Optimal distance-based clustering for tag anti-collision in RFID systems

Waleed Alsalih, Kashif Ali, and Hossam Hassanein
School of Computing
Queen's University
Kingston, Ontario, Canada, K7L 3N6
{waleed, kashif, hossam}@cs.queensu.ca

***Abstract*—Tag collisions can impose a major delay in Radio Frequency IDentification (RFID) systems. Such collisions are hard to overcome with passive tags due to their limited capabilities. In this paper, we look into the problem of minimizing the time required to read a set of passive tags. We propose a novel approach, the distance-based clustering, in which the interrogation zone of an RFID reader is divided into equal sized clusters (discs), and tags of different clusters are read separately. The novel contributions of this paper are the following. First, we provide a mathematical analysis to the problem and derive a closed-form formula relating delay to the number of tags and clusters. Second, we devise a method to efficiently find the optimal number of clusters. The proposed scheme can be augmented with any tree-based anti-collision scheme, and substantially improve its performance. Simulation results show that our approach makes significant improvements in reducing collisions and delay.**

I. INTRODUCTION

Radio Frequency IDentification (RFID) is an emerging identification technology which has recently received significant attention. RFID has several advantages over traditional identification technologies (e.g., barcodes): RFID does not require line-of-sight for communication and RFID tags can be read at longer distances. RFID is relatively fast and many tags can be read simultaneously. RFID technology has the potential to turn everyday objects into a mobile network of nodes, which can then be used to track objects, trigger events, and take actions [22].

A typical RFID system consists of RFID tags and an electromagnetic reader [8]. The reader uses radio waves to send requests to tags within its interrogation zone, and tags reply with their IDs. The tags can be either active or passive. A passive tag does not have a power supply and obtains its operating and communication power from the electromagnetic signal it receives from the reader. A collision occurs when more than one tag send their signal to the reader at the same time. Tag collisions bring about an extra delay and energy consumption to the interrogation process [15]. While reader collisions are also possible, they can be dealt with using conventional medium access control techniques; in fact, a reader is a more powerful device and can detect collisions and communicate with other interfering readers. A passive tag can only transmit data by reflecting the reader transmitted electromagnetic waves, and hence can not detect nor communicate with the neighboring tags. The energy received by a tag is usually less than 100μW; accordingly, conventional medium access protocols are not practical for passive tags' anti-collision [18], [25], [26].

While several schemes have been proposed to deal with tags' collisions in RFID systems (see Section II for details), a significant amount of collisions can be avoided by partitioning the interrogation zone spatially into smaller clusters, and interrogating tags in different clusters separately (i.e., one cluster at a time). Any existing anti-collision scheme can be used to resolve collisions in a single cluster. That should reduce the number of collisions as it reduces the number of tags that may respond at the same time. To the best of our knowledge, our work in [2] was the first to follow such an avoidance approach. However, while the viability and efficiency of this approach was shown in [2], it was not clear how to find the best partitioning. Indeed, many clusters may result in many empty clusters, which is an extra overhead, and few clusters may result in having crowded clusters; both situations affect the performance of that approach significantly. In this paper, we present a mathematical delay analysis to the problem and provide a closed-form formula for the delay as a function of the number of tags and the number of clusters. Moreover, we give an efficient method to find the optimal number of clusters. This enables the clustering approach to choose the optimal number of clusters and to adapt to different environments with different tag densities. The ns-2 simulator [1] is extended to simulate the RFID system. Our simulation results show that our approach can provide significant improvements in terms of the number of collisions, reading rates, and delay as compared with other approaches.

The rest of the paper is organized as follows. Section II surveys existing anti-collision protocols. Section III briefly explains the distance-based clustering approach. In Section IV, we present our optimal distance-based clustering approach. Section V shows our simulation settings and results. Finally, section VI concludes our work.

II. LITERATURE SURVEY

Anti-collision algorithms are generally divided into probabilistic and deterministic algorithms [12]. Probabilistic algorithms are based on the framed ALOHA scheme where the reader sends the frame length and each tag picks out a slot for its data transmission. In the framed slotted ALOHA [23],

978-1-4244-2413-9/08/$25.00 ©2008 IEEE

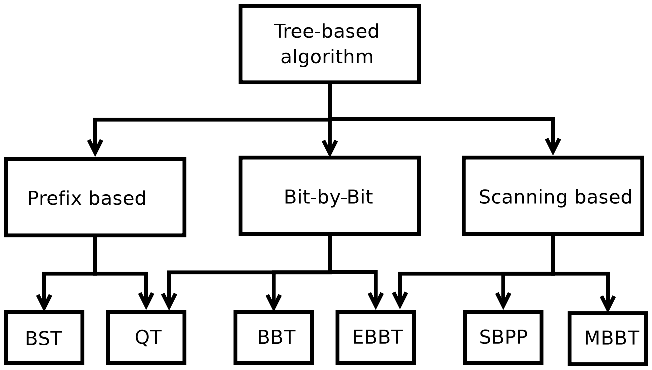

Fig. 1. Taxonomy of tree-based anti-collision protocols [2]

a passive tag randomly selects a slot number in the frame and uses the selected slot to respond to the reader. The probability of a collision is proportional to the number of tags in the interrogation zone. To alleviate this problem, an enhanced dynamic framed slotted ALOHA algorithm was proposed in [15]. In this algorithm, the frame size is dynamically adjusted according to a tag count estimation to keep the collision probability below a certain threshold.

