CatBoost-based Intrusion Detection Method for the Physical Layer of Smart Agriculture

Zizhong Wei[1, 2,†], *Fanggang* Ning[3, †], *Kai* Jiang[2], *Yang* Wang[2], *Zixiang* Bi[2], *Qiang* Duan[2], *Jichen* Zhang[2], and *Rui* Li[2, *]

[1]Key Laboratory of Agricultural Blockchain Application, Ministry of Agriculture and Rural Affairs, Beijing, China
[2]Inspur Academy of Science and Technology, Jinan, Shandong, China
[3]Inspur Software Co. Ltd, Jinan, Shandong, China

Abstract. Agriculture holds a pivotal role in the progress of human society. The challenges stemming from a burgeoning population, land degradation, water scarcity, and urbanization have intensified the need for more efficient agricultural production. While smart farming brings significant benefits to farmers and agricultural output, it also introduces complex cybersecurity risks to agricultural production. The security of the physical layer in smart agriculture is intricately tied to crop growth and yield, with indirect implications for the security of the network and application layers. This paper introduces a novel intrusion detection scheme based on CatBoost for the physical layer and evaluates its effectiveness using the publicly available ToN_IOT dataset. In binary classification results, the scheme achieves a remarkable recognition accuracy of 99.94%, along with a precision and recall of 99.88%. In multi-classification results, the scheme outperforms other existing solutions across all metrics. The experimental findings clearly illustrate the exceptional recognition accuracy of this implemented method against physical layer attacks within the domain of smart agriculture. Furthermore, the system's implementation ensures the security of input data for the smart agriculture network layer, cloud, and blockchain applications.

1 Introduction

Agriculture has been integral to human development. Firstly, agriculture's emergence ensured sufficient food supplies to meet basic subsistence needs. The enhancement of agricultural yield and quality played critical roles in preventing famines and guaranteeing societal stability. Secondly, agriculture has served as the economic foundation for several nations. Agriculture supports significant employment opportunities and contributes to regional economic growth through the production and trade of agricultural goods. Moreover, agricultural development is inextricably linked to land, water resources, and ecosystems. Implementing scientific methods in agriculture can prevent soil erosion, lessen water pollution and mitigate ecosystem damage, promoting sustainable development over time.

Nevertheless, numerous real-world challenges demand increased efficiency in agricultural production. Worldwide, 500 million people remain undernourished, and 821 million experience famine – a situation that will exacerbate with population growth [1]. Furthermore, overexploitation of land and pollution of water have hindered crop yields. Urbanisation has resulted in a decrease in arable land and agricultural labour, leading to a significant impact on the effectiveness and scope of agricultural production [2]. Consequently, finding a solution to the food supply problem brought about by population growth in the context of a diminishing number of farmers and limited agricultural production resources has become a primary concern in agricultural development today.

Increasing agricultural productivity is crucial for solving current industry challenges. To enhance efficiency and reduce operational expenses, agriculture has seen an increased adoption of cutting-edge and innovative technologies, including artificial intelligence, Internet of Things (IoT), blockchain, and cloud computing [3]. These implementations have formed a sophisticated, smart agriculture system, as shown in Fig. 1. The smart agriculture system's physical layer comprises drones and IoT devices, which gather and monitor data related to crops, soil, weather, and livestock. When certain conditions are met, the relevant controllers are prompted to carry out automated management actions, such as irrigation or automatic feeding. In addition, IoT devices merge collected data into edge or fog nodes to form various network types at the network layer. Cloud computing and blockchain are utilised at the application layer for modelling and trusted storage of agricultural data. Predictive analytics based on agricultural data and management decision-making with regards to agricultural production are enabled via cloud computing. Transparency and traceability of the agricultural supply chain is provided by blockchain.

