

Article

The Latency Performance Analysis and Effective Relay Selection for Visible Light Networks

Baozhu Yu  and Xiangyu Liu *

School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110158, China; yubz19@mails.jlu.edu.cn

* Correspondence: liuxy@sylu.edu.cn

Abstract: In visible light communication (VLC), the precise latency evaluation of wireless access networks and the efficient forwarding strategy of core networks are the crux for end-to-end reliability provisioning. Leveraging martingale theory, an elegant latency-bounded reliability analysis framework is studied for the VLC network. Considering the characteristic that VLC links are easy to block, the martingale of the service process is constructed. Based on the time shift features, the martingale process related to latency is proposed for the VLC system, which models the influence of entanglement between aggregate arrivals and random service on latency. A stopping time event about latency is defined. Renting the stopping time theory, a tight upper bound of the unreliability with regard to latency is derived. In the core network, a dynamic forward backhaul framework is proposed, which uses the relay selection algorithm based on back-pressure theory to improve data transmission quality. The theoretical latency-bounded reliability matches the simulation results well, which verifies the effectiveness of the proposed analysis framework, and the proposed relay selection algorithm can also improve network performance under data-intensive transmission.

Keywords: latency; statistical reliability; martingales; aggregate traffic; forwarding scheme



Citation: Yu, B.; Liu, X. The Latency Performance Analysis and Effective Relay Selection for Visible Light Networks. *Sensors* **2024**, *24*, 2748. <https://doi.org/10.3390/s24092748>

Academic Editor: Hazer Inaltekin

Received: 27 February 2024

Revised: 19 April 2024

Accepted: 22 April 2024

Published: 25 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In 5G and beyond, massive businesses with strict reliability requirements surge, which expects extensive bandwidth provisioning. This will lead to more severe spectrum shortages. Visible light communication (VLC) is a potential paradigm to ease the dilemma of radio frequency (RF) resources in the next-generation communication systems [1,2]. Benefiting from the characteristics of light, VLC can provide high transmission rates, ultra-wide frequency bands, high-energy efficiency, low cost, and indoor full coverage. It can be deployed as indoor high-speed data links for personal area networks and RFs in non-friendly environments. In critical scenarios, such as remote control in medical operating rooms and factory automation under strong magnetic interference, VLC is expected to embrace massive services with traffic possessing strong randomness, heterogeneity, and burstiness [3]. The data require the reliability of latency within the millisecond level to reach 99.99–99.999% [4,5]. For traffic featuring complex characteristics transmitted on visible light channels, we should explore an effective and precise evaluation method of reliability with regard to latency, which can reveal the influence of entanglement between visible light data traffic and visible light transmission schemes. Meanwhile, an efficient routing strategy needs to be explored for the VLC network.

In the VLC wireless network, the light-emitting diode (LED) access point (AP) carries packet flows from different services, which possess heterogeneous characteristics [6,7]. The convergence of heterogeneous visible light data flows intensifies the uncertainty of the aggregate traffic. In [8], the aggregate traffic was composed of multiple 2-state Markov-modulated Bernoulli processes (MMBP-2) and was tackled as one MMBP-2 by the Kronecker product operation. The proposed method is not of universal applicability and is

unable to provide a theoretical result of the latency-bounded reliability. The stochastic network calculus (SNC) theory, effective bandwidth (EB), and effective capacity (EC) theories provide classical analysis methods of latency-bounded reliability. The union-bound inequality, as the key enabler in SNC, only supports capturing a loose upper bound of unreliability [9]. In the EB/EC framework, the stochastic traffic is tackled as a constant flow [10,11], which always triggers rougher theoretical results of latency performance, especially for bursty traffic.

Martingale, a random process throughout modern probability theory, has demonstrated great superiority in the precise analysis of the statistical reliability regarding latency. The theoretical upper bounds of unreliability derived in [9,12] are more precise than the existing conclusions. This martingale-based analysis framework is a milestone. In [13,14], the latency-bounded unreliability of the Markov-modulated on-off (MMOO) arrival process was analyzed in the ALOHA access scheme. Based on the results, the bandwidth resources were allocated reasonably. More importantly, martingales have been explored to support the evaluation of the end-to-end reliability of latency. In [15], a precise analysis framework of the end-to-end latency-bounded unreliability was introduced, where a multi-hop routing path was considered and the traffic was modeled as a Markov-modulated process. Then, the proposed method was extended in the 6G network scenarios [16]. For the system adopting the THz wireless access scheme, the end-to-end latency performance was analyzed for the traffic generated from virtual reality. In [17], we constructed an analysis framework of latency performance for the aggregate traffic, which was constituted by the heterogeneous Markov-modulated flows. The martingale of the queuing length was defined, and the latency-bounded unreliability was derived.

These remarkable conclusions have inspired our work on the statistical latency QoS analysis based on martingale theory. The framework in [17] is theoretical without specific application scenarios. In this paper, the latency analysis in the VLC network is focused on. The differences with [17] can be summarized as follows. (1) The cross-layer service process is modeled, which is the improvement of the proposed method in [17]. Specifically, in the physical layer, the Lambertian sources are considered and the channel gain is depicted. The Shannon Theorem is leveraged to model the achievable transmission rates, which achieves the mapping between the physical layer and the data link layer. We focus on the characteristic that the VLC links are easy to block. Blocking the line of sight (LoS) between the VLC AP and the targeted terminal at each time slot is considered a random variable, which triggers the stochastic features of the service process and further impacts the latency of the aggregate traffic. (2) Meanwhile, the martingale process related to latency is a proposed novelty, which is another way to analyze latency different from [17]. It introduces the statistical characteristics of the service scheme and data traffic into martingale parameters, which reveals the influence of random blocks of the VLC link and the burstiness of aggregate traffic on the latency. (3) The stopping time event about latency is defined, which is the time point when the system latency exceeds the threshold. Applying the stopping time theory to the latency-related martingale, the upper bound of the unreliability with regard to latency is captured. (4) More importantly, the end-to-end reliability provisioning is investigated in this paper, not just the air interface latency analysis.

Further, for the aggregate traffic forwarded by the VLC AP, an efficient routing algorithm is proposed based on the back-pressure theory in the core network. Some studies have investigated the utilization of back-pressure theory to improve network performance, specifically within the Transmission Control Protocol domain [18]. Additionally, alternative research has concentrated on employing back-pressure algorithms to facilitate relay selection in wireless sensor networks and wireless multi-hop networks [19]. Although the source node (LED access point) can adopt the classic back-pressure algorithm to promote the network capacity, the classic algorithm results in significant packet delays and poor node energy. Therefore, combining the latency performance analysis, we design a new link weight calculation method to make routing and scheduling decisions, which considers the data packet's recent node record, the neighbor node's remaining energy status, and the

queue latency. Meanwhile, we note that no prior research has explored the integration of back-pressure theory in visible light network systems.

To summarize, in this paper, considering the VLC APs serve the aggregate traffic with burstiness, a precise analysis framework of the latency-bounded reliability is proposed. In the core network, an efficient routing algorithm is introduced for the complex aggregate traffic. The contributions can be concluded as follows.

