Joint Optimization of Data Freshness and Fidelity for Selection Combining-Based Transmissions

Zhengchuan Chen¹,²,*, Mingjun Xu¹, Min Wang³ and Yunjian Jia¹,*

¹ School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China; xmj@cqu.edu.cn
² State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China
³ School of Optoelectronics Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; wangm@cupt.edu.cn
* Correspondence: czc@cqu.edu.cn (Z.C.); yunjian@cqu.edu.cn (Y.J.)

Abstract: Motivated by big data applications in the Internet of Things (IoT), abundant information arrives at the fusion center (FC) waiting to be processed. It is of great significance to ensure data freshness and fidelity simultaneously. We consider a wireless sensor network (WSN) where several sensor nodes observe one metric and then transmit the observations to the FC using a selection combining (SC) scheme. We adopt the age of information (AoI) and minimum mean square error (MMSE) metrics to measure the data freshness and fidelity, respectively. Explicit expressions of average AoI and MMSE are derived. After that, we jointly optimize the two metrics by adjusting the number of sensor nodes. A closed-form sub-optimal number of sensor nodes is proposed to achieve the best freshness and fidelity tradeoff with negligible errors. Numerical results show that using the proposed node number designs can effectively improve the freshness and fidelity of the transmitted data.

Keywords: wireless sensor network; age of information; sensor node deployment

1. Introduction

With the emergence of the Internet of Things (IoT) and continuous development of communication techniques, increasingly more facilities are connected to the Internet. At the same time, a flood of information swarms into the fusion center (FC) waiting to be processed.

Wireless sensor network (WSN) is a typical data-driven application in IoT, where multiple sensor nodes are deployed dispersedly aiming at observing the information source. The observed information is then transmitted through wireless channels, reaching the FC for further data gathering and recovery. Typical examples of WSN include forest salinity monitoring, smart parking, intelligent transportation safety supervising, and so on.

In WSN, the freshness of the data is essential for taking prompt actions. For example, in a pedestrian detection system, timely data from sensors in-vehicle and infrastructure is critical to avoid collision and ensure the safety of pedestrians. In smart parking, the occupancy of the parking lot is vital for the users. In environment supervising, fresh data from the sensors is urgently wanted to better monitor the air pollution condition [1].

The above examples show the significance of the freshness of the received information in wireless sensor networks since it is transmitted to operate, supervise, and monitor the systems.

Aside from the freshness, ensuring the fidelity, or accuracy of the received data is also of paramount importance. In wireless communication systems, diversity techniques are applied as useful tools for combating fading effects and improving the quality of the received signal [2–6]. The fundamental diversity combining is the selection combining (SC) scheme, where there is a signal to noise ratio (SNR)-based SC scheme [3] and log-likelihood
ratio (LLR)-based SC scheme [5]. For the LLR-based SC scheme, it selects the branch based on the LLR. For SNR-based SC scheme, it selects the branch that provides the largest SNR from several diverse branches. This approach can effectively improve the fidelity of the ultimately received data.

Based on the background information, we can conclude that when accounting for these two metrics, on the one hand, better freshness can be achieved by waiting for observations from fewer sensor nodes, yet it leads to a decline in the fidelity as there might be insufficient observations; On the other hand, a higher fidelity of transmitted data can be obtained when FC receives more observations from sensors. However, the data may become stale since there may be a long time for receiving enough observations. Motivated by the opposite trend of two metrics affected by the number of sensor nodes, we aim to find the optimal number of sensor nodes by carefully characterizing the data freshness and fidelity using the age of information (AoI) and minimum mean square error (MMSE) metric.

1.1. Contributions

Contributions of this paper can be summarized as follows.

1. We establish a theoretical model that simultaneously takes into account the data freshness and fidelity for jointly optimizing two metrics through adjusting the number of sensor nodes. Explicit characterizations of freshness and fidelity were given by applying the average AoI and MMSE metrics.

2. By endowing the weighting factor to the average AoI and MMSE, the optimal number of sensor nodes optimization problem is formulated. Through appropriate approximations on both metrics, we specify the optimal number of sensor nodes by proposing a sub-optimal solution, which jointly minimizes data freshness and fidelity.

3. Numerical results validate the correctness of the proposed solution and it can achieve the best freshness-fidelity weighted-sum tradeoff with negligible errors.

1.2. Organizations

The remainder of this paper is organized as follows. The related works are demonstrated in Section 2. In Section 3, we introduce the system model. In Section 4, we formulate the problem by introducing the freshness and fidelity metric. Closed-form sub-optimal solution is given based on appropriate approximations in Section 5. Numerical results are provided in Section 6. Finally, conclusions are drawn in Section 7.

2. Related Works

In this Section, we provide a thorough overview of the current academic and state-of-the-art works, to provide a broader perspective and deeper understanding of our work.

