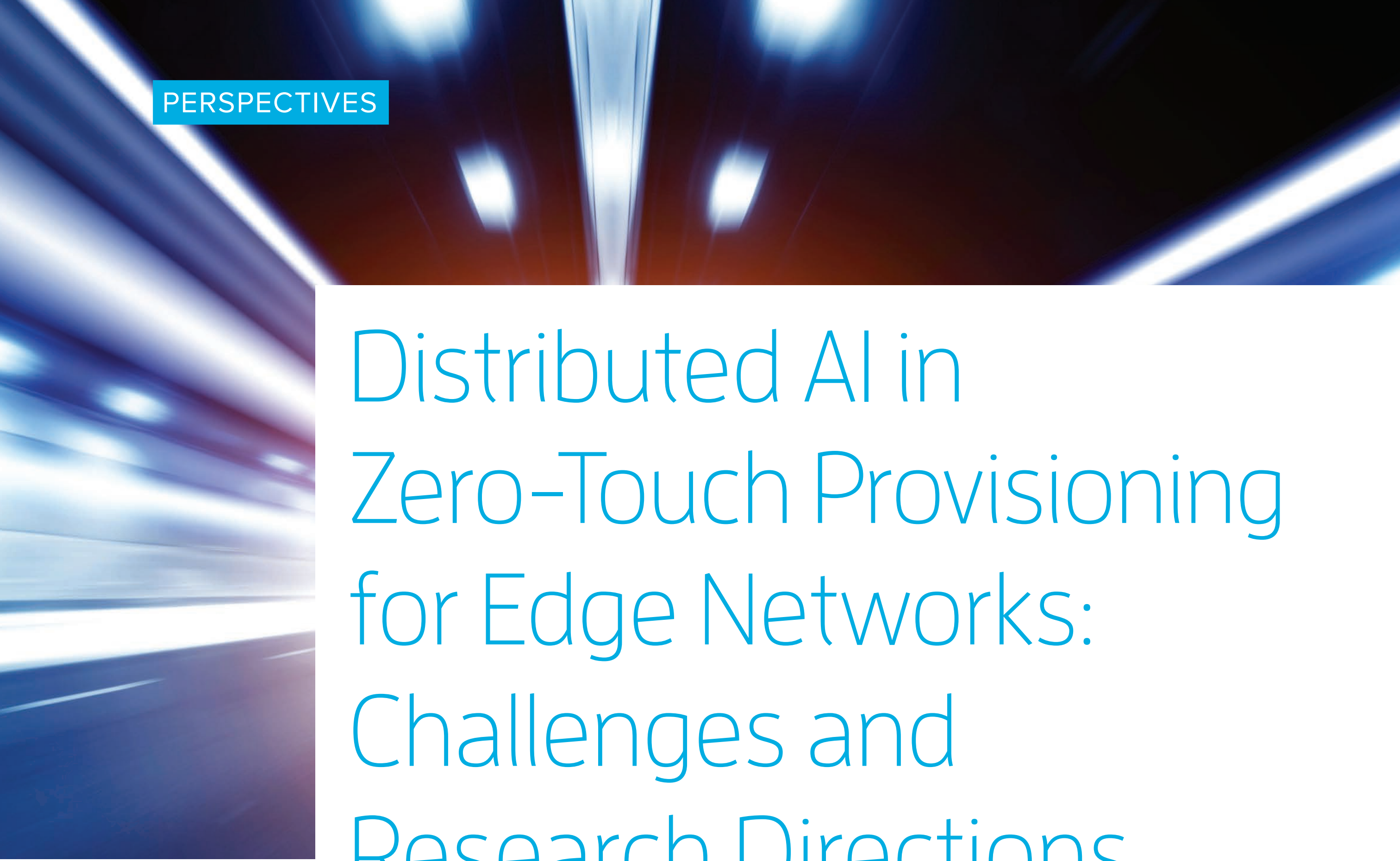

PERSPECTIVES

Distributed AI in Zero-Touch Provisioning for Edge Networks: Challenges and Research Directions

Abhishek Hazra, Indian Institute of Information Technology, Sri City and National University of Singapore

Andrea Morichetta and **Ilir Murturi**, TU Wien

Lauri Lovén, University of Oulu

Chinmaya Kumar Dehury, University of Tartu

Víctor Casamayor Pujol, **Praveen Kumar Donta**, and **Schahram Dustdar**, TU Wien

This article combines Distributed Artificial Intelligence (DAI) with zero-touch provisioning (ZTP) for edge networks. Several advantages are also highlighted that come with incorporating DAI into ZTP in the context of edge networks. Further, we draw potential research directions to foster novel studies in this field.

The Internet of Things (IoT) has witnessed a rapid rise in recent years. Connected devices have integrated with various areas of our societies, such as electric grids, transportation, and industries.[1] This paradigm change leads to rethinking the processes related to IoT systems. Indeed, IoT systems that deal with people, connected devices, and data produced in their respective environments have to guarantee the seamless management and integration of all these actors.

