

批量到达物联网状态更新系统的决策年龄与决策失真分析

刘磊^① 靳文凯^① 张清清^② 李渝舟^{*③} 江帆^①

^①(西安邮电大学通信与信息工程学院 西安 710121)

^②(浙江工业大学信息工程学院 杭州 310023)

^③(华中科技大学电子信息与通信学院 武汉 430071)

摘要: 针对具有批量到达特征的物联网(IoT)状态更新系统, 该文以决策年龄(AuD)和决策失真(DaD)作为性能指标研究了系统决策的新鲜度和失真度。利用排队论, 该文在批量大小服从一般分布的条件下推导出平均AuD与平均DaD的解析表达式。在此基础上, 考虑具有典型随机波动特征的几何分布批量大小, 采用交替迭代优化算法对批到达率、平均批量大小和决策阈值进行联合优化, 从而实现最小化平均AuD和平均DaD的加权和。仿真结果验证了理论分析的正确性。该研究发现, 在一般分布的批量大小条件下, 系统决策的新鲜度和失真度之间存在明显的权衡关系; 在批量大小服从几何分布的典型场景下, 该文所设计的算法能够在平均AuD和平均DaD之间实现帕累托最优权衡。

关键词: 物联网; 状态更新系统; 批量到达; 决策年龄; 决策失真

中图分类号: TN926

文献标识码: A

文章编号: 1009-5896(2026)00-0001-10

DOI: 10.11999/JEIT260359

CSTR: 32379.14.JEIT260359

1 引言

状态更新贯穿物联网(Internet of Things, IoT)“感知-传输-决策”全链路, 保障状态更新的信息新鲜度是支撑垂直行业应用关键场景落地发展的技术前提。然而, 自终端设备生成后, 状态更新信息需要经历排队、传输与计算等多个环节, 诸如时延等传统性能指标难以有效衡量所获取状态更新的信息新鲜度^[1]。针对这一问题, 研究人员将接收端的当前时刻与所接收到最新状态更新的生成时刻之间的时间差定义为信息年龄(Age of Information, AoI), 并广泛用于刻画IoT状态更新场景中的信息新鲜度^[2,3]。在面向垂直行业的IoT应用中, 系统的可靠稳定运行高度依赖各类型决策的质量。如果利用陈旧信息进行决策将会诱发风险, 因此决策质量高度依赖决策时刻的瞬时AoI。例如, 智能电网台区故障隔离决策仅在故障告警触发时启动, 尽管智能电表与监测终端会批量上传计量与工况等数据, 但只有在决策时刻该批量数据的AoI才会直接影响故障隔离范围与恢复时长。鉴于此, 研究人员在对AoI的研究基础之上利用决策年龄(Age upon Decisions, AuD)指标表征在决策时刻的信息新鲜度^[4]。

利用排队论工具能够刻画AoI和AuD的动态演

化规律, 已有工作在不同队列模型下对状态更新系统的信息新鲜度性能展开了分析研究。文献[1]推导了先到先服务(First Come First Served, FCFS)规则下M/M/1, M/D/1和D/M/1队列的系统平均AoI, 并探究了实现平均AoI的最优负载配置方式。文献[5]通过设计混合抢占和非抢占的队列服务规则, 可进一步改善队列系统的平均AoI。利用兼具灵活性与解析性的位相型分布, 文献[6]在PH/PH/1/1与M/PH/1/2队列模型中, 解析了抢占与非抢占服务规则下的AoI分布特征。文献[7]关注更新的到达过程和决策间隔对AuD性能的影响, 并发现采用随机决策时, 周期性的状态更新可使平均AuD达到最小。在无缓冲队列系统中, 服务过程的统计特性对AuD性能影响更为显著。为此, 文献[8]延展关注服务过程的影响, 具体在不同服务时间分布下推导了指数与确定决策间隔对应的平均AuD, 并发现确定性服务时间在降低平均AuD方面优于均匀分布与指数分布服务时间。文献[9]研究了FCFS和后到先服务(Last-Come-First-Serves, LCFS)两种服务规则下系统的AuD性能, 发现在指数决策间隔下, 相较于周期性到达和均匀到达过程, 泊松到达过程在FCFS下对应的平均AuD最大, 而在LCFS下则表现出更优的AuD性能。文献[10]关注工业物联网(Industrial IoT, IIoT)应用场景, 通过构造马尔可夫链建模终端队列状态的转移过程, 在伯努利到达过程与几何分布决策间隔条件下推导得出了平均AuD的闭式解, 并通过调整短包块长与传输功率进一步降低平均AuD。

收稿日期: 2026-xx-xx; 改回日期: 2026-06-24; 网络出版: 2026-07-02

*通信作者: 李渝舟 yuzhouli@hust.edu.cn

基金项目: 国家自然科学基金(62471195, 62331010, 62201456)

Foundation Items: The National Natural Science Foundation of China (62471195, 62331010, 62201456)

尽管已有研究工作考虑了具有不同到达间隔分布的到达过程,但其建模对象均局限于单个状态更新的逐个到达。然而,在如前所述的智能电网故障决策应用场景中,终端的多条具有不同属性的状态更新具有批量到达特征。此外,这一现象在IIoT集群监测等其他场景中也普遍存在,即终端设备生成并传输一批状态更新,形成具有显著时间聚集性的批量到达业务。针对批量到达特征,文献[11]引入周期性批量采样模型,分析了批量到达下系统的平均AoI性能,并揭示了批量采样频率对信息新鲜度的影响。文献[12]研究了成批到达业务流对峰值AoI(Peak AoI, PAoI)的影响,指出当状态更新呈现成批到达特征时,PAoI的统计特性将显著发生改变。然而,上述研究主要关注批量到达对AoI及其相关指标的影响,对于批量到达特征对AuD的影响仍缺乏系统研究。针对批量到达业务的决策触发往往依赖于聚合在同一批内状态更新特征的统计分布而非单个更新本身。忽略批量到达特点将对此场景中队列状态跃迁演变的刻画产生偏差,从而影响到AuD性能的分析。因此,有必要在具有批量到达的队列模型中研究AuD性能。

