RFID Anti-Collision Technology in Big Data Environment

Buying Chen
Xingtai Polytechnic College, Xingtai, China
chenmo62@163.com

Abstract—Along with the rapid development of Web of Things, the RFID technology is widely applied in every field, but today, the great challenge we face is how to avoid information conflict and collision in the process of acquisition and treatment of massive information. It is the keystone of the study. This paper conducts a comparative analysis on different Tree-based algorithms as improved, and integrated with multiple sub-cycle response mechanism, a Multi-Response Collision Tree algorithm is proposed. And beyond that, this paper simulates and analyzed this algorithm and other improved ones. The results reveal that, compared to other algorithms, MRCT algorithm features a better performance, less recognition cycles, least query time slots on average, and ceiling throughput rate.

Keywords—RFID; anti-collision; Tree-based algorithm; Manchester code

1   Introduction

The popularity of Web of Things attributes to the spread of Internet. WoT enables information interaction between humans and things, things and things, people and spaces with perceptive and wireless communication technologies [1-3]. As a plurality of applications have sprung up, such as 3D printing, smart wearable devices and unmanned technologies, the WoT has ushered in a new era of rapidly developed 2.0 version. This technology began to develop on a large scale, collaboratively and intelligently, and has infiltrated into all fields of our daily life, so that it plays a strategic role in the development of China's economy. However, the first level of the WoT, i.e. the perceptive technology, faces a severe challenge [4]. The performance of perceived information is poor under the complex scale of the WoT, while the energy and time consumptions of data transmission hit the mark. Whence, how to improve the reliability and availability of perceptive technology in big data environment has become the focus the current WoT technology should conquer [5-6].

In relation to other recognition technologies, Radio Frequency Identification (RFID) technology is finding wider and wider application due to its advantages of non-contact, multi-target recognition and strong anti-interference. In 2003, the United States began to vigorously spread FRID; after 2007, the FRID industry showed a development in super express and on a large scale; in 2016, China's market volume

reached 53 billion. More and more retail industries were also deploying FRID programs, including some technology companies who also turned their eyes to the development of RFID technology. Credited with these, China has been armed with technological power in the fields of its industrial structure and economic development with strategic significance [7].

In the era of rapid development of the WoT, Japan and South Korea, as neighbors of China, took the lead in planning and implementing the Internet and the WoT. In the E.U., CERP-LOP program team has also been set up, and 38 relevant projects have been completed since its inception. The development of WoT technology involves a string of standards and specifications pertaining to FRID technology, and the current standards are the EPC Global proposed by North America, EU, AIM, ISO, UID led by the EU and Japan, and the IP-X led by Africa and Asia. China started relatively late in the RFID field. In 2006, China issued a white paper on the policy of radio frequency identification (RFID) and funded 19 FRID-relevant projects through the 863 Program. In the field of FRID anti-collision technology, the constant investigation on its application in big data environment has greatly stimulated the development of WoT 2.0, for example, the acquisition and the treatment of a huge mass of information reveal the features of FRID in WoT technology. In the process of information acquisition, the information conflicts and collision occur when multiusers access to target frequently. This directly affects system identification performance and throughput rate. The current solutions mainly include the Time Division Multiple Access (TDMA), the Space Division Multiple Access (SDMA), the Frequency Division Multiple Access (FDMA) and the Code Division Multiple Access (CDMA), but these mainstream technologies often run short of ways in dealing with huge mass of information. To meet this challenge, the multi-label identification for big data is proposed as the keystone of present study [8].
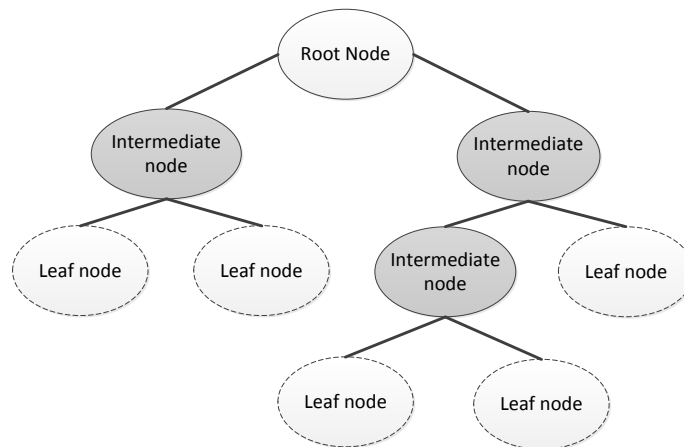
2 Tree algorithm

Fig. 1. Tree shape algorithm two forked tree structure

FRID system label collision refers to unrecognized information by labels due to signal superposition when a number of labels send the message in parallel. In order to reduce collisions and conflicts, an anti-collision algorithm is required to optimize FRID system for the purpose of improving system identification efficiency and anti-interference capacity.

Tree-based algorithm is to differentiate the labels to which collision has occurred into various subsets constantly, until the collision-free labels are identified. It gets its name for the process seems like a tree [9-10], as shown in Fig. 1. The initial node is the root one, the collision subset is the intermediate node, the identified is leaf node and the free time slot node is also the leaf one.