INTRODUCTION

The primary driving force in fostering this digital transformation is, thus, the integration of multiple functional systems to provide faster, steadier, more cost-effective, and overall better services. In particular, there's a need for automating resource provisioning in complex and broad scenarios like the device-edge-cloud computing continuum. Currently, most technologies, especially cloud-based services, focus on centralized strategies.[2] However, centralized approaches cannot work in these broad scenarios.

Digital Object Identifier 10.1109/MC.2023.3334913
Date of current version: 1 March 2024

In this regard, zero-touch provisioning (ZTP) represents an appealing direction. This class of approaches aims at providing methods to seamlessly and automatically manage the devices in a network, adapting to changes without requiring direct human intervention.[2] Such approaches are primarily implemented with machine learning (ML) methods.

The challenge is how to let artificial intelligence (AI) approaches govern such complex systems.[3] Research on Distributed AI (DAI) goes in this direction. DAI has seen waves of popularity over the decades.[4] With AI research focus recently shifting toward decentralized and widespread systems, DAI is now studied with renewed interest. In particular, when managing networks with limited or no human contact, it is essential to separate knowledge and learning mechanisms over the infrastructure. This separation optimizes the system, makes it work at scale, and preserves privacy. DAI methods can be crucial in providing this separation, ensuring that individual nodes in the network can compute with local data only.[5] Moreover, DAI methods can allow more accurate and robust prediction models by combining knowledge from many data sources separated by computational limits, administrative boundaries, or privacy restrictions. Furthermore, distributing AI minimizes the resources or costs needed for moving the information.

THE INTEGRATION OF DAI FOR ZTP AT EDGE NETWORKS HAS INTRIGUING IMPLICATIONS, ESPECIALLY CONSIDERING THE ADVANCES PROPOSED BY 5G AND 6G TECHNOLOGIES.

Still, there are open issues in letting multiple distributed computing agents communicate and produce effective solutions in real, complex, and heterogeneous scenarios such as ZTP.[6] The limited computational capabilities of devices in edge networks need novel methods to learn where to compute, how and when to distribute the data, how to guarantee optimal model management to keep the performance adequate, and how to be security aware.

The integration of DAI for ZTP at edge networks has intriguing implications, especially considering the advances proposed by 5G and 6G technologies.[7] In more detail, ZTP can foster several relevant advantages,[8] as highlighted in Figure 1. In summary, automating the logical setup of the network reduces the effort that IT teams have to put into the configuration phase. Most of the remaining work involves cabling and booting devices. Furthermore, when dealing with large and widely distributed networks, the autonomy provided by ZTP reduces the time needed to operationalize the networks. Moreover, ZTP leads to less complex and more effective management at runtime, reducing

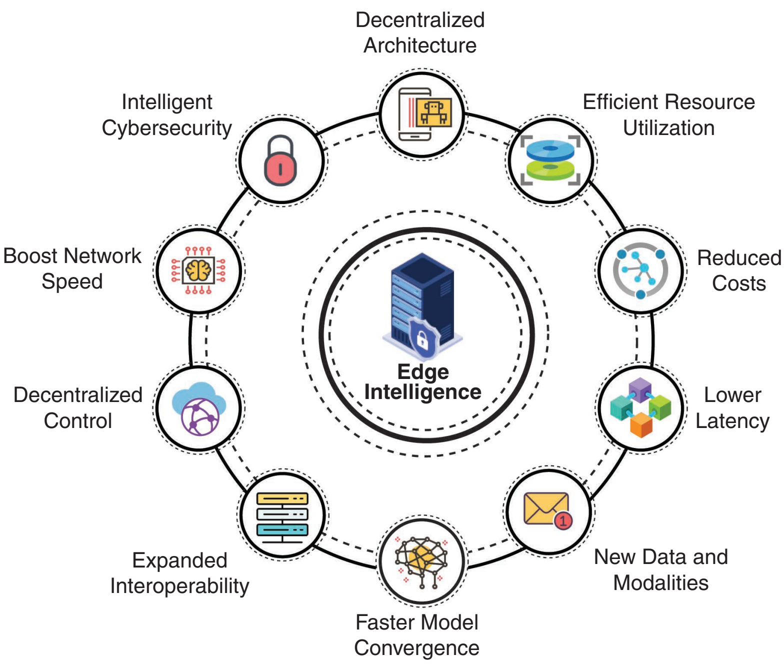

FIGURE 1. The benefits of ZTP in the computing continuum.

the probability of human error and enabling faster updates. Overall, it provides the ability to exploit large-scale computation and efficiently utilize spatially distributed computing resources in a decentralized manner with low operating costs, low latency, faster model convergence, and decentralized control.[9]

On the other hand, software-defined management is not risk free. Misconfigurations caused by automated processes may be challenging to detect, leading to complex error detection. Furthermore, managing a network on a high abstraction level means exchanging data and potentially sensitive information, introducing security threats. Consequently, the ML methods must emphasize distributed and privacy-preserving intelligence.[10] Finally, harmonically managing independent or partially correlated agents that make decisions in complex heterogeneous systems requires studying new ways of coordination.