In 2011, a metric termed the age of information (AoI) was widely applied to characterize the freshness of data delivered in a communication system. It is defined as the time elapsed since the generation of the latest received update [7]. Various systems characterized by different queueing models such as M/M/1, D/M/1, M/D/1, M/G/1, and G/G/1/1 adopted AoI to evaluate the freshness of the received data [7–9]. Authors in [10] considered the multi-source preemptive queuing model and investigated the optimal generation rate of each source achieving the best information freshness. The optimal generation rate of update which induced the minimal violation probability is found in [11]. Apart from the average AoI metric, The metric of peak AoI (PAoI) was introduced to capture information about the maximum value of the AoI [12,13]. Based on the violation probabilities of AoI and peak AoI, the authors in [14] analyzed the optimal arrival rate of status update in the sight of asymptotical optimality. Authors in [15] proposed a new metric termed age upon decisions (AuD) to evaluate the freshness of received updates at some decision epochs.

In WSN, the mean square error metric is widely applied to characterize the fidelity of the system. In [16], techniques for designing robust pre-coders and combiners to the linear decentralized estimation of unknown vector parameters in a coherent multiple-input multiple-output (MIMO) network with multiple sensors under imperfect channel state
information (CSI) were presented. Their proposed techniques were based on the criteria that minimize the MSE of the estimated signal at the FC limited by total network power constraint. The authors in [17] proposed a framework of joint collaboration-compression for sequential estimation of a random vector parameter in WSN. By alternatively minimizing the sequential MMSE, they designed near-optimal and linear compression strategies under power constraints. Following from [16], the authors continued to develop optimal precoders minimizing the sum-MSE for the scenario of transmitting quantized observations in [18].

There have been typical works focusing on freshness and fidelity metrics in WSN [19–22]. An adaptive monitoring framework was introduced in [19] to achieve a balance between efficiency and accuracy on Internet-enabled physical devices. The authors in [20] determined the age-optimal policies for the update request and processing times subject to a maximum allowed distortion constraint on the updates. In [21], it was proved that there is an optimal number of quantization bits that optimizes the tradeoff of AoI and the MMSE. In [22], the AoI and MSE metrics were explicitly characterized by deriving the closed-form expressions for the block lengths and accuracy levels, after which they optimized the coding schemes by demonstrating a reachable region of AoI and MSE.

The above-mentioned works focused on the impact of request time, quantization bits, and the block length, respectively, on system freshness and fidelity. While our work is distinguished from them in that we try to optimize system performance through adjusting the number of sensor nodes, which is of paramount significance at the very beginning deploying phase.

3. System Model

We consider a wireless sensor network consisting of \( N \), \( N \in \mathbb{N} \) sensor nodes, which observe a common metric \( X \) (state update information such as temperature, humidity, and speed) characterized by a Gaussian process with zero-mean and variance \( \sigma_X^2 \) (the assumption is based on the pervasiveness that a large number of phenomena are distributed in normal form, as is assumed in [23–25]), as illustrated in Figure 1. Each node’s observation is transmitted through an orthogonal channel, which follows block Rayleigh fading. For ease of manipulation, we assume that the observation experience independent channel gain denoted as \(|h_i|^2 \sim \text{exp}(1)\), channel noise \( Z_i \) with zero-mean and variance \( \sigma_{Z_i}^2 \). The received signal for the \( i \)-th node \( Y_i \) is expressed as:

\[
Y_i = h_i X + Z_i,
\]

where \( h_i \) can be obtained by transmitting pilot signal which is used for channel estimation.

![Figure 1. Signal model of the proposed system.](image)

Then, the observations are transmitted to the FC. Considering that the ability of signal processing, collecting, and transmission are different among the facilities, we assume the transmitting time from sensor nodes to the FC obeys exponential distribution with mean \( 1/\nu \) and is independent and identically distributed (i.i.d.) across the sensor nodes. Specifically, the transmission is organized in rounds where the FC waits for observations
from $N$ sensors in each round, and it begins a new round at the exact moment when the previous round is complete. At the FC, selection combining (SC) technology is employed to better control the fidelity of the received data.

3.1. Data Freshness Metric

We adopt the AoI of the FC as the freshness metric of the system. It is defined as the difference between the current time and the time when the most recently received observation is generated. Formally, at a given time $t$, AoI is expressed as:

$$\Delta_t := t - u_t,$$

where $u_t$ represents the generating time of the most recently received observation. Accordingly, as is described in [7], the average AoI in the interval of $(0, T)$ is given as:

