER) firstly. The AER can effectively utilize historical incomplete signal maps collected and learn the nonlinear temporal features therein and effectively reconstruct the signal map; Secondly, we propose an active crowdsourcing scheme for better performance of AER. This method can reveal the more valuable measurement sites for reconstruction algorithm and effectively reduce the reconstruction error with lower crowdsourcing budget. Finally, we also propose a more realistic signal map model with the description of the signal dynamics in the same location over time, and correspondingly, an extended AER algorithm is proposed to solve the reconstruction problem on this model. The simulation experiments results demonstrate the effectiveness of our solution.