Deterministic algorithms are typically based on the slotted ALOHA scheme where the reader identifies a set of tags that are allowed to transmit in a given slot. Tree-based algorithms [3], [4], [6]–[8], [10], [11], [13], [14], [16], [17], [21], [24]–[26] fall under this category. These algorithms require a precise tag timing synchronization to determine the position of the collision bits. Tree-based anti-collision schemes are categorized in Fig. 1 (adapted from our work in [2]).

The Binary Search Tree (BST) is a prefix-based scheme that was proposed in [8]. The BST relies on the ability to determine the position of the colliding bits for all tags' IDs. An RFID reader starts with broadcasting an inquiring request in which all bits are set to 1. In case of a collision, another inquiring request is created and the most significant colliding bit is set to 0 in that request. Thereby, only those tags with an ID value less than or equal to that of the request will respond. This process repeats until a single tag is identified. Once a tag is being read successfully, it is put in a sleep mode. During the sleep mode, the tag will not respond to any further requests until it is reset by the reader. This process repeats until all the tags are identified. In the BST scheme, inquiring requests and tags' responses involve complete IDs. To overcome this, the dynamic binary search algorithm, which is proposed in [8], gets the tags to send only the least significant bit (LSB) starting from the last colliding bit. This reduces the number of bits transferred between the tags and the reader; however, the number of identification cycles is equivalent to that of the BST scheme. An Enhanced BST (EBST) with backtracking is proposed in [19]. In the EBST scheme, the reader sends the location of the most significant colliding bit rather than sending a complete ID to compare with. When a tag receives such a location, it responds only if it has the bit corresponding to that location set to 0. When a tag is identified successfully by the reader, the reader backtracks to previous unsuccessful requests. The EBST scheme reduces the amount of data sent by the reader and lowers the tag reading delay, as the singulation process no-longer starts from the root of the tree as is the case with the conventional BST algorithm.

The bit-by-bit binary tree (BBT) scheme uses a bit-wise arbitration scheme [5], [9], [10]. In the BBT scheme, a tag maintains a pointer to the last sent bit . Initially this pointer points to the most significant bit (MSB) of a tag ID. The reader broadcasts an inquiring bit, either 0 or 1. A tag whose pointed bit matches with the inquiring bit sends the next least significant bit (LSB) and updates the pointer. A non-matching tag will go into the sleep mode. On the reader's side, if the reader receives a bit without collision, it will use this bit as the next inquiring bit. In case of a collision, the reader will use 0 as the next inquiring bit. A tag is identified once the pointer reaches the LSB. The reader resets the sleeping tags only after a single tag is properly identified and read.

The Modified Bit-by-bit Binary Tree (MBBT) scheme, which is proposed in [6], requests the bits in sequence, starting from the LSB. In case of a collision at the k^{th} bit, the reader inactivates all the tags whose k^{th} bit is 1. In the MBBT scheme, if tags' IDs are sequential, the identification time is reduced, otherwise, it shows the same performance as that of the BBT. The Enhanced Bit-by-bit Binary Tree (EBBT) algorithm [6] is proposed to overcome this. In the EBBT scheme, the reader initially requests tags to send their IDs. The reader keeps track of the position of the collision bits while saving the non-collided bits. The reader, using the bit-by-bit approach, sequentially requests the bits at the collision positions only. The EBBT scheme is energy efficient as it reduces the data transmission between the tags and the reader. The EBBT scheme falls under the bit-by-bit and scanning-based categories (Fig. 1), as it uses both approaches.

The adaptive memoryless protocol was proposed in [17] to effectively handles mobility (i.e., some tags stay and some leave the interrogation zone) without requiring extra memory. To reduce tags' collisions, the scheme utilizes the information obtained from the previous identification process. The ID patterns resulting into idle cycles (i.e., no tag response) and readable cycles are saved for the next identification process. The tags' collisions are reduced only when the tags' IDs do not change significantly between two consecutive identification processes.