$$\overline{\Delta} = \lim_{T \to \infty} \frac{1}{T} \int_0^T \Delta_t \, dt.$$  

Figure 2a gives a typical example of the AoI evolution curve where $t^{(1)}$ represents the generation time of the first update, and $t^{(1)'}$ is the time it arrives at the FC. The AoI curve keeps growing until a new update is received by the FC, thus it exhibits a sawtooth shape. However, recall that in our proposed model, the FC works in rounds (round index shown in the form of superscript), where, in each round, there are $N_j$ sensors (shown in the form of subscript) observing the source, and a new round of data collecting begins at the exact moment when the previous round is finished, i.e., $t^{(1)'}_N = t^{(2)}_N$, etc. As shown in Figure 2b, the average AoI can be obtained through calculating the difference between the isosceles triangle expressed as $\frac{1}{2} [(R_{N_j}^{(j)} + R_{N_j}^{(j+1)})^2 - (R_{N_j}^{(j+1)})^2]$. Specifically, considering $I$ rounds of sensing in total, the time average AoI is:

$$\overline{\Delta} = \lim_{I \to \infty} \mathbb{E} \left[ \left( \frac{1}{I} \sum_{j=1}^{I} \left( \frac{1}{2} R_{N_j}^{(j)} + R_{N_j}^{(j)} R_{N_j}^{(j+1)} \right) \right) \right].$$

where $\mathbb{E}[:,]$ represents the mathematical expectation operator. Notice that $\mathbb{E}[R_{N_j}^{(j)}]$ and $\mathbb{E}[R_{N_j}^{(j+1)}]$ are independent from each other as round index approaches infinity, we have $\mathbb{E}[R_{N_j}^{(j)}] \mathbb{E}[R_{N_j}^{(j+1)}] = \mathbb{E}[R_{N_j}^{(j)}]^2$. Notice that the statistical characteristics of $R_{N_j}^{(j)}$ are the same, and $R_{N_j}^{(j)}$ is actually related to the number of sensor nodes $N$, so we remove the superscript $j$ and specify the number of sensor nodes $N$ as a subscript of the interval random variable, i.e., $R_N$. Then, the average AoI is:

$$\overline{\Delta}(N) = \mathbb{E}[R_N] + \frac{\mathbb{E}[R_N^2]}{2\mathbb{E}[R_N]},$$  

Figure 2. (a) Typical AoI evolution. (b) AoI evolution of our proposed system.
3.2. Data Fidelity Metric

We use selection combining technology at the FC, i.e., we select the largest $|h_i|$ among all sensor nodes denoted as $|h_{\text{max}}|$, i.e., $\max \{ |h_i|, 1 \leq i \leq N \} = |h_{\text{max}}|$. And that results in:

$$Y = |h_{\text{max}}|X + Z_i. \quad (5)$$

Accordingly, the distortion between the source $X$ and the received signal $Y$ is denoted as $D$, which is expressed as

$$D = \mathbb{E}[kY - X]^2 = \mathbb{E}[(k|h_{\text{max}}| - 1)X + kZ_i]^2, \quad (6)$$

where $k$ is the signal estimating coefficient.

4. Problem Formulation

In this section, we firstly give explicit characterizations of two metrics. After that, a problem of jointly optimizing data freshness and fidelity of the sensing system is formulated.

4.1. Average AoI of the Proposed Model

**Lemma 1.** The average AoI observed at the FC is:

$$\Delta(N) = \frac{3H_N}{2v} + \frac{G_N}{2vH_N}, \quad (7)$$

where $H_N = \sum_{i=1}^{N} \frac{1}{i}$ and $G_N = \sum_{i=1}^{N} \frac{1}{i^2}$.

**Proof of Lemma 1.** Recall that the general expression on average AoI we derive in Equation (4) requires us to figure out $\mathbb{E}[R_N]$ and $\mathbb{D}[R_N]$. And the transmitting time from sensor nodes to the FC obeys exponential distribution with mean $1/v$ and is i.i.d. across the sensor nodes. In terms of order statistic [26], the mathematical expectation and variance of the time difference denoted as $D[\cdot]$ between the first arriving observation and the start point is given by:

$$\mathbb{E}[R_1 - R_0] = \frac{1}{Nv}, \quad \mathbb{D}[R_1 - R_0] = \frac{1}{N^2v^2}, \quad (8)$$

as we treat this observation as the first arriving observation among $N$ sensors. As we set $\mathbb{E}[R_0]$ and $\mathbb{D}[R_0]$ to 0, $\mathbb{E}[R_1]$ and $\mathbb{D}[R_1]$ can be calculated as:

$$\mathbb{E}[R_1] = \frac{1}{Nv}, \quad \mathbb{D}[R_1] = \frac{1}{N^2v^2}. \quad (9)$$

The mathematical expectation and variance of the time difference between the second and first observation is:

$$\mathbb{E}[R_2 - R_1] = \frac{1}{(N-1)v}, \quad \mathbb{D}[R_2 - R_1] = \frac{1}{(N-1)^2v^2}, \quad (10)$$

as we treat this observation as the first arriving observation among the remaining $N - 1$ sensors. Take Equation (9) into Equation (10), $\mathbb{E}[R_2]$ and $\mathbb{D}[R_2]$ can be figured out as:

$$\mathbb{E}[R_2] = \frac{1}{Nv} + \frac{1}{(N-1)v} = \frac{H_N - H_{N-2}}{v},$$

$$\mathbb{D}[R_2] = \frac{1}{N^2v^2} + \frac{1}{(N-1)^2v^2} = \frac{G_N - G_{N-2}}{v^2}. \quad (11)$$

