SCIENCE CHINA
Information Sciences

• **Supplementary File** •

On Efficient Key Tag Writing in RFID-Enabled IoT

Pengfei Zhang[1], Hao Liu[1] & Jihong Yu[1*]

[1]*Beijing Institute of Technology, Beijing* 100081, *China*

Appendix A Proof of Theorem 1

Theorem 1. When the time efficiency η_j of round j reaches its maximum value, the corresponding f_j must fall into the interval $[1, \frac{(K_j+n')^2}{0.5K_j+n'}]$ where $K_j \leqslant k$ is the number of active key tags at the beginning of round j, and $n' \leqslant n$ is the number of active normal tags after Phase 1.

Before the proof of Theorem 1, we first state three auxiliary lemmas.

Lemma 1. Given the original vector size f_j at the round j, the expected number of the pure slots in the vector should be

$$S_j = f_j \sum_{i=1}^{G'} \big(1-(1-\frac{1}{f_j})^{g_{i,j}}\big)(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}. \tag{A1}$$

Proof. Consider an arbitrary group i. A slot is pure with respect to this group if and only if it is chosen by only key tags of this group. That is to say, at least one tag of a group selects this slot and the tags of the other groups do not. The probability of this event is $\big(1-(1-\frac{1}{f_j})^{g_{i,j}}\big)(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}$. Since there are G' groups and f_j slots, we have the final result of S_j as shown in (A1).

Lemma 2. Given the original vector size f_j at the round j, the expected number of key tags written within this round should be

$$w_j = \sum_{i=1}^{G'} g_{i,j}(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}. \tag{A2}$$

Proof. We also consider an arbitrary key group j. The probability that a slot is chosen by only x key tags of this group is $(\frac{1}{f_j})^x(1-\frac{1}{f_j})^{g_{i,j}-x}\cdot(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}$. As there are $\binom{g_{i,j}}{x}$ kinds of possible combinations and x ranges from 0 to $g_{i,j}$, we have the expression $\sum_{x=0}^{g_{i,j}}\binom{g_{i,j}}{x}(\frac{1}{f_j})^x(1-\frac{1}{f_j})^{g_{i,j}-x}\cdot(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}$. Following the properties of the binomial distribution, we have $\frac{g_{i,j}}{f_j}(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}$. In addition, there are G' groups and f_j slot, we can thus get the result of the lemma and complete the proof.

Lemma 3. Given the original vector size f_j at the round j, the expected length of the compressed filter can be derived as follows where log denotes the logarithm to the base 2:

$$f'_j = S_j \log\left(1+\frac{1}{\sum_{i=1}^{G'}\big(1-(1-\frac{1}{f_j})^{g_{i,j}}\big)(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}}\right). \tag{A3}$$

Proof. Because the compressive filter uses the number of consecutive zeros as a hint to indicate the positions of pure slots (i.e., '1's), the filter size is decided by the number of '1's and the length of the longest consecutive zeros. To make analysis feasible, we use its expectation to approximate the exact value. We have obtained the expected number of pure slots, the compressive filter length can be derived as $S_j\log(1+\frac{f_j}{S_j})$. The lemma follows from algebraic operation.

From Lemma 1, Lemma 2 and Lemma 3, η_j can be rewritten as

$$\eta_j = \frac{\sum_{i=1}^{G'} g_{i,j}(1-\frac{1}{f_j})^{-g_{i,j}}}{f_j\sum_{i=1}^{G'}\big(1-(1-\frac{1}{f_j})^{g_{i,j}}\big)(1-\frac{1}{f_j})^{-g_{i,j}}\cdot\left(\log\left(1+\frac{1}{\sum_{i=1}^{G'}\big(1-(1-\frac{1}{f_j})^{g_{i,j}}\big)(1-\frac{1}{f_j})^{K_j+n'-g_{i,j}}}\right)+d\right)t_1}. \tag{A4}$$

* Corresponding author (email: jihong.yu@bit.edu.cn)

As η_j is a function of f_j, we should carefully configure f_j to maximize η_j. Yet, it is challenging to directly derive the optimum f_j, we instead find a range of f_j so that f_j out of this range will lead to a smaller η_j, and thus the optimum f_j must falls into this range.

Proof. **Proof of Theorem 1**. We now start to prove Theorem 1. It is difficult to derive the exact f_j that maximizes the time efficiency η_j, thus a feasible method is to find a rang where the optimum f_j falls. The key of the proof lies in finding such an upper bound f_j^* of f_j that $\eta_j(\hat{f}_j) \leqslant \eta_j(f_j)$ for any $\hat{f}_j \geqslant f_j^*$. If we can find such f_j^*, then the optimum f_j must falls into $[1, f_j^*]$.

To this end, we first revise (A4) and get

$$\frac{1}{\eta_j} = \frac{t_1 f_j \sum_{i=1}^{G'} \left(1-(1-\frac{1}{f_j})^{g_{i,j}}\right)(1-\frac{1}{f_j})^{-g_{i,j}}}{\sum_{i=1}^{G'} g_{i,j}(1-\frac{1}{f_j})^{-g_{i,j}}} \times \left(\log\left(1+\frac{1}{\sum_{i=1}^{G'}\left(1-(1-\frac{1}{f_j})^{g_{i,j}}\right)(1-\frac{1}{f_j})^{K_i+n'-g_{i,j}}}\right)+d\right).$$

Then the objective changes to find f_j^* so that $\frac{1}{\eta_j}$ is nondecreasing for all $f_j \geqslant f_j^*$. It is easy to check through derivative analysis that $\log(1+\frac{1}{\sum_{i=1}^{G'}\left(1-(1-\frac{1}{f_j})^{g_{i,j}}\right)(1-\frac{1}{f_j})^{K_i+n'-g_{i,j}}})+d$ is increasing with f_j if $f_j \geqslant \frac{(K_j+n')^2}{0.5K_j+n'} \triangleq f_j^*$.