The Query Tree (QT) algorithm [13] is based on the BBT scheme and is categorized under both the prefix-based and the bit-by-bit approach. The major difference between the QT and the BBT schemes is that the QT scheme is a memoryless protocol. The QT scheme does not require a tag to maintain any inquiring history (e.g., a bit pointer). During each interrogation cycle the reader broadcasts a prefix. Only those tags whose IDs matches the received prefix send the remaining of their ID bits to the reader. If the reader detects a tag collision, it appends either 0 or 1 to the recent prefix and broadcasts the new prefix. This process continues until there

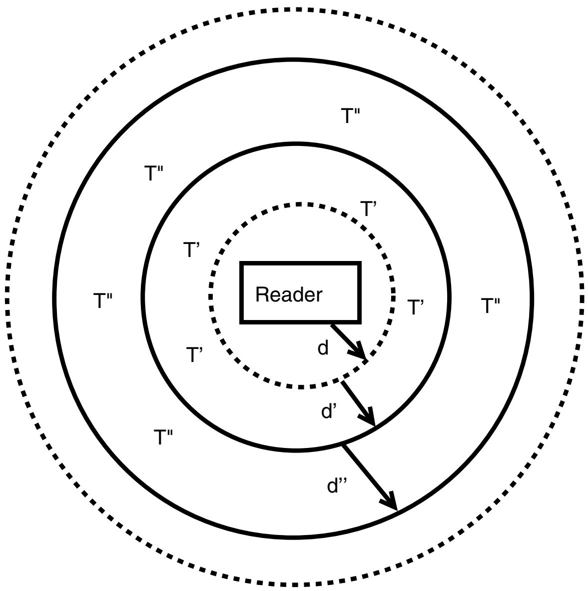

Fig. 2. Clustering based on the reader to tag distance, using reader power-levels to adjust its interrogation range

are no collisions and all the tags have been identified. Several variations of the QT algorithm exist, e.g., the work in [11], [25], [26] and [20].

The scanning-based sequential search algorithm is proposed in [16]. Initially, the reader finds all the positions of colliding bits for the initial IDs sent by the tags. The reader generates, from the LSB to the MSB, 2^n prefixes for the $n-$bit collision. This scheme can identify two tags in a cycle, as the prefix are generated from the LSB to the MSB and the tags' IDs are usually unique. The scanning-based pre-processing scheme [14] avoids requesting and scanning of the non-colliding IDs' bits. The tags' IDs are initially scanned to find the position of the collision bits. A bit position map, consisting of the position of the collision bits, is sent to all tags. This mechanism condenses a tag ID length and its value to the length and value of the bit position map. The BBT or the QT protocol is used for arbitrating collisions of the condensed IDs. This scheme lowers data transmission and increases throughput at the cost of extra memory requirement of the tags.

III. Power-based Distance Clustering

In [2], we proposed the power-based distance clustering anti-collision scheme. The main idea in that scheme is to divide the interrogation zone into smaller clusters based on the distance to the reader. Tags in different clusters are to be read separately. Since the number of tags in a cluster is less than that in the whole interrogation zone, the number of collisions should be intuitively reduced. Partitioning the interrogation zone can be achieved by controlling the reader's antenna power level. An example of such a partitioning is shown in Fig. 2, where the interrogation zone is divided into three clusters: d, $d^{'}$, and $d^{''}$. When the reader sends a request, only those tags in the current cluster respond. For instance, assume the reader had read the tags from cluster d and has just sent a request to cluster $d^{'}$. Then only tags from cluster $d^{'}$, i.e., tags marked as $T^{'}$, will respond to that request. After all tags, marked as $T^{'}$, have been read, they are put into a sleep mode. Then, the reader range is increased to $d^{''}$ and a new request is sent. Now only the tags from cluster $d^{''}$, i.e., labelled as $T^{''}$, will respond to the reader requests. This identification process continues until the reader reaches its maximum interrogation range. Any anti-collision scheme can be used to resolve collisions in a single cluster. The scheme is sensitive to the 'step' value. A small step value in sparse tag environments yields longer delay because of idle cycles, whereas a large step value may result in having too many tags in one cluster which renders the scheme ineffective, specially in dense tag environments.

IV. Optimal Distance-based Clustering

In this section we introduce our optimal distance-based clustering approach. We show a mathematical analysis to the delay, and give an efficient algorithm to find the optimal number of clusters.

A. System model and problem definition

We consider an RFID system consisting of an RFID reader and n passive tags. The interrogation zone is modeled as a circle centered at the RFID reader with a radius of R units; the RFID reader can communicate with, and read, all tags located within its interrogation zone.
The interrogation zone is divided into k equal sized clusters, and tags in different clusters will be interrogated separately (existing protocols use $k = 1$). In general, any anti-collision protocol can be used to resolve collisions in a particular cluster. To read tags in one cluster, the RFID reader and tags in that cluster will go through several cycles; in each cycle the RFID reader sends a request and zero, one, or more tags respond to the request by sending their IDs. An idle cycle is one in which no tag responds, a successful cycle is one in which exactly one tag responds, and a collision cycle is one in which two or more tags respond. Upon detecting a collision, the RFID reader will send a more restricted request which excludes some of the colliding tags. Eventually, a tag will be read successfully. This process will be repeated until all tags are read successfully. Each empty cluster causes one idle cycle. Now, the problem can be defined as follows.

