Federal Register Notice: 89 FR 51554, Federal Register :: Networking and Information Technology Research and Development Request for Information on Digital Twins Research and Development, June 18, 2024.

Request for Information on the National Digital Twins R&D Strategic Plan

Assefaw Gebremedhin, Associate Professor

Monowar Hasan, Assistant Professor

Digital Twins for Protecting Agricultural Cyberinfrastructure

RFI Response: Digital Twins R&D Plan

Assefaw Gebremedhin, Associate Professor
School of Electrical Engineering and Computer Science, Washington State University
Email: [redacted]

Monowar Hasan, Assistant Professor
School of Electrical Engineering and Computer Science, Washington State University
Email: [redacted]

RFI Response Focus: The U.S. agriculture sector is being infused with various smart technologies and cyber components. However, the cybersecurity of these digital assets is still an afterthought. Digital twins can be useful for analyzing the security posture of applications in precision agriculture and for improving preparation and protection of this sector against cyberattacks. Building effective digital twin frameworks requires concerted, cross-domain research, thoughtful synthesis of ideas and methods, and investment in workforce development.

1. Introduction

Our increased reliance on digital technologies and networked distributed systems across various industrial sectors has made critical infrastructure cybersecurity one of the top priorities for national security. The food and agriculture sector is identified as one of the 16 critical national infrastructures in the U.S., accounting for roughly one-fifth of the nation's economic activity [1], [2]. The adoption of innovative digital technologies -- such as connected sensors, embedded computers, smart tractors, and drones -- and Artificial Intelligence (AI) in agriculture has led to increased production, nutritional value, disease resistance, and, in recent times, maintaining agricultural productivity in the face of climate change. While these technologies have enabled unprecedented gains in innovation and productivity, they expose modern agriculture to various cyber vulnerabilities impacting precision agriculture operations. Due to the high value of agricultural sectors to adversaries, limited or no cybersecurity defense, and lack of domain-specific security understanding of those systems, digital agricultural systems are increasingly becoming

targets for cyber breaches. Examples of recent major cyber incidents in the U.S. agricultural sector include ransomware attacks on JBS facilities [3], attacks by Russian hackers targeting an Iowa grain co-op [4], and attacks on a Minnesota grain handler [5]. In fact, the food and agricultural sectors have been increasingly targeted (over 30 major cyber incidents in recent times), leading to significant financial losses. Attacks originating from other critical sectors, such as the utility grid and industrial control systems, could also threaten the food supply chain due to close interdependence, as indicated by the U.S. Department of Energy [6]. Any cyberattack targeting agriculture or closely related critical infrastructure could jeopardize the nation's agricultural production, exports, food security, and ultimately, national security.

A critical gap in our understanding of security vulnerabilities in the agriculture sector is the lack of domain knowledge and misconceptions about adversarial capabilities. As a concrete target for this RFI, we select critical infrastructure associated with agricultural weather and irrigation support. Weather data is a key input attribute needed for agricultural decision-making. Uninterrupted and reliable weather data, obtained at farm scale, is necessary for a variety of real-time tasks, including irrigation, disease and pest management, fruit-growth prediction, and frost mitigation, among others [7]. We envision that **building a "digital twin" framework will be vital for analyzing the security posture of distributed agricultural networks**.

2. Digital Twin for Precision Agriculture Cybersecurity

The correctness, resiliency, and efficiency of smart agriculture systems, especially those that rely on weather and irrigation data, can be improved by using digital twin architectures. The digital twin will be a network representation of active weather and water stations, including associated services for collecting field data, storing data in the cloud or local server, and retrieving data to perform farm decision-making. This digital representation will make it possible to formulate and evaluate various relevant cybersecurity scenarios at the network scale. Some of the benefits of using a digital twin for agriculture cybersecurity include:

a) weather and irrigation data monitoring at different time granularity (e.g., hour, day, month, season) and real-time analysis of misbehaviors or anomalies,
b) early detection of potential vulnerabilities (viz., root cause analysis), and
c) the ability to simulate and test different malicious behaviors and defense measures before implementing them in the production environment.