By using Mathematical induction, the $N$-th arriving observation is given by:

$$\mathbb{E}[R_N] = \frac{H_N}{v}, \quad \mathbb{D}[R_N] = \frac{G_N}{v^2}. \quad (12)$$

Note that $\mathbb{E}[R_N^2]$ can be expressed as

$$\mathbb{E}[R_N^2] = \mathbb{E}^2[R_N] + \mathbb{D}[R_N]. \quad (12)$$
By taking Equations (11) and (12) into Equation (4), the proof is completed.

4.2. MMSE of the Proposed Model

Recall that in our model each node transmits its observation through block Rayleigh fading channel, where the channel gain within a block remains the same while it varies independently across different blocks. Based on this model, let us use the MMSE metric to evaluate the data fidelity of the system.

**Lemma 2.** The MMSE $\mathcal{D}(N)$ by using SC technique is:

$$
\mathcal{D}(N) = \sum_{i=1}^{N} (-1)^{i-1} \binom{i}{N} i \Gamma(0, \frac{i}{\gamma_0}) e^{i/\gamma_0},
$$

where $\Gamma(a, b) = \int_{b}^{\infty} u^{a-1} e^{-u} du$ represents the incomplete Gamma function.

**Proof of Lemma 2.** Firstly, one can establish objective problem of minimizing the distortion as shown in Equation (6), which is

$$
\min_k \mathbb{E} [(|k| h_{\max} | - 1) X + k Z_i]^2,
$$

where $\mathbb{E}[X]^2 = \sigma_X^2$ and $\mathbb{E}[Z_i]^2 = \sigma_Z^2$. Notice the irrelevance between source $X$ and the noise $Z_i$, the objective problem can be further elaborated as

$$
\min_k (|h_{\max}|^2 \sigma_X^2 + \sigma_Z^2) k^2 - 2|h_{\max}| \sigma_X^2 k + \sigma_X^2.
$$

With some manipulations, the distortion, i.e., the optimal value of Equation (15) can be written as

$$
D(N) = \frac{\sigma_X^2 \sigma_Z^2}{|h_{\max}|^2 \sigma_X^2 + \sigma_Z^2}.
$$

In our model, we assume the noise power is the same among all the sensor nodes, i.e., $\sigma_Z^2 = 1$. Denote the SNR of each node as $\gamma_0 = \sigma_X^2 / \sigma_Z^2$, then the distortion is

$$
D(N) = \frac{\gamma_0}{|h_{\max}|^2 \sigma_Z^2 + 1} = \frac{1}{|h_{\max}|^2 + 1 / \gamma_0}.
$$

Recall $|h_i|^2 \sim \exp(1)$ and $\max\{|h_i|^2, 1 \leq i \leq N\} = |h_{\max}|$. Now we need to figure out the probability density function (pdf) of $|h_{\max}|^2$ to calculate the MMSE. For conciseness we denote $|h_{\max}|^2$ and its corresponding pdf as $g$ and $p(g)$, respectively. Given $N$ i.i.d. exponential distribution random variables with parameter $1/\theta$, the pdf of the maximum variable among them can be derived:

$$
p(g) = \frac{N}{\theta} e^{-g/\theta} (1 - e^{-g/\theta})^{N-1}
= \frac{N}{\theta} e^{-g/\theta} \sum_{i=0}^{N-1} \binom{i}{N-1} (-e^{-g/\theta})^i
= \sum_{i=0}^{N-1} (-1)^i N \binom{i}{N-1} \frac{1}{\theta^i} e^{-g(i+1)/\theta}
= \sum_{i=1}^{N} (-1)^{i-1} \binom{i}{N} \frac{1}{\theta^i} e^{-g/\theta}
$$

(18)
where in our work we have $1/\theta = 1$. After that, the MMSE can be computed as:

$$D(N) = \int_0^{\infty} \frac{1}{g + \frac{1}{\gamma_0}} p(g) dg$$

$$(a) \sum_{i=1}^{N} (-1)^{i-1} \left( \begin{array}{c} i \\ N \end{array} \right) \left( \int_{\frac{1}{\gamma_0}}^{\infty} e^{-\omega} d\omega \right)e^{i/\gamma_0}$$

$$\sum_{i=1}^{N} (-1)^{i-1} \left( \begin{array}{c} i \\ N \end{array} \right) i \Gamma(0, \frac{i}{\gamma_0}) e^{i/\gamma_0},$$

where (a) follows from $\omega = (g + \frac{1}{\gamma_0})i$ and $d\omega = idg + \frac{i}{\gamma_0}$. \qed

4.3. Tradeoff between Data Freshness and Fidelity

Following the analyses in the above two subsections, we can see that the number of sensor nodes $N$ is essential in affecting both freshness and fidelity metrics. Based on the results we derive so far, the objective problem is formulated by importing a weighting factor $\alpha$, to jointly optimize the two metrics simultaneously through adjusting the number of sensor nodes and the weighting factor:

$$P_1 : \min_{N \in \mathbb{N}} f(N) = \alpha \Delta(N) + (1 - \alpha) D(N),$$

where $\Delta(N)$ and $D(N)$ are given by Equations (7) and (13).