2.1 Query Tree protocol (QT)

The Query Tree Protocol (QT) consists of a sequence of queries and responses [11-13]. It, also known as a memory-free algorithm, first stores the tree structure in a reader in the absence of memory function. At the onset, the reader sends a query string prefixed with 0 and 1 to the label, if matched, the ID of the label returns to the reader; if not, it does not make response. when multiple labels match, this is the case of a collision, the string will be added with 0 or 1 to generate a new prefix, it re-queries until no collision occurs.

For example, QT algorithm finishes the identification of a set of labels (1010, 1100, 1011) within 9 time slots. as shown in Fig. 2.
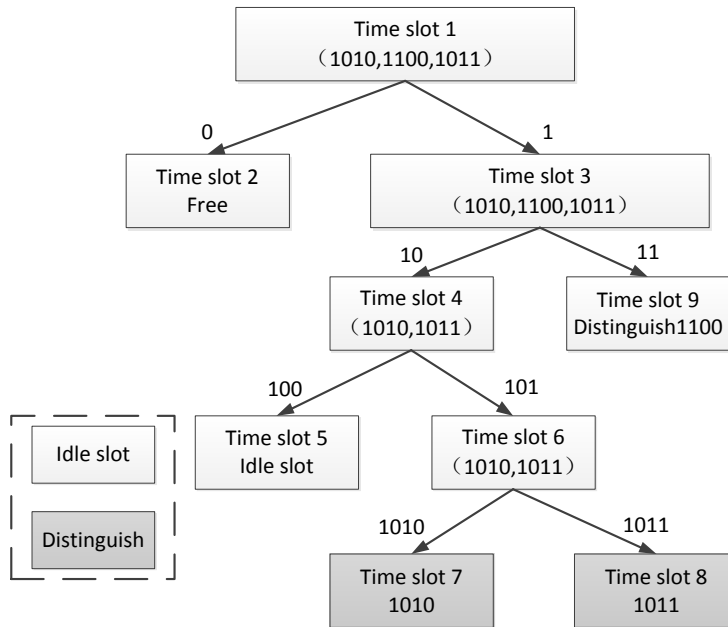

Fig. 2. Tree graph of QT algorithm

In Fig. 2, we can see that the algorithm is a binary tree. When the prefix Q0 is sent, the label makes no response. Then collision must occur when sending prefix Q1. The algorithm can be improved to send Q10 and Q11 directly, for which reduces the waste of time slot caused by sending Q1.

2.2 Jumping and dynamic searching algorithm

In traditional FRID system, non-return-to-zero code is often used, i.e. 1 and 0 denote high level and low level, respectively. When the sequence AB sends the code sync, and any bit of the information and the number received by the reader is 1, it is unlikely to determine whether there is a collision in the bit, as shown in Fig. 3.
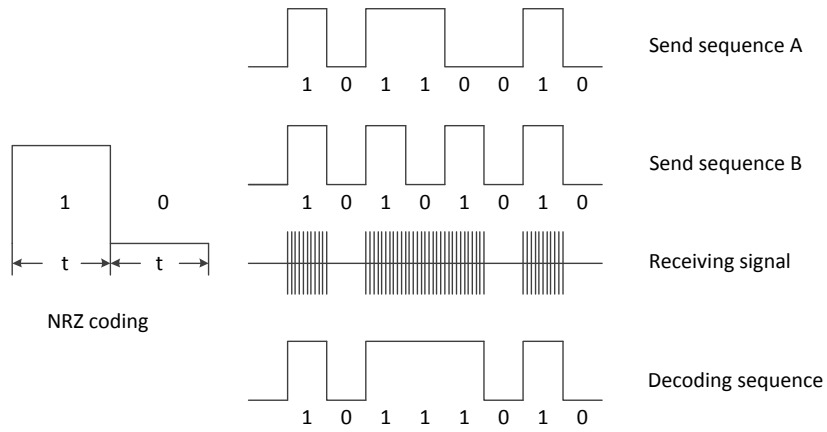

Fig. 3. Coding and decoding diagram of non return to zero
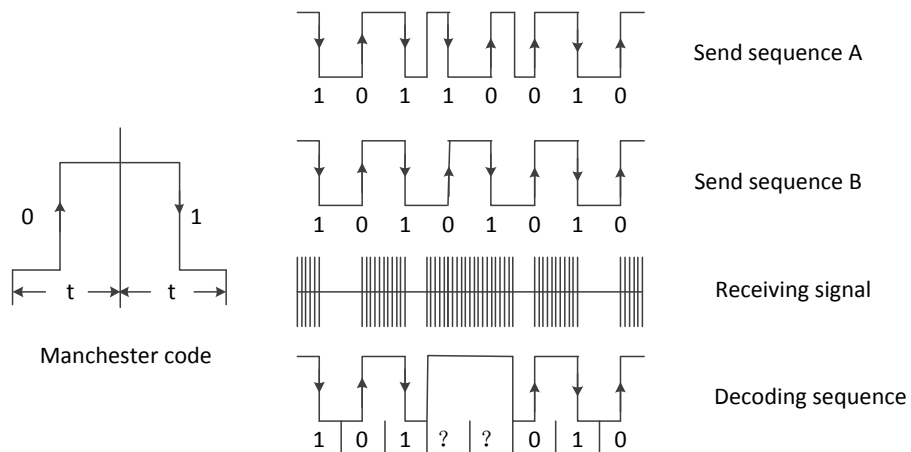

Fig. 4. Manchester coding and decoding diagram

Manchester code [14] is encoded by phase, but there is hop in the signal acquired. When the code is 1, the high level transfers to a low one; when the code is 0, the low level changes to a high one, that is, the level transition is based to judge whether collision occurs in the bit, as shown in Fig. 4.