AuD虽然能够有效刻画决策时刻的信息新鲜度,但决策质量不仅取决于状态更新是否新鲜,还与决策时所利用信息的完整性和准确性密切相关。为了刻画这一性能,本文进一步引入决策失真度(Distortion at Decisions, DaD)作为衡量决策质量的另一关键指标。在状态更新系统中,过早触发决策可能导致可用信息不足而增大决策失真,而依赖更多累积更新则会因信息陈旧而降低决策价值,从而在AuD与DaD之间形成权衡关系。已有相关工作围绕新鲜度与失真度之间的关系开展研究。文献[13]在状态更新系统中引入常数及年龄相关失真约束,分析了不同失真条件对AoI性能的影响。文献[14]从联合采样与压缩设计的角度研究了AoI与重构失真之间的相互作用,揭示了采样频率与压缩精度的系统选择将同时影响AoI与失真性能。文献[15]面向实时决策场景构建了同时考虑AuD与失真的性能分析框架,指出两者对决策性能具有协同影响。文献[16]在无线传感器网络场景中,基于不同信息收集策略刻画了AoI与数据失真之间的权衡关系,进一步揭示了信息时效性与失真度对系统性能的联合影响。

当考虑批量到达特征时,信息新鲜度与失真度之间的权衡关系更加突出。具体而言,为了降低失真度,需要尽可能完整地利用到达批量中的所有状态更新,但批量到达对排队与服务过程的累积效应

会导致AuD恶化;相反,为了保持较高的新鲜度,需要更早触发决策,但这会限制对批内状态更新的利用数量从而导致DaD升高。针对这一问题,本文面向具有批量到达特征的IoT状态更新系统,引入AuD和DaD分别作为刻画决策时信息新鲜度和失真度的性能指标。在此基础上,在一般批量大小分布条件下推导得出了平均AuD和平均DaD的表达式。进一步地,针对具有显著随机波动与突发特性的几何批量大小分布这一典型特例,通过联合优化批到达率、平均批量大小与决策阈值,实现最小化平均AuD与平均DaD的加权和。仿真结果验证了理论分析的准确性以及所提方法的有效性。

2 系统模型

如图1所示,本文考虑一种具有批量到达特征的IoT状态更新系统。该系统由传感器和决策控制单元两部分构成,其中,传感器采集状态更新信息,并以批量形式发送至网络边缘侧的决策控制单元,决策控制单元对批量到达的状态更新进行处理,并基于处理结果做出相应决策。相邻批的到达时间间隔服从参数为 λ 的指数分布,每批中状态更新的数量,即批量大小 N ,服从一般分布。此外,第 k 批中的第 j 个状态更新的服务时间 $X_{k,j}$ ($j = 1, 2, \dots, N$)相互独立且服从参数为 μ 的指数分布。

考虑到同一批中状态更新包含具有不同属性的数据信息,本文以随机权重反映同一批内不同状态更新的差异性,并通过状态更新的权重的分布描述其统计特征。鉴于Beta分布定义在 $[0, 1]$ 区间,且可通过调节形状参数灵活刻画多种分布形态,因此,假设批内各状态更新的归一化权重取值服从Beta分布。由于不同状态更新信息具有不同的决策价值,本文采用基于权重阈值的决策触发机制,即当某个完成服务的状态更新的权重超过阈值时,决策控制单元触发一次决策。令 $\theta \in [0, 1)$ 表示决策阈值,则对于批量内的任意状态更新其触发决策的概率可表示为

$$q(\theta) \triangleq \mathbb{P}\{\omega > \theta\} = \int_{\theta}^1 f_{\omega}(x) dx \quad (1)$$

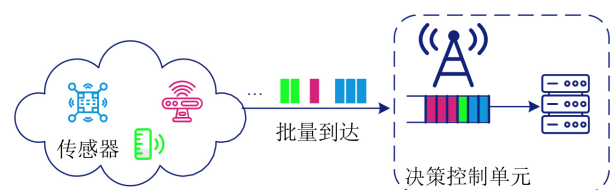


图1 系统模型

其中, $f_\omega(x) = x^{\alpha-1}(1-x)^{\beta-1}\Gamma(\alpha+\beta)(\Gamma(\alpha)\Gamma(\beta))^{-1}$ 为Beta分布的概率密度函数(Probability Density Function, PDF), α 和 β 为Beta分布的两个形状参数, $\Gamma(\cdot)$ 表示Gamma函数。

在批量到达特征下, 系统AoI反映了当前时刻与最近一个完成整批服务的状态更新批量的生成时刻之间的时间差, 其演化与整批的服务过程相关。当同批内所有状态更新均完成服务后, AoI才会根据该批的生成时刻进行刷新。令 $\Delta(t)$ 表示 t 时刻的AoI, 根据AoI与AuD之间的关系, 系统在第 l 次决策时的AuD可被定义为

$$\Lambda_l \triangleq \Delta(\tau_l) = \tau_l - U(\tau_l) \quad (2)$$

其中, τ_l 表示第 l 次决策发生的时间, $U(\tau_l)$ 表示截至决策时刻 τ_l , 最近已完成整批服务的状态更新批量的到达时刻。系统的平均AuD, 记为 $\bar{\Lambda}$, 定义为决策时刻AuD的长期时间平均, 即

$$\begin{aligned} \bar{\Lambda} &= \lim_{T \rightarrow \infty} \frac{1}{M_T} \sum_{j=1}^{M_T} \Lambda_j \\ &= \mathbb{E} \left[\sum_{j=1}^N \Lambda_j \chi_j \right] \left(\mathbb{E} \left[\sum_{j=1}^N \chi_j \right] \right)^{-1} \end{aligned} \quad (3)$$

其中, M_T 表示在时间段 T 内共有 M_T 次决策发生, $\chi_j \in \{0, 1\}$ 表示二元指示函数, 当第 j 个状态更新发生决策时其值为1, 否则为0。

传感器对第 k 批量源状态 x_k 进行观测, 其中第 j 个状态更新携带的观测信息可表示为

$$y_{k,j} = x_k + n_{k,j} \quad (4)$$

其中, 噪声服从均值为0, 方差为 $\sigma_{k,j}^2$ 的高斯分布。观测噪声的方差通常用于刻画观测精度。基于对批量内状态更新信息差异的建模, 假设观测噪声方差与状态更新权重成反比, 即 $\sigma_{k,j}^2 = \sigma_0^2 / \omega_{k,j}$ 。其中, σ_0^2 为单位权重下的基准噪声方差。可见, 权重越大的状态更新对应更小的观测噪声方差, 即更高的观测精度。

当第 k 批中第 j 个状态更新完成服务并触发决策时, 决策控制单元利用该批量中已完成服务的前 j 个状态更新进行联合估计。由于观测噪声相互独立, 因此在最佳线性无偏估计(Best Linear Unbiased Estimator, BLUE)^[17]下产生的DaD可表示为

$$\Upsilon_{k,j} = \sigma_0^2 \left(\sum_{i=1}^j \omega_{k,i} \right)^{-1} \quad (5)$$

