Advances in Engineering Research, volume 197

Proceedings of the 2020 9th International Conference on Applied Science, Engineering and Technology (ICASET 2020)

Research on Missing Tag Iceberg Query for Multi-Category RFID Systems With Non-Ideal Channel

Zhu Qingqing[1], Song Juan[1], Sun Qiang[2]

[1]*College of Intelligent Manufacturing, Qingdao Huanghai College, Qingdao, Shandong 266427, China*

[2]*Offshore Oil Engineering (Qingdao) Co., Ltd, Qingdao, Shandong 266550, China*

ABSTRACT

Large-scale multi-category RFID system can be widely applied in many applications, such as localization, inventory management and smart logistics, etc. Missing tag iceberg query is one of the key techniques in the above application scenarios. Based on the existing missing tag iceberg query scheme called MAC-SSZE, this paper propose a novel missing tag iceberg query scheme, which adopts the signal integrity check method ECC to solve the problem of transmission error with the non-ideal communication channel. The proposed missing tag iceberg query scheme is based on the framed slotted Aloha protocol, and is able to conduct the self-correction by adding the check data into the broadcasted segments. The simulation results illustrate that the missing tag iceberg query scheme can achieve satisfied performance when there exists transmission error.

Keywords: *Radio Frequency Identification, Missing Tags; Iceberg Query, Signal Verification, ECC.*

1. INTRODUCTION

The Internet of things widely uses various communication sensing technologies such as intelligent sensing, RFID and pervasive computing for terminal sensor communication, among which RFID is a new sensing technology that can make the device "speak". RFID system tags store some important equipment parameter information set in advance, such as the tag's equipment code and other information. The RFID reader terminal collects the information in the tag through the wireless communication network and feeds it back to the central processing system to judge and identify the status of the device. Then it can further realize the "transparent" monitoring of the device over a long distance through the Internet.

RFID system is widely used in many aspects, such as logistics management, equipment monitoring and object positioning [1]. How to estimate the number of active tags in real time and accurately is a key problem in these applications? Because the server of RFID system keeps the backup data of all tags in the whole system, the number of tags lost in the system can be obtained indirectly by monitoring the number of active tags in real time. Moreover, by setting the security threshold of the number of tags lost, the system can alarm when the number of tags lost suddenly exceeds the security threshold. At the same time, in a large-scale RFID system, there are different tags in different states. How to monitor these tags in real time is of great practical significance?

In addition, in the research of RFID system, including the number of system tags detection [2], unknown tag identification[3]and the number of lost tags statistics[4], most of the researchers are based on the ideal transmission channel for system modeling, but the actual transmission channel of the signal has noise and electromagnetic interference, so for the non-ideal channel for system modeling It is necessary to carry out further research later.

In wireless sensor networks, although the network bandwidth continues to improve, the phenomenon of bit error and packet loss in the network has not been improved[5]. To solve this problem, the most common method is to judge whether the data stream has error code and packet loss exception by the receiver. If there is an exception, the receiver will request the sender to resend the packet. If the receiver judges that the packet is lost again when the sender resends the packet, it needs to continue to request the resend packet. But in some applications, this simple data transmission scheme will bring great problems. Especially in some applications with high real-time requirements, multiple communication between sender and receiver caused by bit error or packet loss will bring high network delay and increase network load. Especially in a multicast network environment, such as RFID system, if the above mechanism of packet loss and re-transmission is adopted to deal with the unstable channel environment, then in the case of serious packet loss rate, a large number of receiver nodes request re-transmission packets from sender nodes, which will bring disastrous consequences to the multicast network environment.

The practical application environment of RFID system is full of all kinds of electromagnetic interference, so it is of great practical significance to study how to carry out data transmission in multicast network under non ideal channel, and how to deal with the problem of bit error and packet loss in multicast network.