To the best of our knowledge, this article is the first work targeting DAI in ZTP for the computing continuum. The key objective is to introduce a novel edge computing architecture that combines DAI and ZTP into one platform and offers better services to the users. The major contributions of this article are as follows:

- We design a ZTP-enabled edge computing architecture to support intelligent service provisioning while enhancing computation, communication, and storage functionalities to the users.
- We aim to highlight the challenges that come with DAI, ZTP, and their combination in the context of edge networks.
- We also discuss the network and service management challenges while offering computation and resource management solutions with ZTP in edge networks.
- Finally, we introduce potential research directions that can foster novel studies in this field and overcome the current limitations.

ZTP LEADS TO LESS COMPLEX AND MORE EFFECTIVE MANAGEMENT AT RUNTIME, REDUCING THE PROBABILITY OF HUMAN ERROR AND ENABLING FASTER UPDATES.

The remaining sections are structured as follows. The "DAI in ZTP for Edge Networks" section discusses the advantages of DAI in ZTP-enabled edge networks, where we introduce a novel edge computing architecture in the computing continuum. The "Potential Challenges and Future Directions" section highlights several potential challenges that still need to be addressed for the proper deployment of ZTP in edge networks. Finally, we conclude our discussion.

DAI IN ZTP FOR EDGE NETWORKS

Traditional edge computing brings cloud services to the network's edge. However, edge computing, while offering computation, communication, storage, and resource solutions, has several evolving challenges related to end-to-end service and network management. To address such ripening challenges, we propose a new collaborative service management model, combining DAI, edge resource federation, and ZTP concepts. The edge-enabled ZTP framework is a trusted collection of services and related resources that intelligently integrates the infrastructures of distributed computing continuum service providers and ultrareliable communication technologies to achieve low latency, scalability, and cost-efficient edge data transmission and processing, as shown in Figure 2. This framework consists of the following two major units:

- *Edge intelligence for zero-touch networks:* Edge intelligence plays a crucial role in realizing the concept of zero-touch networks. Fundamentally, this technology facilitates data processing at a local level, granting edge devices the ability to independently assess and respond to data without relying on centralized decision making. The use of distributed decision-making processes reduces latency, optimizes network resources, and guarantees real-time responsiveness, as illustrated in Figure 2. Edge intelligence enables the deployment of ML models at the edge, facilitating the implementation of predictive maintenance,