However, building digital twins for such systems involves many challenges. Some of the key challenges are outlined below.

I. <u>Data Collection and Framework Development Challenges</u>

<u>Data Heterogeneity</u>: Agricultural weather and water networks involve various data sources, including weather stations, satellite imagery, soil sensors, and irrigation systems. How can a digital twin framework be designed to integrate and analyze these heterogeneous data types?

Real-Time Data Acquisition: Ensuring real-time data collection from diverse sources is critical for an accurate digital twin. How do we ensure reliable and continuous data streaming, which can be hindered by connectivity issues, especially in remote agricultural areas?

Data Quality and Consistency: The quality and consistency of data from different sources can vary, impacting the accuracy of the digital twin. Therefore, how can we address issues such as missing data, noise, and inaccuracies due to cyberattacks? Besides, weather and water systems are dynamic and continuously changing. The digital twin must be capable of adapting to these changes in real-time, requiring sophisticated algorithms for dynamic updating and real-time analytics. This is crucial for reliable simulations and predictions.

High-Resolution Modeling: Creating high-resolution models for weather patterns and water distribution systems requires further research. Growers use real-time weather data and associated forecasts tied to crop phenology and cold hardiness models [8], [9] as decision support to actuate resource-intensive active frost damage mitigation methods, i.e., heaters, wind machines, and over- and under-tree sprinkler irrigation. However, what if a *compromised* weather sensing ecosystem and pertinent malicious weather data drive the inversion forecasting and associated wind mixing decision support? For instance, "no actuation" in critical times would result in significant crop loss, and "over actuation" would mean excess use of natural resources and economic burden on the farmer. Hence, how do we develop high-fidelity models to predict misbehaviors? Can a digital twin ecosystem trigger the *early detection* of anomalies?

Scalability and Continuous Integration: The digital twin framework must be scalable to accommodate large agricultural regions with diverse environmental conditions and complex weather and water networks. How do we develop scalable algorithms and leverage high-performance computing resources to handle large data volumes and intricate simulations? In addition, establishing feedback loops between the digital twin and its physical counterpart is crucial for refining and improving the models. Real-time data from the physical system should be used to update and validate the digital twin, ensuring that it remains accurate and relevant. The challenge is: how do we build a continuous feedback mechanism that helps maintain alignment between the digital twin and the physical system?

Dealing With Legacy Systems: Another challenge is that many agricultural operations (as well as other critical cyberinfrastructure such as power grid and control systems) rely on legacy systems and equipment that may not be easily integrated into a modern digital twin framework. Hence, how can solutions be built to bridge the gap between old and new technologies?

Software-defined Digital Twins for Better Resource Management and Resiliency: In current practice, digital agricultural decision-making is often ad-hoc. Further new research is needed to develop a "software-defined" digital twin approach that brings together fault detection, isolation, and system reconfiguration to ensure robust and resilient digital agriculture operations through the identification of servicing needs among potential clients, careful long-term resource management, and cross-checking fidelity of field nodes (e.g., weather stations). For instance, how can software-

defined networking (SDN) [10] capabilities be leveraged to provide a "global view" of the distributed infrastructure and use this in the digital twin ecosystem to ease resource management and misbehavior detection?

Another challenge is building *adaptable* digital twin architectures that can be easily transferred and applied to various contexts while preserving their security measures. Furthermore, the end-users (growers and ag-tech companies) can use simulation data from the digital twin and aid in informed decision-making (such as predicting frost mitigations and fruit surface temperatures).

Privacy Concerns: Farmers and stakeholders may have concerns about model and data privacy and the potential misuse of their information. Establishing clear data privacy policies and secure data handling practices is necessary to build trust and encourage participation.

II. Standards and Interoperability Challenges

The need for robust standards and interoperability is paramount. This ensures that digital twins can securely interact with various systems and data sources, maintaining the integrity, confidentiality, and availability of the information they process. Establishing and adhering to standards is crucial for fostering trust and reliability in digital twin technology.