We can see that when the weighting factor $\alpha$ is large, the average AoI part is dominant. This condition can be applied in scenarios such as daily news updating and music hits ranking where timeliness is more important. On the contrary, when $\alpha$ is relatively small, the fidelity of the system becomes the major concern while the freshness requirement becomes relatively low. Such systems like fuel consumption reporting and indoor temperature monitoring are in this case since the state of these systems often changes slowly over time.

In order to find the optimal $N^*$, we first analyze the monotonicity of $\Delta(N)$, $D(N)$ and the tradeoff $f(N)$ respectively.

4.3.1. Monotonicity of Average AoI

As is derived from Lemma 1, the monotonicity of $\Delta(N)$ can be figured out by calculating the difference between $\Delta(N)$ and $\Delta(N - 1)$:

$$\Delta(N) - \Delta(N - 1) = \frac{3N + H_{N-1} - NG_{N-1}}{2N^2}.$$  \hspace{1cm} (20)

It can be easily verified that Equation (20) is positive, which indicates $\Delta(N)$ monotonically increases w.r.t. the number of sensor nodes $N$.

4.3.2. Monotonicity of $D(N)$

It is not difficult to conclude from Equation (13) that $D(N)$ is a monotone decreasing function w.r.t the number of sensor nodes $N$ by using Mathematical induction:

$$D(2) - D(1) = \Gamma(0, 1)e - 2\Gamma(0, 2)e^2 < 0;$$

$$D(3) - D(2) = \Gamma(0, 1)e - 4\Gamma(0, 2)e^2 + 3\Gamma(0, 3)e^3 < 0;$$

by applying the same method, it yields:

$$D(N) - D(N - 1) < 0.$$
4.3.3. Monotonicity of the Weighted-Sum Tradeoff

Since the average AoI $\overline{\Delta}(N)$ monotonically increases w.r.t. $N$, and $\overline{D}(N)$ monotonically decreases w.r.t. $N$. The conclusion can be drawn that the weighted-sum is most time a unimodal function that decreases first and increases then, even it may fluctuate sometimes, we can still treat the first local optimum as global optimum since it saves system cost for deploying sensors. The monotonicity analysis paves the way for finding a sub-optimal number of sensor nodes which we will elaborate on in the next section.

5. Sub-Optimal Number of Sensor Nodes

In this section, we give a closed-form sub-optimal number of sensor nodes to achieve the best system timeliness and fidelity. By applying proper approximations on two metrics, a closed-form sub-optimal number of sensor nodes can be derived.

**Theorem 1.** The sub-optimal number of sensor nodes can be approximately given by:

$$N^* = \arg \min \{ f(\lfloor \tilde{N} \rfloor), f(\lceil \tilde{N} \rceil) \},$$

where $\tilde{N} = \exp \left( -\frac{1}{\gamma_0} + \sqrt{\frac{2v(1-\alpha)}{3\alpha}} \right)$.

**Proof of Theorem 1.** Firstly, recall that $\overline{\Delta}(N) = \frac{3H_N}{2\nu} + \frac{\gamma_0}{2\nu H_N}$, in fact, the latter part $\frac{\gamma_0}{2\nu H_N}$ is small enough to omit, then we have $\overline{\Delta}(N) \approx \frac{3H_N}{2\nu}$. For MMSE, we try to use $1/\mathbb{E}[g + 1/\gamma_0]$ to approach $\overline{D}(N)$. That is

$$\mathbb{E}[g + \frac{1}{\gamma_0}] = \int_{0}^{\infty} (g + \frac{1}{\gamma_0}) p(g) dg = \frac{1}{\gamma_0} + \sum_{i=1}^{N} (-1)^{i-1} \frac{(i)}{N} \int_{0}^{\infty} i g e^{-ig} d\gamma_{g} + \int_{0}^{\infty} \frac{1}{\gamma_0} e^{-ig} d\gamma_{g}$$

where (b) follows from:

$$\sum_{i=1}^{N} (-1)^{i-1} \frac{(i)}{N} \int_{0}^{\infty} i g e^{-ig} d\gamma_{g} = H_N.$$

Then, the objective problem can be further reformulated as:

$$f(N) = \frac{3\alpha H_N}{2\nu} + \frac{1 - \alpha}{H_N + \frac{1}{\gamma_0}}.$$  