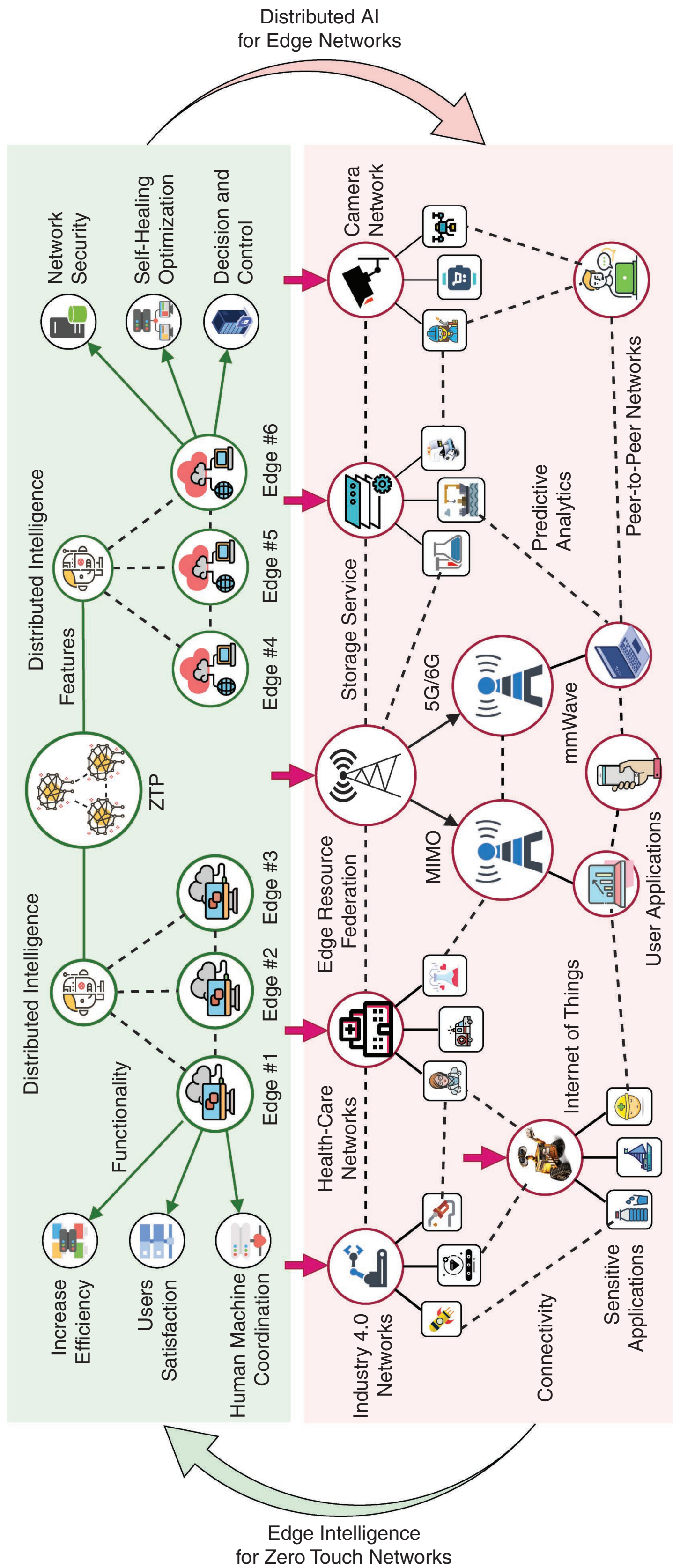

FIGURE 2. A zero-touch network in the computing continuum. MIMO: multiple-input, multiple-output.

anomaly detection, and dynamic load balancing. These features enable networks to function effectively, reduce the need for human involvement, and smoothly adjust to dynamic circumstances.

- *DAI for edge networks:* DAI facilitates the deployment of AI capabilities to the periphery of network infrastructures. Edge devices equipped with AI models can make real-time decisions, process data locally, and function independently. Recent advances in DAI exhibit high scalability, making it well suited to networks that experience an increasing quantity of devices and services. AI models have the capability to be tailored to fulfill specified criteria, hence guaranteeing self-governance and the protection of edge operations. AI deployment at the edge of networks has several advantages, including improved efficiency, decreased reliance on centralized control, and the fulfillment of zero-touch network and service management objectives.

In summary, edge computing and ZTP technology can present new business opportunities to network operators, heterogeneous IoT users, and cloud service providers. Further combining DAI and edge resource federation in the ZTP networks, customers can experience rapid data accessibility, seamless network coverage, interoperable data migration, and innovative services, which will eventually help to enhance user happiness.[11] The key techniques of edge-enabled ZTP networks are discussed next.

Edge resource federation

Standard edge computing and cloud computing models, delivering services to end users, suffer from inflated resource mismanagement. There is a need for a unique and collective service provisioning strategy, where overcrowded edge devices interoperably communicate with nearby underloaded edge devices or cloud servers and share the excessive workload.

Edge federation, also known as *edge resource federation*, is a combined resource provisioning strategy for edge networks. Edge federation manages the resources of the different edge devices offered by service providers and brings the edge resources into one platform.[12] In essence, edge federation aims for low latency, scalability, and cost-effectiveness by seamlessly integrating edge-to-edge and edge-to-cloud resources into one platform. Network function virtualization, software-defined networking, and containerization and container orchestration as well as multiaccess edge computing are anticipated as critical enablers for automated edge resource federation.

The edge federation model has two key advantages. First, it has the capability to gather edge services under one platform and handle dynamic service requests coming from different users while optimizing network resources and service delay. Second, edge federation combines different edge infrastructures and resources offered by different service providers by optimizing service deployment costs. As expected, combining DAI and edge federation techniques can introduce new responsive service assistance models, which could be a win-win solution for edge infrastructure providers, edge service providers, and end users. Overall, we can summarize some of the benefits as follows:

- reliable interconnection between edge and cloud
- moving computing resources to the network edge
- consistent user satisfaction ratio
- building an edge hierarchy model.
- easy knowledge sharing among user devices.

Distributed intelligence

In contrast to cloud AI, centralized edge AI trains ML models in nearby suitable computing devices and then deploys the models across distributed end devices, endowing the devices with local decision-making strategies.[13] However, centralized edge AI faces a number of challenges, including a lack of coordination among edge devices, a lack of global knowledge, and limited scope for edge federation. The present edge networks must be updated to use distributed intelligence, where edge devices can communicate and share end-device data models.[14]

DAI can solve complex understanding, learning, and decision-making problems by modeling them as multiagent systems. The agents, or edge nodes in the DAI network, can operate independently and communicate asynchronously to combine partial solutions. Owing to the large data scale, DAI systems are resilient, flexible, and by definition, loosely connected. In contrast to monolithic or centralized AI systems, which have tightly connected and geographically close processing nodes, DAI systems do not require all relevant data to be gathered in a single location. Instead, many DAI systems work with small subsets of data, making them easy to employ. In Table 1, we have briefly presented the advantages of incorporating DAI in edge networks compared to traditional AI.

One of the most critical challenges in distributed edge computing is data gravity. *Data gravity* refers to the capability of a rich source of data to attract applications and services. Edge networks can be considered as such rich data sources, with ZTP attracting users for edge services and applications while ensuring high throughput and optimized latency.

