Digital Twin and Federated Learning: Enhancing and Securing Critical Infrastructure

by Niccolò De Carlo, Ciro Romano, Gianluca Granero, Fabrizio D'Amico, Enrico Cappelli, Gianluca Fabbri

The integration of Digital Twin technology with Federated Learning is revolutionizing the monitoring of critical infrastructure by improving operational efficiency and data security. How do these technologies enable predictive maintenance and optimize resource management?

Critical infrastructures such as bridges, transportation networks, roads and railways, and power plants form the backbone of our society. Efficient management and maintenance of these systems are crucial to ensure safety, operational continuity, and economic sustainability. As technology advances, new tools are emerging to tackle these challenges. Among them, the concept of Digital Twin and Federated Learning techniques are proving to be particularly promising.

Digital Twin concept and its Benefits

A Digital Twin is a digital replica of a physical entity or system. This digital twin is connected in real-time to the physical object through sensors that continuously collect data. These data are then used to simulate, analyze, and predict the performance of the physical entity. In the context of critical infrastructure, a Digital Twin can represent anything from a single component to an entire structure.

The Benefits of Using a Digital Twin in Infrastructure Monitoring:

• *Continuous Monitoring*: Sensors provide real-time data, enabling constant monitoring of operating conditions.
• *Predictive Maintenance*: By analyzing the collected data, potential failures can be predicted and addressed before they occur, reducing management risks, downtime, and repair costs.
• *Resource Optimization*: Data can be used to optimize resource utilization, improving operational efficiency.
• *Simulation and Analysis*: Digital models allow for the simulation of various scenarios and analysis of the impact of different factors, supporting informed decision-making.

Federated Learning: a new paradigm in Machine learning

Federated Learning is a machine learning technique that enables the training of models on distributed data without the need to centralize the data. Instead of transferring raw data to the cloud, models are trained locally, and only model updates, such as parameters or weights, are shared.
The key advantages of Federated Learning include:

• *Data Privacy*: Data remains local on the devices, reducing risks associated with transmission and centralized storage.
• *Transmission Efficiency*: Minimizing the amount of data transferred conserves bandwidth and improves response times.
• *Collaboration Without Data Sharing*: Entities can collaborate to improve machine learning models without having to share their sensitive data.

Synergy Between Digital Twin and Federated Learning

The integration of Digital Twin with Federated Learning represents a powerful solution for monitoring critical infrastructure. This synergy leverages the benefits of both technologies, optimizing the monitoring, maintenance, and management of infrastructure.

Local Data Processing
Sensors installed on infrastructures continuously collect data. This data is processed locally using computing capabilities integrated into the infrastructures themselves or nearby devices such as edge nodes. Local processing reduces the need to transfer large volumes of data to the cloud, preserving bandwidth and improving latency.

Knowledge Transfer
Through Federated Learning, model updates are shared among different infrastructures without transferring raw data. This enables the creation of a global model that incorporates knowledge gained from each infrastructure. The global model can then be deployed to local infrastructures, enhancing their ability to predict failures, optimize maintenance, and efficiently manage resources based on shared experience.

Practical Applications: Bridge monitoring

Imagine a bridge monitoring system that utilizes sensors to gather data on vibrations, deformations, and environmental conditions. This data is processed locally to detect any anomalies. Through Federated Learning, model updates are shared with other bridges, creating an interconnected monitoring system that enhances predictive maintenance and structural safety.
This approach establishes a networked monitoring system where each bridge contributes its data to optimize predictive models. Aggregating data from multiple bridges improves the accuracy of predictive maintenance models, as they learn from a larger volume of data than any single bridge could generate.
Similarly, an exact digital replica of the bridge can be used to simulate and predict the structure's behavior under different stress conditions. The digital twin, fueled by real sensor data, enables engineers to conduct virtual tests and identify potential critical points before real issues arise. For example, simulating the effect of an earthquake or increased heavy traffic on the bridge allows for evaluating potential consequences and planning preventive measures.
The integration of digital twin technology with Federated Learning enables much more effective and timely predictive maintenance. Engineers can intervene precisely, minimizing downtime and maintenance costs while ensuring maximum safety for users. This interconnected system not only improves the individual management of each bridge but also enhances the overall resilience of the national road infrastructure. Optimized maintenance strategies and information can be shared and implemented on a broader scale, benefiting the entire infrastructure network.