By searching for the zero-point of its first-order derivative and omit the unnecessary root, the sub-optimal number of sensor nodes can be figured out through

$$3\alpha \frac{N}{2\nu} - \frac{1 - \alpha}{N \left( H_N^2 + \frac{2H_N}{\gamma_0} + \frac{1}{\gamma_0} \right)} = 0.$$  

Noticing that $H_N \approx \ln N$, it yields

$$3\alpha \left( \ln^2 N + \frac{2\ln N}{\gamma_0} + \frac{1}{\gamma_0} \right) - (1 - \alpha)2\nu = 0.$$  

After some manipulations, $\tilde{N} = \exp \left( -\frac{1}{\gamma_0} + \sqrt{\frac{2v(1-\alpha)}{3\alpha}} \right)$ can be calculated where another unnecessary root is omitted.

At last, by comparing $f(\lfloor \tilde{N} \rfloor)$ with $f(\lceil \tilde{N} \rceil)$, the solution $N^*$ can be obtained by choosing the integer resulting in smaller data freshness and fidelity weighted-sum tradeoff. $\square$
One can conclude from Theorem 1 that the optimal number of sensor nodes $N^*$ is strongly affected by SNR $\gamma_0$, the weighting factor $\alpha$ and the arrival rate $v$.

6. Numerical Results

In this section, we firstly perform simulations to validate the correctness of theoretical results. After that, we investigate the effects of system parameters on the overall system freshness and fidelity. The realistic setup is shown in Table 1.

| System Parameter                  | Value                   |
|----------------------------------|-------------------------|
| Single-sided noise spectral density $N_0$ | $-173$ dBm/Hz           |
| Available bandwidth $W_{max}$     | $1$ MHz                 |
| Path loss model $10 \log(a_k)$    | $35.5 + 37.6 \log(d_k)$ [27] |
| Maximum transmit power of sensor  | $13$ dBm (20 mW) [27]   |

Figure 3 shows the simulation and analysis results of the average AoI $\Delta(N)$ when the arrival rate of the updates are $v = 4$ and $v = 6$, i.e., in each time slot, there are 4 and 6 updates arriving at the FC respectively. As can be seen intuitively, the average AoI increases while the number of sensor nodes $N$ grows. This accords with the analysis in Section 4.3.1. It is also validated that the average AoI is overall larger when the arrival rate $v$ is smaller (cf. Equation (7)). Most importantly, by comparing the curves with the markers we can conclude that our analysis result on average AoI (cf. Lemma 1) is identical to the simulation result in which we invoked 100,000 groups of arrays obeying exponential distribution with parameter $\nu$ to calculate the average AoI.

![Figure 3](image)

Figure 3. Simulation and analysis results of the average AoI $\Delta(N)$ w.r.t. the number of sensor nodes $N$ under different arrival rate $v$.

Figure 4 shows the simulation and analysis results of the MMSE $\mathcal{D}(N)$. The realistic setup in this simulation is introduced as follows. The signal power $\sigma_x^2$ is set to 10 dBm (10 mW) and 13 dBm (20 mW) respectively, the channel noise power $\sigma_Z^2$ is $N_0B = -173 \times 10^6$ dBm. For the path loss model, we adopt $35.5 + 37.6 \log(d_k)$, where we assume the distance between the sensors and the FC $d_k$ is 100 m. Then we have the corresponding SNR $\gamma_0$ as 12.5 dB and 15.5 dB respectively. Firstly, it can be seen from the figure that our theoretic analysis on MMSE (Lemma 2) accords with the simulation in which we invoked 100,000 groups of random variables obeying $|h_i|^2 \sim \exp(1)$ and chose the largest among them as the received signal to further calculate the MMSE. Secondly, it is shown that the MMSE declines with the growth of sensor nodes number $N$, which also echoes with the analysis in Section 4.3.2.
Figure 4. Simulation and analysis results of the average MMSE $\overline{D}(N)$ w.r.t. the number of sensor nodes $N$ under different SNR $\gamma_0$.

Figure 5 shows the varying trend of the data freshness-fidelity weighted-sum tradeoff $f(N)$ with respect to the number of sensor nodes $N$ under different SNR $\gamma_0$ and weighting factor $\alpha$. The realistic setup of the system is the same as in Figure 4. In this case, we consider the same arrival rate as $v = 8$ (eight updates arrive at the FC in each time slot). Overall, all curves well demonstrate the monotonicity of the weighted-sum tradeoff which decreases first and goes up then. Comparing circular dotted curve with a circular solid curve which differs in SNR $\gamma_0$, we can conclude that a smaller SNR $\gamma_0$ results in smaller MMSE and so is the tradeoff, whereas the optimal number of sensor nodes $N^* = 4$ remains unchanged. Comparing the circular solid curve with the square solid curve, it is seen that a larger weighting factor $\alpha$ leads to a smaller optimal number of sensor nodes $N^*$. This can be ascribed to the fact that a larger weighting factor $\alpha$ means a larger AoI, hence the curve begins the upward trend earlier. By further reducing the weighting factor $\alpha$ to 0.45, the optimal number of sensor nodes becomes even larger, see the dotted square curve, reaching $N^* = 12$. Most importantly, it is obvious that our proposed sub-optimal number of sensor nodes $N^*$ is identical to the optimal $N^*$. This well validates the effectiveness of our proposed sub-optimal number of sensor nodes.

