

A Physical Layer Algorithm for Estimation of Number of Tags in UHF RFID Anti-Collision Design

Zhong Huang¹, Jian Su², Guangjun Wen¹, Wenxian Zheng³, Chu Chu¹, Yijun Zhang^{4,*} and Yibo Zhang⁵

Abstract: A priori knowledge of the number of tags is crucial for anti-collision protocols in slotted UHF RFID systems. The number of tags is used to decide optimal frame length in dynamic frame slotted ALOHA (DFSA) and to adjust access probability in random access protocols. Conventional researches estimate the number of tags in MAC layer based on statistics of empty slots, collided slots and successful slots. Usually, a collision detection algorithm is employed to determine types of time slots. Only three types are distinguished because of lack of ability to detect the number of tags in single time slot. In this paper, a physical layer algorithm is proposed to detect the number of tags in a collided slot. Mean shift algorithm is utilized, and some properties of backscatter signals are investigated. Simulation results verify the effectiveness of the proposed solution in terms of low estimation error with a high SNR range, outperforming the existing MAC layer approaches.

Keywords: UHF RFID, anti-collision, cluster algorithm.

1 Introduction

Anti-collision algorithms are carried out in multi-access UHF RFID systems to reduce collisions as well as to increase channel efficiency. As one of the most popular anti-collision algorithms for RFID system, dynamic framed slotted ALOHA (DFSA) algorithm employs a mechanism similar with time division multiple access (TDMA). Synchronized frame is divided into several time slots for random access. Reader dynamically adjusts the number of time slots based on the estimated number of tags. It is well known that DFSA reaches the optimal system throughput when the number of slots equals to the number of tags waiting to be identified. There are two main problems in DFSA design, estimation of the number of tags and frame length adjustment [Chen, Liu,

¹ School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China.

² Nanjing University of Information Science & Technology, Nanjing, 210044, China.

³ Shenzhen Intellifusion Technologies Co., Shenzhen, 518000, China.

⁴ School of Foreign Languages, University of Electronic Science and Technology of China, Chengdu, 611731, China.

⁵ Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada.

* Corresponding Author: Yijun Zhang. Email: zhangyj@uestc.edu.cn.

Ma et al. (2018)].

Various researches are proposed to improve the estimation accuracy so as to improve the channel efficiency of RFID system. The number of tags is estimated by multiplying the number of collided slots in a frame by the expected number of tags per collided slot ($=2.39$), which is a constant for all frames regardless of number of tags and frame size [Schoute (1983)]. In the mechanisms proposed by Vogt et al. [Vogt (2002); Chen and Lin (2006); Cha and Kim (2005)], estimation of the number of tags is based on successful, collided, and empty probability in a frame. Probability theory is utilized in these works. The number of tags is estimated in Chen [Chen (2009)] by multiplying the number of collided time slots by a well-defined factor, which is found by an iterative algorithm. A posteriori probability distribution-based method is proposed in Eom et al. [Eom and Lee (2010)] to further improve the accuracy. However, the computation complexity is higher than the others. A Bayesian method [Annur, Srichavengsup, Nakpeerayuth et al. (2015)] is used to update the posterior probability distribution of number of users' slot by slot so as to estimate the number of tags.

Since RFID readers usually adopt in-phase and quadrature information of tag signals, tag recovery and estimation methods of the number of tags based on signal processing are developed to solve collision problem. Tag recovery is able to turn a collided slot into a successful slot while estimation of the number of tags based on physical layer process is able to enhance performance of anti-collision algorithm in upper layer. Most researches require multiple antennas receiver [Angerer, Langwieser and Rupp (2010)], specific tag signal strength [Fyhn, Jacobsen, Popovski et al. (2011)], modified coding mechanisms [Parks, Liu, Gollakota et al. (2014)], etc. Bipartite Grouping (BiGroup) [Ou, Li and Zheng (2017)] is the first one proposed to parallelly decode multiple CTOS tags with the help of both time domain and constellation domain information in physical layer. Algorithms based on signal processing in time domain are proposed to reduce SNR requirements. Collided signals are transformed to time-scale domains and LS criterion is utilized for tag signal separation in Zeng et al. [Zeng, Wu, Yang et al. (2017)]. An edge transition scheme is proposed to recover collision and decode tag signals in Benbaghdad et al. [Benbaghdad, Fergani and Tedjini (2016)]. These works focus on tag signal recovery issues in physical layer and pay less attention on MAC layer design. Tan et al. [Tan, Wang, Fu et al. (2018)] proposed a collision detection and signal recovery method and combine them with DFSA algorithm. Optimal frame length is calculated based on a collision recovery probability coefficient, which is obtained by simulations of its physical layer design. A novel closed-form solution is further proposed in Ahmed et al. [Ahmed, Salah, Robert et al. (2018)] for optimal FSA frame length decision, in which collision recovery probabilities are provided arbitrarily. An accurate tag number estimation algorithm is still needed and not addressed.