利用 $\bar{\Upsilon}$ 表示系统的平均DaD, 其定义为决策时刻DaD的长期时间平均, 即

$$\begin{aligned} \bar{\Upsilon} &= \lim_{T \rightarrow \infty} \frac{1}{M_T} \sum_{j=1}^{M_T} \Upsilon_j \\ &= \mathbb{E} \left[\sum_{j=1}^N \Upsilon_j \chi_j \right] \left(\mathbb{E} \left[\sum_{j=1}^N \chi_j \right] \right)^{-1} \end{aligned} \quad (6)$$

3 平均决策年龄与平均失真分析

为获得平均AuD解析式, 首先推导 $\mathbb{E} \left[\sum_{j=1}^N \Lambda_j \chi_j \right]$ 。根据全期望公式, 可将 $\mathbb{E} \left[\sum_{j=1}^N \Lambda_j \chi_j \right]$ 表示为

$$\mathbb{E} \left[\sum_{j=1}^N \Lambda_j \chi_j \right] = \mathbb{E} \left[q(\theta) \sum_{j=1}^n \mathbb{E}[\Lambda_j | N = n] \right] \quad (7)$$

考虑到第 k 批中第 j 个状态更新的AuD由批到达时间间隔、批等待时间和批服务时间三部分构成。结合时间递推关系, 当第 j 个状态更新发生决策时, 其AuD可表示为

$$\Lambda_j = \begin{cases} Y_k + W_k + S_{k,j}, & 1 \leq j < n \\ W_k + S_{k,j}, & j = n \end{cases} \quad (8)$$

其中, Y_k 表示第 $k-1$ 批与第 k 批状态更新的到达时间间隔, W_k 表示第 k 批状态更新的等待时间, $S_{k,j} = \sum_{i=1}^j X_{k,i}$ 表示批量中第1个状态更新到第 j 个状态更新的服务时间总和。

因此, 将式(8)代入到式(7)中, 并注意到 $\mathbb{E}[Y] = 1/\lambda$ 和 $\mathbb{E}[S] = \mathbb{E}[N]\mathbb{E}[X]$, 可得

$$\begin{aligned} \mathbb{E} \left[\sum_{j=1}^N \Lambda_j \chi_j \right] &= q(\theta) \left(\frac{\mathbb{E}[N] - 1}{\lambda} + \mathbb{E}[N]\mathbb{E}[W] \right. \\ &\quad \left. + \frac{(\mathbb{E}[N^2] + \mathbb{E}[N])}{2\mu} \right) \end{aligned} \quad (9)$$

由于批量服务时间 S 的二阶矩可表示为 $\mathbb{E}[S^2] = \mathbb{E}[N]\mathbb{E}[X^2] + (\mathbb{E}[N^2] - \mathbb{E}[N])(\mathbb{E}[X])^2$, 因此, 根据Pollaczek-Khinchine公式可得

$$\mathbb{E}[W] = \frac{\lambda(\mathbb{E}[N^2] + \mathbb{E}[N])}{2\mu(\mu - \lambda\mathbb{E}[N])} \quad (10)$$

注意到 $\mathbb{E} \left[\sum_{j=1}^N \chi_j \right] = q(\theta)\mathbb{E}[N]$ 。从而, 将式(9)和式(10)代入到式(3)可得平均AuD为

$$\begin{aligned} \bar{\Lambda} &= \frac{\mathbb{E}[N] - 1}{\mathbb{E}[N]\lambda} + \frac{\lambda(\mathbb{E}[N^2] + \mathbb{E}[N])}{2\mu(\mu - \lambda\mathbb{E}[N])} \\ &\quad + \frac{\mathbb{E}[N^2] + \mathbb{E}[N]}{2\mu\mathbb{E}[N]} \end{aligned} \quad (11)$$

为获得平均DaD解析式, 首先推导 $\mathbb{E} \left[\sum_{j=1}^N \Upsilon_j \chi_j \right]$ 。根据全期望公式, 可将 $\mathbb{E} \left[\sum_{j=1}^N \Upsilon_j \chi_j \right]$ 表示为

$$\begin{aligned} & \mathbb{E} \left[\sum_{j=1}^N \Upsilon_j \chi_j \right] \\ &= \mathbb{E} \left[\sum_{j=1}^n \mathbb{E} \left[\sigma_0^2 \chi_j \left(\sum_{i=1}^j \omega_i \right)^{-1} \middle| N = n \right] \right] \quad (12) \end{aligned}$$

首先关注条件期望项 $\sum_{j=1}^n \mathbb{E} \left[\sigma_0^2 \chi_j \left(\sum_{i=1}^j \omega_i \right)^{-1} \middle| N = n \right]$ 。考虑到批量大小 N 与权重 ω 之间相互独立, 结合条件期望的定义可得

$$\begin{aligned} & \sum_{j=1}^n \mathbb{E} \left[\sigma_0^2 \chi_j \left(\sum_{i=1}^j \omega_i \right)^{-1} \middle| N = n \right] \\ &= \sum_{j=1}^n q(\theta) \sigma_0^2 \mathbb{E} \left[\left(\sum_{i=1}^j \omega_i \right)^{-1} \middle| \omega_j > \theta \right] \quad (13) \end{aligned}$$

接下来, 对式(13)中的条件期望项 $\mathbb{E} \left[\left(\sum_{i=1}^j \omega_i \right)^{-1} \middle| \omega_j > \theta \right]$ 进行推导。令 $\Omega_j = \sum_{i=1}^j \omega_i$, 则有

$$\mathbb{E} \left[\Omega_j^{-1} \middle| \omega_j > \theta \right] = \int_0^{\infty} \mathbb{E} \left[\exp(-s\Omega_j) \middle| \omega_j > \theta \right] ds \quad (14)$$

考虑独立同分布的状态更新权重¹⁾, 从而有

$$\begin{aligned} \mathbb{E} \left[\exp(-s\Omega_j) \middle| \omega_j > \theta \right] &= \mathbb{E} \left[\exp(-s\Omega_{j-1}) \right] \\ &\quad \cdot \mathbb{E} \left[\exp(-s\omega_j) \middle| \omega_j > \theta \right] \\ &= \left(\mathbb{E} \left[\exp(-s\omega) \right] \right)^{j-1} \\ &\quad \cdot \mathbb{E} \left[\exp(-s\omega_j) \middle| \omega_j > \theta \right] \quad (15) \end{aligned}$$