2. RELATED WORK

A large number of RFID system related research literature, including RFID active tag and lost tag number estimation, unknown tag identification and other algorithms [6][7][8], are mostly based on the ideal wireless transmission channel as the model, without considering that there may be electromagnetic interference in the information transmission between the reader and the tag, and there may be abnormal problems of data bit reversal and data packet loss [9]. Sze et al. Considered the unreliability of the channel, then proposed the MSMD re and MSMD be methods [10][11]. The author thinks that there may be some abrupt abnormal time slots in the actual channel due to noise and other electromagnetic interference, for example, the noise of the channel may turn a space-time slot into a single time slot or a conflict time slot. The author reconstructs the system model and considers two models: random noise model and burst noise model. At the same time, in order to eliminate the possible noise impact, the data format between reader and tag is reconstructed, and CRC verification is introduced. This method can check the data frame with error or packet loss, and then apply for re-transmission to improve the stability of communication.

If only one or two bits in the 96bit data frame are reversed, it will be very time-consuming if the re-transmission mechanism is still required. At this time, an algorithm can be used to automatically correct the error data with one or two bits reversed, so as to avoid multiple applications for re-transmission and improve efficiency.

3. ICEBERG CLASSFICATION METHOD OF MISSING LABEL

3.1. System Model

At present, RFID system has been widely used in cargo monitoring, tracking and intelligent logistics. These tags can be divided into two categories: passive tags and active tags. Generally speaking, the communication distance of the passive tag is less than 20 feet, and the active tag can provide a longer communication distance due to the battery power supply, generally reaching 300 feet [12][13].

The label model selected in this chapter is the active label because some more complex methods can be performed. The reader uses the frame time slot protocol as the protocol to communicate with the tag. The reader first broadcasts system parameters $< f, R >$, which f is called frame length, and R are random number seeds used by tags for hash functions. When the tag receives the parameters broadcast by the reader, it first uses its own ID and random number R to calculate the hash value $H(ID, R)$, and then leaves the hash value with the frame length f to get the time slot index for response. The time length of the tag response slot depends on which kind of information the tag responds to. The EPC global C1G2 standard specifies three time slots t_{tag}, t_l and t_s, which respectively means that the tag sends a 96bit ID, the tag sends a long response and a short response. This chapter sets $t_{tag} = 2.4ms$, $t_l = 0.8ms$, $t_s = 0.4ms$. RFID system includes a background server, a reader and several active tags. The background server has a strong ability of data storage and calculation. The reader can communicate with the background server through a high-speed path. In order to simplify the system model, the background server and the reader are considered as a whole, and the reader is used instead of the background server. Because labels can be pasted on different kinds of articles, suppose there are labels of different kinds, which are expressed as C_1, C_2, …And C_n. In this paper, the first S $(1 < S < 96)$ bit of the 96bit tag ID is used to represent the class ID of different class tags. The number of tags in the whole RFID system is n, the number of class i tags n_i $(1 < i < l)$, $n = \sum_i^l n_i$ which is obvious.The number of lost tags in each type of tag is defined as m_i $(1 < i < l)$, and the number of lost tags in the whole RFID system is m, obviously $m = \sum_i^l m_i$. It is also assumed that the reader has stored the ID information of all category tags, because the background server has maintained all the system information.

The iceberg classification problem of missing tags in multi-category RFID system can be described as: give a preset threshold T, error $\varepsilon \in (0,1]$ and classification accuracy $\delta \in [0,1)$, the classification method needs to meet the following two constraints:

$$\forall C_i \in \Gamma, \hat{m}_i \geq T, \Pr[m_i \geq (1-\varepsilon)T] \geq \delta \qquad (1)$$

$$\forall C_i \notin \Gamma, \hat{m}_i < T, \Pr[m_i \leq (1+\varepsilon)T] \geq \delta \qquad (2)$$

The expression C_i in the formula means the number of missing tags estimated in the category. For example, the constraint condition (1) represents the label category belonging to the set Γ, the estimated number of lost labels shall not be less than the threshold T, and in fact, the probability that the number of lost labels is greater than or equal to $(1-\varepsilon)T$ is at least δ. The constraint condition, as shown in formula (2), indicates the label category not belong to set Γ, which is less than the threshold T, and the probability that the number of lost tags is less than or equal to $(1+\varepsilon)T$ is at least δ.