- The service process of the VLC network and the aggregate arrival process are modeled in the martingale domain. For the VLC link with a random block, the i.i.d. service process provided by the LED AP is considered. Through mapping the statistical features of random block behavior in the martingale parameters, the martingale of the service is constructed. Leveraging the relation between the spectral radius and the eigenvector in Markov chains, martingales of Markov-modulated arrival processes are defined.
- We propose a martingale process related to latency, which models the impact of entanglement between complex arrivals and random service on the latency. A stopping time point at which the latency violates the threshold is considered. Based on the stopping time theory, the complementary cumulative distribution function of latency is obtained for the VLC network. We evaluate the latency in terms of burstiness, which is measured by the squared coefficient of variation.
- When the LED AP node data reach the core network, a dynamic back-pressure algorithm based on energy and latency from aggregate traffic is proposed, which designs a new link weight calculation method to make routing and scheduling decisions. The modified back-pressure algorithm trades off three factors, including the data packet's recent node record, the neighbor node's remaining energy status, and the queue latency, to boost the network performance and data transmission quality. Meanwhile, our scheme is the first research to apply back-pressure theory in the visible light network.

2. The Network Model and the Queuing System

We consider a multi-service uplink communication scenario, as shown in Figure 1. The data flow from different services (business) is statistically heterogeneous. That is, diverse models are adopted to describe the randomness of flows. The data packets belonging to different flows could access the same VLC AP, where these flows compose the aggregate traffic. The targeted AP supplies the transmission rates for the aggregate traffic according to a specific scheme. Because of the randomness of the VLC channel and the block, the transmission process provided for the aggregate traffic is stochastic. Thus, the backlog and latency are triggered in the buffer of the targeted VLC AP.

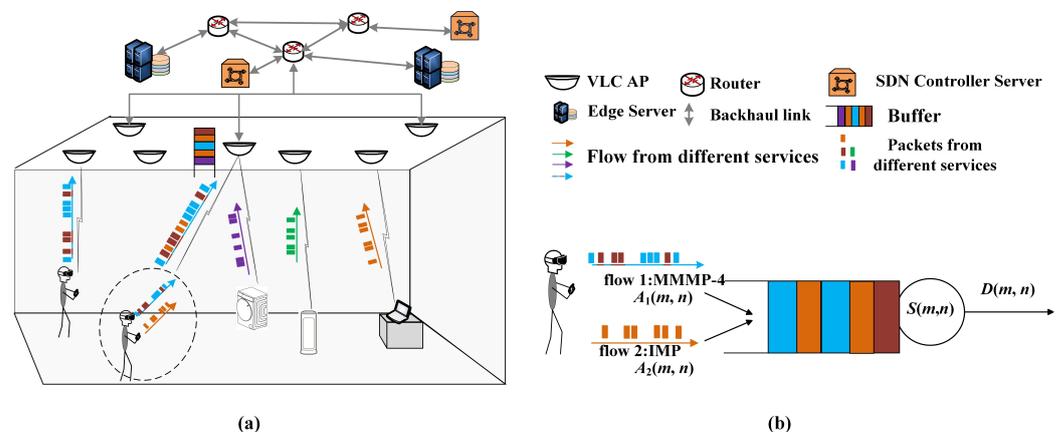


Figure 1. Scenario diagram. (a) is the network scenario. (b) is the corresponding queuing model.

A. The queuing system of aggregate traffic

In the VLC access network, the target terminal is associated with the nearest VLC AP. This terminal carries two services where the generated packet flows are heterogeneous. Thus, the aggregate traffic is transmitted from this terminal, as shown in Figure 1a. We consider the block of LoS between the VLC AP, and this terminal obeys a Bernoulli distribution of parameter $1 - p$. Then, the channel gain can be modeled as

$$hv = \begin{cases} \frac{(b+1)Ap}{2\pi d^2} \cos^b(\varphi^{ir}) g_{of} f(\varphi^{in}) \cos(\varphi^{in}), & p \\ 0, & 1 - p \end{cases}, \quad (1)$$

where b is the Lambertian index, and $b = -1/\log_2(\cos(\vartheta_{1/2}))$. $\vartheta_{1/2}$ is the half-intensity radiation angle. Ap is the area of reviewer photodetector (PD). d is the distance between the AP and the terminal. g_{of} is the gain of the optical filter. φ^{in} and φ^{ir} are the angle of incidence and irradiance between VLC AP and the terminal, respectively. $f(\varphi^{in})$ is the optical concentrator gain, which is a function of φ^{in}

$$f(\varphi^{in}) = \begin{cases} \frac{r^2}{\sin^2(\Theta)}, & 0 \leq \varphi^{in} \leq \Theta \\ 0, & \varphi^{in} > \Theta \end{cases} \quad (2)$$

where r is the refractive index, and Θ is the semi-angle of the field of view (FoV) of the PD.

Based on the Shannon theorem, the achievable transmission rate between AP and the terminal, which is defined as $R^{VLC}(n)$, can be given as

$$R^{VLC}(n) = W_{VLC} \times \log_2 \left(1 + \frac{p(n)hv^2}{\sigma_{VLC}^2} \right) [bits/s], \quad (3)$$

where W_{VLC} is the system bandwidth, and $p(n)$ is the transmission power. σ_{VLC}^2 is the noise power.

The access process of the packets from aggregate traffic can be modeled as a queuing system, as is presented in Figure 1b. These two flows constituting the aggregate traffic are independent and described as the four-state Markov-modulated multinomial process (MMMP-4) and the interrupted multimomial process (IMP), respectively, which are introduced in part B in detail. These are the arrival processes of the queuing system. The access process is modeled as the service process provided by the AP. Because latency is often described in units of packets, the arrival process and the service process are modeled from the perspective of packets/slots. Without loss of generality, the length of packets is assumed to be fixed. The flows from different services have the same priority. Thus, the first-in-first-out (FIFO) scheduling scheme is adopted in the buffer.

In Figure 1b, $A_i(m, n)$, $i = 1, 2$ is the accumulated arrival processes in $[m, n]$ of flow i . $A_i(m, n)$ can be written as $A_i(m, n) = \sum_{k=m}^n a_i(k)$, where $a_i(k)$ is the number of arrival packets of flow i at time slot k . The arrival process $\{a_1(n), n \geq 0\}$ is modeled as an MMMP-4, and $\{a_2(n), n \geq 0\}$ is modeled as an IMP, which are bursty and heterogeneous. Let $A(m, n)$ denote the aggregate accumulated arrival process from slot m to slot n . Similarly, the accumulated service process $S(m, n)$ provided by the VLC AP is also a bivariate process, $S(m, n) = \sum_{k=m}^n s(k)$, where $\{s(n), n \geq 0\}$ is the service process. $s(n)$ is the number of the served packets at slot n . In this paper, $s(n)$ is a random variable following a Bernoulli distribution with successful probability p and service rate C . $s(n)$ can be given by

$$s(n) = \begin{cases} C & p \\ 0 & 1 - p \end{cases} \text{ packets/slot}, \quad (4)$$

where $C = \lfloor \frac{R^{VLC}(n)T}{L} \rfloor$. T s/slot is the duration of a slot and L bits/packet is the length of a packet. To simplify the subsequent analysis, C is assumed as a constant.

