

Research Article

Research on the RFID Anticollision Strategy Based on Decision Tree

Hui Zhang,¹ Ling Gao,¹ Han-guang Luo ,² and Yannan Zhai¹

¹Electrical and Electronic Teaching and Research Section, Basic Department, Aviation University of Air Force, Changchun 130022, China

²Research Center for Industrial Internet, Zhejiang Lab, Hangzhou 311121, China

Correspondence should be addressed to Han-guang Luo; luohg@zhejianglab.com

Received 30 March 2022; Revised 23 April 2022; Accepted 25 April 2022; Published 5 May 2022

Academic Editor: Kalidoss Rajakani

Copyright © 2022 Hui Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

RFID can automatically read the information stored on the tag. When multiple tags send data to the reader and writer at the same time, multiple tag collision will occur, resulting in the reader and writer being unable to successfully identify the tag. To solve this problem, an RFID multitag identification method based on decision tree is proposed. The Manchester code is used to represent different level states of voltage. The reader extracts the collision bits in the tag ID by using the Manchester code, and these collision bits form the new tag ID. In the process of decision tree search, the “XOR” command operation is added, and the query prefix is selectively sent to the tag to mark all collision bits. The process will continue until one tag is identified; then, reset the remaining unrecognized tags and continue to participate in a new round of identification. The simulation results show that the RFID multitag identification method based on decision tree can reduce the time complexity and communication overhead and improve the throughput, so it has high search efficiency.

1. Introduction

Radio frequency identification (RFID) technology uses radio waves and electromagnetic fields to automatically read the information stored on the tag, and the tag is attached to the object for automatic identification. RFID electronic tag has the advantages of small volume, low cost, and light weight. It can work normally for several years in extreme environments. Therefore, it appears in more and more application fields. RFID tag has been widely used in storage management, logistics management, unmanned supermarket, and other fields because of its low cost and remote batch reading. At the same time, RFID technology has the advantages of reusable electronic tags, high security, strong environmental adaptability, rapid identification of moving objects, and so on. It is also widely used in campus cards, national defense and military positioning and tracking, identity recognition, and vehicle Internet [1]. Generally, the RFID application system consists of a reading device called

card reader and one or more tags. Card readers are usually powerful devices with sufficient memory and computing resources. The electronic tag is used to attach to the object to be identified. When an object enters the reading area of the card reader, the electronic tag generates energy through magnetic field induction for response operation in the process of automatic recognition. In many of the above applications based on RFID technology, the system deploys a large number of passive RFID electronic tags. All tags have a unique tag ID number, which is used to store various data information of the article. Such tags have low reception sensitivity and weak decoding ability, so they cannot communicate with each other, and all tags can only receive signals from readers and writers. The reader/writer needs to quickly identify the tags deployed in the system in a short time [2]. In the RFID application system, when multiple tags access the same channel, the problem of allocating channel capacity to different users usually occurs, which will lead to signal overlap, which is usually called conflict [3]. In order to

reduce the error of RFID multitag identification, a lot of research [4, 5] has been carried out in related fields, and some results have been achieved.

Reference [4] proposed a multibranch tree RFID tag identification method based on bit detection. This method proposes a multibranch tree (BMT) label recognition protocol based on bit detection. This method proposes a bit detection tag number estimation method, which uses a small number of time slots to preliminarily estimate and pregroup the tags. The idle time slot is eliminated by a multibranch tree label recognition strategy. At the same time, the hash silence method is used to silence the identified tags and mark hidden tags to reduce the communication overhead. This method has high accuracy in RFID tag recognition, but there are few studies on tag collision in RFID tag recognition, which has some limitations. A parallel UHF RFID anticollision recognition algorithm is proposed in literature [5]. The algorithm is aimed at the problems of low system throughput and low system recognition rate of the traditional dynamic frame slot ALOHA multilabel anticollision algorithm (DFSA). By combining the code division multiple access technology based on the orthogonal variable spreading factor (OVSF) code as spreading code and DFSA protocol, it breaks through the limitation that only one label can be recognized in a single time slot in the traditional algorithm. It realizes the transformation from label collision to code collision. Based on theoretical analysis and simulation experiments, the performance of the algorithm in terms of system throughput, total number of time slots, and system recognition rate are explored. This method has a large throughput in system collision recognition, but there are some limitations in the accuracy of recognition.