3.2. MAC-SSZE Algorithm Model

The algorithm of MAC-SSZE (Missing tag iceberg query using segmented enhanced scheme which adapts the Singleton-Ze) includes w round robin,selecting the k-th cycle arbitrarily. The reader will first broadcast the system parameters $< f_k, R_k >$, f_k indicating the frame length, R_k indicating the random number seed. These parameters will be used by the tag to calculate the hash value, which is used to indicate the index of its response slot at the frame. Note that in MAC-SSZE algorithm, the reader uses the frame length parameter f_k at the beginning of k-th the loop, which is no longer equal to the total number of tags in the system. It changes dynamically with the reader loop estimation. The specific calculation process is discussed below. In consideration of the direct communication between the reader and the background server, the ID information of all the grouped tags can be obtained. Because the reader also knows the system parameters $< f_k, R_k >$ of the frame, the reader can calculate the hash value of all the tags in the background server in advance, and use the hash value as the slot index of the frame. The frame generated by the calculation in advance is called the expected frame.

In MAC-SSZE algorithm, the reader can generate a vector of f_k bit length based on the expected frame. If the state of a slot in the expected frame is a single slot state, the bit value at the vector V corresponding index is 1; if the state of a slot in the expected frame is a space slot state or a collision slot state, the bit value at the vector V corresponding index is 0. As shown in Figure 1, the reader divides the vector V into $\lceil \frac{f_k}{96} \rceil$ segments, each segment is96 bit long, and then the reader broadcasts each segment in turn.

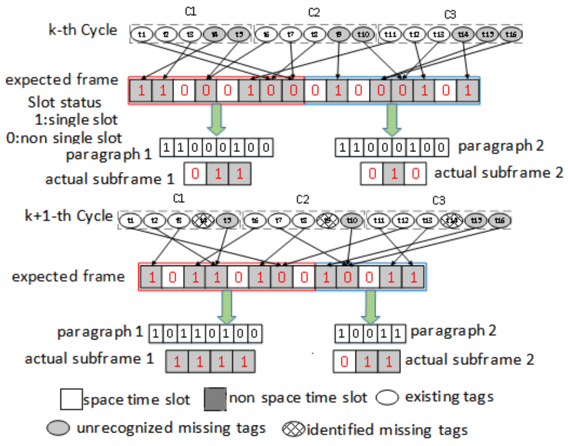


Figure 1 Illustration of the missing tag iceberg query procedure of the MAC-SSZE scheme

When the reader broadcasts each segment, the reader will enter the state of detecting subframe, in which the reader will wait for the response of the tag. After receiving the system parameters $< f_k, R_k >$ broadcast by the reader, the tag first calculates which segment it will respond to according to the stored ID. If the tag calculates that the segment to be

answered matches the segment broadcast by the reader, the tag will receive the 96bit information of the next reader broadcast. If the corresponding bit value is 1, then the slot state to be answered by the tag is single slot, and the tag will respond in the monitoring subframe state of the reader, otherwise the tag will not respond.
When the tag determines to respond to the 96bit information broadcast by the reader in this segment, the tag needs to count the number of 1 values between the beginning of the segment and the corresponding response bits in the 96bit information. For example, if the number of 1 is s, then the tag responds to a short response in the $s+1$ time-slot of the response in the subframe state monitored by the reader, otherwise the tag will not participate in the reader monitoring Response to subframe status. When the reader monitors the subframe status, the reader will monitor the status of each time slot. After the reader monitors the subframe status, the reader will get a subframe. If the status of a time-slot in the frame is 0, the corresponding label of the time-slot is lost. Otherwise, the label corresponding to the time slot is not lost. If the reader can be sure that some tags must be lost after the end of the k -cycle, the lost tags will not participate in the $k+1$ -cycle.
At the end of the k -cycle, the reader estimates the number of lost tags for each type of tag by comparing the difference between the expected frame and the subframe composed of 96 bits of information. When the k-cycle ends, the reader completes the task of lost tag classification on the premise of satisfying the constraints.