Based on SNC, the departure process $D(0, n)$ can be defined by $A(0, n)$ and $S(0, n)$ using $(\min, +)$ convolution. The backlog process $\{Q(n), n \geq 0\}$ in the buffer is defined as

$$Q(n) = \sup_{n \geq 0} \{A(0, n) - S(0, n)\}, \quad (5)$$

and the latency $W(n)$ at slot n is

$$W(n) = \inf\{k \geq 0 : A(0, n - k) \leq D(0, n)\}. \quad (6)$$

B. The heterogeneous arrival models

The MMMP-4 is proposed to describe the interweaving arrival characteristics of packet flows from different services in a terminal. A 2D Markov chain is considered, which is denoted as $\{X_1(n), Z_1(n)\}$. $X_1(n) = \{U_1, V_1\}$ represents the service that generates the packets. $Z_1(n) = \{N, Y\}$ models whether there are packets generated at time slot n . $Z_1(n) = N$ means that no packet is arriving, while $Z_1(n) = Y$ means that packets are generated. Corresponding to U_1 and V_1 , packet arrival probabilities at each slot are λ_1 and γ_1 , respectively. If service U_1 generates packets at slot n , the number of packets is R_{M1} , and in the other case, it is R_{M2} . Thus, the state space of the MMMP-4 model is $\{(U_1, N), (U_1, Y), (V_1, N), (V_1, Y)\}$. The state transition probabilities between two services, U_1 and V_1 , are p_1 and q_1 , respectively. Based on [17], the state transition matrix of the 2D Markov chain is defined as \mathbf{T}_1 .

The IMP model is proposed to embody the bursty and sporadic arrival characteristics of the small data service. It is also a 2D Markov chain, which is denoted as $\{X_2(n), Z_2(n)\}$. $X_2(n) = \{U_2, V_2\}$. V_2 represents the service that generates the packets. In state U_2 , there are no packets generated. $Z_2(n) = \{N, Y, I\}$ describes whether there are packets at slot n . $Z_2(n) = N$ represents that no packet arrives in state V_2 , and $Z_2(n) = Y$ represents that R_{I2} packets arrive in state V_2 . It is worth noting that $Z_2(n) = I$ means that no packets arrive in idle state U_2 with probability one. The state space of the IMP model is $\{(U_2, I), (V_2, N), (V_2, Y)\}$. The state transition probabilities between U_2 and V_2 are p_2 and q_2 , respectively. In V_2 , the packet generation process is stochastic and follows a multinomial distribution with the arrival probability γ_2 and the number of arrival packets R_{I2} . The state transition matrix of the IMP is \mathbf{T}_2 . The MMMP-4 model and IMP model are shown in [17].

3. Martingale-Based Latency-Bounded Reliability Analysis

The arrival and service processes are described by martingales in this part. The upper bound of the latency-bounded unreliability is analyzed based on the stopping theory.

A. Martingale constructions of arrival and service processes

Firstly, we introduce the definition of martingales.

Definition 1 (martingales). Let $(\Omega, \mathcal{F}_\infty, P)$ be a probability space and $\{\mathcal{F}_t\}_{t \geq 0}$ be a filtration, i.e., a non-decreasing sequence of σ -fields of \mathcal{F}_∞ , $\mathcal{F}_t \subset \mathcal{F}_{t+1}$, $t \geq 0$. A random sequence $\{X(t), t \geq 0\}$ is adapted to $\{\mathcal{F}_t\}_{t \geq 0}$. That is, $X(t)$ is measurable with respect to \mathcal{F}_t , $\forall t \geq 0$. If

- (1) $E[|X(t)|] < \infty, \forall t \geq 0$
- (2) $E[X(t+1)|\mathcal{F}_t] = X(t), \forall t \geq 0$,

holds, then $\{X(t), t \geq 0\}$ is a martingale sequence.

Further, if $E[X(t+1)|\mathcal{F}_t] \leq X(t)$, then $\{X(t), t \geq 0\}$ is a supermartingale.

As a powerful mathematical method, martingales can provide us with a framework to depict heterogeneous arrivals with burstiness. Martingale processes are inclusive, which shows that other random processes are transformed as martingale processes by martingale construction. Meanwhile, martingale can provide a modularized description method for arrival and service processes, which contributes to the model, analyzes the features of the backlog process (the difference between two random processes), and further enables the

analysis of latency for aggregate traffic. It is noted that the inequality formulary of martingales supports the upper bound analysis of unreliability with regard to latency naturally. We construct the martingale processes of arrival and service, respectively, as follows.

Definition 2 (Martingale of arrival). Consider a Markov process $\{x(n), n \geq 0\}$, which is \mathcal{F}_t -measurable. The state space is $\mathcal{N} = \{1, 2, \dots, N\}$. The state transition matrix is $\mathbf{T} = [\mathbf{T}_{i,j}]_{N \times N}$. $\mathbf{T}_{i,j} = P\{x(t+1) = j | x(t) = i\}$. The arrival process is Markov modulated, i.e., $a(t) = f(x(t))$ and $A(m, n) = \sum_{j=m}^n f(x(j))$, $n \geq m \geq 0$. Define a θ -transform for \mathbf{T} as $\mathbf{T}^\theta = [\mathbf{T}_{i,j}^\theta]_{N \times N}, \forall \theta > 0$. $\mathbf{T}_{i,j}^\theta = \mathbf{T}_{i,j} e^{\theta f(x(t+1))}$. The spectral radius of \mathbf{T}^θ is $sp(\mathbf{T}^\theta)$, $sp(\mathbf{T}^\theta) > 0$, and the corresponding eigenvector is $\mathbf{V}^a = [\mathbf{V}_i^a]_{N \times 1}, \mathbf{V}^a \succ 0$. The random process

$$M^a(m, n) = \mathbf{V}_{x(n)}^a e^{\theta[A(m,n) - (n-m)Ka]}, \quad (7)$$

is constructed as a martingale relative to $\{\mathcal{F}_t\}_{t \geq 0}$. Then, $M^a(m, n)$ is admitted as a martingale of arrival $\{a(n), n \geq 0\}$. The function Ka is

$$Ka = \frac{\log sp(\mathbf{T}^\theta)}{\theta}. \quad (8)$$

We use the definition of martingales to prove Definition 2.

