IET Communications

Research Article

ISSN 1751-8628
Received on 31st May 2014
Accepted on 4th November 2014
doi: 10.1049/iet-com.2014.0541
www.ietdl.org

# On maximising tag reading efficiency of a multi-packet reception capable radio frequency identification reader

*Subodh Pudasaini*[1], *Kyung Sup Kwak*[1], *Seokjoo Shin*[2] ✉

[1]*UWB Wireless Communication Research Center, Inha University, Incheon, Korea*
[2]*Department of Computer Engineering, Chosun University, Gwangju, Korea*
✉ *E-mail: sjshin@chosun.ac.kr*

**Abstract:** Maximising the tag reading rate of a reader is one of the most important design objectives in radio frequency identification (RFID) systems as it is inversely proportional to the time required to completely read all the tags within the reader's radio field. To this end, numerous techniques have been independently suggested so far and they can be broadly categorised into pure advancements in the link-layer tag anti-collision protocols and pure advancements in the physical-layer RF signal reception model. This study shows by rigorous mathematical analysis and Monte-Carlo simulations that how those two independent approaches can be coupled to maximise the tag reading efficiency in an RFID system, considering a slotted Aloha-based dynamic link-layer anti-collision protocol at tags and a multi-packet reception capable RF reception model at the reader.

## 1 Introduction

Radio frequency identification (RFID) is a rapidly evolving automatic identification and tracking system. Even though the basic operating principles of modern RFID systems have been known for several decades, their adoption in numerous industrial and consumer applications (such as supply chain management, inventory control, supermarket checkout process and toll collections) has been proliferated recently because of the ability now to build miniaturised RFID components at low cost [1].

Typically, an RFID system consists of two components: a reader and tags. Each tag has a unique ID stored in its memory. The reader should read (interrogate) IDs of all the tags within its radio field, and for this purpose it broadcasts interrogation RF signal periodically. If an RFID tag finds itself within the RF-field of the reader, it backscatters (i.e. transmits back) a signal containing its unique ID [2]. When more than one RFID tags backscatter their IDs using a common chunk of the shared wireless channel (in terms of frequency, time, space or code), signal from one tag interferes the signals from others, and the reader might not be able to decode IDs of the backscattering tags. Such phenomenon is commonly known as tag-collision. Occurrence of such tag-collision events triggers the collided tags to retransmit their IDs in the subsequent interrogation rounds and thus elongates tag identification delay (or in other words reduces the tag reading rate) at the reader. Many link-layer (more precisely medium access control sub-layer) anti-collision protocols have been developed so far to address the tag-collision problem [3]. These protocols not only reduce the frequency of occurrence of tag-collision events but also help to recover from such events as quickly as possible.

In a broad sense, time division multiple access RFID anti-collision protocols are classified as either deterministic or probabilistic protocols based on how tags are allocated a fraction of the shared channel resource (a time slot) to transmit their IDs. The former type of protocols is based on binary tree (BT) where the collided tags are split into two subsets. The tags in the first subset transmit their IDs in the next slot, while the tags in the other subset wait until the first subset of tags are successfully identified. This process is repeated recursively until all tags are recognised. The performance of tree-based anti-collision protocols deteriorates with the increase in the number of tags. This can be attributed to the fact that even though colliding tags are successively grouped into two subsets, each subset may still contain many tags resulting in collision [4]. On the other hand, in probabilistic protocols such as framed slotted Aloha (FSA), the channel time is split into frames and a single frame is further divided into several time slots. During each frame, each tag randomly chooses a time slot and transmits its ID to the reader in that slot. The unidentified tags will transmit their IDs in the next frame. It has been shown that the probabilistic FSA can achieve smaller tag identification delay than its deterministic counterpart [5].

In the literature there exist many works that have been independently developed by different researchers and engineers to enhance the tag identification performance of RFID systems. Some of the representative works are available in [6–9]. Based on the scope of their design, they can be categorised into (i) pure advancement in the link-layer anti-collision protocols and (ii) pure advancement in the physical layer RF reception models. The fundamental approach behind the first category of enhancements is to dynamically adjust the frame length of the probabilistic FSA protocols to its optimal value in each interrogation round (resulting in new protocol referred to as dynamic FSA or DFSA [6]), or to optimise tree search algorithm in the deterministic BT protocols by taking advantage of inherent correlatedness among the tag IDs [7]. The latter category of enhancement, on the other hand, uses multiuser multiple-input multiple-output (MU-MIMO) technique along with efficient blind signal separation algorithms to realise multi-packet reception (MPR) capable of RF reception model at the reader [8, 9]. Owing to the MPR capability at the reader, simultaneously transmitted signals from several tags can be separated and the transmitting tags can be correctly identified (which would have been otherwise treated as being collided).