令 $\mathcal{L}_\omega(s) = \mathbb{E} \left[\exp(-s\omega) \right]$ 表示权重 ω 的拉普拉斯-斯蒂杰斯变换(Laplace-Stieltjes Transform, LST), $\mathcal{L}_\omega^{\text{cond}}(s; \theta) = \mathbb{E} \left[\exp(-s\omega_j) \middle| \omega_j > \theta \right]$ 表示在 $\omega_j > \theta$ 条件下权重 ω 的条件LST。因此, 式(15)可进一步表示为

$$\mathbb{E} \left[\exp(-s\Omega_j) \middle| \omega_j > \theta \right] = \left(\mathcal{L}_\omega(s) \right)^{j-1} \mathcal{L}_\omega^{\text{cond}}(s; \theta) \quad (16)$$

将式(13)、式(14)和式(16)代入到式(12)中, 可得

$$\begin{aligned} \mathbb{E} \left[\sum_{j=1}^N \Upsilon_j \chi_j \right] &= q(\theta) \sigma_0^2 \int_0^{\infty} \mathcal{L}_\omega^{\text{cond}}(s; \theta) \\ &\quad \cdot \mathbb{E} \left[\sum_{j=1}^N \left(\mathcal{L}_\omega(s) \right)^{j-1} \right] ds \quad (17) \end{aligned}$$

接下来需要推导式(17)中的期望 $\mathbb{E} \left[\sum_{j=1}^N \left(\mathcal{L}_\omega(s) \right)^{j-1} \right]$, 进一步可将其表示为

$$\begin{aligned} \mathbb{E} \left[\sum_{j=1}^N \left(\mathcal{L}_\omega(s) \right)^{j-1} \right] &= \mathbb{E} \left[\frac{1 - \left(\mathcal{L}_\omega(s) \right)^N}{1 - \mathcal{L}_\omega(s)} \right] \\ &= \frac{1 - G_N(\mathcal{L}_\omega(s))}{1 - \mathcal{L}_\omega(s)} \quad (18) \end{aligned}$$

其中, $G_N(z) = \mathbb{E} [z^N]$ 是关于批量大小 N 的概率母函数(Probability Generating Function, PGF)。

注意到 $\mathbb{E} \left[\sum_{j=1}^N \chi_j \right] = q(\theta) \mathbb{E} [N]$, 因此, 将式(17)和式(18)代入到式(6)中可得平均DaD为

$$\bar{\Upsilon} = \frac{\sigma_0^2}{\mathbb{E} [N]} \int_0^{\infty} \frac{\mathcal{L}_\omega^{\text{cond}}(s; \theta) (1 - G_N(\mathcal{L}_\omega(s)))}{1 - \mathcal{L}_\omega(s)} ds \quad (19)$$

4 平均决策年龄与平均失真权衡

在批量状态更新系统中, 决策时刻的信息新鲜度与失真度之间存在内在冲突: 利用更多已完成服务的状态更新进行联合估计可以降低DaD, 但同时会延迟决策触发, 从而导致决策时刻AuD增大。因此, 在降低失真度的同时维持较低的新鲜度, 成为系统设计中的关键问题。为刻画二者之间的权衡关系, 本节构建了以平均AuD与平均DaD为目标的加权和优化问题。在状态更新系统中, 批量大小通常受数据采集和汇聚过程的影响而呈现随机波动性。考虑到批量大小的随机波动特征, 本节在一般批量大小分布性能分析的基础上, 进一步选取几何分布作为典型批量大小分布开展优化分析。

在批量大小服从几何分布的条件下, 系统的平均AuD可表示为

$$\bar{\Lambda}^{\text{Geo}} = \frac{m-1}{\lambda m} + \frac{\lambda m^2}{\mu(\mu - \lambda m)} + \frac{m}{\mu} \quad (20)$$

其中, $m = \mathbb{E} [N]$ 表示平均批量大小。系统的平均DaD可表示为

$$\bar{\Upsilon}^{\text{Geo}} = \sigma_0^2 \int_0^{\infty} \frac{\mathcal{L}_\omega^{\text{cond}}(s; \theta)}{m(1 - \mathcal{L}_\omega(s)) + \mathcal{L}_\omega(s)} ds \quad (21)$$

基于此, 本节旨在通过联合优化批量到达特征参数(批到达率 λ 和平均批量大小 m) 以及决策阈值 θ , 最小化系统平均AuD与平均DaD的线性加权和。该优化问题可表述为

¹⁾ 该条件主要用于获得解析表达式, 且已被文献[18]用于对状态更新权重的建模与分析之中。当批内状态更新的权重存在相关性时, 可在本文的分析框架下考虑联合分布进行扩展分析。

$$\begin{aligned}
\text{P1: } \min_{\{\lambda, m, \theta\}} J &= \eta \bar{\Lambda}^{\text{Geo}} + (1 - \eta) \bar{\Upsilon}^{\text{Geo}} \\
\text{s.t. C1: } \lambda m q(\theta) &\geq D_{\text{th}} \\
\text{C2: } 1 \leq m &< \mu \lambda^{-1} \\
\text{C3: } \lambda_{\min} \leq \lambda &< \lambda_{\max} \\
\text{C4: } 0 \leq \theta &< 1
\end{aligned} \quad (22)$$

其中, $\eta \in [0, 1]$ 表示权重因子, 用于平衡系统对新鲜度与失真度的需求。约束条件C1为系统最低决策率约束, D_{th} 表示单位时间内最小决策次数, 用于防止由于决策阈值过高导致系统长期处于无决策状态; 约束条件C2为在满足队列稳定性条件下平均批量大小的可行区间约束; 约束条件C3为在满足队列稳定性条件下批到达率的可行区间约束; 约束条件C4为决策阈值的可行区间约束。

定理1 当 $\lambda \in (0, \mu m^{-1})$ 时, 加权和 J 是关于 λ 的凸函数。此外, J 是关于 θ 的单调递减函数。

证明 由式(20)和式(21)可以看出, 批到达率 λ 仅影响平均AuD。因此, 加权和目标函数 J 关于 λ 的优化问题分析可以转化为平均AuD关于 λ 的分析。下面分析平均AuD关于 λ 的性质。

根据式(20), 平均AuD关于 λ 的二阶导数为

$$\frac{\partial^2 \bar{\Lambda}^{\text{Geo}}}{\partial \lambda^2} = \frac{2(m-1)}{m\lambda^3} + \frac{2m^3}{(\mu - \lambda m)^3} \quad (23)$$

当 $0 < \lambda < \mu m^{-1}$ 时, $\partial^2 \bar{\Lambda}^{\text{Geo}} / \partial \lambda^2 > 0$, 因此平均AuD是关于 λ 的凸函数, 进而目标函数 J 也是关于 λ 的凸函数。