In this paper, we propose to estimate the number of tags in a collided time slot in physical layer. Different from the existing work, we focus on improving estimation accuracy under CTOS tag assumption. Estimation error is reduced compared with MAC layer design and SNR range is expanded compared with current physical layer design. In this scheme, the number of tags in single time slot is determined by the number of clusters located in constellation domain. Signal samples are firstly scaled based on the baseband

noise level. After that, mean shift algorithm is utilized to divide the data into several clusters. In the end, a cluster adjustment process is carried out to get better performance. Simulation results verify the effectiveness of the proposed solution in terms of low estimation errors and high SNR range, outperforming the existing MAC layer-based approaches. This paper is organized as following. Section II describes signal model of UHF RFID system and drawback of clustering algorithms. After that, design consideration and cluster-based algorithm is proposed in Section III, which is evaluated with numeric simulations in Section IV. Finally, conclusion follows.

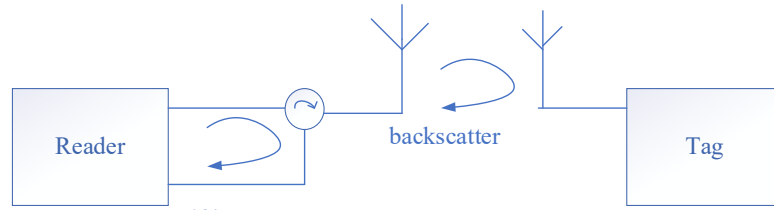
2 Signal model and cluster algorithm

2.1 Backscatter signal and its distribution

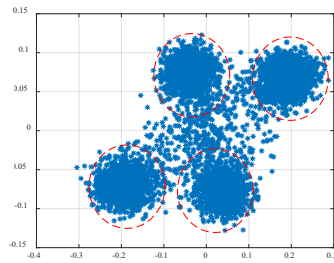
In a UHF RFID system, reader energizes tags by transmitting continuous wave. Passive tags backscatter the radio to communicate with the reader. After that, backscatter signals are down converted in the reader side. As shown in Fig. 1(a), baseband signal in receive path of a reader consists of three components, backscatter signal of tags, self-jammer and noise. Eq. (1) shows the detailed format of those signals.

$$s(t) = \sum_{i=1}^n A_i e^{-j\theta_i - j\varphi(t)} + \mu e^{-j\theta_0 - j\varphi(t)} + n(t) \tag{1}$$

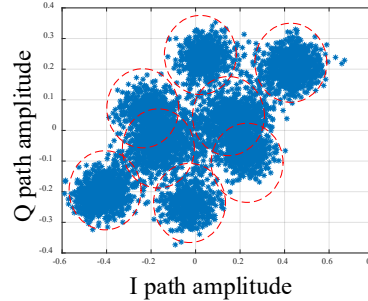
The first one indicates signals of multiple tags, the second one indicates the signals of self-jammer caused by RFID system. Where i indicates the index of collided tags in the same time slot, A_i and θ_i are respective transmitted data and initial phase of the i -th tag signal. θ_0 and μ is the amplitude and phase of self-jammer. $n(t)$ represents white noise in the receiver.



(a) Example of RFID system



(b) signal samples of two tags



(c) signal samples in low SNR scenario

Figure 1: RFID framework and backscatter signal model

Backscatter signals have both in-phase and quadrature components, with which we could plot the signal samples in a two-dimension coordinates. Every sample is represented by a point on the I-Q plane. Samples of the same transmitting status are dispersed and scattered around a centroid position, forming a cluster. As shown in Fig. 1(b), there are 4 clusters, representing 4 transmitting status of 2 tags. The number of clusters is decided by the number of tags. Every tag has two transmission status, the size of the full status space is 2^n , where n is the number of tags. Three properties could be obtained from Eq. (1).

