Tag Cardinality Estimation using Expectation-Maximization in ALOHA-based RFID Systems with Capture Effect and Detection Error

Chuyen T. Nguyen, Van-Dinh Nguyen, and Anh T. Pham

***Abstract*—Tag cardinality estimation is one of the most crucial issues in radio frequency identification (RFID) technology. The issue, however, usually faces with challenges in wireless fading environments due to the presence of the so-called capture effect (CE) and detection error (DE). The aim of this study is to provide an efficient and accurate estimation method to cope with the CE and DE using Expectation-Maximization algorithm and the standard Aloha-based protocol. We show that the proposed method gives more accurate estimates than a conventional one. Thanks to this fact, the Aloha frame size used for the tag identification process can also be optimally selected so that the identification efficiency can be improved. Computer simulations are presented to confirm the merit of the proposed method.**

***Index Terms*—RFID, Aloha, capture effect, detection error, EM algorithm, estimation.**

I. Introduction

TAG cardinality estimation holds a crucial task in Radio Frequency Identification (RFID) technology with many practical applications such as intelligent transportation, indoor stadium, and warehouse systems. The task has been investigated in several previous works [1], [2] with a frame slotted Aloha (FSA) protocol, which is originally and standardly used to detect RFID tags' Identity (ID) [3]. In those works, tags randomly transmit their IDs in a frame of time slots. Then, observations of the number of responses in each slot, i.e, *no response*, *one response*, and *multiple responses* [4], can be utilized for the tag cardinality estimation. The tags' ID identification process can be significantly improved with an accurate estimate of the cardinality.

On the other hand, under effects of wireless channel impairments, the observations of time slots may not accurately reflect the real number of responses. Indeed, due to the channel fading phenomenon, a tag might be detected with a probability in a multiple-response slot, which is well-known as the capture effect (CE) [5]. The observed state of the slot, in this case, is assumed to be *singleton* to differentiate from one-response. In addition, a tag might not be detected with a probability in the corresponding one-response slot, which is referred to as the detection error (DE) [6]. Similarly, the observed state of the slot is called *empty*, while in other cases, the state is observed as *collision*. These phenomena has been extensively studied in the literature of RFID both in theoretical and experimental aspects [6]. They are usually hidden from the reader, and therefore, affect the estimation accuracy of conventional methods.

Several works have been proposed to deal with the cardinality estimation in the presence of the CE [7]-[9]. The method in [7], i.e., capture-aware backlog estimation (CMEBE), estimates the tag cardinality and the CE probability by minimizing the norm-2 distance between theoretical and observed vectors of empty, singleton and collision slots. In [8], they are found by using Bayesian approach. Also, in [9], the capture probability is analyzed in a more accurate approach by considering the number of contention tags in a time slot and physical layer parameters. Thanks to the approach, a closed-form solution of the optimal frame length is found, which then improves the identification performance. The key limitation of those works, nevertheless, is that the DE is completely ignored. On the other hand, the cardinality estimation in the presence of both the CE and DE was recently studied in [10] based on the maximum likelihood (ML) approach. In the approach, an approximation of the likelihood function of the tag cardinality, CE and DE probabilities, for given observations of slots is determined. Nevertheless, to maximize the likelihood function, the method adopted an exhaustive search algorithm to check all possible values of the tag cardinality and the probabilities. This approach thus resulted in a very high computational complexity, and affected the overall performance of the identification process. Although, the complexity could be reduced with a simple transmission model such as flat Rayleigh fading where a deterministic relation between CE and DE probabilities could be obtained [10], it would be much more challenging in practical ones.

In this work, we propose a new method employing the Expectation-Maximization (EM) algorithm [11] and FSA protocol to efficiently and accurately estimate the tag cardinality in presence of both CE and DE. The method includes iterative estimation rounds. In each round, the cardinality is first estimated by ML approach given expected values of hidden data/observations caused by the CE and DE. The CE and DE probabilities can then be found in closed-forms for the given estimate, which significantly reduces the computational complexity in comparison with the method in [10]. Simulation results also confirm the effectiveness of our proposed method compared to the conventional methods.

This work is partly supported by the JSPS Kakenhi project 18K11269.

Chuyen T. Nguyen is with Hanoi University of Science and Technology, Hanoi, Vietnam. E-mail: chuyen.nguyenthanh@hust.edu.vn.