In order to improve the reliability and efficiency of identification, this paper designs an RFID multitag identification method based on decision tree. The designed method uses tag data coding to divide the colliding tags and probes them according to the received string. Simulation results verify the effectiveness of the proposed algorithm.

2. RFID Multitag Identification Method Based on Decision Tree

2.1. RFID Multitag Data Coding. The basic premise of RFID is to use tags to mark the items to be identified. These tags contain transponders that send information that can be read by a dedicated RFID reader. Most RFID tags store some kind of identification number. RFID tag is embedded or pasted on the identified object, which is the carrier of data. As the core part of the tag, the integrated circuit is responsible for responding to the signal transmitted by the reader and writer, taking the unique identifier of the tag as the response, and modulating and demodulating the RF signal [3]. The reader retrieves the information about the ID number from the database and operates it accordingly. The RFID structure is shown in Figure 1.

RFID tags can also contain writable memory, which can store information for transmission to RFID readers in different locations. This information can track the movement of marked items so that each reader can obtain this informa-

tion. Integrated circuits generally include modem, codec module, power module, memory, logic control module, clock module, etc. Modems are used to modulate and demodulate RF signals, and codec modules are used for encoding and decoding information. Coding is an important work in the RFID system. In order to ensure the stability, accuracy, and security of signal transmission between the reader and electronic tag, the data needs to be encoded [4]. The encoding and modulation process of the reader and tag is shown in Figure 2.

Digital coding in the RFID system uses binary numbers "0" and "1" to represent different level states of voltage. Data coding can reduce the malicious modification of the signal by the outside world, provide security protection for the transmitted data, prevent the signal from receiving the interference of the environment, and improve the transmission quality. Different coding methods will have different effects on the reliability of data transmission, error processing ability, data transmission speed, and so on. The RFID system uses automatic identification communication technology, which needs to encode, modulate, and decode the signal. If there is no tag collision, the data information sent by the tag shall be subject to security authentication, decoding, and other operations, and then, the tag information shall be sent to the background application system through the middleware to complete the reading of the tag; in case of tag collision, the anticollision algorithm will be called to process the collision tag before subsequent operations [5]. Although NRZ coding and Miller coding have certain advantages in remote power supply, the detection of data interaction between the electronic tag and reader is more important. Therefore, Manchester coding is used to simplify the coding with jump in symbol time. Manchester coding can use 0 and 1 to represent the jump change between symbols, and the value of a bit is represented by the jump of the level during half a bit period in the bit length. In Manchester coding, the jump from high level to low level is expressed as 1, and the jump from low level to high level is expressed as 0. When the tag collides, the high and low phases are superimposed and offset to form a level state without jump. Based on this, the reader and writer can judge that the data position collides, so as to facilitate further processing of the collision bit. According to the characteristics of the Manchester code, it is easier to detect the conflict of the electronic tag, so the Manchester code is used in the communication link where the electronic tag sends ID information to the reader.

2.2. Design of the Multitag Anticollision Algorithm Based on Decision Tree. In the RFID system, the medium of information exchange between the electronic tag and reader is the channel. They send the signal with information to the receiver through the channel. After receiving it, the receiver processes it and then sends the feedback information of the signal through the channel. In the previous anticollision algorithms, the tag is not processed, and the recognition process is carried out directly. In practical application, there are a large number of tags within the working range of the reader, and some bits of these tags are the same. If they are identified directly, there will be a large number of