由式(20)和式(21)可以看出, 决策阈值 θ 仅影响平均DaD。因此, 加权和目标函数 J 关于 θ 的优化问题分析可以转化为平均DaD关于 θ 的分析。根据式(21), 平均DaD关于 θ 的一阶导数为

$$\frac{\partial \bar{\Upsilon}^{\text{Geo}}}{\partial \theta} = \sigma_0^2 \int_0^\infty \frac{\partial \mathcal{L}_\omega^{\text{cond}}(s; \theta)}{\partial \theta} \frac{1}{m(1 - \mathcal{L}_\omega(s)) + \mathcal{L}_\omega(s)} ds \quad (24)$$

注意到, 上式仅有 $\mathcal{L}_\omega^{\text{cond}}(s; \theta)$ 这一项与 θ 有关。进一步地, 条件LST可表示为

$$\mathcal{L}_\omega^{\text{cond}}(s; \theta) = \frac{\int_0^1 \exp(-sx) f_\omega(x) dx}{\int_0^1 f_\omega(x) dx} \quad (25)$$

根据式(25), $\partial \mathcal{L}_\omega^{\text{cond}}(s; \theta) / \partial \theta$ 可表示为

$$\begin{aligned}
\frac{\partial \mathcal{L}_\omega^{\text{cond}}(s; \theta)}{\partial \theta} &= \frac{f_\omega(\theta)}{\left(\int_0^1 f_\omega(x) dx \right)^2} \\
&\cdot \left(\int_0^1 (\exp(-sx) - \exp(-s\theta)) f_\omega(x) dx \right) \quad (26)
\end{aligned}$$

对于积分区间内任意 $x \in (\theta, 1]$, 有 $\exp(-sx) - \exp(-s\theta) < 0$, 从而可得 $\partial \mathcal{L}_\omega^{\text{cond}}(s; \theta) / \partial \theta < 0$ 。此外, 由于 $0 < \mathcal{L}_\omega(s) < 1$ 且 $m > 0$, 可得 $m(1 - \mathcal{L}_\omega(s)) + \mathcal{L}_\omega(s) > 0$ 。因此, $\partial \bar{\Upsilon}^{\text{Geo}} / \partial \theta < 0$, 即平均DaD关于 θ 单调递减。从而, J 也关于 θ 单调递减。综上, 定理1得证。

由定理1可知, 目标函数 J 是关于批到达率 λ 的凸函数。因此, 在给定平均批量大小 m 和决策阈值 θ 的条件下, 关于 λ 的子问题构成凸优化问题, 可采用黄金分割搜索法求得最优 λ^* 。同时, J 是关于 θ 的单调递减函数, 因此最优解必然在约束边界处取得, 根据约束条件C1可得到最优的决策阈值为 $\theta^* = q^{-1}(D_{\text{th}}(\lambda m)^{-1})$ 。进一步地, 在给定 λ 和 θ 时, 结合约束条件C2可通过枚举法遍历所有可行取值来得到最优平均批量大小 m^* 。基于上述结构性性质, 本文提出了一种交替迭代优化算法来联合求解最优参数。在所提迭代优化算法中, 每次迭代过程均利用当前变量求解相应子问题。其中, 批到达率通过黄金分割法进行搜索, 决策阈值根据目标函数的单调性在约束边界处更新, 平均批量大小则通过枚举有限可行集合获得使目标函数最小的取值。由此可见, 每个变量的迭代更新均不会使加权目标函数值增大。若记第 r 次迭代后的目标函数值为 $J^{(r)}$, 则有 $J^{(r+1)} \leq J^{(r)}$ 。同时, 由于平均AuD和平均DaD均为非负量, 因此, $\{J^{(r)}\}$ 构成有下界的单调非增序列, 故目标函数值可迭代收敛。该算法的时间复杂度为 $\mathcal{O}(R(M + \log(\epsilon_{\text{GS}}^{-1})))$, 其中, R 表示外层迭代次数, M 表示平均批量大小 m 的搜索空间大小, ϵ_{GS} 表示黄金分割法的收敛精度。具体步骤如**算法1**所示。

5 仿真结果

本节通过数值仿真结果来展示批量到达特征下状态更新系统的批到达率、平均批量大小、决策阈值等参数对平均AuD和平均DaD性能的影响。在无特别说明下, 批量内状态更新的权重服从形状参数为 $\alpha = 5$, $\beta = 2$ 的Beta分布, 基准噪声方差为 $\sigma_0^2 = 1$, 单位时间内最小决策次数为 $D_{\text{th}} = 0.3$ 。

图2给出了在给定平均批量大小时, 不同批量

大小变异系数(Batch-size Coefficient Of Variation, BCOV)下的平均AuD随批到达率的变化关系。BCOV定义为批量大小的标准差与均值之比,可用于表征批量大小的波动程度。由图可见,平均AuD随着批到达率的增大呈现出先下降后上升的趋势。当 λ 较小时,批量到达较为稀疏,决策控制单元处于相对空闲的状态,状态更新能够及时获得服务,从而降低系统的平均AuD。随着 λ 的不断增大,系统负载逐渐加重,状态更新在队列中造成堆积导致排队时延增大,进而使平均AuD上升。此外,在相同批到达率下,随着BCOV的升高,平均AuD逐渐增大。这是因为较大的BCOV意味着批量规模的波动性增强,更容易出现大规模批量的突发到达,从而导致队列长度和等待时间增加,最终使得AuD性能劣化。此外,当 λ 较小时,不同BCOV条件下的平均AuD差异较小;而随着 λ 的增大,各曲线之间的差距逐渐扩大。这是因为在高负载条件下,队列积压加剧,此时批量到达的随机波动性对排队过程的影响被显著放大,所带来的服务时间不稳定性更加显著,从而放大BCOV对平均AuD的影响。

图3给出了在给定批到达率和决策阈值时,平均AuD和平均DaD随平均批量大小的变化关系。平均批量大小的变化使得难以将BCOV作为独立变量进行对照。因此,图3选取分别对应低变异性和高变异性的两类批量大小分布,进行对照。可以看出,随着平均批量大小的增大,平均AuD不断增加。这是因为批量规模的增大使得每批量到达带来的瞬时负载显著增加,从而导致系统队列长度和排队时延增大,进而导致平均AuD上升。然而,随着平均批量大小的增加,系统平均DaD逐渐降低。这是由于较大的批量能够提供更多已完成服务的状态更新,通过联合估计可有效降低失真度。此外,相比于确定性分布,几何分布由于具有更高的变异性,更容易产生大批量到达,从而加剧排队延迟,致使决策时新鲜度性能恶化;但同时,其更大的批量波动也带来了更多的状态更新,有助于抑制决策失真。