Proof. It is obvious that $sp(\mathbf{T}^\theta) < \infty$ and $\mathbf{V}_{x(t)}^a < \infty$. Thus, $E[|M^a(m, n)|] < \infty, \forall n \geq 0$ holds. We can write that

$$\begin{aligned} & E[\mathbf{V}_{x(n+1)}^a e^{\theta[A(m,n+1) - (n+1-m)Ka]} | \mathcal{F}_n] \\ &= e^{\theta[A(m,n) - (n-m)Ka]} e^{-\theta Ka} \\ & \quad \times \sum_{x(n+1) \in \mathcal{N}} \mathbf{V}_{x(n+1)}^a e^{\theta f(x(n+1))} P\{x(n+1) | x(n)\} \\ & \stackrel{(a)}{=} e^{\theta[A(m,n) - (n-m)Ka]} e^{-\theta Ka} \\ & \quad \times \sum_{x(n+1) \in \mathcal{N}} \mathbf{V}_{x(n+1)}^a \mathbf{T}_{x(n), x(n+1)}^\theta \\ &= e^{\theta[A(m,n) - (n-m)Ka]} e^{-\theta Ka} (\mathbf{T}^\theta \mathbf{V}^a)_{x(n)}^a \\ & \stackrel{(b)}{=} e^{\theta[A(m,n) - (n-m)Ka]} e^{-\theta Ka} sp(\mathbf{T}^\theta) \mathbf{V}_{x(n)}^a \\ &= M^a(m, n) \end{aligned} \quad (9)$$

(a) is from the basic operation of conditional expectation and the definition of $\mathbf{T}_{i,j}^\theta$. $(\mathbf{T}^\theta \mathbf{V}^a)_i$ is the i th element of vector $(\mathbf{T}^\theta \mathbf{V}^a)$. Then, (b) holds due to the relation between the eigenvector and eigenvalue, i.e., $\mathbf{T}^\theta \mathbf{V}^a = sp(\mathbf{T}^\theta) \mathbf{V}^a$, and the definition of Ka . \square

For the MMMP-4 and IMP models, the martingale parameters are $\mathbf{V}_{x_1(t)}^a, Ka_1$, and $\mathbf{V}_{x_2(t)}^a, Ka_2$ respectively, which are related to \mathbf{T}_1 and \mathbf{T}_2 . The corresponding specific formulas can refer to [17].

Similar to the martingale of arrival, the martingale of the i.i.d. service process is constructed as described in Definition 3.

Definition 3 (Martingale of service). The service process is $\{s(n), n \geq 0\}$, which is \mathcal{G}_n -measurable. The moment generating function is $\Psi(\theta) = E[e^{\theta s(n)}] < \infty, \forall \theta > 0$. The accumulated service process of $s(n)$ is $\{S(m, n), n \geq m \geq 0\}$. Then, the random process

$$M^s(m, n) = \mathbf{V}_{s(n)}^s e^{\theta[(n-m)Ks - S(m,n)]}, \quad (10)$$

is constructed as a martingale relative to $\{\mathcal{G}_n\}_{n \geq 0}$. $M^s(m, n)$ is admitted as a martingale of service $\{s(n), n \geq 0\}$. The function Ks is

$$Ks = -\frac{\Psi(-\theta)}{\theta}. \quad (11)$$

Proof. According to $\Psi(\theta) < \infty, \forall \theta > 0$, we can get $E[|M^s(m, n)|] < \infty, \forall n \geq 0$. Meanwhile,

$$\begin{aligned} & E \left[V_{s(n)}^s e^{\theta[(n+1-m)Ks - S(m, n+1)]} | \mathcal{G}_n \right] \\ & \stackrel{(a)}{=} e^{\theta[(n-m)Ks - S(m, n)]} e^{\theta Ks} E \left[V_{s(n+1)}^s e^{-\theta s(n+1)} \right] \\ & \stackrel{(b)}{=} e^{\theta[(n-m)Ks - S(m, n)]} e^{\theta Ks} V_{s(n)}^s \left[e^{-\theta s(n+1)} \right] \\ & = M^s(m, n). \end{aligned} \quad (12)$$

(a) is because of the i.i.d. property. To support formula (b), $V_{s(n)}^s$ is set as a constant. The proof is completed based on $\Psi(-\theta) = E[e^{-\theta s(n+1)}]$. \square

Based on the characteristics of martingale processes, the backlog process can be modeled in the martingale domain, which facilitates the analysis of the latency performance.

Definition 4 (Martingale of backlog). *The martingales of arrival and service are defined as Definition 2 and 3. Define the product of (7) and (10) as the martingale of backlog $M^L(n)$ by the restriction of θ^* ,*

$$M^L(n) = \prod_{i=1}^n \mathbf{V}_{x_i(n)}^a V_{s(n)}^s e^{\theta^*(A(m, n) - S(m, n))}, \quad (13)$$

where $\theta^* = \sup\{\theta > 0 : Ka_1 + Ka_2 \leq Ks\}$.

B. The latency-bounded unreliability

Theorem 1. *Consider a link for the aggregate traffic. The service process is modeled as $\{s(n), n \geq 0\}$. The aggregate traffic is $\{a_1(n) + a_2(n), n \geq 0\}$. The martingale of arrival is defined by Definition 2, and the martingale of service is in Definition 3. Then, the latency-bounded unreliability of the link holds for $\forall k > 0$,*

$$P\{W(n) \geq k\} \leq \frac{E[\mathbf{V}_{x_1(0)}^a] E[\mathbf{V}_{x_2(0)}^a] E[V_{s(0)}^s]}{H} e^{-\theta^* Ks k}, \quad (14)$$

where $H = \min\{\mathbf{V}_{x_1(n)}^a \mathbf{V}_{x_1(n)}^a V_{s(n)}^s : a_1(n) + a_2(n) - s(n) > 0\}$. k is the latency threshold.

The proof of Theorem 1 is based on the stopping time theory of martingale, as shown in Theorem 2 of [17].

Proof. Define a stopping time T for the martingale of backlog $M^L(n)$, which is the first time that the backlog exceeds threshold σ , i.e., $T = \min\{n : A(0, n) - S(0, n) \geq \sigma\}$. Let $T \wedge n = \min\{T, n\}$. For $M^L(n)$, Theorem 2 is used. The probability of system backlog exceeding threshold σ can be obtained

$$P\{Q(n) \geq \sigma\} \leq \frac{E[\mathbf{V}_{x_1(0)}^a] E[\mathbf{V}_{x_2(0)}^a] E[V_{s(0)}^s]}{H} e^{-\theta^* \sigma}. \quad (15)$$

The detailed derivation process can refer to [17]. Based on the definition of the departure process, the latency-bounded probability, $P\{W(n) \geq k\}$ can be transformed as follows.

$$\begin{aligned} & P\{W(n) \geq k\} = P\{A(0, n - k) \geq D(0, n)\} \\ & \leq P\left\{ \max_{n-k \geq l} \{A(l, n - k) - S(l, n)\} \geq 0 \right\} \\ & = P\left\{ \max_{n-k \geq l} \left\{ A_1(l, n - k) - (n - k - l)Ka_1 + A_2(l, n - k) - (n - k - l)Ka_2 \right. \right. \\ & \quad \left. \left. + (n - l)Ks - S(l, n) \right\} \geq kKs \right\} \end{aligned} \quad (16)$$

According to the time shift feature of martingales, it is obvious that

$$M_l^{L-s}(n-k) = \prod_{i=1}^2 \mathbf{V}_{x_i(n-k)}^a V_{s(n-k)}^s e^{\theta^*(A(l, n-k) - S(l, n))}, \quad (17)$$

is also a martingale process. $M_l^{L-s}(n-k)$ can be regarded as a martingale related to latency. To complete the latency analysis, the stopping time event about latency is defined as $SW = \{W(n) \geq k\}$ and can be transformed as

$$SW = \{A_1(l, n-k) - (n-k-l)Ka_1 + A_2(l, n-k) - (n-k-l)Ka_2 + (n-l)Ks - S(l, n) \geq kKs\} \quad (18)$$

Then, the first time slot that SW occurs is the stopping time T_W .

$$T_W = \inf\{n > 0 : A_1(l, n-k) - (n-k-l)Ka_1 + A_2(l, n-k) - (n-k-l)Ka_2 + (n-l)Ks - S(l, n) \geq kKs\} \quad (19)$$

Combined with the derivation of (15) and renting the stopping time theory for the martingale process $M_l^{L-s}(n-k)$, we can derive that

$$E[M_l^{L-s}(0)] = \prod_{i=1}^2 E[\mathbf{V}_{x_i(0)}^a] E[V_{s(0)}^s] \geq He^{\theta^* Ksk} P\{T_W \leq n\} \quad (20)$$

When $T_W \rightarrow \infty$, we can obtain $P\{W(n) \geq k\} = P\{T_W \geq \infty\}$. The proof is completed. \square

4. Relay Selection for Visible Light Network

We assume that the visible light network is represented by $G = (V, L)$. V is the set of all nodes in the network, and L is the set of all links in the network. The nodes are static and the communication links are bidirectional. The topology changes as nodes die or links fail. All visible light nodes have the same configuration.