Problem definition: Find the optimal number of clusters that minimizes the total number of cycles required to read all tags.

B. Assumptions

To resolve collisions within a single cluster, we will use the Query Tree (QT) protocol [13]. Nevertheless, any existing tree-based protocol can be used for that purpose. We will also assume that tags are uniformly distributed over the interrogation zone. However, a very similar analysis can be performed to deal with other distributions. We also assume that the total number of tags is known to the reader; several schemes in the literature (e.g., [12]) are able to find a precise estimation to the number of tags in a negligible time.

C. Delay analysis

Let $f(n)$ denote the number of cycles required to read a set of n tags, and let $E[f(n)]$ denote the expected value for $f(n)$. For the QT protocol, it has been shown in [13] that for $n \geq 4$,

$$2.881n - 1 \leq E[f(n)] \leq 2.887n - 1 \tag{1}$$

To generalize our analysis to other tree-based anti-collision schemes (e.g., the recent work in [19][1]), we will have:

$$E[f(n)] = c\,n - 1, \tag{2}$$

where c is a constant.
Now, assume that the interrogation zone is divided into k equal sized clusters, and tags are uniformly distributed over the interrogation zone. Let $g(n,k)$ denote the number of cycles required to read a set of n tags uniformly distributed over k clusters and using distance-based clustering. Thereby, we deduce the following lemma.

Lemma 1: If we have n tags, and the interrogation zone is divided into k equal sized clusters, then

$$E[g(n,k)] = c\,n - k + 2s, \tag{3}$$

where s is the number of empty clusters (i.e., clusters which do not contain any tag).

Proof: We have s empty clusters and t nonempty clusters. Let n_i denote the number of tags in the i^{th} nonempty cluster. For each empty cluster, we will have a single idle cycle; this results in a total of s cycles. The expected number of cycles required to read tags in the i^{th} nonempty cluster is $c\,n_i - 1$. Therefore, the expected number of cycles required to read tags in all nonempty clusters is $c\ n - t$. Therefore, the expected total number of cycles is $c\,n - t + s$, but $t = k - s$, so we get $c\,n - k + 2s$. ■

Now, we deduce the following lemma for the expectation of the number of empty clusters.

Lemma 2: If tags are uniformly distributed over the interrogation zone, then

$$E[s] = k\left(\frac{k-1}{k}\right)^n \tag{4}$$

[1]The scheme proposed in [19] has $f(n) = 2\,n - 1$.

Proof: Let P_i denote the probability that the i^{th} cluster is empty (i.e., no tag is located in that cluster). Then,

$$\begin{aligned} E[s] &= \sum_{i=1}^{k} P_i && (5)\\ &= \sum_{i=1}^{k}\left(\frac{k-1}{k}\right)^n && (6)\\ &= k\left(\frac{k-1}{k}\right)^n && (7) \end{aligned}$$

■

The following theorem is a direct result from Lemma 1 and Lemma 2.

Theorem 1: If we have n tags uniformly distributed over k equal sized clusters, then

$$E[g(n,k)] = c\,n - k + 2k\left(\frac{k-1}{k}\right)^n \tag{8}$$

D. Optimizing the delay

The objective of this sub-section is to find the optimal number of clusters (i.e., the value of k that minimizes $E[g(n,k)]$). We start with the following lemma.

Lemma 3: $E[g(n,k)]$ is a convex function over the interval $[1,\infty)$.

Proof: $E[g(n,k)]$ is twice differentiable with respect to k over the interval $(0,\infty)$, and

$$\frac{d^2}{dk^2}\,E[g(n,k)] = \frac{2n}{k^3}(n-1)\left(\frac{k-1}{k}\right)^{n-2} \tag{9}$$

Since $n \geq 1$ and $k \geq 1$, $\frac{d^2}{dk^2}\ E[g(n,k)]$ is non-negative. Therefore, $E[g(n,k)]$ is convex over the interval $[1,\infty)$. ■

From lemma 3, we know that the optimal value for k is the one at which $\frac{d}{dk}E[g(n,k)] = 0$. Therefore, we need to solve the following equation.

$$-1 + 2\left(\frac{k-1}{k}\right)^n + \frac{2n}{k}\left(\frac{k-1}{k}\right)^{n-1} = 0 \tag{10}$$

Let x_{opt} denote the solution to equation 10. In general, there is no closed-form solution to equation 10. However, it can be solved by numerical methods. Moreover, since the number of clusters is integer, we just need to find an integer x_{int}, such that $x_{int} \leq x_{opt} \leq x_{int} + 1$; the optimal number of clusters is either x_{int} or $x_{int} + 1$. The following lemma shows that $1 \leq x_{int} \leq n - 1$.