4. ICEBERG CLASSIFICATION METHOD BASED ON NON-IDEAL CHANNEL

For the iceberg classification problem of the number of lost tags in the RFID system of multi-category tags, MAC-SSZE can ensure a high time efficiency under the pre-set accuracy of the system. But the algorithm model is based on the ideal channel. In fact, the electromagnetic interference in the transmission signal may cause some data bit errors. Therefore, this paper modifies MAC-SSZE lost label iceberg classification algorithm based on non ideal channel, so that it can complete the iceberg classification of the number of lost labels in multi-category RFID system under the transmission condition of non ideal channel.

4.1. Mac-ssze Algorithm Based on Non-ideal Channel

In the RFID system model, firstly, it is assumed that the short response of the tag to the reader does not need to be verified, while the 96bit data sent by the reader to the tag, in order to ensure the stability of data transmission, the tag end needs to provide the verification function. In MAC-SSZE method, readers need to continue to broadcast 96bit data frames to tags after broadcasting system parameters to tags. The reader uses this frame to mark the tag to index the response of single time slot and similar time slot, so that the tag in RFID system will respond according to the index of 1 in the frame broadcast by the reader, and the index time slot of 0 will not respond. Because the proportion of the number of frames broadcast by the reader with a median value of 1 is limited, the tag does not need to respond to all the frame slots, so it can improve the time efficiency of the algorithm.
However, when the reader broadcasts a frame of f slot length, there may be other electromagnetic interference in the communication channel between the reader and the tag, which may result in packet loss or bit error. For the MAC-SSZE algorithm without auxiliary verification, because there is no branch of the label program to deal with the abnormal situation, it still misjudges the abnormal situation as the normal situation for processing, which reduces the recognition accuracy of the algorithm and affects the accuracy and reliability of the algorithm.
If CRC cycle check is added to the tag program, the tag can accurately determine whether the current frame has packet loss or error, and whether it is necessary to apply to the reader to resend the 96bit frame until it passes the CRC check of the reader. This verification method can determine the abnormal frame caused by environmental interference. However, no matter how many bits of the 96bit data are wrong, the tag needs to apply to the reader for re-transmission. This cycle of re-transmission takes a long time, although it will not affect the classification accuracy, but it will extend the operation time.
If ECC algorithm is added to the tag program, when the tag detects 1 bit data error in the 96bit data, the tag can calculate which bit of the frame has error by itself and correct it to the correct data. If the tag detects that there are 2 bit data errors in the 96bit data frame, the tag can calculate which bit of data has errors, but it cannot correct itself. The tag requests the reader to resend the data frame. After that, the processing mode of ECC is similar to that of CRC.

4.2. MAC-SSZE Algorithm with ECC Check

The MAC-SSZE algorithm using ECC check is implemented by appending check bits to the verified data bits. For example, when the data bits are 8 bits, 7 check bits need to be added for ECC check, when the data bits are 16 bits, 8 check bits need to be added for ECC check, and so on. When the data bits are 64 bits, 12 check bits need to be added for ECC check. In the non ideal channel, the lost label iceberg query method based on MAC-SSZE is shown in Figure 2. In this section, the model of MAC-SSZE method is modified so that it can effectively use ECC check method to check data bits and improve the stability of its signal transmission.

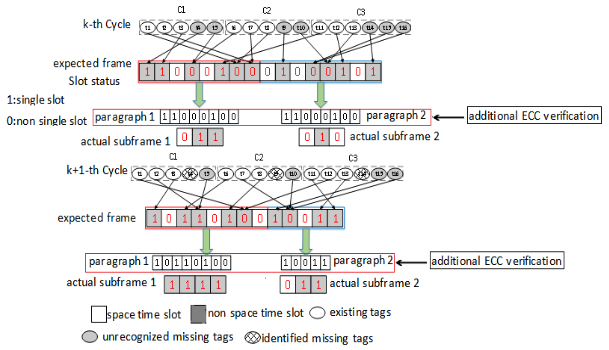


Figure 2 Illustration of the missing tag iceberg query procedure of the MAC-SSZE scheme with non-ideal channel