Van-Dinh Nguyen is with Soongsil University, Seoul, South Korea. E-mail: nguyenvandinh@ssu.ac.kr

Anh T. Pham is with the University of Aizu, Japan. E-mail: pham@u-aizu.ac.jp

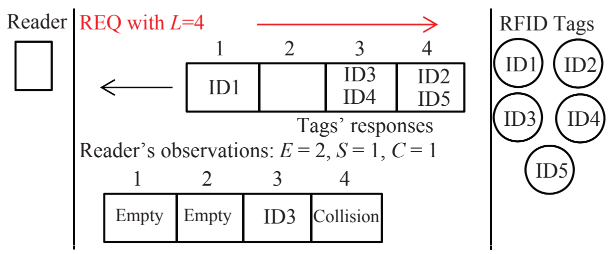

Fig. 1. A reading round of Aloha-based identification operation.

II. Protocol Description and Proposed Method

A. Protocol Description

Our considered RFID system consists of a reader and n tags in its communication range. The FSA protocol is implemented, in which the reader first broadcasts a request consisting of a time slotted frame size of L. Then, each tag responds to the reader by its identity (ID) randomly in one of the L slots. The reader tries to estimate the tag cardinality based on the observed numbers of empty, collision, and singleton slots, denoted as E, S, and C, respectively [4].

Practically, the DE and CE may happen in any slot with tags' responses. To focus on the tag cardinality estimation, we use the similar model as in [10], in which, each one-response slot is assumedly detected as an empty one with an average DE probability of β, while a multiple-response slot is recognized as singleton with an average CE probability of α. The inaccurate cardinality estimation problem due to the CE and DE can be illustrated as in Fig. 1, which presents a reading round in an Aloha-based RFID system with a reader and 5 tags. In this example, a request $L = 4$ is initially used and the correct observation should be $E = 1$, $S = 1$, $C = 2$. Nevertheless, due to the impact of the CE and DE, tag 3 is detected in slot 3 (CE), while tag 1 is not detected in slot 1 (DE). The reader therefore has a wrong observation with $E = 2$, $S = 1$, $C = 1$; and consequently, it will inaccurately estimate the cardinality of tags in the system.

B. Proposed Method

In our method, we first denote p_E, p_S and p_C as the average probabilities of observing an empty, a singleton and a collision slot, respectively; and they can be expressed as

$$p_E \approx p_0 + \beta p_1,\ p_S \approx (1-\beta)p_1 + \alpha p_2,\ p_C \approx (1-\alpha)p_2, \quad (1)$$

where p_0, p_1 or p_2 is a probability that a slot is, respectively, no-response, one-response or multiple-response, i.e., $p_0 \approx \left(1-\frac{1}{L}\right)^n$, $p_1 \approx \frac{n}{L}\left(1-\frac{1}{L}\right)^{n-1}$, and $p_2 \approx 1 - p_0 - p_1$ [10]. It is noted in (1) that the DE is assumed to not happen in multiple-response slots due to signal diversity, which has also been validated in [10] under the assumption of a simple Rayleigh fading channel model. More practical models of the DE, CE, and status of each slot should be investigated in future works.

Then, the estimates of n, α, and β denoted by $\hat{n}$, $\hat{\alpha}$ and $\hat{\beta}$, respectively, can be approximately found as

$$\begin{aligned}\left(\hat{n},\hat{\alpha},\hat{\beta}\right) &= \arg \max_{n\in\mathbb{N},\ \alpha,\ \beta\in[0,1]} f(E,S,C|n,\alpha,\beta) \\ &\approx \arg \max_{n\in\mathbb{N},\ \alpha,\ \beta\in[0,1]} \frac{(E+S+C)!}{E!S!C!} p_E^E p_S^S p_C^C, \quad (2)\end{aligned}$$

where $f(E,S,C|n,\alpha,\beta)$ is the likelihood function of n, α and β, given E, S and C. It should be also noted in (2) that the likelihood function has been approximately modeled as a multinomial distribution with L repeated independent trials, where each trial has one of three outcomes: empty, singleton, or collision. Although this approximation does not reflect the exact likelihood function [12], it results in accurate estimates as ML ones especially when L is large, which has been numerically validated in [2] and [10]. Since there is no guarantee on the convergence of the (pseudo) likelihood function, it could be possible to solve (2) by an exhaustive search algorithm or finding a deterministic relation between the two probabilities in an assumed fading channel model [10]. Nevertheless, while the former costs a very high computational complexity, the latter is difficult to obtain for practical fading models.