Lemma 4: When there are n tags, the optimal number of clusters is at most n.

Proof: To prove lemma 4, it suffices to show that the expected number of cycles with n clusters is less than that with $n+i$ clusters, where $i \geq 1$, (i.e., we need to show that $E[g(n,n)] < E[g(n,n+i)]$). With $n+i$ clusters, we are certain (with probability 1) that at least i clusters are empty, which results in i idle cycles. Thus,

$$\begin{aligned} E[g(n,n+i)] &= E[g(n,n+i)|\text{at least } i \text{ clusters are empty}] \\ &= i + E[g(n,n)] \\ &> E[g(n,n)] \end{aligned} \tag{11}$$

■

Since $E[g(n,k)]$ is convex and the optimal number of clusters is at most n, $\frac{d}{dk}E[g(n,k)]$ is a non-decreasing function. Therefore, one can find the optimal solution through binary search over the set $\{1,2,\cdots,n\}$. Algorithm 1 finds the optimal number of clusters in $O(\log n)$.

Algorithm 1: Finding the optimal number of clusters

Function Derivative(n,i)
Input: n: the number of tags.
i: an integer between 1 and n.
Output: $\frac{d}{dk}E[g(n,k)]$ at $k=i$.
begin
  return $-1+2\left(\frac{i-1}{i}\right)^n + \frac{2n}{i}\left(\frac{i-1}{i}\right)^{n-1}$;
end

Function FindOptimal(n)
Input: n: the number of tags.
Output: the optimal number of clusters.
begin
  $left = 1$;
  $right = n$;
  while *right-left>1* **do**
    if Derivative($n, \lfloor\frac{left+right}{2}\rfloor$) < 0 **then**
      $left = \lfloor\frac{left+right}{2}\rfloor$;
    else
      $right = \lfloor\frac{left+right}{2}\rfloor$;
    end
  end
  if |Derivative($n, left$)| $<$ |Derivative($n, right$)| **then**
    optimal $= left$;
  else
    optimal $= right$;
  end
  return *optimal*;
end

E. Discussion

Dividing the interrogation zone into clusters is achieved by controlling the power level of the RFID reader's signal. Since RFID readers may have a finite number of power levels they can use, that may put a limit on the maximum possible number of clusters. The only change we need to make in order to accommodate such a limitation is to limit the search space in Algorithm 1 accordingly (i.e., we start with $right = \text{MIN}(n, \text{maximum possible number of clusters})$ instead of $right = n$).

It is also important to notice that Algorithm 1 is very efficient in practise as it runs in $O(\log n)$ time; even if we have 10^6 tags, it will take less than 20 computation steps to find the optimal solution, which can be done in a negligible time with today's standard personal computers.

Finally, it is important to recall that the mathematical analysis is not limited to the uniform distribution; a very similar analysis can be performed to handle other distributions, and that is left to an extended version of this paper. The same thing applies to the anti-collision scheme used within single clusters; any existing tree-based scheme can be used. In fact, even non-tree-based schemes can be used as long as their delay can be expressed as a function of the number of tags.

V. Simulation and Results

In this section, we compare the following schemes: the Query Tree (QT) [13] without clustering, the Enhanced Binary Search Tree (EBST) [19] without clustering, the QT and the EBST schemes with static clustering, and the QT and the EBST with optimal clustering. We expect to see similar results if other tree-based schemes are used.

A. Simulation environment

We extended the ns-2 simulator [1] to implement RFID. In the setup, tags are uniformly distributed in a 20×20 m^2 region. A single reader is located at the center of the region. The maximum range of the reader is 10 m. Each tag has a 32-bit randomly generated ID. Simulations are made to run until all tags are successfully identified. In evaluating the different schemes, we use performance metrics, such as number of cycles required to identify all tags, number of bits transferred by the tags to the reader, number of responses, and number of bits colliding at the reader. We show the normalized improvement of our schemes; the normalized improvement is the ratio of the results of an original scheme (without clustering) to the results of the same scheme with clustering (with optimal clustering and static clustering). For static clustering, we use our algorithm in [2] with a stepping value of 0.5. Our results represent the average of 20 randomly generated instances of the system.