1. Since noise in in-phase axis and in quadrature axis are different due to the existence of self-jammer, every cluster is shaped as an eclipse.
2. We assume that noise in both axes obeys gaussian distribution. As a result, around 95% samples of one cluster are located in a circle of radius of two times standard deviation of noise. The higher the noise is, the bigger the circle is. When SNR goes down, the radius gets bigger, and when it is larger than the distance between cluster centers, the clusters overlap. Fig. 1(c) shows an example when clusters are overlapped.
3. The clusters are always located in pairs, symmetrizing to the center of all samples. For every cluster, reverse all the tag status, a symmetric cluster is obtained. Their symmetric center is located at the center of the whole graph. Eq. (2) shows the coordinates of the center point, where n_{sam} , x_{center} and y_{center} denotes the number of samples, x axis coordinate and y axis coordinate.

$$x_{\text{center}} = \frac{\sum_{t=t_1}^{t_N} \text{Rel}\{s(t)\}}{n_{\text{sam}}} \quad y_{\text{center}} = \frac{\sum_{t=t_1}^{t_N} \text{Imag}\{s(t)\}}{n_{\text{sam}}} \quad (2)$$

2.2 Density based clustering algorithm

A straight forward way to find the number of tags is to divide signal samples into clusters with a cluster algorithm. Mean shift algorithm and DBSCAN algorithm are both widely used method to identify multiple dimension data without a prior knowledge of the number of clusters. They both make use of sample density to make decisions. Mean shift algorithm updates every cluster center based on the vector sum of all samples in it until convergence. Bandwidth is setup to decide the range of clusters. On the other hand, DBSCAN choose samples with large number of neighbors as core samples. Connection between core samples are calculated and clusters are divided. Parameters distance is required to decide neighbors and minimum points are required to decide core samples.

Table 1: Parameter settings of cluster algorithms

Algorithm	Parameter	value
Mean Shift	Bandwidth	2 times noise standard error
	SNR	30 dB
DBSCAN	Neighborhood distance	2 times noise standard error
	Minimum points	Number of samples in one byte
	SNR	20 dB

Both algorithms are tested with simulations. Parameter settings are shown in Tab. 1. Fig. (2) shows the cluster division of two algorithms. Samples in one cluster are encircled by a circle. DBSCAN divides the signals into 6 clusters while mean shift divides them into more than 20 clusters.

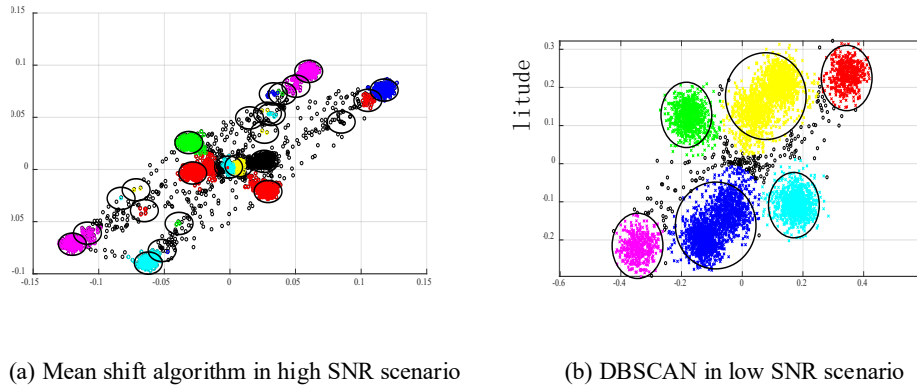


Figure 2: Performance of two algorithm in different SNR scenarios

3 Physical layer algorithm for estimation of the number of tags in single time slot

3.1 Design consideration

Mean shift algorithm and DBSCAN are available for clustering in some cases with proper parameter settings. However, both algorithms are not designed for detection of the number of tags. DBSCAN requires uniform distributed samples in one cluster, which is not preferred in our case. Density of samples in one cluster decreases along with distance to the center. Furthermore, it is not able to identify overlapped clusters. Mean shift algorithm outperforms DBSCAN in low SNR scenario. Cluster centers are updated towards a denser direction, it always finds the densest points. However, isolated samples are identified as a cluster in some cases. In this paper, we make use of samples distribution information and its properties to improve performance of mean shift algorithm. The improvement is based on the following considerations.

First of all, our purpose is to find number of clusters, not accurate cluster division. Wrong assignment to clusters may cause bit error in signal recovery case, but not in number detection case. As a result, we only consider a small core area of clusters, which is defined as samples within distance of two times standard deviation of noise. Find it and we get a valid cluster center. The noise samples are ignored naturally.