A practical example: The SIMICOM project

SIMICOM (Integrated System for Monitoring Critical Infrastructure through Operational Modal Analysis and Smart Damage Detection with Federated Learning

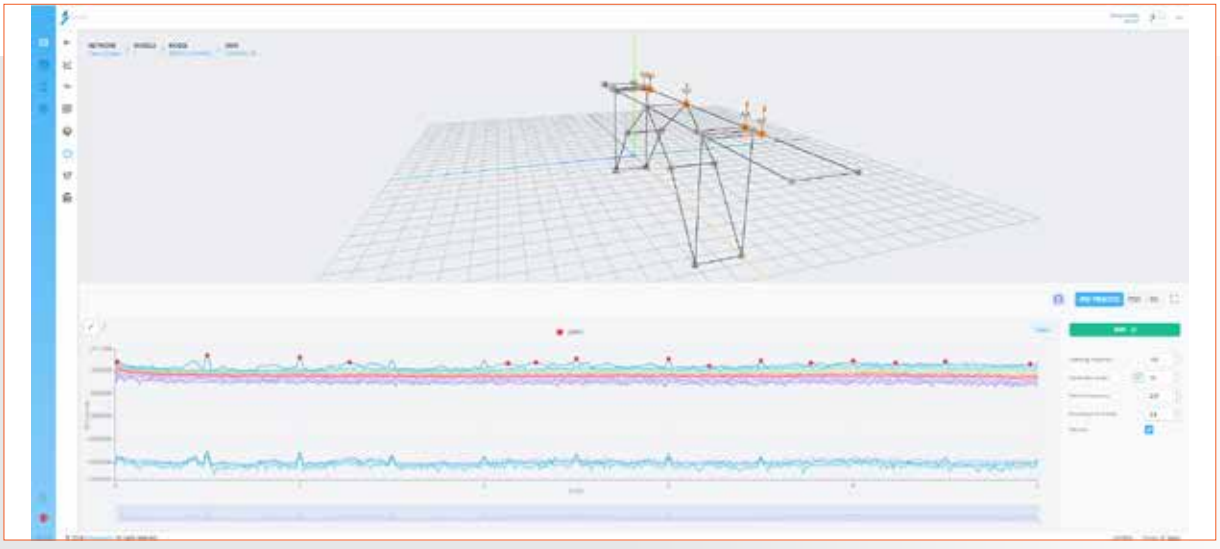
Algorithms) is a research project funded under the Regional ERDF Programme Lazio 2021-2027, born from collaboration between Sensoworks, Bridge Engineering, the University of Roma Tre (Department of Civil Engineering, Computer Science, and Aeronautical Technologies), and Link Campus University.
The main objective of the project is to develop an integrated system for monitoring critical infrastructure through operational modal analysis and smart damage detection using Federated Learning algorithms.
Italy has a significant national heritage of infrastructure often managed with outdated methods: the SIMICOM project aims to create a product capable of overseeing asset management using intelligent programming through a combination of integrated technologies in a single, functional, and efficient platform.
Data shows that Italy is among the European countries most affected by floods, erosion, and landslides, with over 8 million people living in high-risk areas. According to the "Hydrogeological Instability Report in Italy 2021" by ISPRA (Italian Institute for Environmental Protection and Research):

- Compared to EU member states, Italy leads in urbanized territory (7.65% compared to the European average of 4.3%), with over 91.1% of municipalities classified as hydrogeologically at risk.
- In 2021, 1.3 million inhabitants were reported to live in landslide risk areas. Specifically, over the past fifteen years, 82.8% of landslides recorded in Europe occurred in Italy.
- 70% of Italy's territory is at high seismic risk, the highest among European countries, due to its unique geographic position between the African and Eurasian plates and the presence of active faults and seismic structures.

The evolution of transportation also contributes to road instability. The condition of infrastructure highlights the monitoring issue in good management practices: in 2020, Italy ranked twentieth in road quality compared to the EU (Source: Eurostat and World Economic Forum data).
Within the SIMICOM project, an advanced cloud-based platform has been developed. This platform enables functionalities for navigating infrastructure and entering all inspection information for each component using Building Information Modeling (BIM) techniques, creating digital twins, and employing artificial intelligence to assist in detecting, analyzing, and monitoring defects in infrastructure.
The project specifically focused on:

- Calibrating and continuously synchronizing the digital twin with data measured by instrumentation installed on the infrastructure.
- Predictive analysis of the behavior of infrastructure through Federated Learning algorithms.

The main goal of the SIMICOM project is to create an integrated system for monitoring critical infrastructure, aiming to make them more resilient and robust through operational modal analysis and smart damage detection using Federated Learning algorithms for predictive maintenance.



The SIMICOM project.

This theme aligns perfectly with the objective of SDG9 of the 2030 Agenda for Sustainable Development, which seeks to "Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation". The industrial concept behind the project stems from a thorough analysis of the infrastructure in Italy, highlighting concerns about infrastructure monitoring and the promotion of innovation, as well as support for sector research and development.