B. Performance evaluation

1) Collisions: Collision can be defined at either the bit level, the response level, or both. At the bit level, we try to minimize the number of the colliding bits at the reader. Response level measures whether the response from a certain tag collides with other tags' responses. The importance of each collision type depends on the anti-collision scheme. For instance, it is more important to reduce the bit level collisions in the QT scheme which is a bit-by-bit approach, compared to the EBST scheme. The normalized response level collisions and the normalized bit level collisions improvement for the

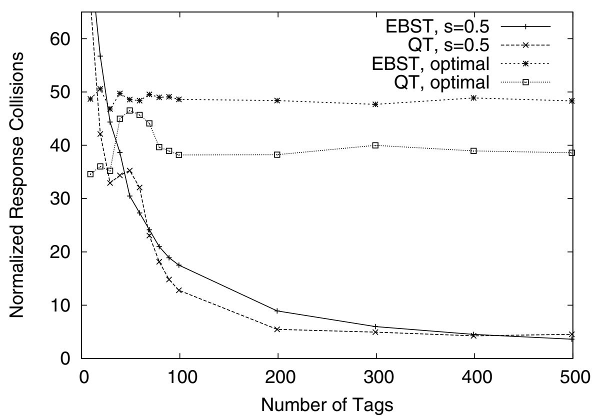

Fig. 3. Normalized Improvement for Response Collision

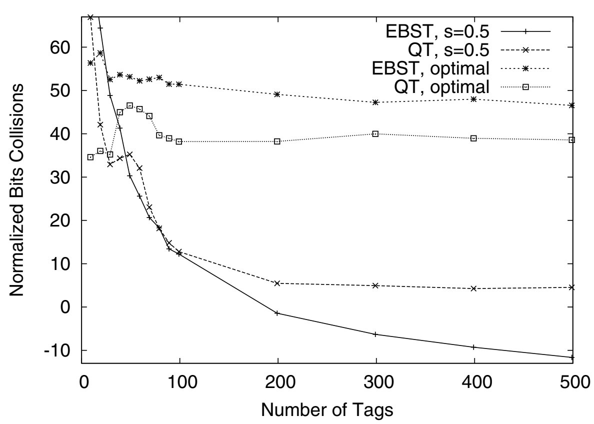

Fig. 4. Normalized Improvement for Bits Collision

optimal and static clustering based QT and EBST are shown in Fig. 3 and Fig. 4, respectively. The normalized improvement in number of collisions for the static clustering scheme when augmented with EBST decreases when we increase the number of tags; this is an effect of the constant stepping value. With constant stepping, the value, which results in near-optimal clustering for a few tags, shows poor performance when the number of tags is large. However, this problem is solved by our proposed optimal clustering scheme which partitions the interrogation zone according to the number of tags, and thereby shows optimal results with any number of tags. For instance, with over 100 tags in the interrogation range, the optimal clustering-based schemes (both QT and EBST) decreased the number of collisions by five times as compared to the static clustering schemes.

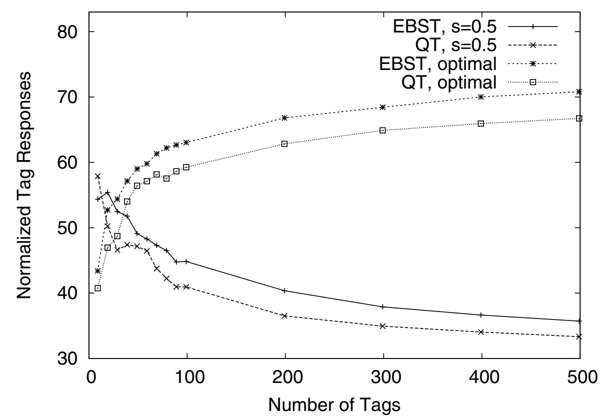

Fig. 5. Normalized improvement in tags responses

2) Communication: Passive tags harvest energy from the reader's generated electromagnetic waves and modulate the reflected signal to send a response to the reader. The normalized improvement in number of responses for the augmented EBST and QT schemes is shown in Fig. 5. The optimal clustering-based scheme results in an improvement of over 70%, whereas the static clustering-based schemes shows normalized improvement of 35%. An interesting observation is that the normalized improvement for the static clustering-based schemes decreases when the number of tags increases. On the other hand, optimal clustering-based schemes results in an incremental improvement even with dense tag deployment. For instance, for the optimal clustering-based EBST, an improvement of 60% and 70% is observed for 100 and 500 tags, respectively.

We also compare the different schemes in terms of the total number of bits transferred during the tags' singulation process. The normalized performance improvements in bits transferred by the tags is shown in Fig. 6. As expected, the optimal clustering-based EBST achieves significant improvement: over 60% as compared to that of the static clustering-based EBST scheme. We have not observed relative improvements in the optimal clustering-based QT scheme. The reason is that the QT scheme is a bit-by-bit approach where the tags reply by sending the requested bit and this process is repeated until all the bits have been transferred and a tag is successfully singulated.