Jumping and Dynamic Searching (JDS) Algorithm [15-17] means that the time slot enquiry sequence when a collision occurs is stored as the start point of next enquiry. In this way, a re-enquiry by returning to the root node is not required after each label identification, so that the subset enquired each time gets smaller and smaller. The identification rate is thus increased. The enquiry sequence in this algorithm include the query prefix and the collision bit pointer, expressed by (Q, P). The identification process is shown in Fig.5.

2.3 New Enhanced Anti-Collision Algorithm (NEAA)

The New Enhanced Anti-Collision Algorithm (NEAA) is a tree algorithm that subdivides labels by the numbers 1 and 0 into groups for identification. This algorithm can identify multiple labels in the one time slot, and the more the labels, the more it recognizes in parallel. It is ideal for the identification in big data scenarios. For example, the labels 0001, 0010, 0100, and 1000 have one number 1 and three numbers 0, and the numbers are transferred by Manchester code. As the reader knows that there are 4 labels in this group and only one number 1 in each label, after all collisions occur, the reader receives the signal xxxx, then the label group can be judged as 0001,0010,0100,1000, that is to say, multiple labels can be identified simultaneously. Take the label group (1010, 1100, 1011) as an example, the identification process is shown in Fig. 6.
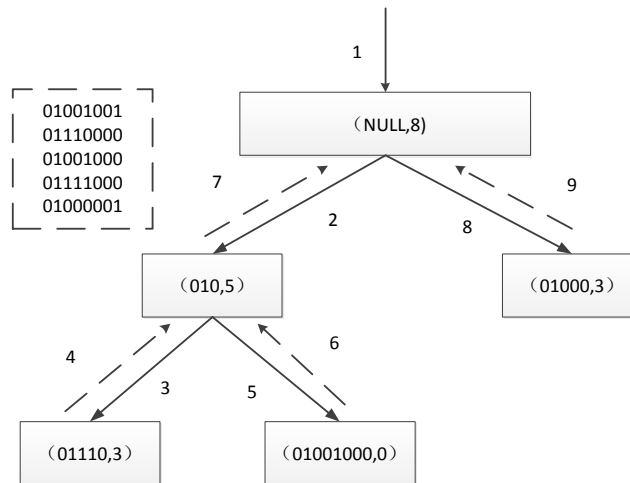

Fig. 5. Example schematic diagram of JDS algorithm
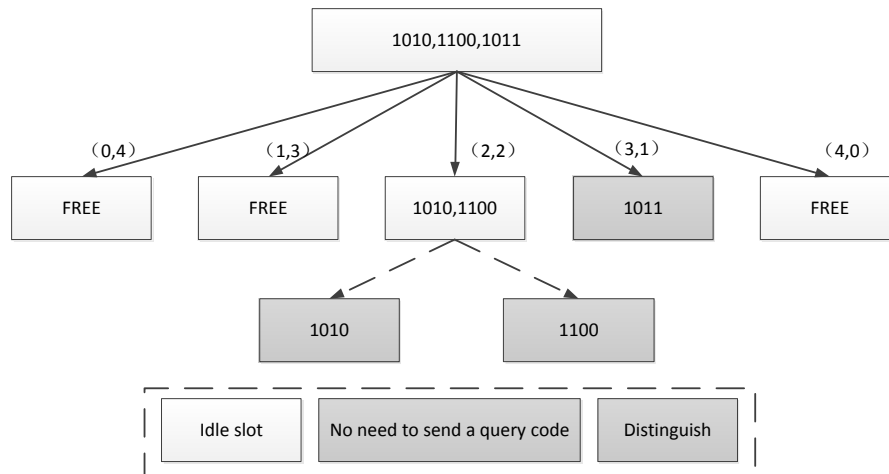
Fig. 6. Example schematic diagram of NEAA algorithm

2.4 Simulation comparison of improved tree algorithm

Aloha and tree algorithms are two key types of anti-collision algorithms in FRID system. In big data environment, due to the randomness of Aloha, some labels are not recognized for a long time, however it is the disclosure of the advantage that tree algorithm can identify multi-labels. Several improved tree algorithms are simulated and compared in several ways such as time complexity, average query time slot, average communication complexity and algorithm throughput. The results are obtained as follows.

Analysis of time complexity simulation algorithm. The time complexity refers to the cycles required for all labels to be identified. Select the label length of 96, the number of labels is 4096, and assume that the labels are not lost, the simulations of QT, JDS, NEAA algorithm are performed with the results as shown in Fig. 7. It can be seen that as the number of labels increases, the total time slots of the three algorithms all go up linearly. However, the growth rate of QT algorithm in time slot is maximum, JDS and NEAA algorithms are roughly equal, and NEAA is slightly smaller.