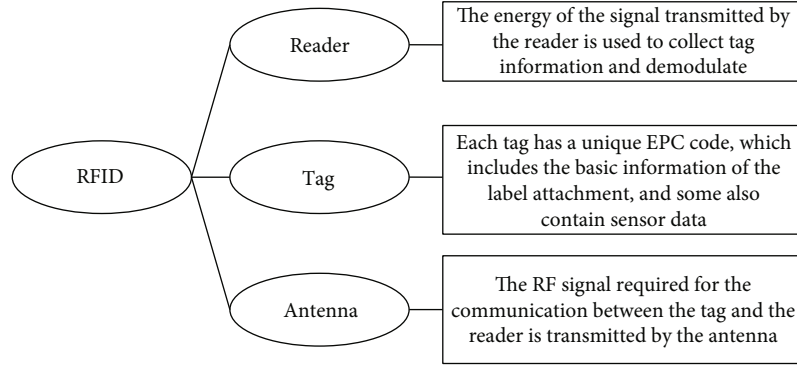


FIGURE 1: RFID structure.

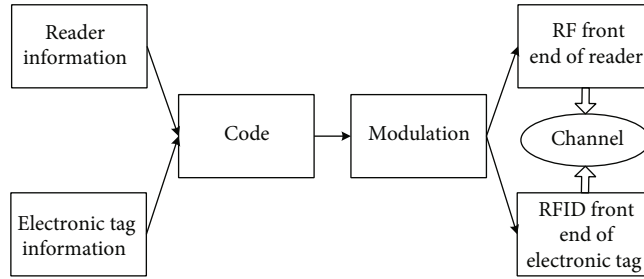


FIGURE 2: Encoding and modulation process of the reader and tag.

unnecessary bits transmitted in the identification process, which will increase the communication consumption and identification time. To solve the above problems, a multitag anticollision algorithm is designed based on the decision tree, and its flow is shown in Figure 3.

In order to reduce the initial tag collision problem, the initial tag set is divided into several small subsets before classification search and recognition using decision tree. However, at the beginning of the recognition process, the reader/writer lacks any information about tags, including the number and distribution of tags. We will analyze the relationship between tag ID, prefix, and tag set and integrate their relationship into a full binary decision tree. The full binary decision tree is just a framework to show all the possibilities of tag ID and prefix [6]. First, the reader extracts the collision bits in the tag ID by using the Manchester code, and these collision bits form a new tag ID. Then, in the following decision tree search process, add the “exclusive or” command operation and selectively send the query prefix to the tag [7]. This avoids a large number of idle time slots and improves the efficiency of the system. The reader/writer sends a search sequence of the same length as the tag ID. After receiving the sequence, the tag compares with its own ID information. If it is less than or equal to this query sequence, the tag responds to the reader. Initialize the query stack and send request (null). If the ID numbers of all tags are less than or equal to it, all tags respond and return to the reader itself. The reader and writer judge whether there is a collision according to the Manchester code. If no collision occurs, a tag is successfully identified. In case of collision, the reader/writer will modify the search sequence, set the highest position of collision to 0, keep other bits

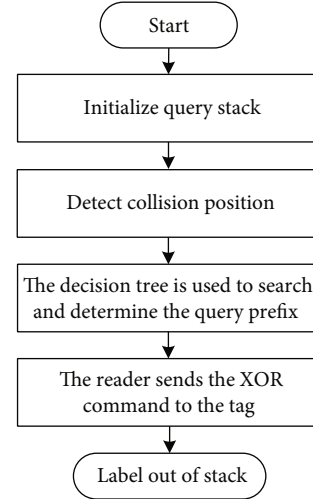


FIGURE 3: Multitag anticollision algorithm flow based on decision tree.