Second, as described in Section II, all clusters are shaped like an eclipse. On the meantime, mean shift algorithm calculates updated vector based on a Euclid distance, which means samples in a perfect circle are all considered. It brings a big performance decrease. It is fitting and proper to adjust the scale of in-phase and quadrature signal magnate by noise level. After that, the core area of clusters is formed as a perfect circle. It is better for identification.

Finally, clusters are distributed in pairs. Every cluster have a symmetrical one with similar number of samples. Their cluster centers are also symmetrical to the whole center of samples. This property makes us able to discard isolated noise cluster identified by mean shift algorithm.

3.2 Algorithm details

The detailed algorithm is shown as Algorithm.1, we first transform samples to make clusters shaped as perfect circle other than eclipse. After that, mean shift algorithm is carried out. Some random points are selected as initial centers. These centers are updated based on samples in their neighborhood. After they converge, an adjustment scheme is carried out to discard isolated noise clusters.

Algorithm 1: Physical layer algorithm for estimation of the number of tags

1. Scale the coordinates with noise level.
2. Initial cluster centers are randomly selected.
3. Calculate sum vector and update the cluster centers until they converge. Eq. (3) shows calculation of sum vectors.

$$M_k = \sum_{x_k \in S_k} (x - x_k) \quad (3)$$

Where M_k denotes the sum vector in kth round, S_k denotes the collection of samples in the cluster center's neighborhood.

4. For every cluster, check if there is another cluster center symmetrical to the whole center, if there is not, discard it. If it is centered at the whole center, divide it into two clusters.
5. Decide number of tags

$$n = \langle \log_2 m \rangle \quad (4)$$

where m denotes the number of clusters and operator $\langle \alpha \rangle$ indicates the smallest integer larger than α .

4 Simulation and performance evaluation

In this section, we evaluate the performance of proposed algorithm under different scenarios. Success rate is firstly proposed for accuracy of detecting number of tags to evaluate performance in single time slot scenario. After that, total success rate and estimation error are derived based on probability to evaluate performance in multiple time slots scenario. Success rate is compared with DBSCAN algorithm and estimation error is compared with probability-based methods.

4.1 Performance in single time slot scenario

The performance of proposed algorithm is evaluated by accuracy or success rate, which is defined as the number of successful experiments over the total number of experiments. In order to improve reliability of simulations, simulations are carried out multiple times in different SNR scenarios. Furthermore, scenarios with different number of tags are evaluated separately due different performance in these cases. Scenarios when the number of tags is larger than 4 is not considered here because it rarely happens in practice.

The simulation runs as the following steps.

In the first step, we initialize the system parameters, i.e., number of tags and SNR.

In the second step, 10000 experiments are executed. In each experiment, pseudo signals are generated based on Eq. (1). Signal strength and initial phase of tags are randomly selected in every experiment. Proposed algorithm and DBSCAN are used to determine the number of tags. Both actual number of tags and determined number of tags are recorded for performance evolution.

In the third step, switch to the next parameter and execute Step 2 for another time.

In the fourth step, success rate in each system parameter set are calculated. Before that, a performance indicator $p(\hat{n}|n)$ based on conditional probability is calculated by real number and determined number, as shown in Eq. (5). Where \hat{n} denotes detected number of tags while n denotes actual number of tags. It is apparent that success rate when the number of tags is n equals probability of $p(n|n)$.

$$p(\hat{n}|n) = \frac{\text{number of experiments that tag number determined as } \hat{n}}{\text{number of experiments that actual tag number as } n} \quad (5)$$

Fig. 3 shows success rate of proposed algorithm in different conditions of SNR and the number of tags. Success rate of proposed algorithm is greater than DBSCAN in all conditions. When SNR is larger than 18 dB, success rate of proposed algorithm is larger than 0.9 in 2 and 3 tags scenarios. Success rate is relatively lower when there are 4 tags, still over 0.8 when SNR is large enough.