It has been shown in [10, 11] that the MPR capability at the reader has the potential to substantially increase the read rate and decrease the identification delay of FSA and BT anti-collision protocols. However, how to ascertain optimal tag reading performance in an RFID system with the MPR capability is remained as an open research problem. To this end, we derive an optimality criterion and present a method to adopt such a criterion in the probabilistic DFSA anti-collision protocol in an RFID system with the MPR-capable reader. To the best of our knowledge, it is the first work in this regard.

The rest of the paper is organised as follows: Section 2 presents the system model, whereas Section 3 presents the analytical derivation of a criterion for achieving optimal tag reading efficiency. Section 4 provides detailed information about simulations environment, performance metric and evaluation methodology. Finally, Section 5 concludes this work.

## 2 System model

We consider an RFID system where $n$ number of tags with single antenna communicate with a reader equipped with an array of $M$ antennas. Under such MU-MIMO setting, it is assumed that spatially multiplexed backscattered signals from multiple tags can be separated using an appropriate signal processing algorithm for blind source separation (i.e. the separation of independent sources from a mixed signal without having knowledge of the mixing process). Mindikoglu and Veen [8] and Dacuna *et al.* [9] have demonstrated the feasibility of such blind signal separation algorithms using zero constant modulus analysis and a modified fast-independent component analysis, respectively, to separate multiple tag signals in an ultra-high frequency RFID system. DFSA is used as the anti-collision protocol. The operation procedures of DFSA at the reader and tag are described below:

***At the reader side***:

1. Set initial frame length.
2. Initiate interrogation round by broadcasting the frame length information.
3. In each slot of the frame, check whether there are any backscattered RF signals from the tags. Mark the slot as an empty slot if no backscattered RF signal is detected. If RF signals are detected, use the advanced signal separation algorithm to separate the multiplexed backscatter RF signals. Based on the outcome of the signal separation operation, mark the slot as a collided slot if none of the transmitting tags are identified, and mark it as a successful slot if any of the tags are identified. Also, record the number of identified tags in the successful slot.
4. After the completion of the frame, check whether any slot within that frame is marked as the collided slot. It is the indication whether any tags are left to be interrogated or not. If none of the slots are marked as the collided slot, terminate the interrogation process. Otherwise, prepare for the next interrogation round.
5. Estimate the total number of contending tags in the last frame using maximum a posteriori (MAP) based estimation method in (14). As to be elaborated in the following section, the MAP estimation mechanism utilises the statistics of the collided, successful and idle slots to perform estimation.
6. Determine the optimal frame length for next interrogation round using (11) and go to step 2. It is noteworthy to mention that we do not restrict frame length value to be a power of 2 as recommended in EPC Global Class-1 Gen-2 standard because this recommendation only offers sub-optimal performance.

***At the tag side***:

1. Wait for interrogation signal from the reader.
2. Obtain the frame length information.
3. Randomly select any of the slot within the frame and backscatter its ID in the selected slot.
4. If the transmission is inferred to be unsuccessful, wait for interrogation signal for the next round.

## 3 Optimal tag reading criterion

In this section, we derive a theoretical criterion for achieving optimal tag reading performance at the reader with the MPR capability and present a method to adopt such a criterion in the RFID systems.

### 3.1 Optimal criterion

Consider the RFID system described in the previous section with $n$ tags to be read. The frame used in an interrogation round initiated by the reader consists of $L$ time slots. So, the probability that $j$ tags among $n$ tags occupy a slot can be expressed by the binomial distribution with parameters $n$ and $1/L$ as