图4给出了在给定平均批量大小时,不同BCOV条件下平均DaD随着决策阈值的变化关系。可以看出,随着 θ 的增大,平均DaD呈现出逐渐下

算法 1 交替迭代优化算法

输入: 权重因子 η , 批到达率 $\lambda^{(0)}$, 平均批量大小 $m^{(0)}$

(1) 初始化迭代索引 $r = 0$, 迭代停止精度 ϵ_{AO}

(2) 执行迭代循环:

(3) 更新 $\theta^{(r+1)} = q^{-1} \left(D_{th}(\lambda^{(r)} m^{(r)})^{-1} \right)$

(4) 固定 $(\theta^{(r+1)}, \lambda^{(r)})$, 利用枚举法解得 $m^{(r+1)} = \arg \min_m J(\lambda^{(r)}, m, \theta^{(r+1)})$

(5) 固定 $(\theta^{(r+1)}, m^{(r+1)})$, 利用黄金分割法解得 $\lambda^{(r+1)} = \arg \min_{\lambda} J(\lambda, m^{(r+1)}, \theta^{(r+1)})$

(6) 计算当前目标函数值 $J^{(r+1)} = J(\lambda^{(r+1)}, m^{(r+1)}, \theta^{(r+1)})$

(7) 更新 $r \leftarrow r + 1$

(8) 直到: $|J^{(r+1)} - J^{(r)}| < \epsilon_{AO}$

输出: $\lambda^* = \lambda^{(r+1)}$, $m^* = m^{(r+1)}$, $\theta^* = \theta^{(r+1)}$

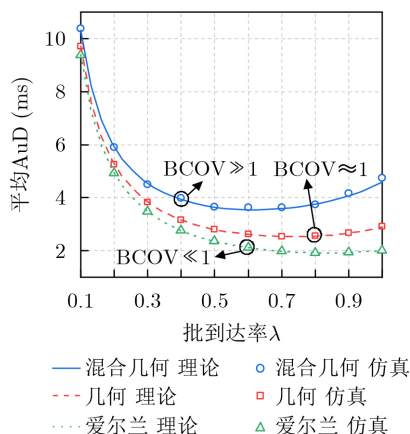


图 2 平均AuD随批到达率变化($\mathbb{E}[N] = 10$)

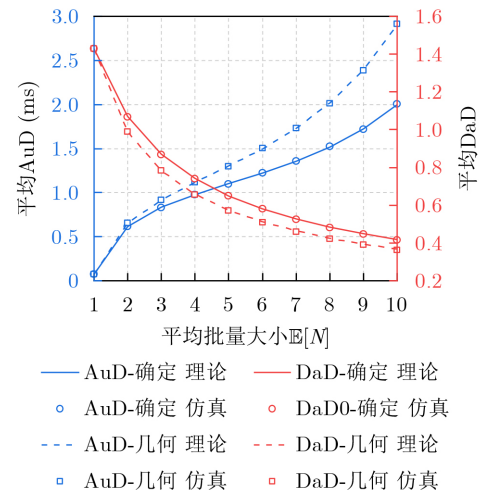


图 3 平均AuD和平均DaD随平均批量大小变化($\lambda = 1, \theta = 0.4$)

降的趋势。这是因为，当 θ 增大时，尽管决策概率降低，但一旦发生决策，决策控制单元能够利用更多已完成服务的状态更新进行联合估计，从而增加估计精度，降低决策失真。进一步比较不同对照组可以发现，在相同决策阈值下，批量大小分布的变异性越大，系统平均DaD越低。其原因在于，较大的批量大小变异性意味着系统更容易出现大规模批量到达。虽然这种随机波动会增强到达过程的突发性，但与此同时，也增加了单次决策前可累积的状态更新数量，从而有利于降低决策失真。相比之下，低变异性的批量分布对应的批量规模更加集中，单次决策前可获得的样本累计效应较弱，因此平均DaD相对更高。

图5给出了平均AuD和平均DaD之间的权衡关系。信息新鲜度与决策失真度之间存在显著的冲突关系；若希望降低平均AuD，就需要更快地触发决策，从而减少可用于联合估计的状态更新积累，导致平均DaD升高；反之，若允许决策发生在更晚的时刻，则决策控制单元能够利用更多已完成服务的

状态更新进行联合估计，从而降低平均DaD，但代价是平均AuD增大。图5表明，利用具有较低复杂度的所提算法可实现落在AuD-DaD的帕累托前沿上的有效求解。此外，还可以看出，不同参数配置会共同影响状态更新的排队积累过程及决策触发过程，从而改变AuD与DaD之间的耦合关系。只有对相关参数进行协同设计，才能有效改善系统在新颖度与失真度之间的性能权衡。

6 结束语

本文针对具有批量到达特征的IoT状态更新系统，以AuD和DaD分别作为决策时新鲜度和失真度的性能指标，在批量大小服从一般分布的条件下推导了系统平均AuD和平均DaD的解析表达式。此外，针对具有典型随机波动特征的几何批量大小分布这一特例，采用交替迭代优化算法，通过联合优化批到达率、平均批量大小和决策阈值以最小化平均AuD与平均DaD的加权和。仿真结果验证了理论分析的正确性，揭示了系统平均AuD和平均DaD随批到达率、平均批量大小和决策阈值等参数的变化规律，并表明本文所提出的参数优化算法能够获得位于AuD-DaD帕累托前沿上的权衡解。本文仅结合状态更新信息的批量到达特征，未来的工作可扩展考虑批量服务机制，深入分析其对AuD与DaD权衡关系的影响；此外，可根据具有不同批量特征业务对新鲜度与失真度的差异化需求，设计灵活的决策机制以实现AuD与DaD之间的自适应权衡。

参考文献

- [1] YATES R D, SUN Yin, BROWN D R, *et al.* Age of information: An introduction and survey[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(5): 1183–1210. doi: [10.1109/JSAC.2021.3065072](https://doi.org/10.1109/JSAC.2021.3065072).
- [2] 陈泳, 蔡跃明, 王萌. 认知物联网短包通信中双向中继系统的信息年龄分析[J]. *电子与信息学报*, 2023, 45(12): 4254–4261. doi: [10.11999/JEIT221377](https://doi.org/10.11999/JEIT221377).
CHEN Yong, CAI Yueming, and WANG Meng. Age of information for short-packet two-way relay system in cognitive IoT network[J]. *Journal of Electronics & Information Technology*, 2023, 45(12): 4254–4261. doi: [10.11000/JEIT221377](https://doi.org/10.11000/JEIT221377). doi: [10.11999/JEIT221377](https://doi.org/10.11999/JEIT221377).
- [3] 肖舒予, 孙兴华, 袁岸珊, 等. 能量收集型短包通信LoRa网络的信息年龄建模与优化[J]. *电子与信息学报*, 2026, 48(1): 145–156. doi: [10.11999/JEIT250814](https://doi.org/10.11999/JEIT250814).
XIAO Shuyu, SUN Xinghua, YUAN Anshan, *et al.* Age of information for energy harvesting-driven LoRa short-packet communication network[J]. *Journal of Electronics & Information Technology*, 2026, 48(1): 145–156. doi: [10.11999/JEIT250814](https://doi.org/10.11999/JEIT250814).