3) Reading Cycles: We compare the different schemes in terms of the total number of cycles required to read all tags. A reading cycle is a query sent by the reader followed by zero, one, or more tag responses. For instance, sending $REQA$ or $SELECT$ command is considered as a cycle, whereas $RESET$ is not, where the reader is not expecting any response from tags. Reading cycles also include idle cycles, i.e., when no tag responds to the reader query. Idle cycles are possible because of empty clusters, i.e., clusters

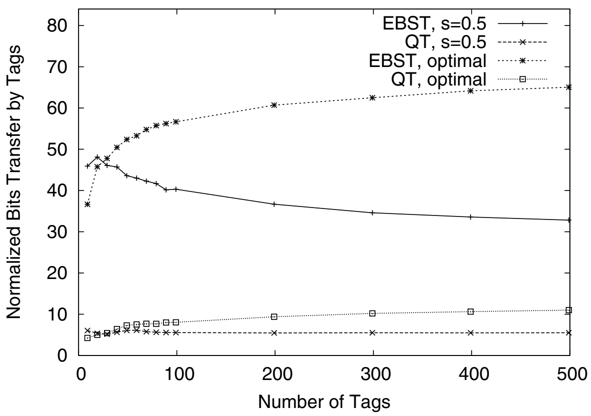

Fig. 6. Normalized improvement in bits send by the tags

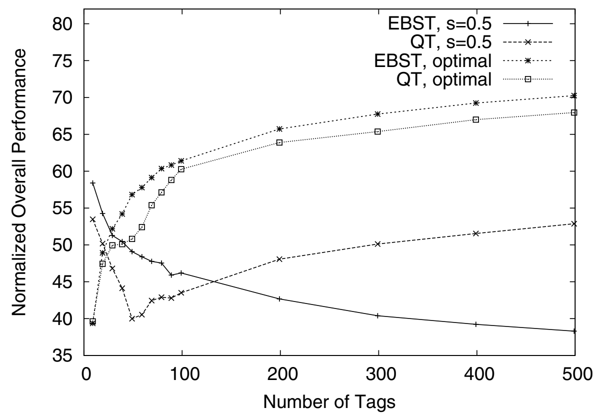

Fig. 8. Overall performance improvement of the optimal-cluster and nonoptimal-cluster augmented anti-collision schemes

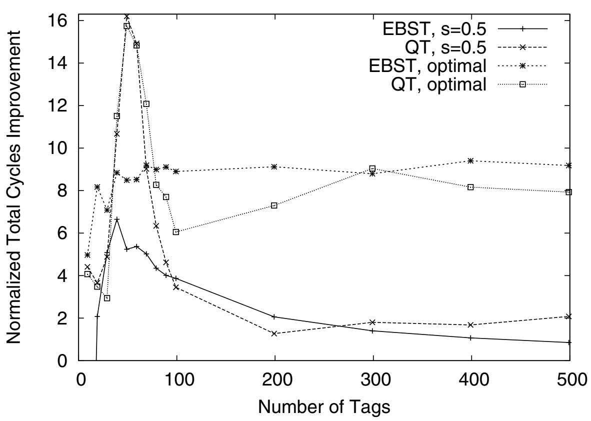

Fig. 7. Normalized Cycles Improvement

with no tags. Fig. 7 shows the normalized improvement in the total number of cycles for the optimal and the static clustering-based schemes. The improvements made by the optimal clustering-based EBST and QT are about five times more than the improvements of the static clustering-based EBST scheme. The optimal scheme is able to reduce the collisions and the number of responses required to singulate a tag, which results in fewer cycles to interrogate all tags.

VI. Conclusion

In this paper we introduce the optimal distance-based clustering approach for RFID passive tags' anti-collision. The main idea in our approach is to divide the interrogation zone into equal sized clusters based on the distance to the reader, and tags in different clusters are read separately. Since the number of tags in a single cluster is less than that in the whole interrogation zone, the likelihood of a collision is reduced. On the other hand, if we have too many clusters, we may end up having many empty clusters, which results in a significant overhead. Our approach finds a balance between reducing the likelihood of collisions and reducing the number of empty clusters. Theoretical analysis and ns-2 simulations have been presented to show the superiority of our approach. Moreover, our approach can be integrated with any existing tree-based anti-collision protocol.

References

[1] The network simulator ns-2.

[2] K. Ali, H. Hassanein, and A.-E. M. Taha. Rfid anti-collision protocol for dense passive tag environments. In *LCN '07: Proceedings of the 32nd IEEE Conference on Local Computer Networks*, pages 819–824, Washington, DC, USA, 2007. IEEE Computer Society.

[3] L. Bolotnyy and G. Robins. Randomized pseudo-random function tree walking algorithm for secure radio-frequency identification. In *AUTOID '05: Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pages 43–48, Washington, DC, USA, 2005. IEEE Computer Society.

[4] J. I. Capetanakis. Tree algorithms for packet broadcast channels. *IEEE Transactions on Information Theory IT-25*, (5):505–515, 1979.

[5] A.-I. center. *Draft Protocol Specification for a Class 0 Radio frequency identification tag*, 2003.

[6] H.-S. Choi, J.-R. Cha, and J.-H. Kim. Fast wireless anti-collision algorithm in ubiquitous id system. In *VTC-2004: Proceedings of the 60^{th} IEEE Vehicular Technology Conference*, pages 4589–4592, 2004.