We set the network operation by time sharing, i.e., $t \in \{0, 1, 2, \dots\}$. At each time slot t , the back-pressure algorithm activates a set of interference-free links in the network. When the new data arrive in the network, each node makes routing and transmission scheduling decisions, and the visible light packets are delivered to the appropriate destinations. The back-pressure algorithm divides the visible light packet types according to the difference in the destination. Each packet in the network records information about the source and destination nodes. $Q_i^d(t)$ is denoted as the total number of packet types on node i at time slot t . If $d = i$, then $Q_i^d(t) = 0$, indicating that node i is the destination node for the d number packets. Each node maintains up to N (total number of nodes) queues for storing packets arriving at different destination nodes.

π is denoted as the set of link scheduling, and Γ denotes the set of all schedulable links under the interference condition. The back-pressure algorithm used in the visible light network is as follows.

STEP-1: Calculating the weights of the links. The link (m, n) is any link in the network, the node n is a neighboring node of the node m within the network communication range, and d is the destination node in the network. We calculate the backlog difference of different kinds of packets between node m and node n , while the value of the backlog difference must be the positive integer, as follows,

$$diff_{mn}^d(t) = \max\{Q_m^d(t) - Q_n^d(t), 0\} \quad (21)$$

where $diff_{mn}^d(t)$ denotes the backlog difference of the packet queue of node d between node m and node n at time slot t . $Q_m^d(t)$ and $Q_n^d(t)$ are the queue length values for node m and node n , respectively.

We select the maximum value of all the queue backlog differences as the weight of the link, as follows,

$$w_{mn}(t) = \max_{d \in D} \{diff_{mn}^d(t)\} \quad (22)$$

where D is the set of all destination nodes. $w_{mn}(t)$ is a maximum value in the set of all different kinds of packet queue backlog differences $diff_{mn}^d(t)$. At time slot t , we make $c_{mn}^*(t)$ as the optimal packet with $w_{mn}(t)$ value for link (m, n) .

STEP-2: Selecting schedulable links. We set the optimization function as follows,

$$\pi(t) = \arg \max_{\pi \in \Gamma} \sum_{m,n} \mu_{mn}(t) w_{mn}(t) \quad (23)$$

where Γ denotes the set of all schedulable links under the interference model. μ_{mn} is the link transmission rate. $\pi(t)$ is the set of optimal transmission links.

STEP-3: Selecting routing paths. At time slot t , link (m, n) under the routing and scheduling policy, we transmit the optimal kind of packet $c_{mn}^*(t)$ from node m to node n .

The back-pressure algorithm for the relay selection in the visible light network is shown in Algorithm 1. We illustrate the workflow of the algorithm with a simple example. Figure 2 shows a visible light communication uploading network. There are three relay nodes C , D , and E , and three different destinations X , Y , and Z . From the figure, we can see the backlog number, i.e., the queue length value, corresponding to each packet type at each node, as shown in Table 1. We assume that the link (C, D) rate is $\mu_{cd} = 4$ packets/slot, and the link (C, E) rate is $\mu_{ce} = 2$ packets/slot. The neighboring node C at time slot t are only D and E .

Algorithm 1 The back-pressure algorithm for the relay selection in visible light network

Input: $G = (V, L)$, m, μ \ \ Web G , Node m , Link rate μ

Output: Routing strategy

```

1: procedure Original back-pressure
  ▷ Calculate Weight
2: for all links  $(m, n) \in L$  do
3:   for all  $d \in D$  do \ \ Packet type  $d$ 
4:      $diff_{mn}^d(t) \leftarrow Q_m^d(t) - Q_n^d(t)$ 
5:   end for
6:    $w_{mn}(t) \leftarrow \max_{d \in D} \{diff_{mn}^d(t)\}$ 
7:
8:    $c_{mn}^*(t) \leftarrow \arg \max_{d \in D} \{diff_{mn}^d(t)\}$  \ \ Packet types with maximum weight
9: end for
  ▷ Link Scheduling
10: for all  $\pi \in \Gamma$  do
11:    $sum_{\pi} \leftarrow \sum_{(m,n) \in L} \mu_{mn}(t) w_{mn}(t)$  \ \ The sum of the products of link weight and rate
12: end for
13:  $\pi(t) \leftarrow \arg \max_{\pi \in \Gamma} sum_{\pi}$ 
  ▷ Data Transferring
14: for all  $(m, n) \in L : (Q_m^d(t) - Q_n^d(t)) > 0$  do
15:   transfer packets  $c_{mn}^*(t)$  from  $m$  to  $n$ 
16: end for
17: end procedure

```

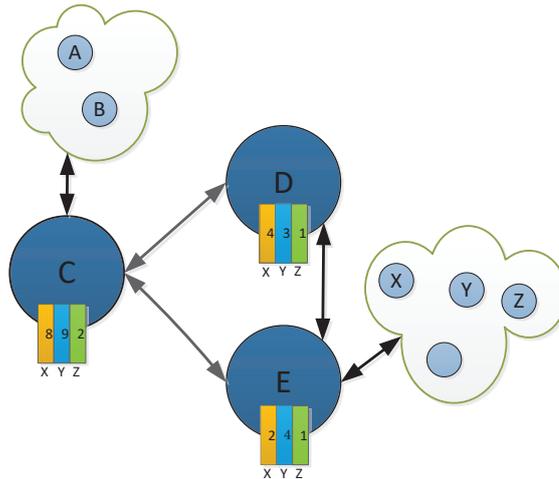


Figure 2. The visible light communication network for relay selection. Both A and B are source nodes, and there are three relay nodes, C, D, and E, and three different destinations, X, Y, and Z.

Table 1. The queue length in the node.

Length Typology	Node	Node		
		X	Y	Z
C		8	9	2
D		4	3	1
E		2	4	1

First, we use Equation (21) to compute the queue backlog difference for different kinds of packets X, Y, and Z for link (C, D) and link (C, E), respectively. The values of queue backlog difference for different kinds of packets are shown in Table 2.

Table 2. Different packet backlog values.