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

If the frame length $L$ is sufficiently large, (1) can be approximated by the Poisson distribution with mean $n/L$. Accordingly, the probabilities that a slot is found to be empty (no tags use the slot), successful ($M$ or less number of tags use the slot) and collided (more than $M$ number of tags use the slot) are given by

$$p_{\mathrm{e}} = B(j=0) \simeq \mathrm{e}^{-n/L} \tag{2}$$

$$p_{\mathrm{s}} = B(1 \le j \le M) \simeq \mathrm{e}^{-n/L}\sum_{j=1}^{M}\frac{(n/L)^{j}}{j!} \tag{3}$$

$$p_{\mathrm{c}} = B(j > M) = 1 - p_{\mathrm{e}} - p_{\mathrm{s}} \tag{4}$$

Based on (3), the expected value of the number of successful slots $S$ in the frame with $L$ slots is

$$E[S] = L\cdot \mathrm{e}^{-n/L}\sum_{j=1}^{M}\frac{(n/L)^{j}}{j!} \tag{5}$$

and, therefore, the expected channel usage efficiency is

$$U = \frac{E[S]}{L} = \mathrm{e}^{-n/L}\sum_{j=1}^{M}\frac{(n/L)^{j}}{j!} \tag{6}$$

To maximise the read rate (i.e. number of successful tags per unit time) of a reader, it should be ensured that the shared channel is used as efficiently as possible. This implies that a criterion that maximises the channel usage efficiency $U$ also maximises the read rate. Since $U$ in (6) is a concave-downward function of $L$, the criterion that maximises $U$ can be obtained by equating the derivative of $U$ (with respect to $L$) to zero as

$$\begin{aligned}\frac{\mathrm{d}U}{\mathrm{d}L} &= \frac{\mathrm{d}}{\mathrm{d}L}\left(\mathrm{e}^{-n/L}\frac{n}{L}\right) + \frac{\mathrm{d}}{\mathrm{d}L}\left(\mathrm{e}^{-n/L}\left(\frac{n}{L}\right)^{2}\frac{1}{2!}\right)\\ &\quad + \cdots + \frac{\mathrm{d}}{\mathrm{d}L}\left(\mathrm{e}^{-n/L}\left(\frac{n}{L}\right)^{M}\frac{1}{M!}\right)\\ &= \left[\frac{n}{L^{2}}\mathrm{e}^{-n/L}\left(\frac{n}{L}-1\right)\right] + \left[\frac{n^{2}}{2!L^{3}}\mathrm{e}^{-n/L}\left(\frac{n}{L}-2\right)\right]\\ &\quad + \cdots + \left[\frac{n^{M}}{2!L^{M+1}}\mathrm{e}^{-n/L}\left(\frac{n}{L}-M\right)\right]\\ &= \frac{1}{L}\mathrm{e}^{-n/L}\sum_{m=1}^{M}\left(\frac{1}{m!}\left(\frac{n}{L}\right)^{m}\left(\frac{n}{L}-m\right)\right) = 0\end{aligned} \tag{7}$$

For finite $n$ and $L$, the factor $(1/L)\mathrm{e}^{-n/L}$ in (7) cannot be equal to zero, so roots of (7) can be obtained by solving

$$\sum_{m=1}^{M}\left(\frac{1}{m!}\left(\frac{n}{L}\right)^{m}\left(\frac{n}{L}-m\right)\right) = 0 \tag{8}$$

As formally proved in Appendix 1, the left-hand side of (8) can be represented as $(((n/L)^{M+1} - M!(n/L))/M!)$. Therefore by solving

$$\frac{(n/L)^{M+1} - M!(n/L)}{M!} = 0 \tag{9}$$

the optimal criterion (i.e. optimal frame length $L$*) that maximises $U$

is found to be

$$L^* = \frac{n}{(M!)^{(1/M)}} \tag{10}$$

By plugging the value of $L^*$ into (6), the optimal value $U$ is found to be $\mathrm{e}^{(-M!)^{1/M}} \sum_{j=1}^{M} \big(((M!)^{1/M})^j / j!\big)$. Table 1 presents the optimal $U$ for some practically interesting values of $M$.

### 3.2 Adoption of the optimal criterion

If the number of tags to be interrogated is known in advance, the optimal frame length that maximises $U$ can be set directly according to (10). The cardinality of tags to be interrogated, however, is generally not known in advance and hence should be estimated on-the-fly for every read cycle $i$ except for the initial read cycle (i.e. for $i > 1$). Thus, the optimal frame length for $i$th read cycle can be adjusted using

$$L_i^* = \frac{\hat{n}_{i-1} - S_{i-1}^{(t)}}{(M!)^{(1/M)}} \tag{11}$$

where the numerator is the estimated remaining number of tags to be interrogated. It is simply the difference between the estimated number of contending tags $\hat{n}_{i-1}$ and the successfully identified tags $S_{i-1}^{(t)}$ in the $(i-1)$th read cycle.