Management of Transport Networks: The Case of Railway Networks

In railway networks, trains and tracks can be equipped with sensors that monitor various parameters such as speed, vibrations, temperature, track wear, wheel conditions, and surrounding environmental conditions. These sensors, strategically positioned along the railway route and on the trains themselves, provide a detailed real-time overview of the health status of both the railway infrastructure and rolling stock.

The collected data is processed locally using machine learning algorithms that analyze patterns and identify any anomalies that could indicate impending issues, such as track defects, signaling system malfunctions, or train engine problems. This enables early diagnosis and timely response, reducing the risk of sudden failures and accidents.

Through Federated Learning, updates and insights gained from a single train or a segment of track are securely and decentralizedly shared across the entire railway network. This approach allows every unit in the network to benefit from experiences and data collected elsewhere, thereby enhancing the predictive and maintenance capabilities of the entire railway network. For instance, if a predictive model detects a specific type of wheel wear tends to occur under certain operational conditions, this information is shared with all trains and track segments, enabling preventive measures to be taken before the issue manifests elsewhere.

The combined use of digital twin with Federated Learning creates an ecosystem of continuous learning and adaptation, where information gathered at one point in the railway network can enhance the entire infrastructure. This interconnected system enables more accurate predictive maintenance, reduces downtime and maintenance costs, and ensures maximum safety for passengers. Additionally, it fosters greater collaboration among different parts of the railway network, standardizing maintenance and operational practices.

Maintenance of Power Plants

Power plants, whether nuclear, thermal, or renewable, can greatly benefit from the integration of Digital Twin and Federated Learning. Sensors installed on the plants gather critical data on the performance of components such as turbines, generators, heat exchangers, and control systems. These sensors monitor essential parameters such as temperature, pressure, vibrations, and operational efficiency, providing a continuous stream of real-time data.

Through Federated Learning, information and insights gained from one power plant are securely and decentralizedly shared with other plants in the network. For example, if one plant detects a pattern preceding a failure, this information can be used by other plants to prevent similar issues. This collective learning process enhances the reliability and operational efficiency of all participating plants.

Similarly, the Digital Twin plays a crucial role in optimizing and simulating operations. It is constantly updated with data collected from sensors, allowing engineers to simulate various operational scenarios and stress tests. For instance, extreme conditions such as peak energy demand, sudden failures, or adverse natural events can be simulated to evaluate system responses and plan targeted interventions.

The combination of Digital Twin with Federated Learning enables power plants to adapt their operations based on demand forecasts and energy consumption patterns, reducing waste and improving sustainability. Moreover, this integration facilitates the management of renewable energies, where environmental variables play a significant and unpredictable role.

Challenges and Considerations

Data Security

Data security is a fundamental concern when discussing Digital Twin and Federated Learning, especially in the context of critical infrastructure. As these infrastructures handle sensitive and often critical data for public safety, it is essential to protect this information from unauthorized access and cyber-attacks.

Local Data Protection: Even though Federated Learning reduces the need to transfer raw data to the cloud, it's crucial to protect data locally. This requires implementing robust security measures such as data

encryption, multi-factor authentication, and continuous threat monitoring.
Model Security: Updates to the model shared across infrastructures must be safeguarded against manipulations. It's important to ensure that only authorized entities can contribute to model updates and that these updates are verified for authenticity and integrity.
Attack Prevention: Systems need to be designed to withstand various types of attacks, including data poisoning and gradient-based attacks. The use of advanced security techniques such as homomorphic encryption and differential privacy protocols can help mitigate these risks.

Interoperability
Critical infrastructures often use a variety of sensors and devices from different manufacturers. Ensuring that these devices can communicate and work together seamlessly is a significant challenge.
Standardization: It is necessary to develop and adopt common standards for data collection, processing, and transmission. Open and interoperable standards can facilitate the integration of diverse technologies and platforms.
Hardware and Software Compatibility: Differences in hardware and software specifications among various devices can create implementation hurdles. Investing in middleware solutions that can bridge different technologies can help overcome these issues.
Coordination among Manufacturers: Promoting collaboration among sensor, device, and software manufacturers is crucial for developing compatible and interoperable solutions. This can be facilitated through industrial consortia and public-private partnerships.

Latency and Computing Capacity
Local data processing requires devices with sufficient computing capacity to handle real-time analysis. Additionally, latency must be minimized to ensure timely responses.
Upgrade Hardware: In some situations, upgrading existing hardware may be necessary to meet the demands of local processing. This can involve significant costs but is essential for ensuring optimal performance.
Edge Computing: Adopting edge computing can help distribute the processing load between local and edge devices, reducing latency and improving efficiency. Implementing a network of edge computing nodes can increase processing capacity without overburdening individual devices.
Software Optimization: Optimizing machine learning algorithms and data processing for execution on devices with limited resources is crucial. This may include using model compression techniques and implementing lighter and more efficient algorithms.