Backlog Difference Typology	Link	Link		
		X	Y	Z
(C, D)		$diff_{CD}^X(t) = 8 - 4 = 4$	$diff_{CD}^Y(t) = 9 - 3 = 6$	$diff_{CD}^Z(t) = 2 - 1 = 1$
(C, E)		$diff_{CE}^X(t) = 8 - 2 = 6$	$diff_{CE}^Y(t) = 9 - 4 = 5$	$diff_{CE}^Z(t) = 2 - 1 = 1$

Then, we calculate the weights of the two links (C, D) and (C, E), respectively, by Equation (22); the results are shown in Table 3.

Table 3. Link weight.

Link	Weight
(C, D)	$w_{CD}(t) = \max_{d \in D} \{diff_{CD}(t)\} = 6$
(C, E)	$w_{CE}(t) = \max_{d \in D} \{diff_{CE}(t)\} = 6$

Next, we derive the schedulable links. At time slot t , for node C, the link scheduling sets (C, D) and (C, E) interfere with each other, they cannot be scheduled simultaneously. The rate of link (C, D) is $\mu_{cd} = 4$ packets/slot, the rate of link (C, E) is $\mu_{ce} = 2$ packets/slot, and $\sum \mu_{CD} w_{CD}(t) = 4 \times 6 = 24$, $\sum \mu_{CE} w_{CE}(t) = 2 \times 6 = 12$. We can find that $\sum \mu_{CD} w_{CD}(t) > \sum \mu_{CE} w_{CE}(t)$, so we select link (C, D). The optimal data packet type is Y and the optimal transmission rate is $\mu_{CD}(t)$.

Finally, at time slot t , the routing decision of node C is that packets of type Y will be transmitted to node D at a rate of 4 packets/slot through link (C, D) for several packets.

5. Simulations and Results Analysis

Firstly, the derived theoretical result is evaluated in part A, and the relay selection algorithm is analyzed in part B.

A. The theoretical upper bound of the unreliability of latency

The dynamic change of packets in the queuing system is simulated, where the arrival process models the aggregate traffic. The simulation results of the latency-bounded unreliability are measured by the ratio of the number of slots where the latency exceeds the threshold to the total observation slots.

In Figure 3, the aggregate arrival process is composed of an MMMP-4 process and an IMP process. The service process is modeled as a geometric distribution. The model parameters are summarized in Nomenclature section. The upper bounds of unreliability regarding latency with different system load ρ is presented, where ρ is defined as the ratio of the average arrival rate to the average service rate. According to the system load, the corresponding service rate C can be determined. The simulation results of unreliability are presented by box plots. It can be shown that the upper bounds of unreliability with regard to latency match the simulated results well with different ρ . The analysis framework can provide effective results of latency performance for the aggregate traffic with heterogeneity and burstiness. Obviously, the system load plays a significant role in the reliability. In the heavy-load system ($\rho = 0.9$), the latency-bounded reliability is terrible such that it cannot support almost any services with statistical reliability requirement even if the latency threshold is large.

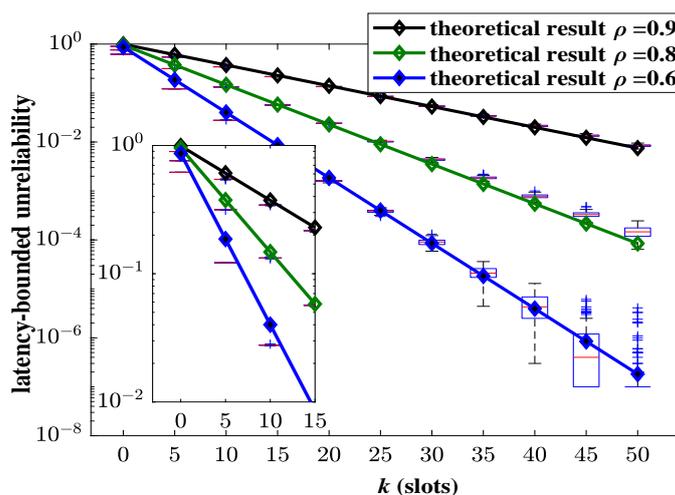


Figure 3. The simulated and theoretical results of latency-bounded unreliability, where the aggregate traffic is composed as a four-state Markov-modulated multinomial process (MMMP-4) and an interrupted multimomial process (IMP), $p_1 = 0.1$, $q_1 = 0.5$, $\lambda_1 = 0.4$, $\gamma_1 = 0.6$, $R_{M1} = 3$ packets/slot, $R_{M2} = 4$ packets/slot, $p_2 = 0.1$, $q_2 = 0.5$, $R_{I2} = 20$ packets/slot, $p = 0.5$.

In order to verify the advantages of the proposed analysis framework, we compare the theoretical results with the state-of-the-art ones based on EB/EC theory in Figure 4. In this part, the aggregate traffic consists of two identical MMOO processes. The parameters of the MMOO model are listed in Nomenclature section. The result based on the EB/EC theory refer to [20]. It can be shown that the upper bound of unreliability obtained in this paper is much tighter than the compared one. The martingale-based theoretical results match the simulations well. The gap between the EB/EC-based results and the simulated values is obvious. Our analysis method can provide a more precise analysis for the latency performance of the Markovian aggregate traffic. The EB/EC theory assumes bursty traffic as the constant fluid, which makes it impossible to reflect the influence of complex random characteristics of aggregate traffic on latency performance.

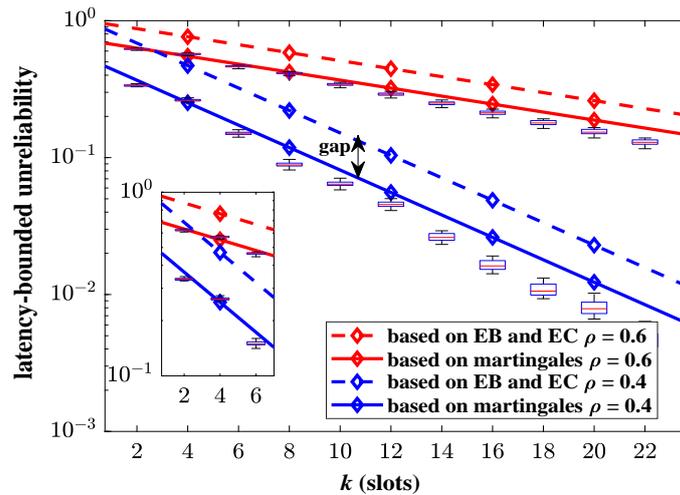


Figure 4. The simulated and theoretical results of latency-bounded unreliability, where the aggregate traffic is composed of two identical Markov modulated on-off (MMOO) processes, $p_a = 0.1$, $q_a = 0.5$, $R = 5$ packets/slot, $p = 0.5$.