Many tag estimation algorithms have been reported so far and they vary in computational complexity and achievable estimation accuracy [12–15]. Most accurate algorithms are computationally heavy while those that require less computations are relatively less accurate. We consider a MAP-based tag estimation method [12] in this work as it offers good tradeoff between the aforesaid traits. Chen's original estimation algorithm did not consider the MPR capabilities in the reader and hence his tag estimation formula is only applicable for single-packet reception model (i.e. $M = 1$). In what follows, we extend the Chen's formula for all possible values of $M$. In a frame with $L$ slots, the joint probability mass function for finding $X$ empty slots, $Y$ successful slots and $Z$ collision slots can be represented using the following trinomial distribution

$$P(X, Y, Z) = \frac{L!}{X!Y!Z!} p_\mathrm{e}^X p_\mathrm{s}^Y p_\mathrm{c}^Z \tag{12}$$

where $p_\mathrm{e}$, $p_\mathrm{s}$ and $p_\mathrm{c}$ are previously defined in (2), (3) and (4), respectively. Hence, when the reader finds $E$ empty slots, $S$ successful slots and $C$ collision slots in a frame, a posteriori probability distribution of having $k$ tags in the system is

$$\begin{aligned} P(k|E, S, C) = {} & \frac{L!}{E!S!C!} (\mathrm{e}^{-k/L})^E \\ & \times \big[\mathrm{e}^{-k/L}\big(T_M(k/L) - 1\big)\big]^S \\ & \times \big[\mathrm{e}^{-k/L}\big(\mathrm{e}^{k/L} - T_M(k/L)\big)\big]^C \end{aligned} \tag{13}$$

where $T_M(k/L)$ is the Taylor polynomial of $\mathrm{e}^{k/L}$ of order $M$. Based on the a posterior probability distribution in (13), for every read cycle except the initial one, the reader determines the total number of estimated tags as

$$\hat{n} = \arg\max_k P(k|E, S, C) \tag{14}$$

Fig. 1 shows the *a posteriori* probability distribution for $n$ tags when one empty slot, six collision slots and three success slots are observed in a frame with 10 slots for three different cases of $M$ (viz. $M = 1$, $M = 2$ and $M = 3$). For each case of $M$, the value corresponding to the peak of the distribution curve is the estimated number of tags.

It is noteworthy to mention that while implementing the MAP-based estimation method in the reader, the first constant factor (involving factorial) in $P(k|E, S, C)$ in (13) can be removed, as it is only responsible in scaling the probability mass function. There will be no difference in the estimation result but significant computation burden from the reader can be reduced, especially when $L$ is large.

**Table 1** Optimal channel efficiency $U$ for different values of $M$

| $M$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $U$ | 0.3679 | 0.5869 | 0.7260 | 0.8167 |

## 4 Performance results

We analysed the performance of the MPR-capable RFID system described in Section 2 for varying $M$, $L$ and $n$ using Monte-Carlo simulations. The average results of 500 simulation trials are presented in terms of the following two metrics: (i) Read rate: Number of tags identified per unit time and (ii) Identification delay: Total time required to read all the tags in the system. We considered the duration of a slot to be a basic unit of time, and hence the read rate is expressed in terms of tag/slot (number of tags per slot) and the identification delay in terms of number of slots.

Fig. 2*a* shows the read rate of a FSA anti-collision algorithm and DFSA anti-collision algorithm with varying MPR capabilities ($M = 1$, 2, 3 and 4) when the initial frame length was set to 128. It is evident from the figure that the read rate substantially increases with the increase in the value of $M$. This is attributed to the reduction in the number of tag-collision events because of MPR capability. The read rate reaches its peak value of 1.9 tags/slot for the case of $M = 4$, which in the conventional single-packet reception-capable reader (i.e. $M = 1$) is caped to be 0.36 tag/slot. Note that DFSA's peak read rate in the single-packet reception-capable reader agrees well to the previously established theoretical network throughput bound of $\frac{1}{e} (\simeq 0.63)$ in any Aloha-based random access systems. In the figure, it is also evident that by merely using FSA it is not possible to attain the read rate closer to $\frac{1}{e}$ in the single-packet reception-capable reader.

