Review

Not peer-reviewed version

Data-Aided Sensing for Semantic Communication: A Brief Survey

Muhammad Awais *

Posted Date: 9 September 2024

doi: 10.20944/preprints202409.0520.v1

Keywords: semantic communication; Data-aided Sensing; Communication

Review

Data-aided Sensing for Semantic Communication: A Brief Survey

Muhammad Awais

Dept. of Electronics and Information Convergence Engineering, Kyung Hee University; mawais@khu.ac.kr

Abstract: Recent advancements in semantic communication (SC) have shown the potential to surpass the limitations of the Shannon theorem, by leveraging an understanding of the meaning of data instead of the data itself. In networks, data-aided sensing (DAS) presents a powerful tool that can significantly reduce communication overhead, aligning perfectly with the goals of SC by providing more powerful semantic messages. This survey explores the synergy between DAS and SC and investigates how their convergence can lead to more efficient and effective communication.

Keywords: semantic communication; Data-aided Sensing; Communication

---

I. Introduction

Recent studies have shown that semantic communication (SC) can efficiently provide high data rates and bandwidth for future 6G communications. In SC, only the meaning of the data is transmitted instead of the entire message, and the receiver decodes it to obtain complete information. SC helps reduce overheads by using only relevant information and discarding irrelevant and redundant data. Additionally, SC improves data reliability by decoding only the message's meaning, resulting in robustness against errors.

SC has garnered significant attention in the research community, with many still exploring its potential applications. For example, joint source and channel coding (JSCC) [3] via SC, proposed to reduce overheads and improving transmission performance through deep learning. By utilizing SC for context-aware resource allocation, we can analyze significance of data, resulting in a more efficient distribution of resources. Moreover, SC has been shown to enhance security and privacy in communication [4] by encrypting semantic messages instead of complete raw data. Furthermore, SC is believed to significantly enhance machine learning performance and enable a better understanding of different data types, resulting in more accurate and efficient performance.

To leverage full potential of SC, proper analysis is required to understand the meaning of received semantic messages, which depends on the quality of the data collected. In traditional IoT sensor networks, this data is collected by fixed or random sequences. However, in practice, data that provides more information is more critical than other data and changes depending on the underlaying cause.

Recently proposed data-aided sensing (DAS) [5] is a cutting-edge approach for collecting large scale sensor data that can revolutionize the field. In DAS strategically selecting sensors based on previously collected data, the base station (BS) can efficiently gather the necessary information with fewer measurements than traditional methods Figure 1. This adaptability ensures that the selection process caters to specific applications, making DAS ideal for optimizing SC analytical performance, especially in situations with limited bandwidth. Moreover, DAS-based data collection leads to better semantic messages due to access to more information from partial sensor nodes, enhancing network performance.

 doi:10.20944/preprints202409.0520.v1

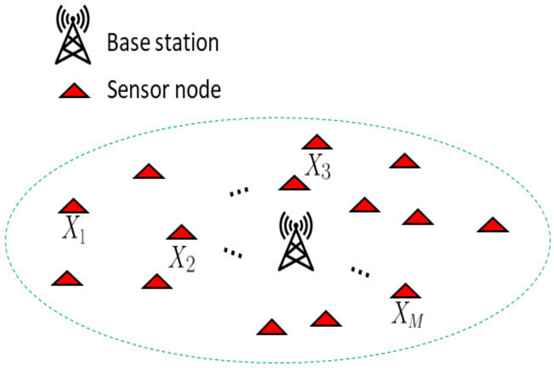

Figure 1. General system model of sensor network for data collection.

Leveraging DAS in SC ensures better network efficiency by providing robust semantic messages containing more critical information from partial sensors data.

II. Data-aided Sensing (DAS) for efficient SC

DAS was initially proposed in [5], where mean square error (MSE) based interpolation was used for sensor node selection from the sparse network while using the compressed sensing based iterative process for data collection, where they showed that compressive random access could increase data collection efficiency. In [6], the mean and covariance matrix are assumed to be known, and using minimum MSE, the next sensor is selected for correlated Gaussian sensor nodes. At the same time, DAS with multiarmed bandit problem was studied for sensor selection in limited available power.

In [7], two types of DAS were studied for large-scale sensors and real-time IoT applications, first centralized DAS, where BS made sensor selection decisions based on collected data, and second distributed DAS, where BS does not have knowledge of local data and following sensor is selected by optimization using random access probabilities with limited bandwidth.

Further, maximum variance [8] is applied for DAS in Gaussian process regression (GPR) to learn complete data from a small subset of total measurements in IoT systems. In [8], authors also generalized DAS for multichannel ALOHA in distributed data uploading from sensors where sensors can make uploading decisions based on the feedback of BS prediction.