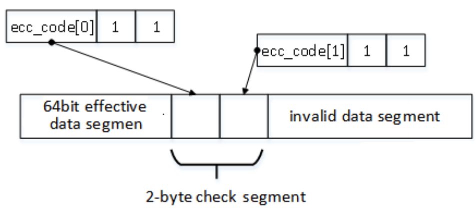


Figure 3 Structure illustration of the 96-bit data

In RFID system, the length of data frame broadcast by reader is 96bit, so we need to modify the frame structure of 96bit to adapt to ECC algorithm. As shown in Figure 3, this paper divides 96bit into two parts, one is 64bit long, the other is 12bit long. The former is the data to be broadcasted by the reader itself. After modification, the effective data of each frame becomes 64bit. The number of frames to be broadcasted is $\lceil \frac{f}{64} \rceil$, and the latter is the 12bit check data of ECC check. The remaining 20bit is the index segment of the data frame, which is not considered in the simulation.

The reader broadcasts 96bit data one frame at a time, including 64bit valid data, 12bit ECC check data and 20bit other data. When a frame of data is broadcast, all tags will first calculate the 12bit ECC check code according to the 64bit of the received 96bit data, and XOR the calculated ECC check code and the received ECC check code bit by bit. If the result is 0, it means that the transmission of the frame is not interfered by the surrounding electromagnetic environment, and the algorithm can proceed to the next step, that is, the tag can follow the and MAC-SSZE algorithm responds to the time slot in a similar way; on the contrary, if the operation result is not all 0, it indicates that there is an error in the transmission of the frame, and if there are 6 ones in the operation result, it indicates that there is a bit error in one of the effective data bits of the frame, and it is necessary to call the correction algorithm to process the data at the error code, and get the correct effective data; otherwise, it indicates that there is more than one bit If there is an error in the effective data bits, and the correct data cannot be obtained when using the ECC correction algorithm, then the data frame can only be applied to the reader for re-sending.

5. SIMULATION RESULTS AND ANALYSIS

Assuming that the tag has the ability to execute CRC and ECC verification codes, this paper divides the classification accuracy into two parts, which are expressed as δ_1 and δ_2. δ_1 represents the ratio between the number of label categories in the set classified as Γ by the algorithm and the number of actually lost labels greater than or equal to the number of $(1-\varepsilon)T$ label categories. δ_2 represents the ratio between the number of label categories classified by algorithm δ_1 as not in the set of Γ and the number of actually lost labels less than or equal to the number of $(1-\varepsilon)T$ label categories.The goal of this paper is to get the shortest algorithm execution time under the premise of satisfying $\delta_1 \geq \delta$ and $\delta_2 \geq \delta$. Where ε is the error, T is the threshold of the number of lost tags, and δ is the classification accuracy.

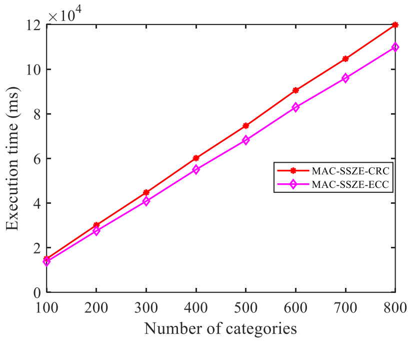
Figure 4 shows the effect of ECC algorithm and CRC algorithm on the performance index of execution time under the condition of different number of label groups l by setting different precision. The precision is set to $\delta = 0.85$, $\delta = 0.9$, and $\delta = 0.95$ respectively, the number of lost labels threshold $T = 50$, error $\varepsilon = 0.2$, and the number of labels in each group is 150. This chapter assumes that the number of missing tags in each group follows the normal distribution $N(\mu, \sigma^2)$, where $\mu = 50$, $\sigma = 20$. The change range of the number of groups of labels is [100,800].