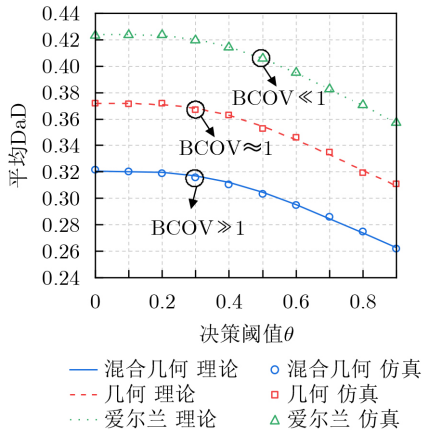


图 4 平均DaD随决策阈值变化($\mathbb{E}[N] = 10$)

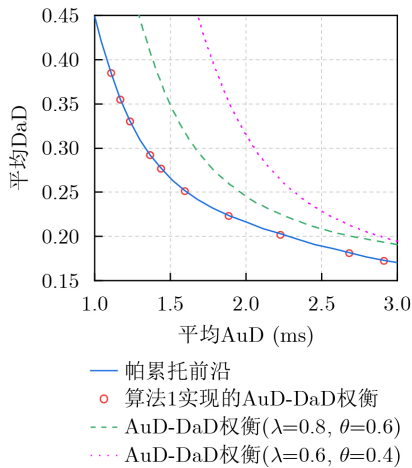


图 5 平均AuD与平均DaD权衡关系

- 11999/JEIT250814.
- [4] DONG Yunquan, CHEN Zhengchuan, LIU Shanyun, *et al.* Age of information upon decisions[C]. 2018 IEEE 39th Sarnoff Symposium, Newark, USA, 2018: 1–5. doi: [10.1109/SARNOF.2018.8720507](https://doi.org/10.1109/SARNOF.2018.8720507).
- [5] FAHIM T E, RABIA S I, ABD EL-MALEK A H, *et al.* Age of information optimization using a hybrid preemptive/non-preemptive discipline[C]. 2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall), Washington, USA, 2024: 1–5. doi: [10.1109/VTC2024-Fall63153.2024.10758006](https://doi.org/10.1109/VTC2024-Fall63153.2024.10758006).
- [6] AKAR N, DOGAN O, and ATAY E U. Finding the exact distribution of (Peak) age of information for queues of PH/PH/1/1 and M/PH/1/2 type[J]. *IEEE Transactions on Communications*, 2020, 68(9): 5661–5672. doi: [10.1109/TCOMM.2020.3002994](https://doi.org/10.1109/TCOMM.2020.3002994).
- [7] DONG Yunquan, CHEN Zhengchuan, LIU Shanyun, *et al.* Age-upon-decisions minimizing scheduling in internet of things: To be random or to be deterministic?[J]. *IEEE Internet of Things Journal*, 2020, 7(2): 1081–1097. doi: [10.1109/JIOT.2019.2950054](https://doi.org/10.1109/JIOT.2019.2950054).
- [8] CHEN Shutong, ZHANG Tianci, CHEN Zhengchuan, *et al.* Minimizing age-upon-decisions in bufferless system: Service scheduling and decision interval[J]. *IEEE Transactions on Vehicular Technology*, 2023, 72(1): 1017–1031. doi: [10.1109/TVT.2022.3202790](https://doi.org/10.1109/TVT.2022.3202790).
- [9] BAO Zhiwei, DONG Yunquan, CHEN Zhengchuan, *et al.* Age-optimal service and decision processes in internet of things[J]. *IEEE Internet of Things Journal*, 2021, 8(4): 2826–2841. doi: [10.1109/JIOT.2020.3020875](https://doi.org/10.1109/JIOT.2020.3020875).
- [10] BAO Zhiwei, HU Yulin, GAN Ming, *et al.* Timeliness of CSMA/CA-based wireless networks with HARQ in the FBL regime: Explicit age characterizations and resource allocation[J]. *IEEE Internet of Things Journal*, 2025, 12(13): 23614–23626. doi: [10.1109/JIOT.2025.3554242](https://doi.org/10.1109/JIOT.2025.3554242).
- [11] GUO Chongtao, LIU Songtao, LIAO Bin, *et al.* AoI-driven power allocation and batch sampling control for V2V status update communications[J]. *IEEE Transactions on Industrial Informatics*, 2024, 20(1): 291–302. doi: [10.1109/TII.2023.3261888](https://doi.org/10.1109/TII.2023.3261888).
- [12] MOLTCHANOV D and GAYDAMAKA A. The impact of tagged and competing traffic properties on peak AoI in 5G systems[J]. *IEEE Internet of Things Journal*, 2025, 12(23): 51761–51771. doi: [10.1109/JIOT.2025.3616508](https://doi.org/10.1109/JIOT.2025.3616508).
- [13] BASTOPCU M and ULUKUS S. Age of information for updates with distortion: Constant and age-dependent distortion constraints[J]. *IEEE/ACM Transactions on Networking*, 2021, 29(6): 2425–2438. doi: [10.1109/TNET.2021.3091493](https://doi.org/10.1109/TNET.2021.3091493).
- [14] LI Jun and ZHANG Wenyi. Asymptotically optimal joint sampling and compression for timely status updates: Age-distortion tradeoff[J]. *IEEE Transactions on Vehicular Technology*, 2025, 74(2): 2338–2352. doi: [10.1109/TVT.2024.3479242](https://doi.org/10.1109/TVT.2024.3479242).
- [15] 王艳芳, 王伟, 董云泉. 面向实时决策的物联网时效与失真性能研究[J]. *浙江大学学报(工学版)*, 2024, 58(4): 664–673,771. doi: [10.3785/j.issn.1008-973X.2024.04.002](https://doi.org/10.3785/j.issn.1008-973X.2024.04.002).
WANG Yanfang, WANG Wei, and DONG Yunquan. Study of timeliness and distortion performance for real-time decision making in IoT[J]. *Journal of Zhejiang University (Engineering Science)*, 2024, 58(4): 664–673,771. doi: [10.3785/j.issn.1008-973X.2024.04.002](https://doi.org/10.3785/j.issn.1008-973X.2024.04.002).
- [16] CHEN Zhengchuan, XU Mingjun, SHE Changyang, *et al.* Improving timeliness-fidelity tradeoff in wireless sensor networks: Waiting for all and waiting for partial sensor nodes[J]. *IEEE Transactions on Communications*, 2023, 71(7): 4151–4164. doi: [10.1109/TCOMM.2023.3277001](https://doi.org/10.1109/TCOMM.2023.3277001).
- [17] ZHANG Yang and HO K C. Localization of transmitters and scatterers by single receiver[J]. *IEEE Transactions on Signal Processing*, 2023, 71: 2267–2282. doi: [10.1109/TSP.2023.3284366](https://doi.org/10.1109/TSP.2023.3284366).
- [18] ZHENG Xi, ZHOU Sheng, and NIU Zhisheng. Urgency of information for context-aware timely status updates in remote control systems[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(11): 7237–7250. doi: [10.1109/TWC.2020.3009881](https://doi.org/10.1109/TWC.2020.3009881).
- 刘 磊: 男, 讲师, 硕士生导师, 研究方向为时敏网络性能分析与优化设计等.
- 靳文凯: 男, 硕士生, 研究方向为宽带无线通信等.
- 张清清: 女, 副教授, 硕士生导师, 研究方向为智能通信、无线网络智能化等.
- 李渝舟: 男, 教授, 博士生导师, 研究方向为智能无线通信、信号处理等.
- 江 帆: 女, 教授, 博士生导师, 研究方向为D2D通信技术、雾计算等.