Model Update
Continuous model updates through Federated Learning require careful management to avoid conflicts and ensure consistency. Implementing effective update mechanisms is essential to synchronize models across different infrastructures.
Synchronization and Consistency: Ensuring that model updates are synchronized across different infrastructures is essential for maintaining system coherence and reliability. This may require the use of distributed consensus algorithms and conflict resolution mechanisms.
Model Versioning: Implementing a model versioning system can help track changes and manage different versions of models. This is useful for analyzing model performance over time and for reverting to previous versions in case of issues.
Validation and Testing: Before deploying model updates, rigorous testing and validation are essential to ensure that the changes effectively improve system performance. This can include simulating various scenarios and analyzing the impact of updates.

Scalability
As the number of monitored critical infrastructures increases, it is essential that the system can scale effectively to handle the growing volume of data and processing.
Scalable Architecture: Designing a scalable architecture from the outset is crucial to ensure that the system can handle a growing number of devices and sensors. This may include using microservices architectures and distributing workload across multiple nodes.
Data Management: Efficient data management is essential to ensure scalability of the system. This can involve implementing data compression strategies, distributed storage, and leveraging big data technologies for data processing and analysis.
Resource Optimization: Employing resource optimization techniques to maximize system efficiency is crucial to ensure scalability without compromising performance. This may include using dynamic resource allocation algorithms and proactive capacity management.

Future Perspectives

The integration of Digital Twin with Federat ed Learning represents an advanced frontier and a paradigm shift in monitoring and managing critical infrastructure. This approach not only enhances efficiency and security but also provides a scalable solution to address future industry challenges. Critical infrastructures, vital to our society, stand to gain immense benefits from implementing these innovative technologies, ensuring a safer and more efficient future.

Adopting these technologies requires initial investments and careful management. However, the long-term benefits in terms of operational efficiency, security, and sustainability far outweigh the costs. As computing capabilities advance and new security protocols are adopted, it will be possible to implement even more sophisticated and integrated systems. New algorithms and learning techniques will enable more efficient data analysis and more accurate event prediction.

The Internet of Things (IoT) will play an increasingly important role in monitoring critical infrastructure. Integrating a growing number of connected devices will allow for gathering even more detailed data and achieving a comprehensive view of operational conditions. This necessitates higher transmission capacities of communication networks through the implementation of 5G and 6G, which must also provide very low latency, facilitating rapid and secure data transfer and model updates.

As these technologies continue to evolve, their impact on monitoring critical infrastructure will become increasingly significant, transforming the sector and enabling smarter and more sustainable management of physical resources. The adoption of these technologies represents a crucial investment for the future, ensuring resilient infrastructures ready to tackle tomorrow's challenges.

REFERENCES

McMahan H. B., Moore E., Ramage D., Hampson S. and y Areas B. A., "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017, February 2016. W364W7352

"Federated Learning: Collaborative Machine Learning without Centralized Training Data - Google AI Blog," [Online]. Available: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html.W364W7352

Konečnt'y J. K., Mcmahan H. B. and Ramage D., "Federated Optimization:Distributed Optimization Beyond the Datacenter," November 2015. W364W7352

Siniosoglou I., Argyriou V., Bibi S., Lagkas T. and Sarigiannidis P., "Unsupervised Ethical Equity Evaluation of Adversarial Federated Networks," ACM International Conference Proceeding Series, August 2021. W364W7352

Konečný J., McMahan H. B., Yu F. X., Richtárik P., Suresh A. T. and Bacon D., "Federated Learning: Strategies for Improving Communication Efficiency," October 2016. W364W7352

KEYWORDS

Digital Twin; Critical Infrastructures; Smart Damage Detection; Machine Learning

ABSTRACT

The article explores the integration of the Digital Twin concept with Federated Learning techniques for monitoring critical infrastructure. This approach allows for local data processing and knowledge transfer between different infrastructures, minimizing the amount of data sent to the cloud. Benefits include enhanced data security, operational efficiency, and more proactive maintenance. Through practical examples, it demonstrates how these technologies can revolutionize the management of critical infrastructure.

AUTHOR

Niccolò De Carlo
niccolo.decarlo@sensoworks.com

Ciro Romano
ciro.romano@sensoworks.com

Gianluca Granero
gianluca.granero@sensoworks.com
Sensoworks s.r.l.

Fabrizio D'Amico,
Università Roma Tre,
fabrizio.damico@uniroma3.it

Enrico Cappelli,
Bridge Engineering s.r.l.,
enrico.cappelli@breng.it

Gianluca Fabbri,
Link Campus University,
g.fabbri@unilink.it