Development of Evaluation Tools: To ensure the cybersecurity of digital twins, comprehensive evaluation tools are essential. These tools should be capable of assessing the entire digital twin ecosystem, including its code base, data handling processes, operational environments, and network connectivity with physical counterparts. Evaluation tools should be designed to identify vulnerabilities, assess risks, and provide actionable insights for enhancing security. Thus, concerted efforts from the research community and ag-tech vendors are required to build open-source tools for better evaluation. Further research is needed to establish guidelines and best practices for conducting threat analysis, risk assessments, and security audits. These should also encompass best incident-response and recovery practices, ensuring that digital twins can effectively withstand and recover from cyber-attacks.

Data Exchange and Encryption Protocols: Establishing standardized data exchange protocols is essential for ensuring interoperability between different digital twin systems. These protocols should define how data is formatted, transmitted, and secured during exchanges between digital twins and their associated physical systems. Secure data exchange protocols help prevent unauthorized access and ensure the integrity of the data being transmitted. In addition, to protect sensitive data within digital twins, common encryption standards must be developed and followed.

Taxonomy and Ontology Development: One challenge in building digital twins for agriculture and other critical infrastructure is the lack of common taxonomy and ontology for ensuring interoperability. Hence, we must define standardized terms and concepts that can be universally understood and applied across different digital twin systems. A common taxonomy facilitates seamless communication and data sharing between digital twins, reducing the risk of misinterpretation and errors.

III. Verification, Validation, and Uncertainty Quantification Challenges

Verification and Validation are critical processes for ensuring the reliability, accuracy, and trustworthiness of digital twins. These processes help establish confidence in the digital twin's ability to accurately represent and predict the behavior of its physical counterpart, thus ensuring its effectiveness in cybersecurity applications.

Code Verification: Code verification ensures that the digital twin's software correctly implements the intended algorithms without errors. This involves rigorous testing to identify and fix bugs, security vulnerabilities, and logic errors within the codebase. How can techniques such as static analysis, formal methods, and automated testing be employed to achieve thorough code verification in this context?

Model Verification and Validation: How do we ensure that the mathematical models and simulations used in digital twins are implemented correctly? We need techniques to check the consistency and correctness of the models against their specifications. Besides, automatic validation tools are required to compare the digital twin's outputs with experimental data and observations from the physical system. We envision further research on statistical validation, benchmark comparisons, and real-world scenario testing, which are essential for validating model accuracy in agricultural digital twins.

Uncertainty Quantification: Another challenge is identifying and quantifying uncertainties in the digital twin's models and predictions, either due to fault or cyberattacks. This includes uncertainties in model parameters, initial conditions, and input data. Can we recontextualize tools such as probabilistic modeling, sensitivity analysis, and Monte Carlo simulations to quantify these uncertainties?

3. Workforce Development and Interdisciplinary Research Environment

Educational Programs: Developing specialized degree programs and certificates focused on digital twin technology and cybersecurity can provide in-depth knowledge and skills to students. These programs should cover areas such as modeling and simulation, data analytics, artificial intelligence, and cybersecurity principles specific to digital twins. We also need to establish innovative programs that provide education opportunities in high school and in undergraduate agricultural programs and hands-on learning experiences for the existing workforce to foster a new generation of professionals equipped with state-of-the-art agriculture and critical infrastructure cybersecurity knowledge and skills related to digital twins.

Cross-Disciplinary Collaboration: Building an end-to-end digital twin ecosystem for protecting crucial cyberinfrastructure requires a coordinated national effort. A tiered "hub and spoke" model could be an effective way to improve communication, adopt new technologies, and mobilize an upskilled workforce [11]. Such models are based on a central hub that collaborates with regional centers that are themselves coordinating hubs for regional or local partners. Establishing a hub and spoke model to address digital twins for critical infrastructure cybersecurity can have significant,

lasting impact. A consortium of land-grant universities with strong research and education programs in agriculture, cybersecurity, cyber-physical systems, and extension are ideally suited to act as regional hubs focused on the challenges and diverse needs of a region's agricultural sector. The role of the regional hub will be to engage regional and local partners, develop testbeds, undertake digital twin and cybersecurity research, promote education and workforce development, and be an essential resource to federal agencies like the USDA NIFA, NSF, FBI, and DHS. Leveraging the regional centers, a "national coordinator" can ensure a unified, equitable, and agile response to cybersecurity challenges across the national agricultural landscape. Together, the national coordinator and regional hubs will serve to advance cyber technology to build a strong, informed digital twin architecture.