THE KEY OBJECTIVE IS TO INTRODUCE A NOVEL EDGE COMPUTING ARCHITECTURE THAT COMBINES DAI AND ZTP INTO ONE PLATFORM AND OFFERS BETTER SERVICES TO THE USERS.

Data gravity poses two fundamental issues. First, end users place tremendous strain on the edge servers to manage all the generated and processed data, resulting in high processing costs for data analysis and training. Data gravity is solved by not collecting all the data from the end devices. Instead, only the essential training data should be collected without noise in the data.

Another issue is the heterogeneity of edge devices. Edge devices are generally made by various infrastructure providers, and services have varying requirements. As a result, a model trained on an edge server will likely not fit all the other edge devices, making it always challenging for distributed edge networks. Therefore, the ZTP considers all such complex network- and system-based challenges into one frame and solves them using the DAI. Further, this framework has the capabilities to bring DAI to innumerable edge devices and allow it to scale across a wide variety of applications. Overall, we can summarize some of the key benefits as follows:

- improve the decision-making capabilities of local devices
- increase user data security and privacy
- reduce data transmission costs to remote servers
- continuously update model and knowledge
- allow training with small and heterogeneous user data.

In the context of Industry 4.0, implementing a smart factory highlights the benefits of utilizing DAI for ZTP. Within this particular environment, the edge devices situated within the factory exhibit AI capabilities that enable them to process sensor data in real time.[15] For example, when a machine sensor detects a possible issue, edge AI promptly recognizes it, implements corrective measures, and reduces the delay in critical decision making. On the other hand, a centralized or cloud-based AI system necessitates data transmission to a distant location for analysis, potentially causing unfavorable delays and operational hazards. This example demonstrates the considerable enhancement of ZTP through the implementation of DAI at the edge, with a special focus on its impact in the context of Industry 4.0. This approach effectively improves production efficiency and reduces downtime by facilitating real-time, localized, and informed decision making.

TABLE 1. The differences between centralized edge AI and distributed edge AI.

Parameters	Centralized edge intelligence	ZTP-enabled distributed edge AI
Model	Traditional supervised learning	Unsupervised and policy-based reinforcement learning
Privacy	No privacy for handling users' data	Supports privacy and security in data handling
Training time	Training on large data exponentially increases the time	Training on local edge devices helps to optimize time
Heterogeneity	Low	High
Scalability	Not scalable	Highly scalable
Applications	Traffic monitoring, data storage, and analysis	Keystroke prediction, smart cities, and autonomous vehicles
Computation cost	Incurs high costs over the edge network	As the model shares only learnable parameters, cost decreases.
Performance	Due to centralized architecture, edge AI suffers from low accuracy.	As the model shares knowledge, the performance of the network increases gradually.
Automation	Medium	High

ZTP

There is a trend toward ever more on-demand offering of storage and resource management capabilities. With the increasing number of resources being managed, delivering and managing dynamic user service requests becomes ever more complex. To overcome this complexity, the European Telecommunications Standards Institute (ETSI) offers the idea of zero-touch network provisioning as a new breed of network management functionality, seeking to integrate network functionality and cutting-edge communication technologies (enhanced mobile broadband, ultrareliable and low-latency communications, and massive machine-type communications) as well as automatically carrying out edge computing processes.

DAI is expected to be a key facilitator of self-learning capabilities, leading to lower operating costs, quicker time-to-value processes, and a smaller chance of human errors. Although there is a rising desire to use DAI in a ZTP network, there may also be limitations and risks associated with doing so. The abilities of ZTP networks are specified on fully combined self-3s lifecycle functions (that is, self-fulfilling, self-serving, and self-assuring) to automatically satisfy and respond to customer resource demands. However, to implement this in real time, we need to take advantage of network controllers and advanced communication technologies such as 5G or 6G.

Inspired by existing cloud service models, such as security as a service, database as a service, etc., ZTP can also be provided to end users as a service model. Under this umbrella, computation, communication, and ephemeral storage can be provided to the end user or other IoT vendors. ZTP providers should be able to personalize the resources the customers would like to avail themselves of based on their needs. For instance, customers could manage the entire lifecycles of edge applications on their IoT devices, including application deployment, configuration, starting, and stopping. This requires managing the computational and storage resources on each device as well as the communication resources for message exchange and data flow among edge application components.

In addition, multiple vendors could install specific IoT devices (with additional computation and storage resources) using ZTP technology. This transformation enables computation execution near the data sources, as presented in Figure 3.[a] In such a multivendor edge infrastructure, idle resources on individual edge devices can be rented out to other vendors as a service. Overall, we can summarize some of the benefits as follows:

- the 100% full automation of network devices
- shorter time for execution on remote servers
- reducing the chances of human errors
- easy-to-fix and auto-upgrade technical programs
- easy upgrade of hardware equipment.