Average query time slot. The average query time slot refers to the time slot required for successfully identify a label. QT, JDS, NEAA three algorithms is performed for simulation with the results as shown in Fig. 8. It can be seen that the average query time slot of QT is significantly greater than that of JDS and NEAA. Except for less labels, NEAA's average query timeslot is slightly higher than that of others, that is, basically approximate to that of JDS. This is because the NEAA algorithm has more empty time slots when grouping in early time, while the JDS adopts the rebound mechanism and does not re-query.

Simulation and Analysis of Average Communication Complexity. The average communication complexity refers to the average transmission quantity of bits required for a label to be identified. Assume the label length is 96, QT, JDS, NEAA algorithms

are adopted for simulation with the results as shown in Fig. 9. It can be seen that the QT algorithm transmits the maximum number of bits, JDS does less while NEAA does least. This is because the QT algorithm increases bits for each search with more repetitive processes, thus features inefficiency and great transmission quantity. The JDS and NEAA algorithms achieve simultaneous identification of multiple labels by packets, thus reducing the number of bits sent.

Analysis of throughput simulation. The algorithm throughput refers to the number of identifiable labels in the unit query timeslot, expressed as S = total number of labels / total query timeslot. QT, JDS, NEAA are performed for simulation with the results as shown in Fig. 10. It can be seen that the QT algorithm has the lowest throughput of 0.35, close to the Aloha. JDS is stable at about 0.5. The NEAA falls within 0.45 to 0.5 when there are less labels, and with the labels increase, it is stable at about 0.5. This is because NEAA generates empty time slots when the number of labels is lower, which worsens the identification efficiency and affects the throughput rate. However, in the big data scenario, the NEAA algorithm will eliminate this defect.
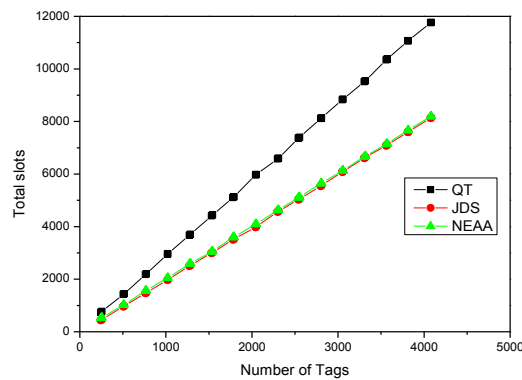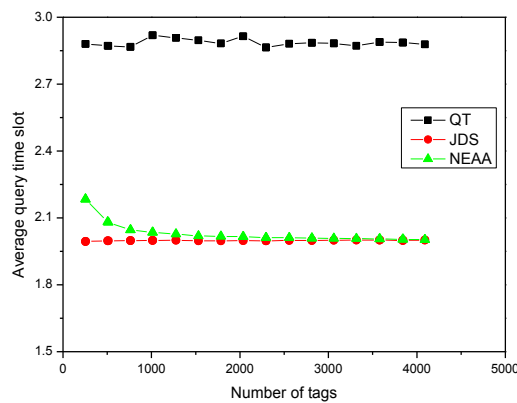

Fig. 7. Total slots contrast diagram

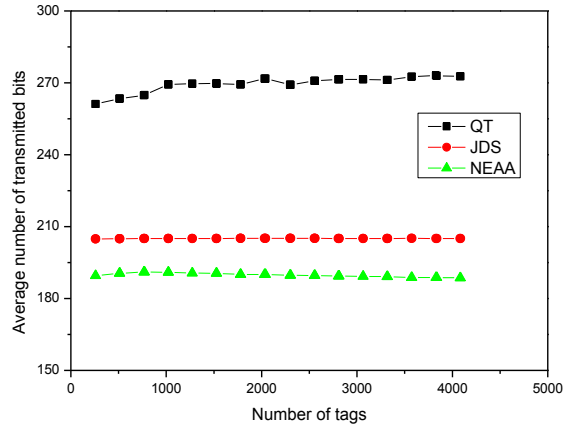**Fig. 8.** Average query time slot contrast diagram

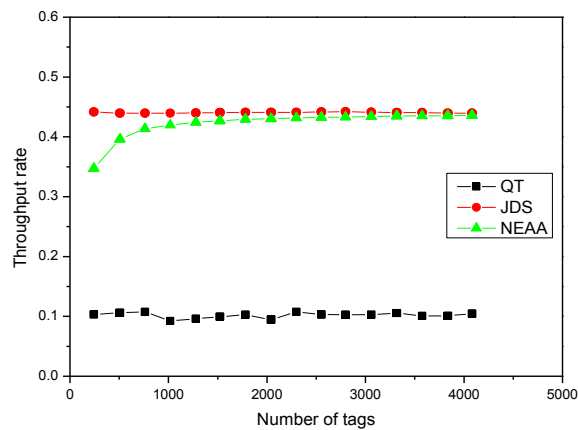**Fig. 9.** Average transmission bit contrast diagram

Fig. 10. Contrast diagram of algorithm throughput

3 Label anti-collision algorithm based on multiple sub-response cycles

3.1 Multiple sub-response cycle mechanism

Since the label ID is unique and belongs to a binary code formed by 0 and 1, the label ID is divided into different finite sets by feature bits. If the collision information is xx10xx01, then in the top bit it must fall in the set prefixed with 00/01/10/11.