In the research presented in [9], J-divergence was employed as a metric for reliable decision-making in DAS when dealing with correlated data. This approach aimed to select the most suitable sensor node for data collection. Meanwhile, the study outlined in [10] utilized entropy criteria to schedule energy-efficient unmanned aerial vehicles (UAVs) for collecting data related to specific queries within a distributed antenna system. A common assumption in both of these works was that the data distribution follows a Gaussian pattern. However, it’s important to note that real-world data may not always conform to this Gaussian assumption.

III. Convergence of SC and DAS via deep learning

Recently, deep learning-based approaches [11,12] for DAS were proposed for multicausal data. In [11] a β-VAE based model is used to find respective cause and complete data representations from partial data using maximum a-posterior (MAP) in latent space. This selected cause/cluster facilitates the process for the next sensor selection in DAS. The work of [11] is extended further in [12] by using a more advanced Generative deep learning network named variational deep embedding (VaDE) and proposed DAS based on mutual information via two methods named reference augmented sensing and causewise augmented sensing. This work utilized the generative capability of the VaDE network using proposed method to reproduce complete data in SC while improving DAS.

A comparison of mentioned DAS methods is provided in the Table 1.

Table 1. A comparison of DAS based methods.

Ref	DAS criteria	Single/Multiple	Performance metric	Methodology
[5]	MSE interpolation	Single cause	Number of errors, MSE	Compressive random access
[6]	Covariance matrix, MMSE	Multicausal	Cause selection probability, MSE	Random access, Multiarmed bandit
[7]	MSE interpolation	Single cause	Error norm, MSE	Centralized/Decentralized DAS (Random access)
[8]	Maximum variance	Single cause	MSE, sum of squired error	Gaussian process regression, Multichannel ALOHA
[9]	MSE/Entropy	Single cause	Log-likelihood ratio	J-divergence
[10]	Entropy	Single cause	Spectral efficiency, amount of collected data	Query based entropy optimization
[11]	MAP, MSE	Multicausal	Cause sensing accuracy, Normalized MSE	Deep learning (β-VAE)
[12]	Mutual information	Multicausal	Cause sensing accuracy, Normalized MSE	Deep learning (VaDE)

IV. Conclusion

This paper investigated the convergence of semantic communication and DAS. By leveraging the strengths of both approaches, combined frameworks have the potential to revolutionize communication systems. SC's ability to exploit data meaning and DAS's strategic data collection can significantly reduce communication overhead, leading to a new paradigm for more effective and resource-conscious information transmission.

References

1. E. C. Strinati, S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications", in *Elsevier Comput. Netw.*, vol. 190, pp. 107930, May 2021.
2. W. Yang et al., "Semantic communications for future internet: Fundamentals, applications, and challenges," in *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 213-250, First-quarter 2023.
3. J. Xu, T. -Y. Tung, B. Ai, W. Chen, Y. Sun and D. Gündüz, "Deep joint source-channel coding for semantic communications," in *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 42-48, Nov. 2023.
4. S. A. Moqurrab, A. Anjum, N. Tariq and G. Srivastava, "Instant_Anonymity: A lightweight semantic privacy guarantee for 5G-enabled IIoT," in *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 951-959, Jan. 2023.
5. J. Choi, "A cross-layer approach to data-aided sensing using compressive random access," in *IEEE Internet Things J.*, vol. 6, no. 4, pp. 7093-7102, Aug. 2019.
6. J. Choi, "Gaussian data-aided sensing with multichannel random access and model selection," in *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2412-2420, Mar. 2020.
7. J. Choi, "Data-aided sensing where communication and sensing meet: An introduction," *IEEE Wireless Commun. Netw. Conf. Workshops* (WCNCW), Seoul, S. Korea, Mar. 2020, pp. 1-6.
8. J. Choi, "Data-aided sensing for Gaussian process regression in IoT systems," in *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7717-7726, May 1, 2021.
9. J. Choi, "Data-aided sensing for distributed detection," in *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1138-1141, May 2021.
10. M. Nemati, S. R. Pokhrel and J. Choi, "Modelling data aided sensing with UAVs for efficient data collection," in *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1959-1963, Sept. 2021.
11. M. Awais, J. Choi, J. Park, Y. H. Kim, "Data aided sensing with variational autoencoder", in *Int. Conf. Artif. Intell. Info. Commun.* (ICAIIC), Osaka, Japan, Feb. 2024, pp. 120–122.
12. M. Awais, J. Choi, J. Park, Y. H. Kim, "Intelligent data-aided semantic sensing with variational deep embedding", in *Elsevier ICT Express*, vol. 10, no. 4, pp. 824–830, 2024.

 doi:10.20944/preprints202409.0520.v1