Fig. 2*b* shows the identification delay of FSA anti-collision algorithm and DFSA anti-collision algorithm with varying MPR capabilities. From the figure one can see that the increased read

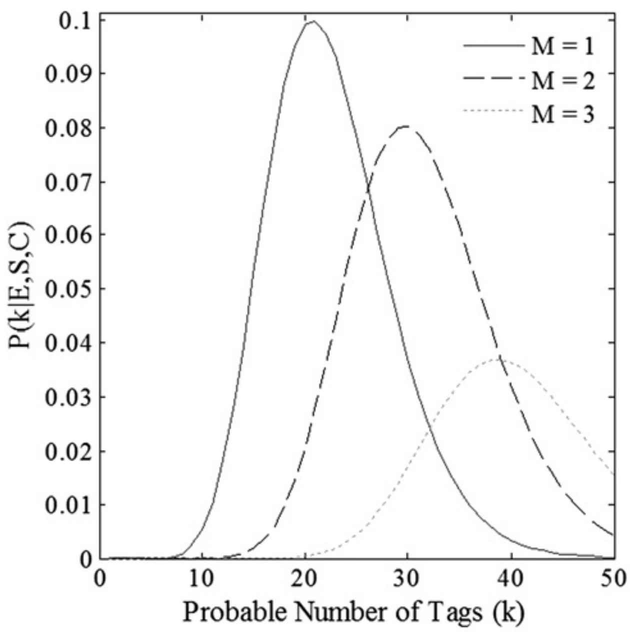


**Fig. 1** *Posteriori probability distribution of the estimated number of tags for different values of M when L = 10, C = 6, S = 3 and E = 1*

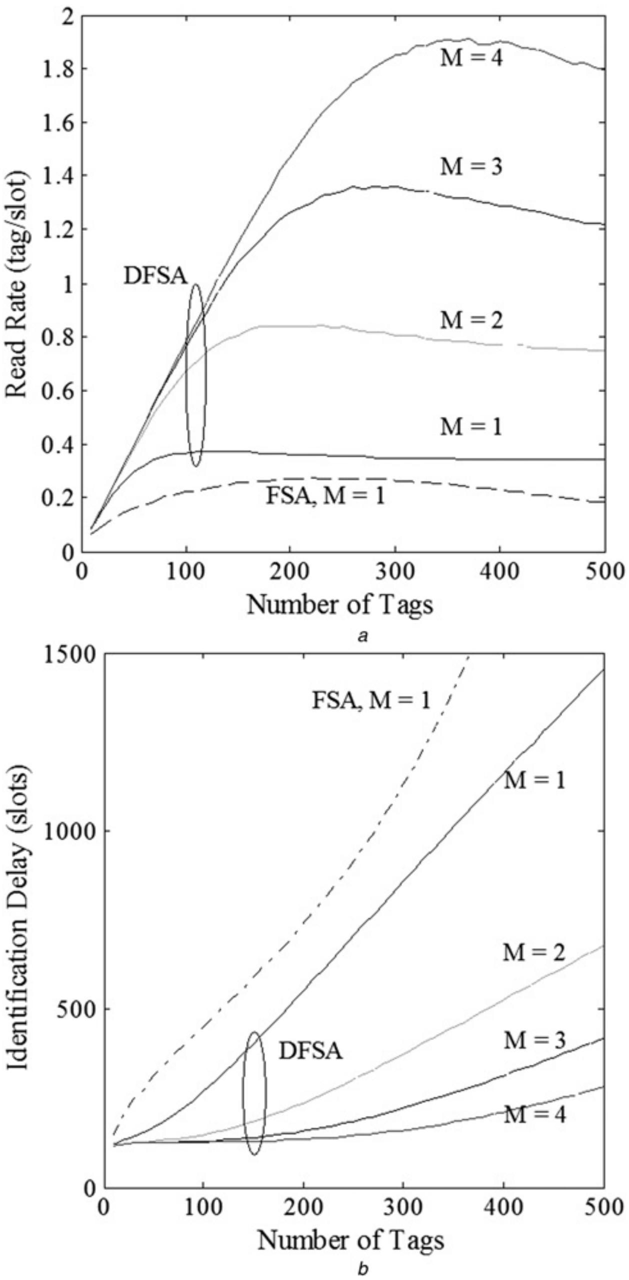


**Fig. 2** *Influence of M on the read rate and identification delay of DFSA anti-collision protocol in an RFID system with the MPR-capable reader*