Figure 5. The weighted-sum of freshness and fidelity tradeoff $f(N)$ w.r.t. the number of sensor nodes $N$ under different SNR $\gamma_0$ and weighting factor $\alpha$. 
Figure 6 depicts the relationship between the data freshness-fidelity weighted-sum tradeoff $f(N)$ and the number of sensor nodes $N$ under different weighting factor $\alpha$ and arrival rate $v$. In the simulation, the signal power $\sigma^2_X$ is fixed to 10 dBm (10 mW) and the noise power remains the same, which means the MMSE distortion remains the same. Observing the circular solid curve and the square solid curve, we can see that under the same weighting factor $\alpha = 0.6$, a larger arrival rate $v$ causes a smaller tradeoff. However, the optimal number of sensor nodes $N^*$ can be comparatively larger: When $v = 10$, $N^* = 8$; When $v = 5$, $N^* = 5$. This is explainable since a larger arrival rate $v$ leads to smaller AoI, which means the curve will start climbing later. The improvement on system overall freshness and fidelity brought by the proposed sub-optimal solution $N^*$ is also validated.

![Figure 6](image-url)

**Figure 6.** The weighted-sum of freshness and fidelity tradeoff $f(N)$ w.r.t. the number of sensor nodes $N$ under different and weighting factor $\alpha$ and arrival rate $v$.

7. Conclusions

Motivated by the contradictory relationship between the freshness and fidelity of the received data, we studied the joint optimization of the two metrics by adjusting the number of sensor nodes. Explicit expressions of average AoI and MMSE were derived, based on which a closed-form sub-optimal solution was obtained via feasible approximations. Numerical results validate that our proposed sub-optimal number of sensor nodes is correct and can achieve the best data freshness and fidelity tradeoff with negligible errors.

Author Contributions: Conceptualization, Z.C., M.X., M.W. and Y.J.; methodology, Z.C. and M.X.; software, M.X.; validation, M.W. and Y.J.; formal analysis, M.X.; investigation, Z.C. and M.X.; resources, all authors; data curation, M.W. and Y.J.; writing-original draft preparation, Z.C. and M.X.; visualization, M.X.; supervision, Z.C.; project administration, Z.C.; funding acquisition, Z.C. All authors approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61901066, Grant 61971077, and in part by Natural Science Foundation of Chongqing, China under Grant cstc2019jcyj-msxmX0575. It is also sponsored by the State Key Laboratory of Integrated Services Networks (Xidian University) under grant ISN22-17.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sonehara, M.; Van Toai, N.; Sato, T. Fundamental Study of Non-Contact Water Salinity Sensor by Using Electromagnetic Means for Seawater Desalination Plants. *IEEE Trans. Magn.* **2016**, *52*, 4001604.

2. Lee, S.; Ko, Y. Joint Selection Combining and Power Loading Transmission With Adaptive M-QAM in Multichannel System. *IEEE Trans. Veh. Technol.* **2009**, *58*, 3819–3825.
3. Al-Suhail, G.A.; Lamahewa, T.A.; Kennedy, R.A. Performance of dual-branch diversity receiver based SR-ARQ in Rayleigh fading channel. In Proceedings of the 3rd International Conference on Signal Processing and Communication Systems, Omaha, NE, USA, 28–30 September 2009; pp. 1–6.

4. Rao, B.D.; Ming, Y. Performance of maximal ratio transmission with two receive antennas. IEEE Trans. Commun. 2003, 51, 894–895. [CrossRef]

5. Young, G.; Sang, W. Optimum selection combining for M-ary signals in frequency-nonselective fading channels. IEEE Trans. Commun. 2005, 53, 84–93.

6. Al-Suhail, G.A.; Kennedy, R.A. Performance of hybrid ARQ on dual-branch diversity receiver in Rayleigh fading channel. In Proceedings of the 2010 4th International Conference on Signal Processing and Communication Systems, Gold Coast, QLD, Australia, 13–15 December 2010; pp. 1–8.

7. Kaul, S.; Yates, R.; Gruteser, M. Real-time status: How often should one update? In Proceedings of the IEEE INFOCOM, Orlando, FL, USA, 25–30 March 2012; pp. 2731–2735.