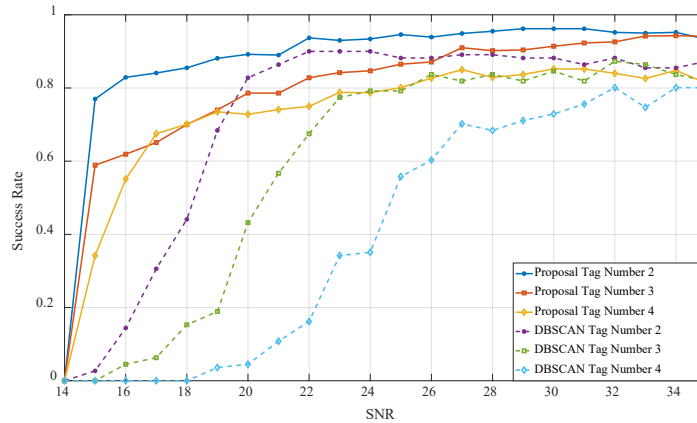


Figure 3: Success rate comparison in different conditions

4.2 Performance of multiple time slots

Performance of proposed algorithm in multiple time slots is evaluated by two indicators, total success rate and estimation error. When there are multiple time slots and unknown number of unidentified tags, number of tags in one time slot follows a binomial distribution, as shown in Eq. (6).

$$B(r) = \binom{n}{r} \left(\frac{1}{L}\right)^r \left(1 - \frac{1}{L}\right)^{n-r} \tag{6}$$

where $B(r)$ denotes the probability of r tags in one slot, n denotes number of tags to be identified in the read range, L denotes frame length, i.e., number of time slots.

Total success rate takes distribution of number of tags in one slot into consideration, which is defined as Eq. (7). Total success rate shows an average performance under specific condition of frame length and the number of tags.

$$p_s = \sum_{r=2}^{\infty} p(r|r)B(r) \approx \sum_{r=2}^4 p(r|r)B(r) \tag{7}$$

Similar with other estimation researches, estimation error is a good indicator for performance evaluation. Here it is defined in a probabilistic way in Eq. (8).

$$error = \frac{|E(n) - E(\hat{n})|}{E(n)} \tag{8}$$

where $E(n)$ denotes average number of tags in one time slot, while $E(\hat{n})$ denotes average estimated number of tags in every time slot. It is calculated by Eq. (9).

$$E(\hat{n}) = \sum_{n=0}^{\infty} p(n) * E(\hat{n} | n) \tag{9}$$

where $E(\hat{n} | n)$ denotes expectation of estimated number of tags on condition of the number of tags in one time slot, shown as Eq. (10).

$$E(\hat{n} | n) = \sum_{i=2}^4 p(i | n) * i \tag{10}$$

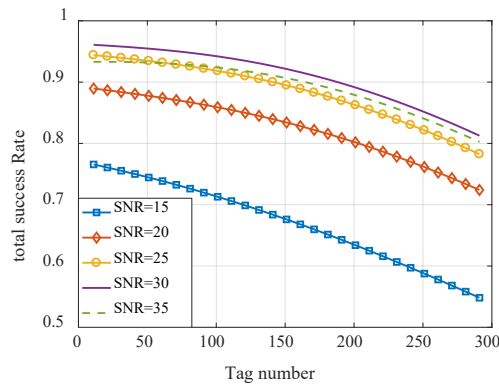


Figure 4: Total success rate in different conditions of the number of tags

Fig. 4 shows total success rate in different conditions of the number of tags when number of time slots are set to 128. Success Rate decreases along with the number of tags increases because possibility is larger when the number of tags is higher. Apparently, SNR effects total success rate. Total success rate decreases to 0.5 when SNR is 15 dB and the number of tags reaches 300. However, total success rate is higher than 0.8 in most high SNR conditions (higher than 20).

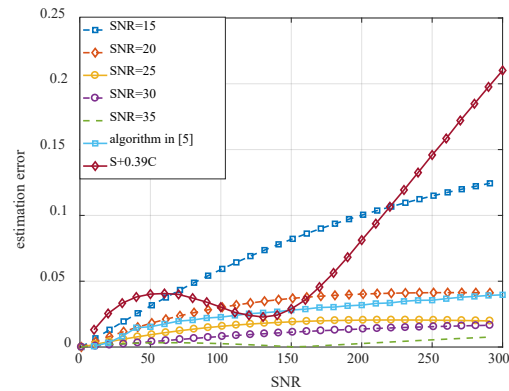


Figure 5: Estimation error under different number of tags and SNR conditions

Fig. 5 shows estimation error comparison between proposed algorithm and two other methods. Estimation error of our proposal decreases while SNR goes up. In 15 dB SNR condition, our proposal beats S+2.39C method when the number of tags is more than 220. In 20dB SNR condition, estimation error of our proposal is close to method proposed in Chen. (2009). Estimation error gets lower when SNR is higher than 20 and outperforms both two methods.