† These authors contributed equally to this work
* Corresponding author: lirui01@inspur.com

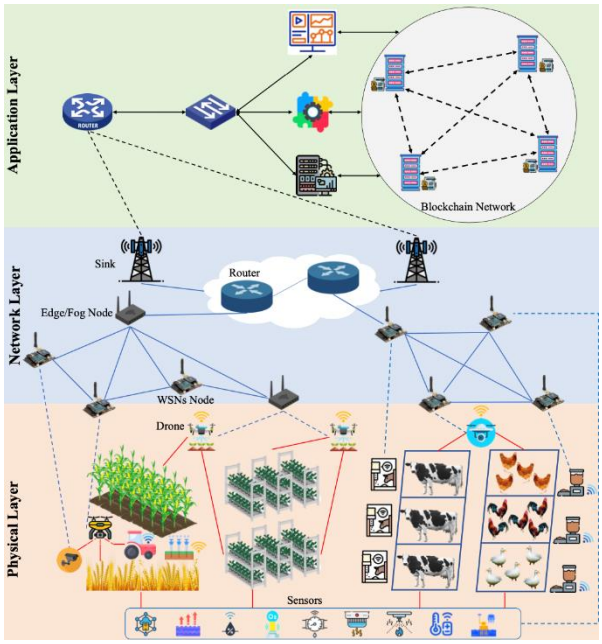

Fig. 1. Smart agriculture system schematic.

The incorporation of technology across all aspects of smart agriculture has augmented the effectiveness of agricultural production but created various cybersecurity obstacles. Cyber threats targeting agriculture, exemplified by the hacking of the Australian wool trading software and remote access to the Florida water supply system, attest to the potential dangers of hacking smart agriculture systems. Physical layer sensors and control devices are highly susceptible to attacks. These devices are deployed directly in the agricultural production environment, making them vulnerable to hijacking by attackers through physical contact or wireless networks. Furthermore, sensing and control devices lack strong security protections against robust cyberattacks. Once an attacker gains control of a physical layer device, they could launch a range of attacks. These include distributed denial-of-service attacks, which prevent access to the network service where the device is located, and false data injections. Attacks on the physical layer could result in the complete destruction of agricultural production. In addition, the spread of malicious data through the wireless sensor network to the network layer and application layer can have adverse effects on the entire system's normal operation.

The cybersecurity sector has been working to identify rapid and efficient resolutions to the security obstacles confronted in agricultural production. Intrusion detection systems, a type of malicious behaviour detection system, is well-suited to the intricate network environment and unpredictable network attacks of agricultural IoT devices. An intrusion detection system is capable of monitoring and evaluating the inputs to both the network and devices. It swiftly detects the presence of intrusion behaviours and subsequently initiates alarms or activates dependable security protection measures to lessen the consequences provoked by the attack behaviours.

Signature-based schemes are the traditional means of detecting intrusions. They are limited to identifying attacks with predetermined attack patterns or specified signatures, and their effectiveness is dependent on the volume and variety of signatures. Unknown intrusions, however, cannot be identified by such schemes. To address unknown and fluctuating network attacks, academia has also proposed anomaly-based intrusion detection methods. Such schemes offer greater flexibility, by constructing models based on intrusion behaviour features to detect and identify anomalies. To enhance the security protection of the smart agriculture system, this article suggests a CatBoost-based intrusion detection scheme for the physical layer of smart agriculture, considering that the physical layer has complex attack surfaces, diverse network structures and large amounts of data. This approach follows the concept of anomaly-based intrusion detection and utilizes the precision and generalizability of the CatBoost algorithm for the prompt and efficient recognition of intrusions in agricultural IoT systems. To validate the efficacy of the proposed scheme in detecting physical layer intrusions amidst diverse and complex security threats, the ToN_IoT dataset serves as the training model and validation source, comprising 9 categories of typical IoT attacks including backdoor, man-in-the-middle, denial-of-service, scanning, and ransomware, among others. This paper's principal contributions are:

1. The physical layer intrusion detection system for smart agriculture presented in this research paper, which utilises CatBoost technology, attains a 99.95% accuracy in identifying anomalies and 99.38% accuracy in identifying attack types in the Ton_IoT dataset.
2. In this study, the CatBoost model's categorical variable processing method was employed to address the issue of over-categorisation in agricultural IoT data. Additionally, the weighted cross-entropy loss function was utilised to tackle the problem of extremely unbalanced attack category distribution.
3. In this study, two intrusion detection models were developed using the CatBoost model. To detect intrusions in physical layer devices that have limited computational resources, a binary classification model has been created. Additionally, a multiclassification model has been designed to detect attack types in data forwarding nodes, thus enabling the implementation of protective measures.

2 Related Work

This section evaluates intrusion detection methods proposed for smart agriculture, as well as those concerning the physical layer. First, the different types of solutions for smart agriculture intrusion detection are analysed chronologically. Then, as the section categorises smart agricultural physical layer devices into three types: IoT devices, drones, and unmanned tractors, it also summarises the corresponding intrusion detection solutions.