Under the premise of three different classification accuracy, the simulation results of MAC-SSZE algorithm with ECC check and CRC check under different tag groups are shown in Figure 4, and the following three conclusions can be drawn:

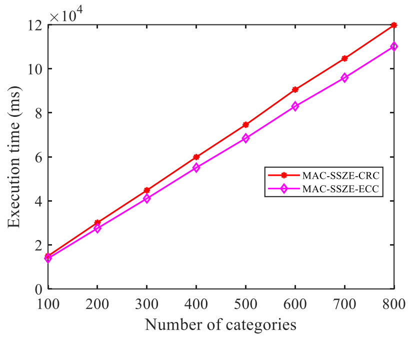
(1)Under the premise of setting different simulation accuracy, with the increase of the number of tag groups in the simulation, the simulation time required to complete the iceberg classification using the MAC-SSZE algorithm of ECC verification and CRC verification respectively increases, because when the reader needs to increase the number of tag groups to be classified, the number of tags to be queried increases, and the two methods need to estimate the lost tags The length of the desired frame slot will increase, and the time interval of each frame slot is a fixed value, so the simulation time will increase.

(2)With the increase of classification accuracy δ, the simulation time of MAC-SSZE algorithm using ECC and CRC will increase gradually. The reason is that for a specific missing label classification test data set, the classification accuracy δ increases from 0.85 to 0.95, and each classification algorithm needs to run for a longer time to reduce the variance of the estimated value to meet the set classification accuracy.

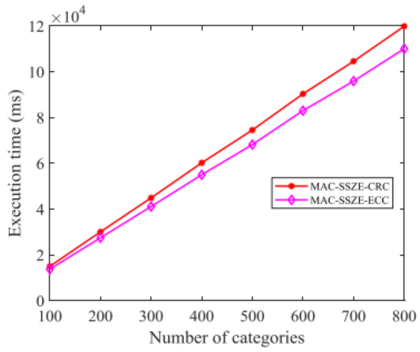
(3)Generally speaking, under the same iceberg classification accuracy δ and the number of label groups, the simulation time trend is MAC-SSZE-CRC>MAC-SSZE-ECC, and MAC-SSZE-ECC classification method has the highest time efficiency under the same conditions. Because the ECC algorithm can correct itself when there is 1 bit data error, but the CRC method can't correct the error, so it can only inform the reader to resend the data frame, which wastes a certain time.



$\delta = 0.85$



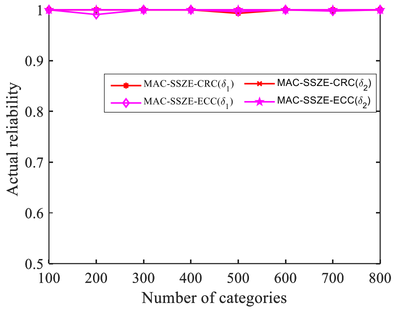
$\delta = 0.9$



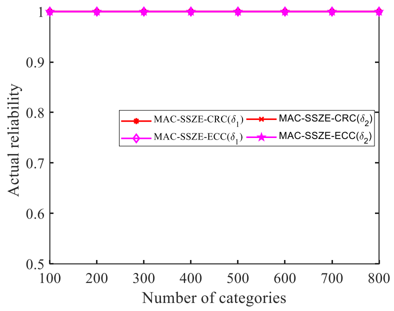
$\delta = 0.95$

Figure 4 Impacts of number of missing tags in each cate gory on the performance of execution time of schemes us ing different checkout algorithms with different required reliability

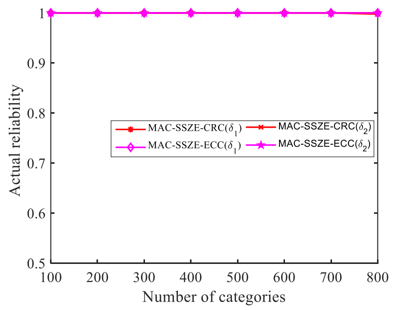
Figure 5 shows the effect of ECC algorithm and CRC algorithm on the performance index of actual estimation reliability under the condition of different tag groups l by setting different precision. The precision is set to $\delta = 0.85$, $\delta = 0.9$ and $\delta = 0.95$ respectively, and the parameter setting of simulation is consistent with the above. The simulation results show that all the algorithms can meet the requirements of $\delta_1 \geq \delta$ and $\delta_2 \geq \delta$, which shows that the classification method using any kind of checking algorithm can meet the precision requirements set in advance.