In what follows, we utilize the EM approach to find the estimates of the tag cardinality and the probabilities. EM is an iterative estimation algorithm, which is especially useful when necessary information/data is hidden/missing. In our model, the hidden data is the number of one-response and multiple-response slots observed as empty and singleton ones denoted by S_1 and C_1, respectively. In particular, each EM iteration includes two steps, namely E-step and M-step. In E-step, expected values of the hidden data S_1 and C_1, which are respectively denoted by $\overline{S}_1$ and $\overline{C}_1$, are estimated. In M-step, the estimates of n, α and β are found for a given complete (observed and hidden) data, i.e., $\left[E,S,C;\overline{S}_1,\overline{C}_1\right]$. This estimation process is repeated until *convergence* that is defined as

$$\epsilon = \sqrt{(\hat{n}^{\{r\}}-\hat{n}^{\{r-1\}})^2 + (\hat{\alpha}^{\{r\}}-\hat{\alpha}^{\{r-1\}})^2 + (\hat{\beta}^{\{r\}}-\hat{\beta}^{\{r-1\}})^2} \le \epsilon_p, \quad (3)$$

where ϵ is the norm-2 distance between two estimated vectors of n, α and β at two consecutive iterations. $\hat{n}^{\{r\}}$, $\hat{\alpha}^{\{r\}}$ and $\hat{\beta}^{\{r\}}$ are, respectively, the estimates of n, α and β at the r-th iteration. ϵ_p is a predefined constant. The two steps are described in details as

1) E-step: From (1), $\overline{S}_1$ and $\overline{C}_1$ are easily found as follows

$$\overline{S}_1 = \beta n\left(1-\frac{1}{L}\right)^{n-1},\ \overline{C}_1 = \alpha L\left(1-\left(1-\frac{1}{L}\right)^n - \frac{n}{L}\left(1-\frac{1}{L}\right)^{n-1}\right). \quad (4)$$

It is noted that, values of n, α and β in (4) are taken from the following M-step, while they can be initially set as $S+2C$, 0.5, and 0.5, respectively.

2) M-step: Given the complete data $\left[E,S,C;\overline{S}_1,\overline{C}_1\right]$, the likelihood function of n, α and β denoted by $f\left(E,S,C,\overline{S}_1,\overline{C}_1|n,\alpha,\beta\right)$ is written as

$$\begin{aligned}f\left(E,S,C,\overline{S}_1,\overline{C}_1|n,\alpha,\beta\right) &= \frac{(E+S+C)!}{(E-\overline{S}_1)!\overline{S}_1!(S-\overline{C}_1)!\overline{C}_1!C!} \\ &\times p_0^{E-\overline{S}_1} p_{01}^{\overline{S}_1} p_{10}^{S-\overline{C}_1} p_{12}^{\overline{C}_1} p_C^C, \quad (5)\end{aligned}$$

Algorithm 1 EM estimation algorithm

Initialization: Generate L for the first request, and observe E, S and C. Set $n = S + 2C$, $\alpha = 0,5$, $\beta = 0.5$
repeat
  E-step:
    $\overline{S}_1 = \beta n\left(1-\frac{1}{L}\right)^{n-1}$
    $\overline{C}_1 = \alpha L\left(1-\left(1-\frac{1}{L}\right)^{n}-\frac{n}{L}\left(1-\frac{1}{L}\right)^{n-1}\right).$
  M-step:
    $\hat{n} = \arg\max_{n\in\mathbb{N}} g_1(n)$
    $\hat{\alpha} = \frac{\overline{C}_1}{\overline{C}_1+C}$, $\hat{\beta} = \frac{\overline{S}_1}{S+\overline{S}_1-C}$,
until Convergence