*a* Read rate
*b* Identification delay

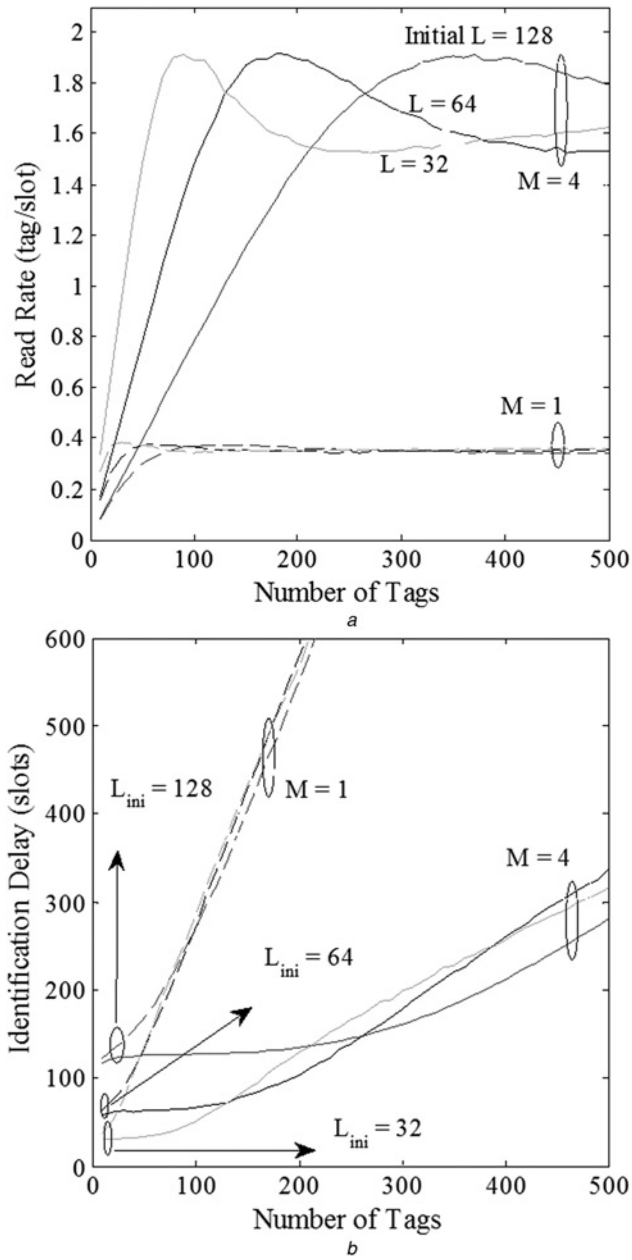


**Fig. 3** *Influence of L on the read rate and identification delay of DFSA anti-collision protocol in an RFID system with the MPR-capable reader*

*a* Read rate
*b* Identification delay

rate because of MPR capabilities (observed in Fig. 2*a*) translates to the reduction in the identification delay. For example, when there were around 350 tags in the RFID system, nearly 5.5-fold decrease in the identification delay (from 1011 slots to 184 slots) was observed when the single-packet reception-capable reader was replaced with the MPR-capable reader at $M = 4$.

Fig. 3 shows that the initial frame length affects the performance of DFSA in terms of read rate and identification delay, especially when the reader has high-order MPR capabilities and the number of tags to be interrogated is small. From the figure it is evident that the read rates for three different cases of initial frame lengths ($L = 32$, 64 and 128) appear to converge to a rate close to the peak read rate with the increase in the number of tags in the system. This implies that the effects of the initial frame length on the read rate tends to vanish with increase in the number of tags. Similarly, the difference in the identification delay for different frame length values shrinks for larger number of contending population size.

Next, we measured the accuracy of the MAP-based tag estimation method used in our previous simulations. For that we calculated the estimation error (in %) as $(|\hat{n} - n|/n) \times 100\%$, where $\hat{n}$ is the estimated number of tags when there were $n$ tags in the system. The lower value of the estimation error corresponds to the higher estimation accuracy. Fig. 4 depicts the estimation errors for four different cases of $M$ (1, 2, 3 and 4) when the frame length was set to 128 in the simulations. From the figure it is evident that the estimation error increases with the increase in the value of $M$, but only up to a certain tag population size. Beyond that tag population size, the estimation error for higher $M$ remains lower. Importantly, for all four different cases of $M$, the estimation error remains lower than 6%, regardless of the number of considered tags.