unchanged, and conduct the next round of search until there is no tag to respond to the query. The reader detects the collision of the tag ID, identifies all collision bits, extracts all collision positions 1 and noncollision positions 0 in the original ID, and uses the “phase and” command to extract the corresponding bits of all collision bits to form a new tag. Assuming that there are tags within the recognition range of the reader/writer, the number of times all tags need to be queried can be expressed as

$$c = \log_2(p!) + p. \tag{1}$$

In formula (1), c is the number of times all tags are queried. The decision tree structure includes three time slots: collision time slot, successful time slot, and idle time slot. The serial number of time slot represents the sequence of time slot occurrence [8]. The decision tree algorithm can generate binary tree, trident tree, and q -ary tree. When $q = 6$, the encoding method will produce 64-bit encoding, which may be longer than the ID encoding of some tags. Therefore, the total number of nodes of the decision tree is determined according to the number of tags [9]. The calculation formula is as follows:

$$u = \frac{pq - 1}{q - 1}. \quad (2)$$

In formula (2), u represents the total number of nodes in the decision tree and q indicates the number of branches of the full fork tree. When the sent query prefix has no tag response and is an idle time slot, the reader/writer discards this prefix and takes out the next prefix for query. When only one tag responds, which is a successful time slot, the reader/writer identifies the tag [10]. When there are multiple tag responses, it is a collision time slot, and the multitag collision problem occurs. At this time, all collision bits are connected together and searched directly by decision tree. After a series of judgment and calculation, the decision tree algorithm generates four new query codes: 00, 01, 10, and 11 [11]. Using the decision tree to query the electronic tag will be faster, which can effectively avoid the idle time slot, so as to improve the recognition efficiency. At this time, the reader first sends the XOR command to the tag and determines the new query command according to the tag return message. The reader sends the XOR command to the tag in the form shown in Figure 4, where a is the highest bit of the current two collision bits.

Push the new query command into the stack and send it to the tag in the order from the bottom of the stack to the top of the stack, which conforms to the tag response of the query command. Determine whether there are query instructions in the query stack. If not, the algorithm ends.

2.3. Establish an RFID Multitag Identification Model. The tag recognition process consists of continuous recognition frames, in which any frame is composed of several frame time slots. The average number of new tags per unit time in the recognition frame is called the tag arrival rate. When the arrival rate is 0, it indicates that there is no static recognition scene in which a new tag arrives in the frame. When the arrival rate is not 0, it indicates that a new tag has arrived in the frame recognition process. Only one query needs to be sent in the query process, because the tag ID has the principle of binary certainty. If there are only two tags in the collision cycle (time slot) and there is only one bit collision, the rest of the IDs of the two tags are exactly the same except the collision bits [12]. The tag ID is composed of "0" and "1," so the two tags can be identified by using only one query request in a query cycle. In the established RFID multitag identification model, the expected time slot of the tag in

the identification frame can be expressed as follows:

$$\gamma = (\tau) \binom{z-1}{p} \left(\frac{1}{\tau}\right)^p \left(1 - \frac{1}{\tau}\right)^{z-p}. \quad (3)$$

In formula (3), γ represents the time slot expectation, τ represents the number of time slots in the identification frame, and z indicates the number of tag transfers. The calculation formula for identifying the number of time slots in the frame is as follows:

$$\tau = \tau_1 + \tau_2 + \tau_3. \quad (4)$$

In formula (4), τ_1, τ_2, τ_3 are the collision time slot, successful time slot, and idle time slot, respectively. The transmission bit length of the reader will be minimized so that the length of the information fed back by the reader to the tag is only 1 bit. The reader only needs to inform the tag whether there is collision or no collision (1: collision, 0: no collision), and there is no need to query the location information of the last collision bit. In the RFID multitag identification process, the reader/writer only broadcasts the select command at the beginning of each frame, and the tag of the current frame receiving the broadcast command can continue to participate in the identification process [13]. Tags that have not received the select command will not participate in the recognition operation of this frame but will be transferred to the next frame for recognition. In the query process, the reader sends one query bit at a time. When the pointing bit of the tag is the same as the query bit, the tag will send the remaining ID bits to the reader, while the tag with different pointing bits and query bits will be converted to the "silent" state, and the remaining query requests will not be answered in this round [14]. The process will continue until one tag is identified; then, reset the remaining unrecognized tags and continue to participate in a new round of identification. The number of tags transferred from the current frame to the next frame is the sum of the number of new qualified tags in the current frame and the number of tags not recognized due to collision. In this state, the process of tag selection time slot can be regarded as the Poisson arrival process by random mathematical knowledge [15]. In the new round of query, only the information of collision location needs to be sent, and the bits without collision no longer participate in the transmission, which reduces the transmission volume by half. This mechanism makes the remaining search process only focus on the previously locked collision bits, which can greatly reduce the redundant transmission of noncollision bits, so as to reduce the total number of identification cycles (time slots). When the ratio of arrival rate to frame recognition efficiency is less than 1, it means that the current number of services is less than the average number of services, and the queuing system composed of the reader, writer, and tag works in a stable state. So far, the design of the RFID multitag identification method based on decision tree is completed.