where $p_{01} = \beta p_1$, $p_{10} = (1-\beta)p_1$, and $p_{12} = \alpha p_2$. Since E, S, C, $\overline{S}_1$ and $\overline{C}_1$ are constants, the estimates can be found by maximizing the function $g(n,\alpha,\beta) = \ln\left(p_0^{E-\overline{S}_1} p_{01}^{\overline{S}_1} p_{10}^{S-\overline{C}_1} p_{12}^{\overline{C}_1} p_C^{C}\right)$, which can be re-written as

$$g(n,\alpha,\beta) = g_1(n) + g_2(\alpha) + g_3(\beta), \tag{6}$$

where $g_1(n) = (E-\overline{S}_1)\ln(p_0) + (\overline{S}_1 + S - \overline{C}_1)\ln(p_1) + (\overline{C}_1 + C)\ln(p_2)$, $g_2(\alpha) = \overline{C}_1\ln(\alpha) + C\ln(1-\alpha)$, $g_3(\beta) = \overline{S}_1\ln(\beta) + (S-\overline{C}_1)\ln(1-\beta)$. Since n, α and β are independent, the estimates are easily obtained by maximizing $g_1(n)$, $g_2(\alpha)$, and $g_3(\beta)$ with respect to n, α, and β, respectively, i.e.,

$$\hat{n} = \arg\max_{n\in\mathbb{N}} g_1(n), \tag{7}$$

$$\hat{\alpha} = \frac{\overline{C}_1}{\overline{C}_1 + C},\ \hat{\beta} = \frac{\overline{S}_1}{S + \overline{S}_1 - C}. \tag{8}$$

It is also noted that (7) can be efficiently solved by the Chen's method [2] where $g_1(n)$ is numerically proven to be converged. We summarize the EM estimation iterations in Algorithm 1. The initial value of n is selected as a lower bound after observing E, S and C. Also, the initial values of α and β can be arbitrary in (0,1). Nevertheless, since we have no knowledge of α and β a priori, they are both initially set as 0.5.

The Aloha frame size can be selected based on the above estimates to improve the performance of tag identification. In particular, the size is found by maximizing the system efficiency which is defined as the average number of detected tags per time slot and denoted by η. Here, η is written as

$$\eta=(1-\beta)\frac{n}{L}\left(1-\frac{1}{L}\right)^{n-1}+\alpha\left(1-\left(1-\frac{1}{L}\right)^{n}-\frac{n}{L}\left(1-\frac{1}{L}\right)^{n-1}\right). \tag{9}$$

By letting the differentiation of η in (9) with respect to L be zero (assuming the continuous relaxation of L), we can find the optimal frame size denoted by L_{opt} as

$$L_{\text{opt}} = n - \frac{\alpha(n-1)}{1-\beta}. \tag{10}$$

In other words, by substituting (7) and (8) into (10), the frame size could be optimally selected.

III. Numerical Results and Discussions

In this section, we evaluate the performance of the proposed estimation method via computer simulations. The frame size L and the predetermined constant ϵ_p are set by 256 and 10^{-4}, respectively. The simulation results are obtained by Monte Carlo method with the number of simulation runs $R = 1000$, and are also compared with those of the conventional CMEBE and Bayesian methods.

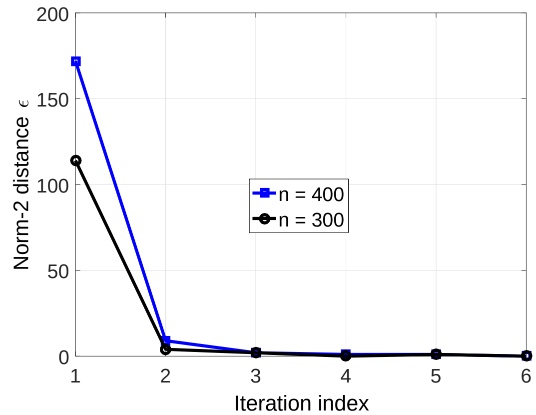

Fig. 2. Convergence behavior of the proposed algorithm with different number of tags (n), for $L = 256$, $\alpha = 0.3$ and $\beta = 0.3$.