$\delta = 0.85$



$\delta = 0.9$



δ=0.95

Figure 5 Impacts of number of missing tags in each cate gory on the performance of actual reliability of schemes using different checkout algorithms with different required reliability

6. CONCLUSION

This paper mainly studies the problem of data transmission in RFID system with non-ideal channel. Firstly, after introducing the previous work based on the ideal channel transmission model, this paper introduces the non-ideal channel transmission model, and then analyzes two label quantity estimation algorithms considering channel noise and electromagnetic interference, i.e. CRC verification algorithm and ECC verification algorithm. Finally, the lost label iceberg classification method using ECC verification has the ability of self-correction by simulation It can save the query time of the reader and improve the working efficiency.

REFERENCES

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications [J]. IEEE Communications Surveys & Tutorials, 2015: 2347-2376.

[2] C. Qian, H. Ngan, Y. Liu, and L. M. Ni. Cardinality Estimation for Large-Scale RFID Systems [J]. IEEE Transactions on Parallel and Distributed Systems, 2011, V22(9):1441-1445.

[3] X. Liu, K. Li, J. Wu, A. X, X. Liu, X . Xie, C. Zhu, and W. Xue. Top-k Queries for Multi-category RFID Systems [C]. In Proc. of The 35th Annual IEEE Conference on Computer Communications, San Francisco, CA, USA, 2016:1-9.

[4] H. Chen, G. Ma, Z. Wang, F. Xia, and J. Yu. Probabilistic Detection of Missing Tags for Anonymous Multi-Category RFID System [J]. IEEE Transactions on Vehicular Technology, 2017:11295-11305.

[5] G. Casella and R. L. Berger. Statistical Inference [M]. 2nd Edition. Duxbury Press.2002.

[6] H. Chen, G. Xue, and Z. Wang. Efficient and Reliable Missing Tag Identification for Large-Scale RFID Systems with Unknown Tags [J]. IEEE Internet of Things Journal, 2017, V4(3):736-748.

[7] L. Xie, H. Han, Q. Li, J. Wu, and S. Lu. Efficient Protocols for Collecting Histograms in Large-Scale RFID System [J]. IEEE Transactions on Parallel and Distributed Systems, 2015, V26(9):2421-2433.

[8] W. Luo, Y. Qiao, and S. Chen. An efficient protocol for RFID multi-group threshold-based classification [J]. IEEE/ACM Transactions on Networking, 2013:890-898.

[9] M. Chen, W. Luo, Z. Mo, S. Chen, and Y Fang. Efficient Tag Search Protocol in Large-Scale RFID Systems With Noisy Channel [J]. IEEE/ACM Transactions on Networking, 2016, V24(2):703-716.

[10] Sze. W. K, Lau. W. C, Yue. O. C. Fast RFID Counting under Unreliable Radio Channels [C]. In Proc.

of the 9th IEEE International Conference on Communications, Dresden, Germany, 2009:1-5.

[11] P. Xu,Y, Su, Design and Implementation of landscape system for East and West Huashi Street in Beijing based on virtual reality technology, Applied Mechanics and Materials, 263(2012)1849-1852.

[12] Honglong Chen, Guolei Ma, Zhibo Wang, Qian Wang, and Jiguo Yu. MAC: Missing Tag Iceberg Queries for Multi-Category RFID Systems. In IEEE Transactions on Vehicular Technology, 2018, V67(10):9947-9958.

[13] Li Yongwei. Research and implementation of RFID positioning system [D]. Xi'an University of science and technology, 2017.

[14] P. Xu, Study on Accurate Measurement Technology for Microscopic Image, Advanced Materials Research, 798(2013)638-642.