Data Quality and the Internet of Medical Things (IoMT)

Arshad Farhad
School of Computing and Mathematics
Keele University, UK
a.farhad@keele.ac.uk

Sandra Woolley
School of Computing and Mathematics
Keele University, UK
s.i.woolley@keele.ac.uk

This paper summarises a programme of PhD research relevant to federated machine learning for IoMT-informed patient care. Sensing performance, data quality and stakeholder perceptions are key to the clinical outcomes that can be achieved in future healthcare systems. However, IoMT data quality challenges are complex and multi-faceted, and could significantly impact clinical decisions that depend on accurate and timely data. The paper outlines the research context and challenges, summarises progress and candidate research questions, and discusses potential solutions, and topics for future research in this area.

Keywords: Internet of Medical Things (IoMT), Data Quality, Federated Learning (FL)

1. INTRODUCTION

The devices connected to the internet are more than ever and this will continue to increase exponentially. The Internet of Things (IoT) can be defined as an ecosystem of identifiable internet devices that collect, transmit, and sometimes store data. In healthcare, devices and sensors that collect health-related data from individuals are considered comprise an Internet of Medical Things (IoMT). These IoMT devices include wearables, mobile phones, smart bands, digital medications, implantable surgical devices, and other portable devices.

The future adoption of IoT in healthcare depends on data being collected and secure transmission for processing. Because of the nature of the healthcare vertical, clinicians are always concerned about the trustworthiness of data. Data trustworthiness refers to the data used as the basis for making decisions that are the right for the right patient. Hence, the data trustworthiness is an important consideration for both, the health experts and the technology experts who believe in future IoT-health innovation. The topic of data quality is the primary concern and possibly 'make and break' for any IoT-based medical solution.

2. RESEARCH CHALLENGES

Clinical decisions can be negatively impacted by poor data quality. Several factors which can degrade the quality of data in an IoT context include deployment scale, resource constraints, fail-dirty, security vulnerability, and privacy preservation processing. These manifest differently at different stages of the AI models which are applied for clinical decision capacity. For example, during data collection, data quality may be affected by sensor fault, or environmental factors; during data communication and pre-processing, network security may impact data quality, and factors such as privacy and security preservation processing affect data quality during storage and use. There is a need to evaluate data quality at each stage and score them, store such scores or add them as metadata and combine them into a single metric that can be made available to data consumers in real-time.

Introducing IoT systems, involving advice from AI to support the clinical decision, requires more than just functionality. There is a need to increase the trust of users in the reliability and accuracy of data and as AI moves from the currently acceptable narrow intelligence directed by clinician-determined action plans to a future in which advice is generated by the IoT system. Our technologically adept participants are not yet ready for this step; research is needed to ensure that technological capability does not outstrip the trust and data quality concerns of the individuals using it.

3. IoT AND DATA QUALITY

In the context of the IoT in healthcare, data quality issues lead to severe consequences on patients' delivery of care and are one of the reasons for its limited adoption in the healthcare industry. We researched and studied the IoT data quality issues
in a healthcare environment and compiled them in the next section which has been referenced by other researchers.

Duplication refers to whether the data sample has been collected more than once due to a defective or faulty device could be the reason among a few others (Perez-Castillo, et al., 2018).

Errors in data could be a result of environmental factors such as temperature, humidity, etc., sensor misplacement and motion artefacts (Ray et al., 2021), signal interference or mistakes in capturing or processing the stream of data. Additionally, the age of the sensor can lead to generating erroneous data (Karkouch et al., 2016a), (Zubair et al., 2019) (Chandola et al., 2009).

Response time is the time to process data. Due to network latency and other factors such as application unresponsiveness or performance-related issues, this may constitute to wrong interpretation of data (Wei et al., 2019).

4. PHD PROGRESS AND FUTURE RESEARCH

A preliminary scoping study identified exponential growth in academic interest in federated learning for IoMT (Farhad, Woolley, Andras, 2021; Farhad, Woolley, Andras, 2021) however, there are remaining challenges in terms of data quality and user/stakeholder trust.

Research questions are being identified for future work. The overarching question is: What are the barriers to successful implementation of federated learning based on IoMT data? More specific candidate research questions are as follows:

- RQ1: How are federated learning and IoMT data perceived amongst stakeholders?
- RQ2: How are data quality and the effects of data quality perceived?
- RQ3: What are the issues related to trust, and how may these issues relate to adoption?

A mixed-methods approach (survey, focus groups and interviews) is proposed for future studies to explore these research questions.

5. POTENTIAL SOLUTIONS TO IoT DATA QUALITY ISSUES

The topic of trusting IoT data concentrates on the clinician's concern about data and its quality. Trust needs to be established before clinical decisions can be based on IoT and sensors. The current lack of standardization is resulting in security and privacy issues. AI can play a vital role in IoT security and to ensure the data quality (Mohanta, et al., 2020; Wei et al.,2019). Blockchain and Swarm learning are also areas of interest in the security and privacy domain (Chen & Lien,2014).

Finally, further research is needed to establish the acceptability and safer utilization among consumers and clinicians of using IoT to improve the care and overall experience. The above viewpoint is a summary of selected literature only and not based on a complete systematic review, we believe that addressing these areas for future research will enable a wider uptake of medical IoT, which can ultimately save lives and improve patient care.

6. REFERENCES

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys (CSUR), 41(3), 1-58

Chen, K.-C., & Lien, S.-Y. (2014). Machine-to-machine communications: Technologies and challenges. Ad Hoc Networks, 18, 3-23

Farhad, A, Woolley, S.I, Andras, A, “A Preliminary Scoping Study of Federated Learning for the Internet of Medical Things”, Medical Informatics Europe Conference (MIE), In Public Health and Informatics (pp. 504-505). IOS Press.2021.

Farhad, A., Woolley, S.I., and Andras, P. “Federated learning for AI to improve patient care using wearable and IoMT sensors”, In 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI), 434-434, IEEE.

Karkouch, A., Mousannif, H., Al Moatassime, H., & Noel, T. (2016a). Data quality in internet of things: A state-of-the-art survey. Journal of Network and Computer Applications, 73, 57-81.

Mohanta, B. K., Jena, D., Satapathy, U., & Patnaik, S. (2020). Survey on IoT Security: Challenges and 14 Solution using Machine Learning, Artificial Intelligence and Blockchain Technology. Internet of Things, 11, 100227.

Perez-Castillo, R., Carretero, A. G., Rodriguez, M., Caballero, I., & Piattini, M. (2018). Data Quality Best Practices in IoT Environments. 2018 International Conference on the Quality of Information and Communications Technology, 272-275.

Ray, D., Collins, T., Woolley, S. and Ponnapalli, P., 2021. A Review of Wearable Multi-wavelength Photoplethysmography. IEEE Reviews in Biomedical Engineering, doi: 10.1109/RBME.2021.3121476.

Wei, X., Liu, Y., Wang, X., Sun, B., Gao, S., & Rokne, J. (2019a). A survey on quality-assurance approximate stream processing and applications. Future Generation Computer Systems, 1062-1080.