Identification cycle refers to the time required for the reader and label to complete a communication, which can be divided into query and response cycles, also into the enquire cycle and multiple response cycle and other sub-cycles according to the

communication time. In doing so, the labels can respond without interference in different sets.

Assume the number of sub-cycles is M, the recognition bit of the label pertains to the value of M. When M is 2, the response period of 0 is S0, and that of 1 is S1. It can also be set as query cycle and multiple response cycles, as shown in Fig. 11.

If M takes 2, 4 and 8 respectively, the recognition bit marches response period, as shown in Table 1.

Different response will be made in different cycles, there is no mutual interference, and sync response can be made in multiple sets of recognition bits to improve the identification rate and efficiency.
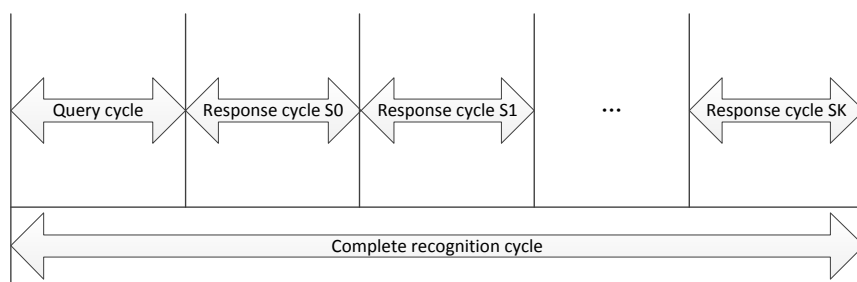
Fig. 11. Periodic diagram of multiple response periodic mechanism

Table 1. M is a matching table for 2, 4, and 8

Response period number	K response	Marked bit
2	1	0
	2	1
4	1	00
	2	01
	3	10
	4	11
8	1	000
	2	001
	3	010
	4	011
	5	100
	6	101
	7	110
	8	111

3.2 Multi-Response collision tree algorithm

Multi-Response Collision Tree algorithm (MRCT) is such an algorithm that transmits with Manchester code based on collision tree algorithm and on the principle of

multi-sub-cycle mechanism. Its basic process includes the reader and label, as shown in Fig. 12.

Take the label group 110011, 100111, 101010, 011011, 011000, and 001001 as an example, when the response sub-period is 2, the identification of the MRCT algorithm is illustrated as shown in Fig. 13. It can be seen that a total of 5 inquiries are performed to complete the identification of all the labels, the specific process is:

1. Send a null string, all labels respond;
2. Send prefix 0 inquiry, there is no collision in the S0 response cycle, and the label 001001 is identified. Collision occurs in S1 response cycle;
3. Continue the enquiry, the labels 011000 and 011011 are identified;
4. Send prefix 1 enquiry, collision occurs in S0 response cycle. There is no collision in the S1 response cycle, and the label 110011 is identified;
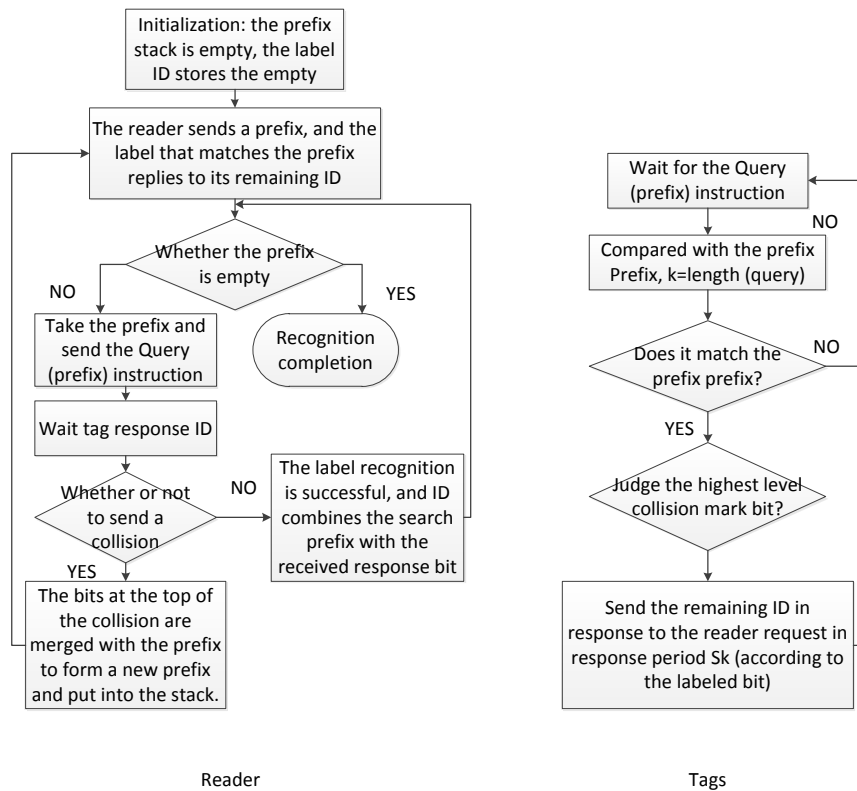5. Continue the enquiry, the labels 100111 and 101010 are identified.