[a]Inspired by "Distributed Artificial Intelligence at the Edge and Beyond," https://engineering.cmu.edu/accelerator/news/2021/03/03-ai-fusion.html.

POTENTIAL CHALLENGES AND FUTURE DIRECTIONS

This section provides a list of challenges and possible future research directions for implementing ZTP and the usability of DAI in edge networks.

Challenges

The ultimate goal of ZTP is to add convenience to edge network management, limiting human intervention. However, using ZTP in edge networks does entail challenges.

- *Cascading failures:* With cascading failures, a low-level failure may also lead to failures on higher levels. ZTP has no mechanism to control such cascades. Instead, ZTP may report a single problem as multiple, making failure analysis more difficult.
- *Anomaly detection:* The ZTP model does not support, maintain, or automate service lifecycles in the entire computing continuum. Moreover, while monitoring services, ZTP has no mechanisms for causing alarms on individual anomalous activities (for example, faulty nodes) and few for responding to them.
- *Data heterogeneity*: The concept of distributed computing continuum data heterogeneity spans a wide range of data types, sources, and spatiotemporal properties. The significance of this lies in its ability to offer extensive perspectives and tailor-made solutions and to facilitate data integration for a more nuanced understanding.[15] Nevertheless, certain issues need to be addressed in distributed networks. These challenges encompass data integration, quality, and scalability complications in broad and diverse environments.
- *Limiting orchestration:* ZTP can automate small tasks and initial setups such as activating licenses, running containerized apps, bootstrapping virtual machines, and even updating device firmware. However,

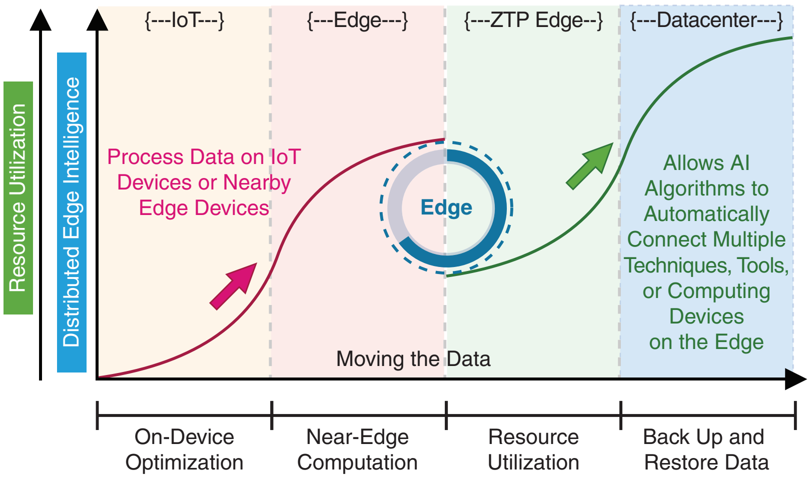

FIGURE 3. The computation strategy in zero-touch edge networks.

current ZTP implementations lack mechanisms for automating processes and workloads. We consider it a challenge because manual work degrades the value of zero touch.

- *Security:* There are a vast number of connected devices with continuous services in the computing continuum. Maintaining security on autonomous systems running on those devices with no human intervention is more challenging. DAI may help to design efficient and dynamic mechanisms across the continuum to detect unforeseen threats or vulnerabilities.

Research directions

This section fills this gap by providing possible open challenges for further research.

- *Lightweight AI/ML:* Resource-constrained end devices and edge nodes need low latency. Lightweight AI/ML algorithms minimize both resource usage as well as the time spent computing without affecting the prediction accuracy. ML model compression, which reduces the amount of redundant data in the models, is one way of achieving lightweight ML models. However, novel methods for lightweight AI/ML algorithms in ZTP are needed to further increase energy efficiency in edge networks.
- *Semantic interoperability*: The computing continuum interconnects a set of devices that are heterogeneous in terms of, for example, technologies, device standards, data formats, etc. This lack of interoperability limits the utility of ZTP in the computing continuum. It is thus necessary to bridge the gap between the ZTP and the computing continuum by developing intelligent interoperable protocols.
- *Privacy*: The IoT, cloud, data centers, gateways, etc., are all generating and exchanging massive volumes of sensitive data. The privacy of these data must be ensured while designing the ZTP for the computing continuum as ZTP precludes human intervention.
- *Low latency:* A number of time-critical use case scenarios such as medical, industry, smart city, etc., require rapid decisions. Designing low-latency mechanisms in ZTP is thus essential for the computing continuum. Future research can focus on developing techniques through intelligent agents that can prioritize time-critical requests and process them autonomously.
- *ZTP for intelligent protocols:* There is an ever-growing number of computing devices in the computing continuum and a vast number of data transmissions between them, so developing adaptive and intelligent data protocols is challenging. In this context, ZTP can help fault diagnosis and autonomous decision-making mechanisms in these protocols. In particular, broker-based publish/subscribe communication patterns may benefit more from ZTP and DAI, which may increase their adaptability and efficiency. There is a huge scope for research into making existing data protocols intelligent with the help of ZTP.
- *Explainability:* ZTP will autonomously select configuration states for large distributed systems, which will determine their behavior. In that regard, it is crucial to develop *sidecar* tools able to explain why that specific configuration was selected. To do that, causality is emerging as a candidate technology to provide explainability for self-adaptive systems.
- *Generative AI for ZTP:* In general, AI or ML techniques can predict issues by analyzing data. However, all these predictions are likely to be expected. In view of the computing continuum's