8. Moltarof, M.; Leinonen, M.; Codreanu, M. On the Age of Information in Multi-Source Queueing Models. IEEE Trans. Commun. 2020, 68, 5003–5017. [CrossRef]

9. Soysal, A.; Ulukus, S. Age of Information in G/G/1/1 Systems: Age Expressions, Bounds, Special Cases, and Optimization. IEEE Trans. Inf. Theory 2021, 67, 7477–7489. [CrossRef]

10. Zhang, T.; Chen, S.; Chen, Z. Internet of Things: The Optimal Generation Rates under Preemption Strategy in a Multi-Source Queueing System. Entropy 2021, 8, 1055. [CrossRef]

11. Hu, L.; Chen, Z.; Dong, Y.; Jia, Y.; Liang, L.; Wang, M. Status Update in IoT Networks: Age-of-Information Violation Probability and Optimal Update Rate. IEEE Internet Things J. 2021, 8, 11329–11344. [CrossRef]

12. Costa, M.; Codreanu, M.; Ephremides, A. Age of information with packet management. Proc. IEEE Int. Symp. Inf. Theor. 2014, 67, 1583–1587.

13. Yang, H.H.; Arafa, A.; Quek, T.Q.S.; Poor, H.V. Optimizing Information Freshness in Wireless Networks: A Stochastic Geometry Approach. Proc. IEEE Trans. Mob. Comput. 2021, 20, 2269–2280. [CrossRef]

14. Hu, L.; Chen, Z.; Jia, Y.; Wang, M.; Quek, T.Q.S. Asymptotically Optimal Arrival Rate for IoT Networks With Aol and Peak Aol Constraints. IEEE Commun. Lett. 2021, 25, 3853–3857. [CrossRef]

15. Dong, Y.; Chen, Z.; Liu, S.; Fan, P.; Ben Letaief, K. Age-upon-decisions minimizing scheduling in Internet of Things: To be random or to be deterministic? IEEE Internet Things J. 2020, 7, 1081–1097. [CrossRef]

16. Rajput, K.P.; Verma, Y.; Venkategowda, N.K.D.; Jagannatham, A.K.; Varshney, P.K. Robust Linear Transceiver Designs for Vector Parameter Estimation in MIMO Wireless Sensor Networks Under CSI Uncertainty. IEEE Trans. Veh. Technol. 2021, 70, 7347–7362. [CrossRef]

17. Cheng, X.; Khanduri, P.; Chen, B.; Varshney, P.K. Joint Collaboration and Compression Design for Distributed Sequential Estimation in a Wireless Sensor Network. IEEE Trans. Signal Process. 2021, 69, 5448–5462. [CrossRef]

18. Rajput, K.P.; Ahmed, M.F.; Venkategowda, N.K.D.; Jagannatham, A.K.; Sharma, G.; Hanzo, L. Robust Decentralized and Distributed Estimation of a Correlated Parameter Vector in MIMO-OFDM Wireless Sensor Networks. IEEE Trans. Commun. 2021, 69, 6984–6990. [CrossRef]

19. Trihinas, D.; Pallis, G.; Dikaikakos, M.D. Low-Cost Adaptive Monitoring Techniques for the Internet of Things. IEEE Trans. Serv. Comput. 2021, 14, 487–501. [CrossRef]

20. Bastopcu, M.; Ulukus, S. Age of Information for Updates With Distortion: Constant and Age-Dependent Distortion Constraints. IEEE/ACM Trans. Netw. 2021, 29, 2425–2438. [CrossRef]

21. Arafa, A.; Seddik, K.B.K.G.; Poor, H.V. Sample, Quantize, and Encode: Temporal Estimation Over Noisy Channels. IEEE Trans. Commun. 2021, 69, 6485–6499.

22. Roth, S.; Arafa, A.; Sezgin, A.; Poor, H.V. Short Blocklength Process Monitoring and Scheduling: Resolution and Data Freshness. IEEE Trans. Wirel. Commun. 2021. [CrossRef]

23. Dong, Y. Distributed Sensing With Orthogonal Multiple Access: To Code or not to Code? IEEE Trans. Signal. Process. 2020, 68, 1315–1330. [CrossRef]

24. Liu, S.; Kar, S.; Fardad, M.; Varshney, P.K. Optimized Sensor Collaboration for Estimation of Temporally Correlated Parameters. IEEE Trans. Signal Process. 2016, 64, 6613–6626. [CrossRef]

25. Liu, S.; Chepuri, S.P.; Fardad, M.; Masazade, E.; Leus, G.; Varshney, P.K. Sensor Selection for Estimation with Correlated Measurement Noise. IEEE Trans. Signal Process. 2016, 64, 3509–3522. [CrossRef]

26. Arnold, B.B.C.; Balakrishnan, N.; Nagaraja, H.N. A First Course in Order Statistics, 2nd ed.; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2008.

27. Study on Scenarios and Requirements for Next Generation Access Technologies, Technical Specification Group Radio Access Network, Release 14, Document 38.913, 3GPP, October 2016. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2996 (accessed on 10 December 2021).