5 Conclusion

This paper focuses on tag number estimation method for RFID anti-collision purpose. A clustering algorithm is proposed to detect the number of tags in physical layer. The proposed algorithm makes full use of in-phase and quadrature information to get better performance in low SNR scenarios. Simulation results show that it is good in a large range of SNR conditions. It is better than DBSCAN in terms of success rate and better than probability-based methods in terms of estimation errors.

Acknowledgement: This work was supported in part by the National Natural Science Foundation of China under project contracts [NOS. 61601093, 61791082, 61701116, 61371047], in part by Sichuan Provincial Science and Technology Planning Program of China under project contracts No. 2016GZ0061 and No. 2018HH0044, in part by Guangdong Provincial Science and Technology Planning Program of China under project contracts No. 2015B090909004 and No. 2016A010101036, in part by the fundamental research funds for the Central Universities under project contract No. ZYGX2016Z011, and in part by Science and Technology on Electronic Information Control Laboratory.

References

- Ahmed, H. A.; Salah, H.; Robert, J.; Heuberger, A.** (2018): A closed-form solution for ALOHA frame length optimizing multiple collision recovery coefficients' reading efficiency. *IEEE Systems Journal*, vol. 12, no. 1, pp. 1047-1050.
- Angerer, C.; Langwieser, R.; Rupp, M.** (2010): RFID reader receivers for physical

layer collision recovery. *IEEE Transactions on Communications*, vol. 58, no. 12, pp. 3526-3537.

Annur, R.; Srichavengsup, W.; Nakpeerayuth, S.; Wuttitsittikulkij, L. (2015): Bayesian method of slotted aloha based anti-collision protocol for rfid systems. *IEEE Twelfth International Symposium on Autonomous Decentralized Systems*, pp. 87-90.

Benbaghdad, M.; Fergani, B.; Tedjini, S. (2016): Toward a new PHY layer scheme for decoding tags collision signal in UHF RFID system. *IEEE Communications Letters*, vol. 20, no. 11, pp. 2233-2236.

Cha, J. R.; Kim, J. H. (2005): Novel anti-collision algorithms for fast object identification in RFID system. *Proceedings of 11th International Conference on Parallel and Distributed Systems*, pp. 63-67.

Chen, H.; Liu, K.; Ma, C.; Han, Y.; Su, J. (2018): A novel time-aware frame adjustment strategy for RFID anti-collision. *Computers, Materials & Continua*, vol. 57, no. 2, pp. 195-204.

Chen, W. (2009): An accurate tag estimate method for improving the performance of an RFID anticollision algorithm based on dynamic frame length aloha. *IEEE Transactions on Automation Science and Engineering*, vol. 6, no. 1, pp. 9-15.

Chen, W. T.; Lin, G. H. (2006): An efficient anti-collision method for tag identification in a rfid system. *IEICE Transactions on Communications*, vol. E89-B, no. 12, pp. 3389-3392.

Eom, J.; Lee, T. (2010): Accurate tag estimation for dynamic framed-slotted aloha in rfid systems. *IEEE Communications Letters*, vol. 14, no. 1, pp. 60-62.

Fyhn, K.; Jacobsen, R. M.; Popovski, P.; Scaglione, A.; Larsen, T. (2011): Multipacket reception of passive UHF RFID tags: a communication theoretic approach. *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4225-4237.

Ou, J.; Li, M.; Zheng, Y. (2017): Come and be served: parallel decoding for cots rfid tags. *IEEE/ACM Transactions on Networking*, vol. 25, no. 3, pp. 1569-1581.

Parks, A. N.; Liu, A.; Gollakota, S.; Smith, J. R. (2014): Turbocharging ambient backscatter communication. *Proceedings of ACM SIGCOMM*, pp. 619-630.

Schoute, F. C. (1983): Dynamic frame length aloha. *IEICE Transactions on Communications*, vol. 31, no. 4, pp. 565-568.

Tan, X.; Wang, H.; Fu, L.; Wang, J.; Min, H.; Min, H.; Engels, D. W. (2018): Collision detection and signal recovery for uhf rfid systems. *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 98-113.

Vogt, H. (2002): Efficient object identification with passive rfid tags. *Proceedings of International Conference on Pervasive Computing*, pp. 98-113.

Zeng, Y.; Wu, H.; Yang, K.; Wang, X. (2017): Dynamic-scale transformation for UHF RFID tag collision separation on physical layer. *4th International Conference on Information, Cybernetics and Computational Social Systems*, pp. 706-711.