IN PARTICULAR, BROKER-BASED PUBLISH/SUBSCRIBE COMMUNICATION PATTERNS MAY BENEFIT MORE FROM ZTP AND DAI, WHICH MAY INCREASE THEIR ADAPTABILITY AND EFFICIENCY.

ABOUT THE AUTHORS

ABHISHEK HAZRA is an assistant professor in the Department of Computer Science and Engineering at the Indian Institute of Information Technology, Sri City 517646, India. His research interests include edge intelligence, 6G, Industry 4.0, and interoperability. Hazra received a Ph.D. from IIT (ISM) Dhanbad, India. Currently, he serves as an editor for *Computer Communications* and *Physical Communication*. He is a Member of IEEE. Contact him at a.hazra@ieee.org.

ANDREA MORICHETTA is a university assistant with the Distributed Systems Group, TU Wien, 1040 Vienna, Austria. His research interests include the intersection of complex distributed systems and machine learning, network monitoring, and security. Morichetta received a Ph.D. degree in machine learning and big data approaches for automatic Internet monitoring from the Telecommunication Network Group, Politecnico di Torino, Turin, Italy. He is a Member of IEEE. Contact him at a.morichetta@dsg.tuwien.ac.at.

ILIR MURTURI is currently a postdoctoral researcher with the Distributed Systems Group, TU Wien, 1040 Vienna, Austria. His research interests include the Internet of Things, edge computing, edge AI, and self-adaptive and cyberphysical systems. Murturi received a Ph.D. from the Faculty of Informatics, TU Wien in 2022. He is a Member of IEEE. Contact him at i.murturi@dsg.tuwien.ac.at.

LAURI LOVÉN is a postdoctoral researcher and the coordinator of the Distributed Intelligence strategic research area in the 6G Flagship research program, at the Center for Ubiquitous Computing (UBICOMP), University of Oulu, 90014 Oulu, Finland. His research interests include edge artificial intelligence (AI), edge intelligence, distributed AI, computing continuum, and spatial artificial intelligence. Lovén received a D.Sc. at the University of Oulu in 2021. He is a Senior Member of IEEE. Contact him at lauri.loven@oulu.fi.

CHINMAYA KUMAR DEHURY is an assistant professor of distributed systems and a member of the Mobile & Cloud Lab, Institute of Computer Science, University of Tartu, 51009 Tartu, Estonia. His research interests include cloud computing, edge-cloud continuum, clustered edge computing, and edge intelligence. Dehury received a Ph.D. from the Department of Computer Science and Information Engineering, Chang Gung University, Taiwan. He is a Member of IEEE and ACM India. Contact him at chinmaya.dehury@ut.ee.

VÍCTOR CASAMAYOR PUJOL is a project assistant (postdoc) in the Distributed Systems Group, TU Wien, 1040 Vienna, Austria. His research interests include self-adaptive methodologies for computing continuum systems, including service level objective based definitions, causal and machine learning inference, and robotics. Pujol received a Ph.D. in information and communications technologies from the Universitat Pompeu Fabra in Barcelona, Spain. He is a Member of IEEE. Contact him at v.casamayor@dsg.tuwien.ac.at.

PRAVEEN KUMAR DONTA is a postdoctoral researcher in the Distributed Systems Group, TU Wien, 1040 Vienna, Austria, since July 2021. His research interests include causal inference, modern learning approaches for distributed systems, computing continuum, the Internet of Things, and wireless sensor networks. Donta received a Ph.D. from the Department of Computer Science and Engineering at IIT (ISM), Dhanbad, India. He is a Senior Member of IEEE and a member of the IEEE Computer Society and IEEE Communications Society. Contact him at p.donta@dsg.tuwien.ac.at.

SCHAHRAM DUSTDAR is a full professor of computer science heading the Research Division of Distributed Systems, TU Wien, 1040 Vienna, Austria. His research focuses on investigating all aspects related to edge-computing, fog-computing, cloud computing, and human-based computing as well as artificial intelligence in the co-evolution of distributed systems for the new decade. Dustdar received a PhD in business informatics from Johannes Kepler Universität Linz. He is the editor in chief of *Computing* (Springer). He is a Fellow of IEEE. Contact him at dustdar@dsg.tuwien.ac.at.

complexity constraints, there is a possibility for unpredictable issues in the future. It is possible to identify or solve unpredictable issues within the systems using recent advances in large language models and generative AI technology. It remains challenging to identify potential computing nodes to perform generative AI in the computing continuum. Also, tracing the accuracy of generative AI decisions on the fly is another challenge. Further research on the use of generative AI for ZTP must provide additional benefits to the computing continuum as a whole.