To put the impact of the estimation error into perspective, we compared the maximum channel efficiency that would be achievable under various tag estimation error settings. The comparison results are given in Table 2. It is important to note that

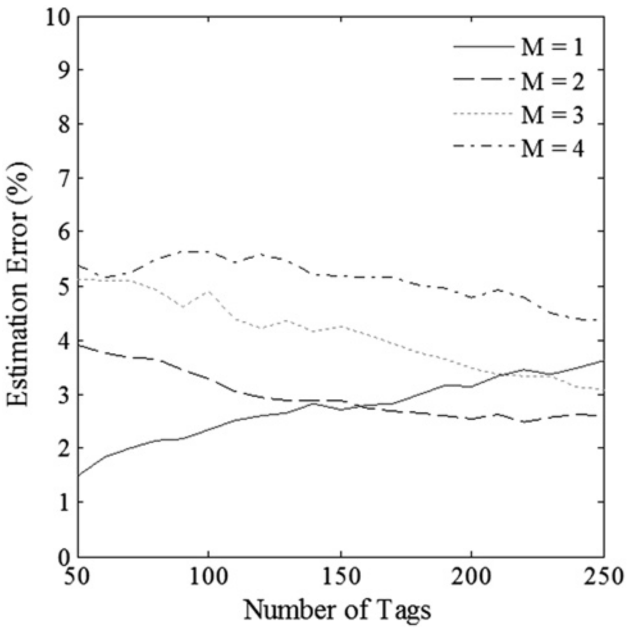


**Fig. 4** *Difference between the real number of tags and the estimated number of tags (expressed in percentage) when L was set to 128*

**Table 2** Influence of the tag estimation error on the maximum achievable channel efficiency *U* for the different values of *M*

| | $M=1$ | $M=2$ | $M=3$ | $M=4$ |
|---|---|---|---|---|
| *U* with 0% estimation error | 0.3679 | 0.5869 | 0.7260 | 0.8167 |
| *U* with 2% estimation error | 0.3678 | 0.5868 | 0.7259 | 0.8165 |
| *U* with 4% estimation error | 0.3676 | 0.5864 | 0.7254 | 0.8160 |
| *U* with 6% estimation error | 0.3673 | 0.5858 | 0.7246 | 0.8151 |

even the maximum tag estimation error reported above (i.e. 6%) has very minimal ( < 1%) impact on the channel efficiency for all considered values of *M*.

## 5 Conclusion

In this paper, we have made a cross-layer analysis (involving physical and link layers) of an RFID system considering jointly the potential of recent physical layer enhancements in MPR technology and a base-line link-layer tag anti-collision protocol. From the analysis, we have derived an optimal operating criterion which relates a link-layer parameter (frame length) to a physical layer parameter (MPR-capability) and the tag population in the RFID system. The criterion is optimal in the sense that it maximises the channel usage efficiency of the RFID system thereby maximising tag reading rate. Implication of the maximised tag reading rate on the performance of RFID applications is straightforward. It reduces tag identification delay which is strongly desirable in many RFID applications, especially the delay-intolerant applications such as autonomous supermarket checkout process or toll collection over highways. To make RFID system with MPR-capable reader able to adaptively operate close to the optimal criterion, we provided further a MAP-based tag estimation method.

From rigorous computer simulation, we show that when an RFID reader is provided with a MPR capability, the tag reading rate of the reader significantly increases. The increase in the read rate depends on the order of MPR capability. In other words, with increase in the order of MPR capability, increased number of tags can be identified for a given target identification delay using the same channel resource.

## 6 Acknowledgment

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (no. 2012R1A1A2006327) and by NRF grant funded by MEST (no. 2010-0018116).