Fig. 12. MRCT algorithm reader and label part algorithm flow
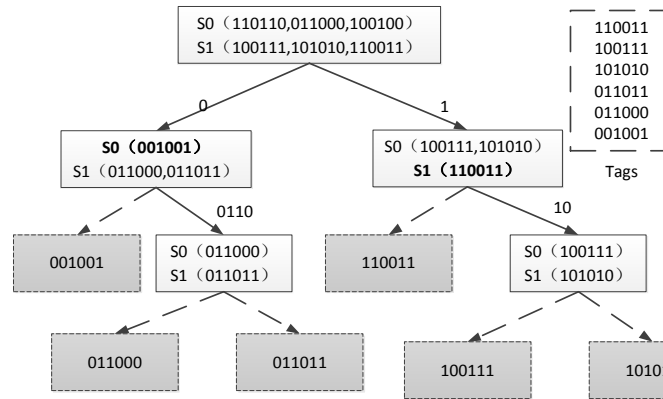
Fig. 13. Label recognition process when M is 2

Take the label group 001001, 011000, 011011, 011001, 011011, 100111, 110001, 110100, 110100, 110111 as an example, when the response subcycle is 4, the identification of the MRCT algorithm is illustrated as shown in Fig.14. It can be seen that a total of 3 inquiries are done to complete the identification of all the labels, the specific process is:

1. For the first inquiry, the label 001001 has no collision response in S0 cycle and is identified; the label 100111 has no collision response in S2 cycle and is identified; S1 and S3 collide;
2. For the second query, the label 011000 has no collision response in the S0 cycle and is identified; the label 011001 has no collision response in the S1 cycle and is identified; the label 011010 has no collision response in the S2 cycle and is identified; the label 011011 has no collision response in the S3 cycle, and is identified
3. For the third query, the label 110001 has no collision response in S0 cycle and is identified; the label 110100 has no collision response in S1 cycle and is identified; the label 110010 has no collision response in S2 cycle and is identified; the label 110111 has no collision response in S3 cycle and is identified;
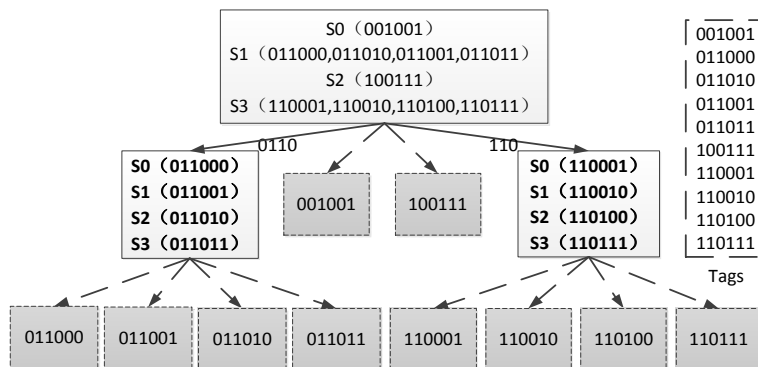

Fig. 14. Label recognition process when M is 4

As above, when the sub-response is more, the more the subgroups of recognition bits, and the more labels to be queried at one time, but the less the number of inquiries.

3.3 MRCT algorithm analysis

The performance indices (recognition time complexity, communication complexity, throughput rate) of MRCT algorithm were analyzed. The number of labels is set as X, the length of the label coding is I, and the search depth is m. The schematic diagram of a tree structure nodes is shown in Fig. 15. The leaf node is set as X_0, the intermediate node set as X2, total number of nodes as N and sub-cycles as M, then the middle node at layer k is m_k.
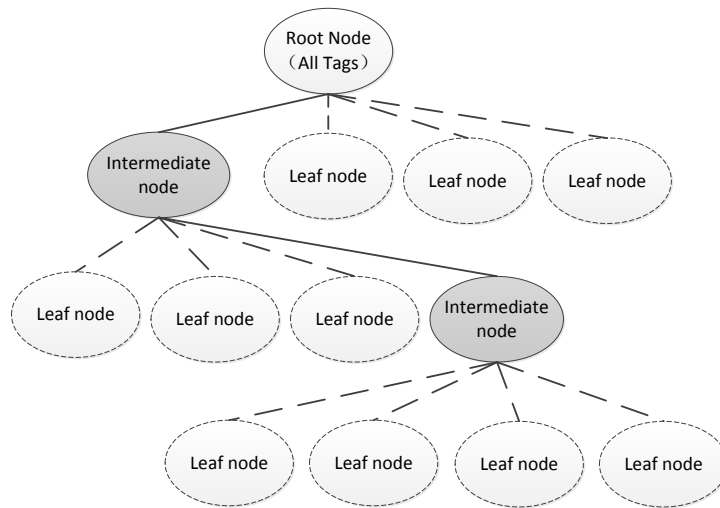