[7] B. Feng, J.-T. Li, J.-B. Guo, and Z.-H. Ding. Id-binary tree stack anticollision algorithm for rfid. In *ISCC '06: Proceedings of the 11th IEEE Symposium on Computers and Communications*, pages 207–212, Washington, DC, USA, 2006. IEEE Computer Society.

[8] K. Finkenzeller. *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*. John Wiley & Sons, Inc., New York, NY, USA, 2003.

[9] M. Jacomet, A. Ehrsam, and U. Gehrig. Contactless identification device with anticollision algorithm. In *CSCC '05: Proceedings of the IEEE conference on Circuits, Systems, Computer and Communications*, 1999.

[10] A. Juels, R. L. Rivest, and M. Szydlo. The blocker tag: selective blocking of rfid tags for consumer privacy. In *CCS '03: Proceedings of the 10^{th} ACM Conference on Computer and Communications Security*, pages 103–111, New York, NY, USA, 2003. ACM Press.

[11] S. H. Kim, M. K. Shin, and P. Park. A new tree-based tag anti-collision protocol for rfid systems. In *CIC '06: Proceedings of the 2006 International Conference on Communications in Computing*, pages 83–86. CSREA Press, June 2006.

[12] M. Kodialam and T. Nandagopal. Fast and reliable estimation schemes in rfid systems. In *MobiCom '06: Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, pages 322–333, New York, NY, USA, 2006. ACM Press.

[13] C. Law, K. Lee, and K.-Y. Siu. Efficient memoryless protocol for tag identification. In *DIALM '00: Proceedings of the 4^{th} International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 75–84, New York, NY, USA, 2000. ACM Press.

[14] C. Law, K. Lee, and K. Y. Siu. Scanning-based pre-processing for enhanced rfid tag anti-collision protocols. In *ISCIT '06: Proceedings of the 2006 International Symposium on Communications and Information Technologies*, pages 75–84, New York, NY, USA, October 2006. ACM Press.

[15] S.-R. Lee, S.-D. Joo, and C.-W. Lee. An enhanced dynamic framed slotted aloha algorithm for rfid tag identification. In *MOBIQUITOUS '05: Proceedings of the The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 166–174, Washington, DC, USA, 2005. IEEE Computer Society.

[16] L. Liu, Z. Xie, J. Xi, and S. Lai. An improved anti-collision algorithm in rfid system. In *Proceedings of the 2^{nd} International Conference on Mobile Technology, Applications and System*, pages 1–5, Nov 2005.

[17] J. Myung and W. Lee. Adaptive splitting protocols for rfid tag collision arbitration. In *MobiHoc '06: Proceedings of the Seventh ACM International Symposium on Mobile Adhoc Networking and Computing*, pages 202–213, New York, NY, USA, 2006. ACM Press.

[18] T. A. Scharfeld. An analysis of the fundamental constraints on low cost passive radio-frequency identification system design. Master's thesis, Massachusetts Institute of Technology, Aug. 2001.

[19] X.-L. Shi, X.-W. Shi, Q.-L. Huang, and F. Wei. An enhanced binary anti-collision algorithm of backtracking in rfid system. *Progress In Electromagnetics Research B*, 4:263–271, 2008.

[20] D.-H. Shih, P.-L. Sun, D. C. Yen, and S.-M. Huang. Taxonomy and survey of rfid anti-collision protocols. *Computer Communications*, 11(29):2150–2166, 2006.

[21] D. Simplot-Ryl, I. Stojmenovic, A. Micic, and A. Nayak. A hybrid randomized protocol for rfid tag identification. In *IEEE International Workshop on Next Generation Wireless Networks*, 2005.

[22] V. Stanford. Pervasive computing goes the last 100 feet with rfid systems. *IEEE Pervasive Computing*, 2(2):9–14, 2003.

[23] H. Vogt. Efficient object identification with passive rfid tags. In *Pervasive '02: Proceedings of the First International Conference on Pervasive Computing*, pages 98–113, London, UK, 2002. Springer-Verlag.

[24] T.-P. Wang. Enhanced binary search with cut-through operation for anti-collision in rfid systems. *IEEE Communications Letters*, 10(4):236–238, 2006.

[25] F. Zhou, , D. Jin, C. Huang, and H. Min. Optimize the power consumption of passive electronic tags for anti-collision schemes. In *Proceedings of the 5th ASICON*, pages 1213–1217, October 2003.

[26] F. Zhou, C. Chen, D. Jin, C. Huang, and H. Min. Evaluating and optimizing power consumption of anti-collision protocols for applications in rfid systems. In *ISLPED '04: Proceedings of the 2004 International Symposium on Low Power Electronics and Design*, pages 357–362, New York, NY, USA, 2004. ACM Press.