Next, we analyze the impact of the burstiness of the aggregate traffic on the latency-bounded unreliability. Two measurement parameters of burstiness are adopted, which are the Hurst and squared coefficient of variation (CV). We change the state transition probabilities of the arrival models to influence the burstiness of aggregate traffic. The Hurst parameter is computed by the method of rescaled adjusted range analysis in this paper. Corresponding to the adjusted state transition probabilities ($\{0.1, 0.01, 0.001\}$), the Hurst parameter is $\{0.5904, 0.8495, 0.9140\}$, respectively. The CV of MMMP-4 can be derived as in (24), where Rav_b is the average rate of the MMMP-4 model. A larger Hurst parameter or CV indicates a stronger burstiness of the traffic. The latency-bounded unreliability versus latency threshold with different Hurst parameters is shown in Figure 5, and the unreliability versus CV is in Figure 6. In Figures 5 and 6, the system loads are 75% and 60%, respectively, which are relatively high. The latency performance decreases sharply with the burstiness. The system will become unsustainable when the system load is high along with a bursty arrival. With the CV rising from 5 to 20, the system could only provide a best-effort QoS in the current service configuration. For the aggregate traffic with complex stochastic features, the impacts of burstiness on the latency-bounded reliability must be considered.

$$CV = \frac{2Rav_b \left[(p_1 + q_1)^2 + (p_1 \lambda_1 R_{M1} + q_1 \gamma_1 R_{M2})(1 - p_1 q_1) \right]}{(p_1 + q_1) [q_1 \lambda_1 R_{M1} + p_1 \gamma_1 R_{M2} + \lambda_1 \gamma_1 R_{M1} R_{M2} (1 - p_1 - q_1)]} - Rav_b - 1. \quad (24)$$

Finally, the impacts of system load on the latency performance are analyzed. Figure 7 presents the results of latency-bounded unreliability versus Rav of the aggregate traffic, $Rav = Rav_b + Rav_i$. When $Rav < 4$ packets/slot, the system load is lower than 66.7%. If the threshold k is relatively loose, the latency-bounded reliability is acceptable. However, for the business with strict latency threshold requirements ($k = 25$ to 70), the unreliability increases linearly with the system load. System load is an essential factor impacting the latency performance. If the system load goes too high, latency performance is unsatisfactory even if the latency threshold is moderate and the average arrival is not too large. From another perspective, even if the load is not high, the strong burstiness of arrivals may cause latency thresholds (those relatively small) to be violated frequently. The dotted lines in Figure 7 support our analysis.

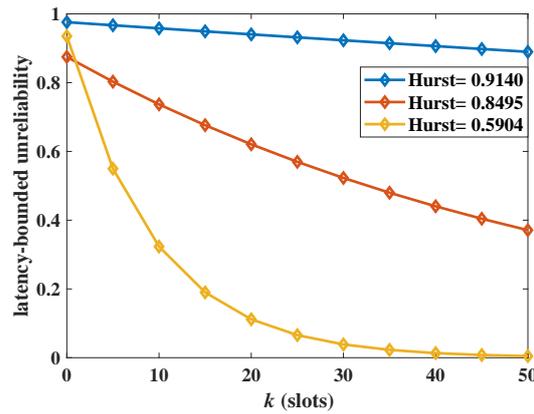


Figure 5. The latency-bounded unreliability of aggregate traffic with different Hurst parameters, where the aggregate traffic is composed as a four-state Markov-modulated multinomial process (MMMP-4) and an interrupted multimomial process (IMP), $p_1 = q_1 = p_2 = q_2 \in \{0.1, 0.01, 0.001\}$, $\lambda_1 = 0.4$, $\gamma_1 = 0.6$, $R_{M1} = 30$ packets/slot, $R_{M2} = 1$ packets/slot, $\gamma_2 = 0.6$, $R_{I2} = 30$ packets/slot, $p = 0.5$, $C = 40$ packets/slot.

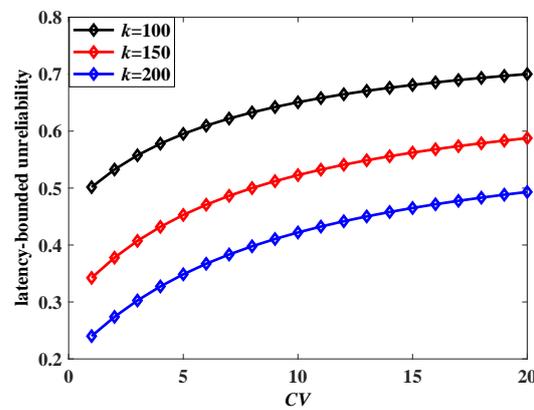


Figure 6. The latency-bounded unreliability of aggregate traffic vs. CV, where the aggregate traffic is composed as a four-state Markov-modulated multinomial process (MMMP-4) and an interrupted multimomial process (IMP), $p_1 = 0.1$, $q_1 = 0.5$, $\lambda_1 = 0.4$, $\gamma_1 = 0.6$, $R_{av_b} = 5$ packets/slot, $p_2 = 0.1$, $q_2 = 0.5$, $\gamma_2 = 0.6$, $R_{I2} = 10$ packets/slot, $p = 0.5$, $C = 20$ packets/slot.

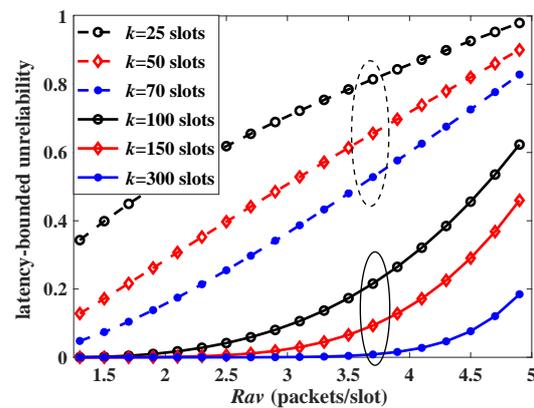


Figure 7. The latency-bounded unreliability of aggregate traffic vs. R_{av} , where the aggregate traffic is composed as a four-state Markov-modulated multinomial process (MMMP-4) and an interrupted multimomial process (IMP), $p_1 = 0.1$, $q_1 = 0.5$, $\lambda_1 = 0.4$, $\gamma_1 = 0.6$, $R_{av_b} \in [0.8, 2.6]$ packets/slot, $p_2 = 0.1$, $q_2 = 0.5$, $\gamma_2 = 0.6$, for the IMP model, $R_{av_i} \in [0.5, 2.3]$ packets/slot, $p = 0.5$, $C = 12$ packets/slot. The dashed circle corresponds to the strict threshold and the solid circle corresponds to the loose threshold.

B. Relay selection algorithm performance

We compare two relay selection algorithms: (1) our proposed relay selection algorithm based on the back-pressure theory, and (2) the traditional relay selection algorithm based on the max–min criterion method. In Figure 8, we show the evaluation of two relay selection algorithms through the average link throughput. From the figure, we can find that the average link throughput continues to increase with the increasing packet arrival rate, under two relay selection algorithms. Our proposed relay selection algorithm always achieves higher throughput than the traditional relay selection algorithm. It shows that the back-pressure theory can alleviate bandwidth limitations.

In Figure 9, we plot the relationship between the signal-to-noise ratio (SNR) threshold and outage traffic flows, under traditional and our proposed relay selection algorithm. If the SNR is still lower than the SNR threshold after the relay selection algorithm optimization, the transmission channel will be interrupted, i.e., outage traffic flows are interrupted. From the figure, we find that when the SNR threshold increases, both relay selection algorithms increase the number of outage traffic flows. However, the number of outage traffic flows from our proposed algorithm is significantly lower than that of the traditional algorithm. The result confirms that our proposed relay selection algorithm can provide better reliability network performance.