Fig. 15. Tree structure node diagram

The formula of time complexity can be expressed as:

$$T_{tag-r}(X) = T(X) * M = X_m * M = \frac{X_0 - M}{M - 1} * M \tag{1}$$

The formula for communication complexity can be expressed as:

$$\overline{C(X)} = C(X)/X \tag{2}$$

$$= (\sum_{I=1}^{T(X)} (l_{com} + l_{p,i} + l_{r,i}) + (M - 1) * \sum_{I=1}^{T(X)} (l_{r,i}))/X$$

$$= (\sum_{I=1}^{T(X)} (l_{com} + l_{ID}) + (M - 1) * \sum_{I=1}^{T(X)} (l_{r,i}))/X$$

$$= \frac{T(X)}{X}(l_{com} + l_{ID}) + (M - 1) * \frac{T(X)}{X}\overline{l_{r,\iota}}$$

$$= \frac{T(X)}{X}((l_{com} + l_{ID}) + (M - 1) * \overline{l_{r,\iota}})$$

The formula of algorithm throughput rate can be expressed as:

$$\beta = \frac{X}{T(X)} = X / \frac{X-M}{M-1} \tag{3}$$

3.4 MRCT algorithm simulation and performance comparison

The number of selectable labels is 4096, the label length is 96. When M takes 2, 4, and 8, the MRCT algorithm is simulated with the results as follows.

As shown in Fig. 16, the query timeslots for different algorithms are compared. It is obvious that the greater the value M, the less the query timeslots. Because there are many response subcycles, the more the subsets divided in label identification bit, and he number of single queries increases, so that the number of queries decreases. In the case when the labels remain unchanged, QT algorithm is maximum, followed by JDS and NEAA, in the total query timeslots, and MRCT algorithm has the least.

The average query timeslots when the number of labels changes is shown in Fig. 17. M falls within 2 - 8, the average query timeslot decreases in turn from 1 to 0.5, which shows that the MRCT algorithm has obvious advantages in the big data environment. With the labels unchanged, the QT algorithm is maximum, JDS and NEAA algorithms are similar, and MRCT algorithm has the least in the average query timeslots.

The average number of bits transmitted with different number of labels is shown in Fig. 18. It is obvious that when M falls within 2 - 4, the MRCT algorithm sends the least number of bits, followed by NEAA, JDS and QT. However, when M takes 8, the MRCT algorithm sends more bits than JDS and NEAA algorithms. This roots in such a fact that the number of bits sent by single label increases when the response subcycles expand.

The comparison of throughput rates of the algorithms is given in Fig. 19. It is clear that as the M increases, the throughput rate goes up, and always greater than that of other algorithms. In addition to the smaller number of labels, the JDS algorithm is greater than the NEAA in throughput rate. The throughput rates in descending order are in turn MRCT (M = 8), MRCT (M = 4), MRCT (M = 2), NEAA and JDS, QT.
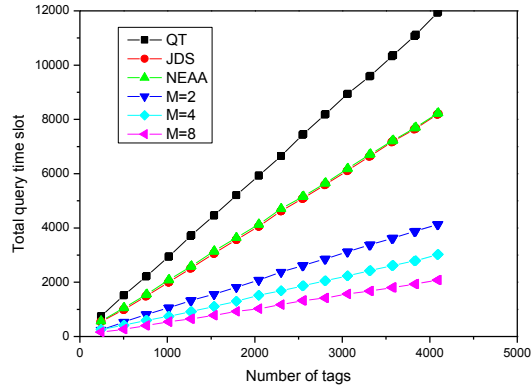
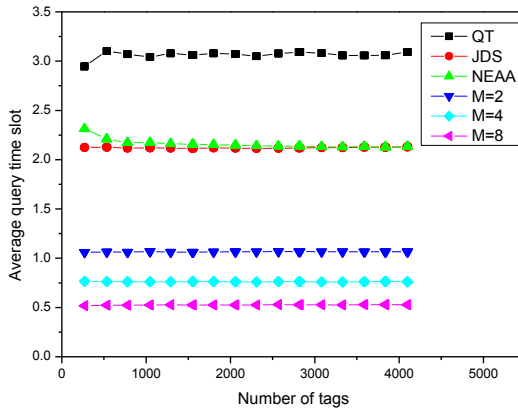**Fig. 16.** Algorithm of algorithm query time slot contrast

Fig. 17. Average query time slot contrast diagram
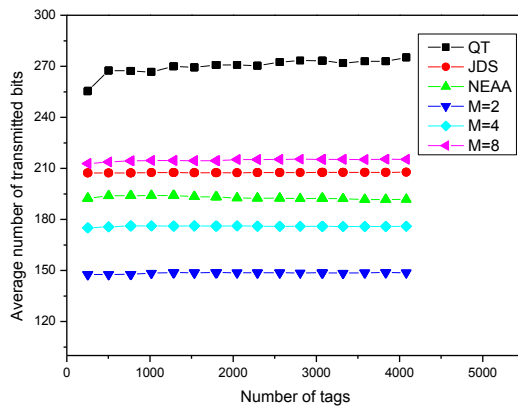

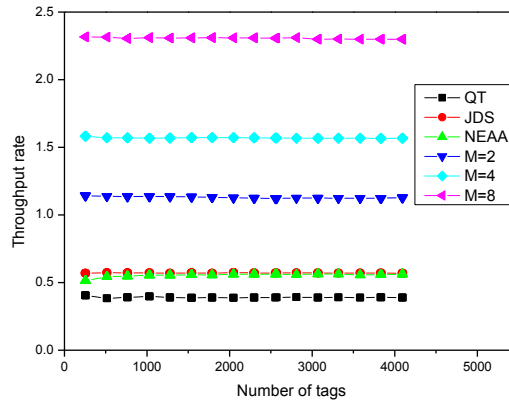**Fig. 18.** A simulation chart of the average number of bits of the label

Fig. 19. Simulation diagram of algorithm throughput

4   Conclusion

Along with the rapid development of Web of Things, the FRID technology is widely applied in the world, but today, the great challenge we face is how to avoid information conflict and collision in the process of acquisition and treatment of massive information, as the keystone of the study. This paper conducts a comparative analysis on different algorithms improved from Tree-based algorithm, and with multiple sub-cycle response mechanism, a MRCT algorithm is proposed here. And beyond that, this paper simulates and analyzed this algorithm, and other improved algorithms.