In this article, we showed the benefits of combining DAI and ZTP in the device-edge-cloud computing continuum. We discussed the pivotal role that DAI approaches maintain in creating ZTP. Moreover, we emphasized the constraints and challenges that may impede the integration of DAI in edge-enabled ZTP networks. We also shed light on several potential research solutions for establishing an intelligent and autonomous edge environment in light of the specified research challenges. ■

ACKNOWLEDGMENT

This research is partially supported by the following projects: 1) the Academy of Finland through the 6G Flagship program (Grant 318927); 2) the European Commission and select member countries through the ECSEL JU FRACTAL project (Grant 877056); and 3) Business Finland through the Neural pub/sub research project (Diary Number 8754/31/2022). Praveen Kumar Donta is the corresponding author.

REFERENCES

1. H. Kokkonen et al., "Autonomy and intelligence in the computing continuum: Challenges, enablers, and future directions for orchestration," 2022, *arXiv:2205.01423*.
2. J. Gallego-Madrid, R. Sanchez-Iborra, P. M. Ruiz, and A. F. Skarmeta, "Machine learning-based zero-touch network and service management: A survey," *Digit. Commun. Netw.*, vol. 8, no. 2, pp. 105–123, Apr. 2022, doi: 10.1016/j.dcan.2021.09.001.
3. S. S. Gill et al., "AI for next generation computing: Emerging trends and future directions," *Internet Things*, vol. 19, Aug. 2022, Art. no. 100514, doi: 10.1016/j.iot.2022.100514.
4. B. Chaib-Draa, B. Moulin, R. Mandiau, and P. Millot, "Trends in distributed artificial intelligence," *Artif. Intell. Rev.*, vol. 6, no. 1, pp. 35–66, 2004, doi: 10.1007/BF00155579.
5. M. A. Ağca, S. Faye, and D. Khadraoui, "A survey on trusted distributed artificial intelligence," *IEEE Access*, vol. 10, pp. 55,308–55,337, May 2022, doi: 10.1109/ACCESS.2022.3176385.
6. B. Chafika, T. Taleb, C.-T. Phan, C. Tselios, and G. Tsolis, "Distributed AI-based security for massive numbers of network slices in 5G & beyond mobile systems," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit)*, 2021, pp. 401–406, doi: 10.1109/EuCNC/6G-Summit51104.2021.9482418.
7. H. Chergui et al., "Zero-touch AI-driven distributed management for energy-efficient 6G massive network slicing," *IEEE Netw.*, vol. 35, no. 6, pp. 43–49, Nov./Dec. 2021, doi: 10.1109/MNET.111.2100322.
8. C. Benzaid and T. Taleb, "AI-driven zero touch network and service management in 5G and beyond: Challenges and research directions," *IEEE Netw.*, vol. 34, no. 2, pp. 186–194, Mar./Apr. 2020, doi: 10.1109/MNET.001.1900252.
9. G. Carrozzo et al., "AI-driven Zero-touch operations, security and trust in multi-operator 5G networks: A conceptual architecture," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, 2020, pp. 254–258, doi: 10.1109/EuCNC48522.2020.9200928.
10. S. Dustdar, V. C. Pujol, and P. K. Donta, "On distributed computing continuum systems," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 4092–4105, Apr. 2023, doi: 10.1109/TKDE.2022.3142856.
11. F. Luque-Schempp et al., "Toward zero touch configuration of 5G non-public networks for time sensitive networking," *IEEE Netw.*, vol. 36, no. 2, pp. 50–56, Mar./Apr. 2022, doi: 10.1109/MNET.006.2100442.
12. X. Cao, G. Tang, D. Guo, Y. Li, and W. Zhang, "Edge federation: Towards an integrated service provisioning model," *IEEE/ACM Trans. Netw.*, vol. 28, no. 3, pp. 1116–1129, Jun. 2020, doi: 10.1109/TNET.2020.2979361.
13. V. C. Pujol, P. K. Donta, A. Morichetta, I. Murturi, and S. Dustdar, "Edge intelligence—Research opportunities for distributed computing continuum systems," *IEEE Internet Comput.*, vol. 27, no. 4, pp. 53–74, Jul./Aug. 2023, doi: 10.1109/MIC.2023.3284693.
14. E. Peltonen et al., "The many faces of edge intelligence," *IEEE Access*, vol. 10, pp. 104,769–104,782, Sep. 2022, doi: 10.1109/ACCESS.2022.3210584.
15. A. Hazra, M. Adhikari, T. Amgoth, and S. N. Srirama, "A comprehensive survey on interoperability for IIoT: Taxonomy, standards, and future directions," *ACM Comput. Surv. (CSUR)*, vol. 55, no. 1, pp. 1–35, 2021, doi: 10.1145/3485130.