The key left is to prove that the left part of the product sign also has the same property. Take the logarithm of this part and derive its derivative, we have

$$\begin{aligned}\frac{\mathbf{d}\ln\frac{1}{\eta_j}}{\mathbf{d}f_j}\cdot\frac{1}{\ln t_1} &= \frac{\sum_{i=1}^{G'}(1-\frac{1}{f_j})^{-g_{i,j}}(1-(1-\frac{1}{f_j})^{g_{i,j}}-\frac{g}{f-1})}{\sum_{i=1}^{G'} f_j((1-\frac{1}{f_j})^{-g_{i,j}}-1)} + \frac{\sum_{i=1}^{G'}\frac{g_{i,j}^2(1-\frac{1}{f_j})^{-g_{i,j}}}{f_j(f_j-1)}}{\sum_{i=1}^{G'} g_{i,j}(1-\frac{1}{f_j})^{-g_{i,j}}} \\ &> \frac{1}{f_j(f_j-1)}\cdot\left(\frac{\sum_{i=1}^{G'}1.5g_{i,j}(1-\frac{1}{f_j})^{-g_{i,j}}}{\sum_{i=1}^{G'} f_j((1-\frac{1}{f_j})^{-g_{i,j}}-1)} + \frac{\sum_{i=1}^{G'} g_{i,j}^2(1-\frac{1}{f_j})^{-g_{i,j}}}{\sum_{i=1}^{G'} g_{i,j}(1-\frac{1}{f_j})^{-g_{i,j}}}\right).\end{aligned} \tag{A5}$$

The inequality is established by the fact that $1-xz < (1-x)^z < 1-xz+0.5zx^2$. Let $y_j = (1-\frac{1}{f_j})^{-g_{i,j}}$, we have

$$\begin{aligned}&\frac{\mathbf{d}\ln\frac{1}{\eta_j}}{\mathbf{d}f_j}\cdot\frac{f_j(f_j-1)}{\ln t_1} > \frac{\sum_{i=1}^{G'}1.5g_{i,j}y_j}{\sum_{i=1}^{G'} f_j(y_j-1)} + \frac{\sum_{i=1}^{G'} g_{i,j}^2 y_j}{\sum_{i=1}^{G'} g_{i,j}y_j} \\ &> \frac{\sum_{i=1}^{G'}(f_j-1.5)g_{i,j}^2y_j^2 - G'f_j\sum_{i=1}^{G'} g_{i,j}^2 y_j + \sum_{i=1}^{G'} f_j g_{i,j}^2 y_j \sum_{r\neq j}^{G'} y_r(1-1.5\frac{g_{i,r}}{f_j g_{i,j}})}{(\sum_{i=1}^{G'} f_j(y_j-1))(\sum_{i=1}^{G'} g_{i,j}y_j)} > 0.\end{aligned} \tag{A6}$$

The result holds due to the fact that $y_j > 1$ and $f_j \geqslant 3$ and $y_r(1-\frac{g_{i,r}}{f_j g_{i,j}}) > 1$ when $g_{i,j} > 1.5$. Note that $g_{i,j} > 1.5$ is true for a large-scale system. When $g_{i,j} = 1$, we can treat this tag of group i as a normal tag and poll it with its group data at last. As a consequence, $\frac{1}{\eta_j}$ is increasing with respect to f_j when $f_j \geqslant f_j^*$. That is to say, η_j is decreasing with f_j when $f_j \geqslant f_j^*$, and we can thus find the optimum f_j within $[1, f_j^*]$.

Appendix B Configuration Rule of h

We search the optimal value of h with the following offline steps: 1) We set an upper bound of h to $\bar{h}$ establishing $n\cdot 0.5^h \leqslant 1$, and choose an $h \in [0, \hat{h}]$. 2) We then find the optimum f_j for each round. 3) We compute w_i and T_2, j, and take the remaining key tags as input to the computation of the next round. 4) We repeat step 2) and 3) until $k - \sum_{j=1}^{R} w_j = 0$, and compute the overall time containing T_1 and T_2. 5) We repeat step 2) and 3) and 4), and fix h to the one minimizing the overall time.

Appendix C Simulation Settings

In the experiments, we use the communication parameters specified in the EPCglobal Gen-2 standard. Specifically, the Tari (Type A reference interval) is set to 25μs corresponding the rate of 40 kbps from the reader to tags. Consequently, it takes 25 μs for the reader to transmit one bit, so we have t_1=25 μs. Moreover, we assume key group data to be written to its member is its group ID and set its length as $\lceil \log_2 G \rceil$ bits where G is the number of key tag groups in the system. Correspondingly, we have group data transmission time 25$*\lceil \log G \rceil$. In addition, we consider the time interval of 302 μs between any two consecutive communications between the reader and tags in the computation of the execution time.