## 7 References

1 Landt, J.: 'The history of RFID', *IEEE Potentials*, 2005, **24**, (4), pp. 8–11
2 Rao, K.V.: 'An overview of back scattered radio frequency identification system'. Proc. Asia Pacific Microwave Conference, November 2009, pp. 746–749
3 Klair, D.K., Chin, K., Raad, R.: 'A survey and tutorial of RFID anti-collision protocols', *IEEE Commun. Surv. Tutor.*, 2010, **12**, (3), pp. 400–421
4 Myung, J., Lee, W., Srivastava, J., Shin, T.K.: 'Tag-splitting: adaptive collision arbitration protocols for RFID tag identification', *IEEE Trans. Parallel Distrib. Syst.*, 2007, **18**, (6), pp. 763–775
5 Chen, W.: 'Performance comparison of binary search tree and framed aloha algorithms for RFID anti-collision', *IEICE Trans. Commun.*, 2008, **E91**, (4), pp. 1168–1171
6 Vogt, H.: 'Efficient object identification with passive RFID Tags'. Proc. Int. Conf. Pervasive Computing, August 2002, pp. 98–113
7 Khandelwal, G., Yener, A., Chen, M.: 'OPT: optimal protocol tree for efficient tag identification in dense RFID systems'. Proc. IEEE Int. Conf. Communications, June 2006, vol. 1, pp. 128–133
8 Mindikoglu, A.F., Veen, A.: 'Separation of overlapping RFID signals by antenna arrays'. Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, September 2008, pp. 2737–2740
9 Dacuna, J., Melia-Segui, J., Pous, R.: 'Multi-tag spatial multiplexing in UHF RFID systems', *IEICE Electron. Express*, 2009, **9**, (21), pp. 1701–1706
10 Lee, J., Kwon, T.: 'RFID anti-collision algorithms with multi-packet reception', in Karmakar, N.S. (Ed.): 'Handbook of smart antennas for RFID systems' (Wiley, 2004, 1st edn.), pp. 573–586
11 Kim, S., Kwack, S., Choi, S., Lee, B.G.: 'Enhanced collision arbitration protocol utilizing multiple antennas in RFID systems'. Proc. Asia-Pacific Conf. Communications, October 2011, pp. 925–929
12 Chen, W.: 'An accurate tag estimate method for improving the performance of an RFID anticollision algorithm based on dynamic frame length ALOHA', *IEEE Trans. Autom. Sci. Eng.*, 2009, **6**, (1), pp. 9–15
13 Vales-Alonso, J., Bueno-Delgado, V., Egea-Lopez, E., Gonzalez-Castano, F.J.: 'Multiframe maximum-likelihood tag estimation for RFID anticollision protocols', *IEEE Ind. Inf.*, 2011, **7**, (3), pp. 487–496
14 Choi, J., Lee, C.: 'Maximum a posteriori (MAP) based tag estimation method for dynamic framed-slotted ALOHA(DFSA) in RFID systems', *Eurasip J. Wirel. Commun. Netw.*, 2012, **2012**, (1), pp. 1–12
15 Shakiba, M., Singh, M.J., Sundararajan, E., Zavvari, A., Islam, M.T.: 'Extending birthday paradox theory to estimate the number of tags in RFID systems', *PLoS One*, 2014, **9**, (4), pp. 1–11

## 8 Appendix 1

Using the principle of mathematical induction, in this appendix we prove that (8) can be simplified to (9). In other words, for any positive integer *M*, we formally prove

$$\sum_{m=1}^{M}\left(\frac{1}{m!}\left(\frac{n}{L}\right)^{m}\left(\frac{n}{L}-m\right)\right)=\frac{(n/L)^{M+1}-M!(n/L)}{M!}$$

*Proof: :* Let us denote the relation in the above equation as $f(M)$. According to the principle of mathematical induction, in order to show that $f(M)$ holds for all positive integers *M*, it is sufficient to establish the following two properties: (i) Base case property: $f(M)$ should hold for the initial value of *M*, and (ii) Induction step property: If $f(M=k)$ is assumed to be hold for arbitrary integer *k*, then $f(k+1)$ should also hold.
For $M=1$, both LHS and RHS of above equation are equal to $(n/L)(n/L-1)$ and thus the first base case property holds. For accessing the second inductive step property, let us first assume that $f(M=k)$ holds, that is

$$\sum_{m=1}^{k}\left(\frac{1}{m!}\left(\frac{n}{L}\right)^{m}\left(\frac{n}{L}-m\right)\right)=\frac{(n/L)^{k+1}-k!(n/L)}{k!}$$

Then, addition of $(1/(k+1)!)(n/L)^{k+1}((n/L)-(k+1))$ on both sides of above equation results in

$$\sum_{m=1}^{k+1}\left(\frac{1}{m!}\left(\frac{n}{L}\right)^{m}\left(\frac{n}{L}-m\right)\right)=\frac{(n/L)^{k+1}-k!(n/L)}{k!}+\frac{1}{(k+1)!}\left(\frac{n}{L}\right)^{k+1}\left(\frac{n}{L}-(k+1)\right)$$

$$=\frac{1}{(k+1)!}\left[(k+1)\left((n/L)^{k+1}-k!(n/L)\right)+\left(\frac{n}{L}\right)^{k+1}\left(\frac{n}{L}-(k+1)\right)\right]$$

$$=\frac{(n/L)^{k+2}-(k+1)!(n/L)}{(k+1)!}$$

which proves $f(k+1)$ also holds. Since both properties are established, $f(M)$ holds for $\forall M$. □