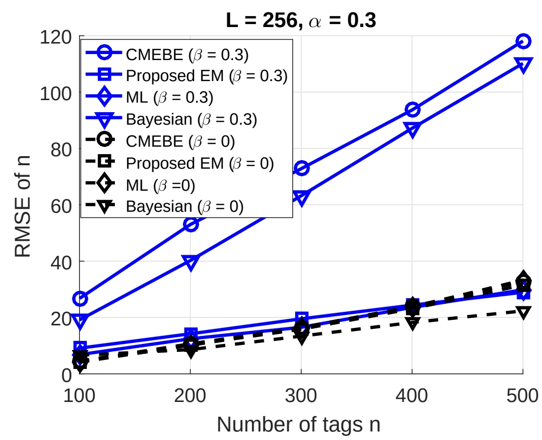

Fig. 3. RMSE of n, for $L = 256$, $\alpha = 0.3$.

First, we investigate the typical convergence behavior of the proposed algorithm by plotting the norm-2 distance ϵ with different numbers of tags $\{n = 300,\ 400\}$ in Fig. 2, for $\alpha = 0.3$ and $\beta = 0.3$. It is seen that the proposed method converges very fast, within only a few iterations in all cases.

Next, Figs. 3 and 4 show the root mean square errors (RMSEs) of n and α (denoted by e_n and e_α, respectively) of the CMEBE, the ML-based method in [10], Bayesian estimate [8], and the proposed method for $L = 256$, $\beta = 0$ or 0.3. Here, the RMSEs are defined as

$$\mathrm{e}_n = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{n}_i - n)^2},\ \mathrm{e}_\alpha = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{\alpha}_i - \alpha)^2}, \tag{11}$$

where $\hat{n}_i$ and $\hat{\alpha}_i$ are, respectively, the i-th estimates of n and α. It should be noted that the estimates in [10] are obtained by maximizing the approximated likelihood function in (2) with an exhaustive search algorithm over all possible values of n, α,

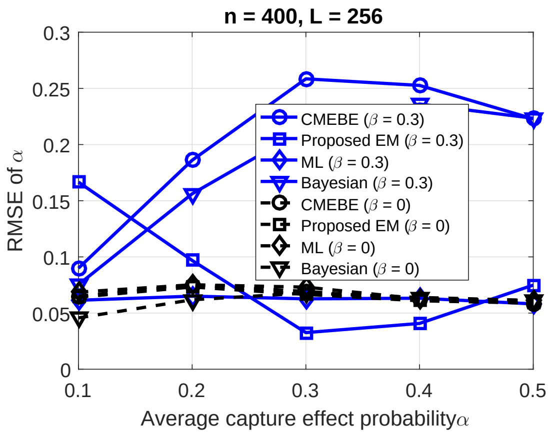

Fig. 4. RMSE of α, for $L = 256$.

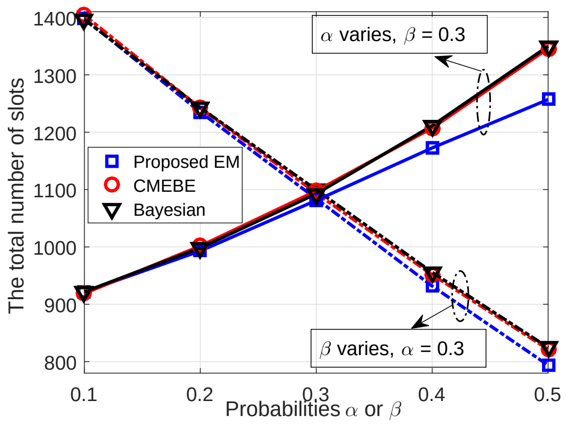

Fig. 5. The total number of slots used to detect $n = 400$ tags with respect to the DE (dash lines) or CE (solid lines) probability.

and β. Therefore, [10] can be approximately considered as a lower bound of all considered methods. We can see that while our method can be comparable with CMEBE and Bayesian when $\beta = 0$, it significantly outperforms them in term of the estimation accuracy when $\beta > 0$. This is because only the CE has been taken into account in CMEBE and Bayesian methods. Nevertheless, the performance of EM-based algorithms greatly depends on the initial values of estimated parameters. Indeed, it is seen in Fig. 4 the degraded performance of the proposed method for small values of α ($\alpha < 0.15$). This fact is also observed for $\alpha = 0.3,\ 0.4$ where the performance of the proposed method is even better than that of [10].