4. Beyond Agriculture

Considering the diverse agriculture domain (plants, crops, livestock, seafood) and external climatic variabilities (weather, water), understanding cybersecurity issues and building techniques using digital twins to bolster security posture requires concerted, multiyear efforts. Although the majority of this RFI Response document focused on digital twin-centered agricultural cybersecurity, accompanying research is needed to investigate how this agricultural domain knowledge can be transferred to understanding and protecting other critical sectors, such as power grids, transportation systems, biomanufacturing, public health, food supply chain, and wildfire management, using the digital twin technologies.

References

[1] "GIAC Cyber Security Discussion Paper." Accessed: Jun. 06, 2024. [Online]. Available: https://www.ams.usda.gov/about-ams/giac-may-2024-meeting/cybersecurity

[2] "Food and Agriculture Sector-Specific Plan - 2015." Accessed: Jun. 06, 2024. [Online]. Available: https://policycommons.net/artifacts/12474227/food-and-agriculture-sector-specific-plan/13370861/

[3] "JBS S.A. ransomware attack," Wikipedia. Accessed: Jun. 15, 2024. [Online]. Available: https://en.wikipedia.org/wiki/JBS_S.A._ransomware_attack

[4] "BlackMatter Hits Grain Cooperative With Ransomware Attack." Accessed: Jun. 15, 2024. [Online]. Available: https://www.itprotoday.com/attacks-breaches/blackmatter-hits-grain-cooperative-with-ransomware-attack

[5] "Minnesota grain handler targeted in ransomware attack." Accessed: Jun. 15, 2024. [Online]. Available: https://www.reuters.com/technology/minnesota-grain-handler-targeted-ransomware-attack-2021-09-23/

[6] US Department of Energy, "Cybersecurity considerations for distributed energy resources on the US electric grid." [Online]. Available: https://www.energy.gov/sites/default/files/2022-10/Cybersecurity%20Considerations%20for%20Distributed%20Energy%20Resources%20on%20the%20U.S.%20Electric%20Grid.pdf

[7] J. P. Boomgard-Zagrodnik and D. J. Brown, "Machine learning imputation of missing Mesonet temperature observations," *Computers and Electronics in Agriculture*, vol. 192, p. 106580, Jan. 2022, doi: 10.1016/j.compag.2021.106580.

[8] A. Saxena, P. Pesantez-Cabrera, R. Ballapragada, K.-H. Lam, M. Keller, and A. Fern, "Grape Cold Hardiness Prediction via Multi-Task Learning," *AAAI*, vol. 37, no. 13, pp. 15717–15723, Jun. 2023, doi: 10.1609/aaai.v37i13.26865.

[9] J. C. Ferguson, M. M. Moyer, L. J. Mills, G. Hoogenboom, and M. Keller, "Modeling Dormant Bud Cold Hardiness and Budbreak in Twenty-Three *Vitis* Genotypes Reveals Variation by Region of Origin," *Am. J. Enol. Vitic.*, vol. 65, no. 1, pp. 59–71, Mar. 2014, doi: 10.5344/ajev.2013.13098.

[10] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, "A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 3, pp. 1617–1634, 2014, doi: 10.1109/SURV.2014.012214.00180.

[11] "Big Data Regional Innovation Hubs: Establishing Spokes to Advance Big Data Applications (BD Spokes)." [Online]. Available: https://new.nsf.gov/funding/opportunities/big-data-regional-innovation-hubs-establishing/505264/nsf17-546/solicitation