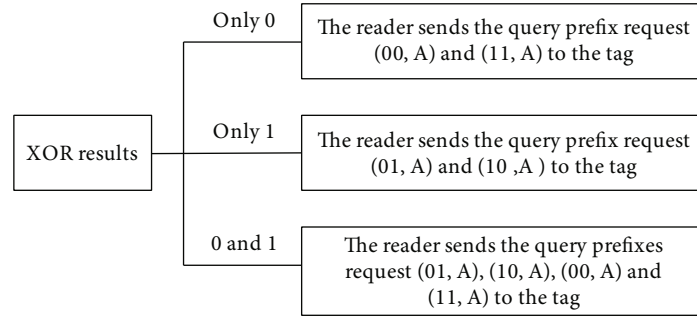


FIGURE 4: The form in which the reader sends the XOR command to the tag.

3. Simulation Experiment

The algorithm is simulated in the MATLAB environment. During the experiment, the overhead of inspection process and control and the necessary communication field length between the reader and the electronic tag are removed, according to the EPCglobal standard [16–18]. The number of electronic tags in the simulation experiment is increased from 50 to 250. The ID bits of each electronic tag are 96 bits and are randomly generated. The experimental results were averaged 50 times. In order to verify the performance of the RFID multitag identification method based on decision tree, the test results are compared with the RFID multitag identification method based on the ALOHA algorithm and perceptual compression.

This experiment mainly analyzes the performance of the RFID multitag identification method based on decision tree from three aspects: time complexity, throughput, and communication overhead. The time complexity is measured by the number of queries, which can accurately reflect the transmission delay when the reader communicates with the electronic tag. The comparison results of query times of each RFID multitag identification method are shown in Table 1.

According to the comparison results in Table 1, the time complexity of each recognition method increases with the increase in the number of tags. Taking the test results when the number of tags is 100 as an example, the number of queries of the RFID multitag identification method based on decision tree is 116.4, which is 84.1 and 110.5 less than that based on the ALOHA algorithm and RFID multitag identification method based on perceptual compression, respectively. The design method in this paper can effectively reduce the query times of the RFID multitag identification method and reduce the time complexity of multitag identification.

Throughput reflects the utilization efficiency of the system to the channel. When tag collision and data transmission error occur on the channel, tag information will be lost and channel resources will be wasted. The throughput can be determined by the ratio of the readable time slot of the electronic tag to the total time slot. The throughput comparison results of various RFID multitag identification methods are shown in Table 2.

According to the comparison results in Table 2, with the increase in the number of tags, the throughput of each method tends to be in a stable value. The throughput of

the RFID multitag identification method based on decision tree is 0.696, which is 0.161 and 0.189 higher than that based on the ALOHA algorithm and perceptual compression. Therefore, the throughput of this design method is maintained in a high range and has strong tag recognition ability.

The communication cost is expressed by the number of transmission bits, which can reflect the energy consumption of the reader and electronic tag in communication. When the number of transmission bits is larger, it means that more instructions should be transmitted between the reader and the electronic tag, which requires more overhead and energy consumption in communication. The comparison results of the number of transmission bits of each RFID multitag identification method are shown in Table 3.