We now provide the worst-case per-iteration complexity analysis of Algorithm 1 and compare to that of [10]. Recall that the per-iteration complexity of the method presented in [10] is $O(n^3)$. The complexity of Algorithm 1 is mostly due to solving (7) (i.e., step 7 of Algorithm 1), which requires the complexity of $O(n)$. This is to say, given the same convergence condition as in (3), the proposed algorithm requires significantly lower complexity, compared to that of [10], especially when n is large.

Finally, we plot the total number of slots used to detect $n = 400$ tags with respect to different values of α (β is set by 0.3) or β (α is set by 0.3) in Fig. 5 to see the impact of estimation methods on identification performance. Here, the frame sizes of the methods are optimally determined by (10) in which β is set by 0 for CMEBE and Bayesian. We can see that the consumed time slots is proportional to α for given β, while inversely proportional to β for given α. The reason is that more tags are detected in multiple-response slots, but more tags are also hidden in one-response slots. Nevertheless, in the both cases, the proposed method takes a smaller number of time slots than conventional ones, especially when the DE and (or) CE are more significant. This is because both the DE and CE have been considered in our estimation scheme thanks to the EM approach.

IV. Conclusion

This paper investigated the issue of tag cardinality estimation with FSA protocol in RFID systems considering impacts of both CE and DE. The EM approach was utilized to iteratively estimate the tag cardinality, the CE, and DE probabilities. Computer simulations confirmed that the proposed method was guaranteed to converge after only a few iterations and provided more accurate estimates than that of the conventional methods. The proposed method was also proven to improve the efficiency of the identification process.

References

[1] W. Gong, J. Liu, K. Liu, and Y. Liu, "Toward more rigorous and practical cardinality estimation for large-scale RFID systems," *IEEE/ACM Tran. on Networking,* vol PP, no. 99, pp. 1-12, Dec. 2016.

[2] W.T. Chen, "An accurate tag estimate method for improving the performance of an RFID anti-collision algorithm based on dynamic frame length aloha," *IEEE Tran. on Automation Science and Engineering*, vol. 6, no. 1, pp. 9–15, Jan. 2009.

[3] "EPC Radio-frequency identity protocols class-1 generation-2 uhf RFID protocol for communications at 860 Mhz-960 Mhz version 1.2.0," available at http://www.gs1.org/sites/default/files/docs/epc/uhfc1g2_1_2_0-standard-20080511.pdf, accessed May 8, 2016.

[4] C. Qian, H. Ngan, Y. Liu, and L. M. Ni, "Cardinality estimation for large-scale RFID systems," *IEEE Tran. on Parallel and Distributed Systems*, vol. 22, no. 9, pp. 1441–1454, Sep. 2011.

[5] H. Salah, H. A. Ahmed, J Robert, and A. Heuberger, "A time and capture probability aware closed form frame slotted aloha frame length optimization," *IEEE Commun. Letters*, vol. 19, no. 11, pp. 2009–2012, Nov. 2015.

[6] C.T. Nguyen, A.-T. H. Bui, V.-D. Nguyen, and A. T. Pham, "Modified tree-based identification protocols for solving hidden-tag problem in RFID systems over fading channels," *IET Commun.*, vol. 11, no. 7, pp. 1132–1142, May 2017.

[7] B. Li and J. Wang, "Efficient anti-collision algorithm utilizing the capture effect for ISO 18000-6C RFID protocol," *IEEE Commun. Letters*, vol. 15, no. 3, pp. 352–354, March 2011.

[8] H. Wu, Y. Wang, and Y. Zeng, "Capture-aware bayesian RFID tag estimate for large-scale identification," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 1, pp. 119–127, Jan. 2018.

[9] H.A. Ahmed, H. Salad, J. Robert and A. Heuberger, "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, March 2011.

[10] C.T. Nguyen, K. Hayashi, M. Kaneko, and H. Sakai, "Maximum likelihood approach for RFID tag cardinality estimation under capture effect and detection errors," *IEICE Tran. on Commun.*, vol. E96-B, no. 5, pp. 1122–1129, May 2013.

[11] G.J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, 2nd ed., Wiley, 2008.

[12] E. Vahedi, V.W.W. Hong, I.F. Blake, and R.K. Ward, "Probabilistic analysis and correction of Chen's tag estimate method," *IEEE Tran. on Automation Sci. and Eng.*, vol. 8, no. 3, pp. 659–663, July 2011.