1. In the era of big data, compared with Aloha algorithm, the tree algorithm can recognize the multiple labels at one time;
2. In different response cycles, no interference occurs for response, and the response is made simultaneously to multiple sets of label bits to improve the recognition rate and efficiency.
3. In the algorithm performance analysis, it is found that the performance of MRCT algorithm is higher than that of other algorithms with less identification cycles, the average query time slots are the least, and the throughput rate is highest.

5   References

[1] Stankovic, J. A. (2014). Research directions for the internet of things. IEEE Internet of Things Journal, 1(1), 3-9. https://doi.org/10.1109/JIOT.2014.2312291
[2] Banerjee, S., Ghosh, A., Mitra, S.K. (2017). A modified mathematical model for lifetime enhancement in wireless sensor network, Mathematical Modelling of Engineering Prolems, 4(2), 84-90. https://doi.org/10.18280/mmep.040204
[3] Wang, T.C., Xie, Y.Z. (2016). BP-GA data fusion algorithm studies oriented to smart home, Mathematical Modelling of Engineering Problems, 3(3), 135-140. https://doi.org/10.18280/mmep.030304

[4] Tsai-Tsu, S. U., Walker, R. M., Xue, L. (2013). Reform and transition in public administration theory and practice in greater china. Public Administration, 91(2), 253-260. https://doi.org/10.1111/padm.12030

[5] Li, M. (2013). Fast tag searching protocol for large-scale rfid systems. Networking IEEE/ACM Transactions on, 21(3), 924-934. https://doi.org/10.1109/TNET.2012.2212454

[6] Wei, R. G., Zhen, J. G., Bao, L. L. (2015). Study on mining big users data in the development of Hubei auto-parts enterprise, Mathematical Modelling of Engineering Problems, 2(4), 1-6. https://doi.org/10.18280/mmep.020401

[7] Tang, R.X., Zhang Z.J., Liu Y.L., Wang Z.P. (2017). Spatial information is exploited to adjust hand movement when hitting moving target, NeuroQuantology, 15(3), 7-11.

[8] Want, R. (2006). An introduction to RFID technology. IEEE Educational Activities Department, 5(1), 25-33. https://doi.org/10.1109/MPRV.2006.2

[9] Chen, Y. H., Horng, S. J., Run, R. S., Lai, J. L., Chen, R. J., Chen, W. C. (2010). A novel anti-collision algorithm in rfid systems for identifying passive tags. IEEE Transactions on Industrial Informatics,6(1), 105-121. https://doi.org/10.1109/TII.2009.2033050

[10] Dirk, K.F,M, Hans, J.H.G. (2017). Consciousness in the universe is scale invariant and implies an event horizon of the human brain, NeuroQuantology, 15(3), 41-79.

[11] Su, J., Wen, G., Hong, D. (2016). A new rfid anti-collision algorithm based on the q-ary search scheme. Chinese Journal of Electronics,24(4), 679-683. https://doi.org/10.1049/cje.2015.10.003

[12] Wu, H., Zeng, Y., Feng, J., Gu, Y. (2012). Binary tree slotted aloha for passive rfid tag anticollision. IEEE Transactions on Parallel & Distributed Systems, 24(1), 19-31. https://doi.org/10.1109/TPDS.2012.120

[13] Liu, C., Wang, C., Kui, F. U. (2013). New group-based processing tag anti-collision algorithm for rfid system. Journal of Computer Applications,33(2), 592-591. https://doi.org/10.3724/SP.J.1087.2013.00592

[14] Forster, R. (2000). Manchester encoding: opposing definitions resolved. Engineering Science & Education Journal, 9(6), 278-280. https://doi.org/10.1049/esej:20000609

[15] Wang, Y. Q., Jiang, G. P. (2010). Anti-collision algorithm based on grouping mechanism and jumping dynamic binary. Acta Automatica Sinica, 36(10), 1390-1400. https://doi.org/10.3724/SP.J.1004.2010.01390

[16] Pan, W., Xie, Q., Li, S., Yang, G., Feng, N. (2012). Anti-collision algorithm based on dynamic grouping mechanism in rfid systems. Journal of Convergence Information Technology, 7(15), 338-346. https://doi.org/10.4156/jcit.vol7.issue15.40

[17] Jia, X., Feng, Q., Ma, C. (2010). An efficient anti-collision protocol for rfid tag identification. IEEE Communications Letters, 14(11), 1014-1016. https://doi.org/10.1109/LCOMM.2010.091710.100793

6   Author

Chen Buying is a teacher from Information Engineering Department of Xingtai Polytechnic College in Xingtai, China. She was engaged as an associate professor in 2009 and her research area is computer science and technology.