责任编辑: 余 蓉

Analysis of Age upon Decisions and Distortion at Decisions in IoT Status Update Systems with Batch Arrivals

LIU Lei^① JIN Wenkai^① ZHANG Qingqing^② LI Yuzhou^③ JIANG Fan^①

^①(School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

^②(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

^③(School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430071, China)

Abstract:

Objective The rapid advancement of the Internet of Things (IoT) has made real-time transmission and processing of status updates essential in modern systems, where the freshness of information at system decision epochs play a critical role. In many IoT applications, including smart grids fault detection and industrial IoT (IIoT) cluster monitoring, the arrival process of status updates typically exhibits pronounced batch characteristics, deviating from the conventional single status update arrival pattern. However, most existing studies on Age of Information (AoI) rely on the assumption of single status update arrivals, making them inadequate for capturing the complex queuing dynamics induced by batch arrivals. Apart from information freshness, the distortion at decision epochs is equally critical, as it directly influences the accuracy of decision outcomes. Taking this into account, a fundamental tradeoff between freshness and distortion emerges: waiting for more complete batch information, i.e., collecting more updates within a batch, incurs additional queuing and transmission delays, thereby degrading information freshness, whereas making timely decisions reduces delay at the cost of increased distortion due to incomplete batch information. To address these challenges, this paper investigates the interplay between freshness and distortion in an IoT status update system with batch arrivals, by utilizing Age upon Decisions (AuD) and Distortion at Decisions (DaD) as performance metrics. Analytical expressions for the average AuD and average DaD are derived under a general batch size distribution. Furthermore, for the representative case of geometric batch sizes, an alternating iterative algorithm is proposed to jointly optimize the batch arrival rate, average batch size, and decision threshold with the goal of minimizing the weighted sum of average AuD and average DaD. The results provide valuable theoretical insights and practical design guidelines for IoT status update systems with batch arrivals.

Methods This paper investigates the information freshness and distortion at system decision epochs in an IoT status update system with batch arrivals. Specifically, AuD and DaD are utilized to characterize the freshness and distortion performance, respectively. By leveraging queueing theory, analytical expressions for the average AuD and average DaD are derived under a general batch-size distribution. Building upon this foundation, a typical scenario with geometrically distributed batch sizes is further examined, where the geometric distribution captures the sequential batch formation process with a memoryless stopping mechanism. An alternating iterative optimization algorithm is then devised to jointly optimize the batch arrival rate, average batch size, and decision threshold aiming at minimizing the weighted sum of the average AuD and average DaD.

Results and Discussions Simulation results validate the correctness of the derived theoretical analysis. The average AuD exhibits a non-monotonic trend with respect to the arrival rate, first decreasing and then increasing as the arrival rate grows. In addition, the batch-size coefficient of variation (BCOV) has a significant impact on the average AuD, where a smaller BCOV leads to improved freshness performance. Furthermore, under high-load conditions, queue backlogs become severe. In such cases, the impact of stochastic fluctuations in batch arrivals on the queuing process is significantly amplified, resulting in greater service-time variability and, consequently, a more pronounced influence of BCOV on the average AuD (Fig. 2). As the mean batch size increases, both the system queue length and queuing delay grow substantially, leading to a higher average AuD. Meanwhile, the decision unit can exploit a larger number of status updates for joint estimation, thereby reducing the average DaD (Fig. 3). Moreover, the average DaD decreases as the decision threshold increases, indicating that more status updates are incorporated into joint estimation, which improves estimation accuracy.

The BCOV also affects the number of status updates available for joint estimation; specifically, a larger BCOV can enhance estimation accuracy (Fig. 4). The joint optimization results show that the solutions obtained by the proposed iterative algorithm lie on the Pareto frontier, demonstrating its effectiveness. In contrast, the performance achieved under fixed arrival rate and decision threshold is significantly inferior to the Pareto frontier, highlighting the advantage of jointly optimizing system parameters (Fig. 5).

Conclusions This paper focuses on an IoT status update system characterized by batch arrivals, where AuD and DaD are leveraged as performance metrics to quantify decision freshness and distortion, respectively. Analytical expressions for the average AuD and average DaD are derived under a general batch-size distribution. Furthermore, for the special case of geometrically distributed batch sizes with inherent stochastic variability, an alternating iterative optimization algorithm is proposed to jointly optimize the batch arrival rate, average batch size, and decision threshold, with the objective of minimizing the weighted sum of the average AuD and average DaD. Simulation results validate the accuracy of the theoretical analysis and reveal the impact of key system parameters, such as batch arrival rate, average batch size, and decision threshold, on the average AuD and average DaD. Moreover, the results demonstrate that the proposed low-complexity iterative algorithm can effectively obtain solutions to the weighted-sum optimization problem that lie on the Pareto frontier of the AuD–DaD tradeoff. This paper considers only the batch arrival characteristics of status updates. Future work can be extended to incorporate batch service mechanisms, enabling a more comprehensive analysis of their impact on the tradeoff between AuD and DaD. In addition, flexible decision-making schemes can be designed to achieve adaptive tradeoffs between AuD and DaD, in accordance with the heterogeneous requirements on freshness and distortion across applications with different batch characteristics.

Key words: Internet of Things (IoT); Status update system; Batch arrival; Age upon Decisions (AuD); Distortion at Decisions (DaD)