According to the results in Table 3, when the number of tags is the same, the number of transmission bits of the RFID multitag identification method based on decision tree is significantly lower than that of the two comparison methods. Taking the test results when the number of tags is 200 as an example, the number of transmission bits of the method based on decision tree is 489.3, which is 1747.1 and 1969.2 lower than that based on the ALOHA algorithm and perceptual compression. With the increase in the number of tags, the growth rate of communication overhead of the design method in this paper is significantly lower than that of the two comparison methods. Therefore, the RFID multitag identification method based on decision tree can reduce the communication overhead. Based on the above simulation results, the RFID multitag identification method based on decision tree has higher search efficiency and performance.

In order to further verify the recognition effectiveness of the proposed method, the experiment further analyzes the method in this paper, the method in literature [4], and the method in literature [5] to recognize the sample RFID multitags and compares the recognition errors of the three methods. The results are shown in Figure 5.

By analyzing the experimental results in Figure 5, it can be seen that there is an obvious gap in the error of sample RFID multitag identification by using the methods in this paper, literature [4], and literature [5]. With the continuous change of the number of iterations, the error of sample RFID multitag recognition by this method, literature [4] method, and literature [5] method has changed significantly. Among them, the identification error of this method is lower than that of the other two methods, and its minimum is about

TABLE 1: Comparison results of query times.

Number of tags	RFID multitag identification method based on decision tree	RFID multitag identification method based on ALOHA	RFID multitag identification method based on perceptual compression
50	50.2	108.6	132.8
100	116.4	200.5	226.9
150	154.8	262.2	284.7
200	204.9	380.4	401.5
250	253.6	475.3	489.2

TABLE 2: Throughput comparison results.

Number of tags	RFID multitag identification method based on decision tree	RFID multitag identification method based on ALOHA	RFID multitag identification method based on perceptual compression
50	0.684	0.534	0.493
100	0.696	0.518	0.486
150	0.715	0.545	0.495
200	0.702	0.556	0.534
250	0.683	0.522	0.526

TABLE 3: Comparison results of transmission bits.

Number of tags	RFID multitag identification method based on decision tree	RFID multitag identification method based on ALOHA	RFID multitag identification method based on perceptual compression
50	117.8	528.4	537.3
100	234.6	1126.8	1145.7
150	362.7	1639.6	1740.4
200	489.3	2236.4	2458.5
250	648.5	2748.2	2871.2

3%. Therefore, it can be seen that the error of RFID multitag identification of samples using this method is low, which verifies the effectiveness of this method.

4. Conclusion

Due to the advantages of fast identification speed, low cost, high security, strong anti-interference, and adaptability to harsh environment, RFID plays an important task of data acquisition in the perception layer of the Internet of things. It has been widely used in logistics, warehousing, urban transportation, medicine, and other fields. The research on RFID electronic tag information collection and recognition is very suitable for the background of the development of the times and has important reference value and practical significance. This paper proposes an RFID multitag identification method based on decision tree. After testing, this method can reduce the time complexity and communication

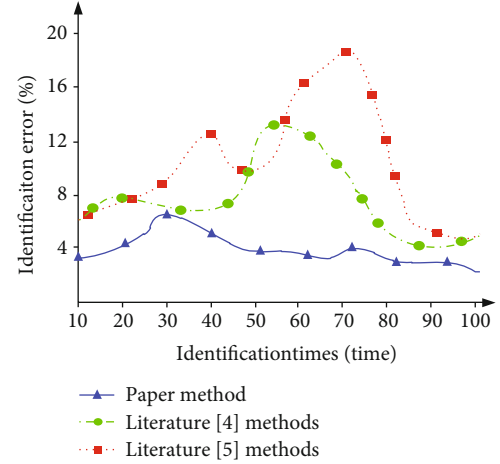


FIGURE 5: Comparison of identification errors of different methods.

overhead and improve the throughput, so it has high search efficiency. The next research should analyze the influencing factors more comprehensively, for example, the automatic adjustment of tag moving speed. The influence mechanism of more factors should also be considered, such as the influence of system energy absorption and signal interference on the performance of the recognition algorithm.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the China Postdoctoral Science Foundation under Grant Nos. 2020M681959 and 2020TQ0291.

References

- [1] K. Fujisaki, "Performance evaluation of table type RFID reader for library automatic book identification," *International journal of web information systems*, vol. 16, no. 1, pp. 65–78, 2019.
- [2] J. Su, A. X. Liu, Z. Sheng, and Y. Chen, "A partitioning approach to RFID identification," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2160–2173, 2020.
- [3] R. Unnikrishnan, O. Rance, N. Barbot, and E. Perret, "Chipless RFID label with identification and touch-sensing capabilities," *Sensors*, vol. 21, no. 14, p. 4862, 2021.
- [4] Z. H. A. N. G. Lijuan, F. A. N. Mingqiu, L. E. I. Lei, Y. Wang, and D. Yuan, "Capture-aware bit-detecting M-ary tree protocol for RFID tag identification," *Journal on Communications*, vol. 42, no. 11, pp. 205–216, 2021.
- [5] Y. He, P. She, L. Zuo, and C. Zhang, "Research on anti-collision algorithm of UHF RFID based on parallelizable

- identification,” *Application Research of Computers*, vol. 37, no. 2, pp. 493–497, 2020.
- [6] T. Dinesh, “Higher classification of fake political news using decision tree algorithm over naive Bayes algorithm,” *Revista Gestão Inovação e Tecnologias*, vol. 11, no. 2, pp. 1084–1096, 2021.
- [7] A. A. Saeed and N. Jameel, “Intelligent feature selection using particle swarm optimization algorithm with a decision tree for DDoS attack detection,” *International Journal of Advances in Intelligent Informatics*, vol. 7, no. 1, pp. 37–48, 2021.
- [8] M. N. Rabbani, A. Yusuf, and D. Rolliawati, “Komparasi Model Prediksi Daftar Ulang Calon Mahasiswa Baru Menggunakan Metode Decision Tree Dan Adaboost,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 10, no. 1, pp. 18–24, 2021.
- [9] P. Kannimuthu and J. Thangamuthu, “Decision tree trust (DTTrust)-based authentication mechanism to secure RPL routing protocol on internet of battlefield thing (IoBT),” *International Journal of Business Data Communications and Networking*, vol. 17, no. 1, pp. 1–23, 2021.
- [10] J. Su, R. Xu, S. Yu, B. Wang, and J. Wang, “Redundant rule detection for software-defined networking,” *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2735–2751, 2020.
- [11] D. Mabuni and S. A. Babu, “High accurate and a variant of k-fold cross validation technique for predicting the decision tree classifier accuracy,” *International Journal of Innovative Technology and Exploring Engineering*, vol. 10, no. 3, pp. 105–110, 2021.
- [12] I. Lim, “Brief paper: simplified tag identification algorithm by modifying tag collection command in active RFID system,” *Journal of Multimedia Information System*, vol. 7, no. 2, pp. 137–140, 2020.
- [13] B. Rahmadya, J. Wang, F. Kong, S. Takeda, K. Kagoshima, and M. Umehira, “Ultra-high frequency band radio frequency identification tag enabling color-change for inventory management systems: a color-change tag,” *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 2, pp. 101–106, 2020.
- [14] W. A. N. G. Li, C. H. E. N. Zhao-xi, and Y. U. Li, “Multi label image classification and recognition method based on conditional random field,” *Computer Simulation*, vol. 37, no. 8, pp. 394–397, 2020.
- [15] L. Wu, P. Sun, Z. Wang, Y. Yang, and Z. Wang, “Toward efficient compressed-sensing-based RFID identification: a sparsity-controlled approach,” *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7714–7724, 2020.
- [16] Y. Yao and J. Su, “An efficient identification algorithm to identify mobile RFID tags,” *Wireless Communications and Mobile Computing*, vol. 2021, 8 pages, 2021.
- [17] J. Su, Z. Sheng, A. X. Liu, Y. Han, and Y. Chen, “Capture-aware identification of mobile RFID tags with unreliable channels,” *IEEE Transactions on Mobile Computing*, vol. 21, no. 4, pp. 1182–1195, 2022.