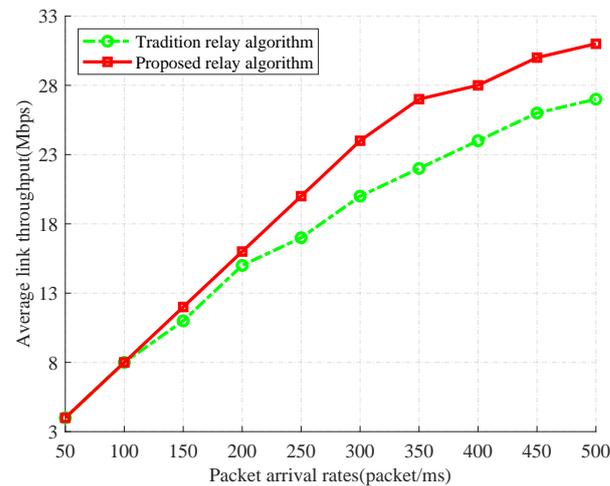


Figure 8. The relationship between packet arrival rates and link throughput under two different relay selection algorithms.

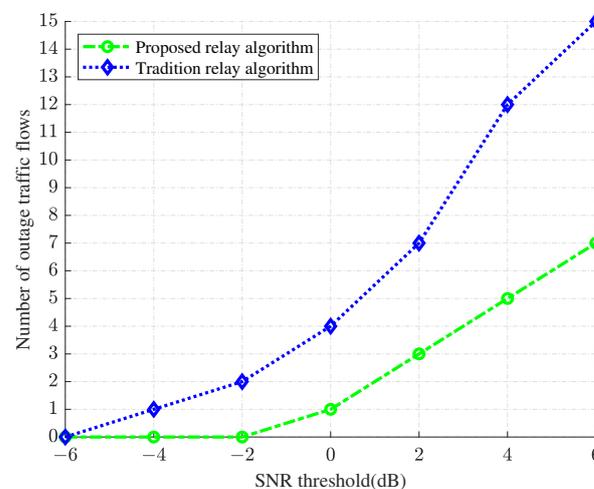


Figure 9. The relationship between signal-to-noise ratio (SNR) threshold and outage traffic flows under two different relay selection algorithms.

6. Conclusions

In this paper, we propose a novel latency-bounded reliability analysis framework based on martingales for the queuing system, where the aggregate traffic with burstiness is considered. The martingale of Markov-modulated arrival processes and the martingale of i.i.d. service processes are constructed, respectively. Leveraging the multiplicative property of martingales, the martingale of the backlog is proposed to model the buffer behavior of aggregate arrivals. Renting the stopping time theory, a tight upper bound of the unreliability with regard to latency for the aggregate traffic is derived. Meanwhile, in the core network, based on the back-pressure theory, a reasonable relay selection algorithm is proposed to boost the network performance and data transmission quality. Simulations verify the effectiveness of the proposed analysis framework. The latency performance is analyzed in terms of burstiness and the system load. When Hurst > 0.8 , the latency-bounded unreliability is up to 10^{-1} level, and the system is unable to provision any reliability guarantee. For the aggregate traffic with sharp burstiness, more resources should be configured.

Author Contributions: Conceptualization, B.Y.; data curation, B.Y.; formal analysis, B.Y.; funding acquisition, B.Y. and X.L.; investigation, X.L.; methodology, X.L.; project administration, X.L.; resources, X.L.; software, B.Y. and X.L.; validation, B.Y. and X.L.; visualization, X.L.; writing—original draft, B.Y. and X.L.; writing—review and editing, B.Y. and X.L. All authors have read and agreed to the published version of the manuscript.

Funding: The work of B. Yu and X. Liu was supported in part by the National Natural Science Foundation of China under Grant 62202215, in part by the General Young Talents Project for Scientific Research grant of the Educational Department of Liaoning Province under Grant JYTQN2023061 and JYTQN2023050, in part by the SYLU-introduced high-level talents scientific research support plan under Grant 1010147001128 and 1010147001205, in part by Shenyang Science and Technology Talent Special Project RC230794.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Nomenclature

The commonly used parameters:

Parameters	Nomenclature
The state transmission probability of the MMMP-4 model	p_1
The state transmission probability of the MMMP-4 model	q_1
The packet arrival probability corresponding to U_1 in the MMMP-4 model	λ_1
The packet arrival probability corresponding to V_1 in the MMMP-4 model	γ_1
The packet arrival rate corresponding to U_1 in the MMMP-4 model	R_{M1}
The packet arrival rate corresponding to V_1 in the MMMP-4 model	R_{M2}
The state transmission probability of the IMP model	p_2
The state transmission probability of the IMP model	q_2
The packet arrival probability corresponding to V_2 in the IMP model	γ_2
The packet arrival rate corresponding to V_2 in the IMP model	R_{I2}
The successful access probability	p
The service rate	C
The state transmission probability of the MMOO model	p_a
The state transmission probability of the MMOO model	p_b
The packet arrival rate corresponding to ON in the MMOO model	R

References

1. Pathak, P.H.; Feng, X.; Hu, P.; Mohapatra, P. Visible light communication, networking, and sensing: A survey, potential and challenges. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 2047–2077. [[CrossRef](#)]
2. Rehman, S.U.; Ullah, S.; Chong, P.H.J.; Yongchareon, S.; Komosny, D. Visible light communication: A system perspective—Overview and challenges. *Sensors* **2019**, *19*, 1153. [[CrossRef](#)] [[PubMed](#)]
3. Matheus, L.E.M.; Vieira, A.B.; Vieira, L.F.M.; Vieira, M.A.M.; Gnawali, O. Visible light communication: Concepts, applications and challenges. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 3204–3237. [[CrossRef](#)]
4. Popovski, P.; Stefanovic, C.; Nielsen, J.J.; de Carvalho, E.; Angjelichinoski, M.; Trillingsgaard, K.F.; Bana, A.-S. Wireless Access in Ultra-Reliable Low-Latency Communication (URLLC). *IEEE Trans. Commun.* **2019**, *67*, 5783–5801. [[CrossRef](#)]
5. Kim, K.S.; Kim, D.K.; Chae, C.-B.; Choi, S.; Ko, Y.-C.; Kim, J.; Lim, Y.-G.; Yang, M.; Kim, S.; Lim, B.; et al. Ultrareliable and Low-Latency Communication Techniques for Tactile Internet Services. *Proc. IEEE* **2019**, *107*, 376–393. [[CrossRef](#)]
6. Qian, L.; Chi, X.; Zhao, L. Hybrid Access Algorithm for eMBB Terminals with Heterogeneous QoS in MPR aided VLC system. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018; pp. 1–6.
7. Hammouda, M.; Akin, S.; Vegni, A.M.; Haas, H.; Peissig, J. Link Selection in Hybrid RF/VLC Systems Under Statistical Queueing Constraints. *IEEE Trans. Wirel. Commun.* **2018**, *107*, 2738–2754. [[CrossRef](#)]
8. Dong, W.; Chi, X.F.; Liu, Z.C.; Zhu, Y.H. Queueing theory based analysis for packet jitter of mixed services. *J. China Univ. Posts Telecommun.* **2014**, *21*, 71–76. [[CrossRef](#)]
9. Ciucu, F.; Poloczek, F.; Schmitt, J. Sharp per-flow delay bounds for bursty arrivals: The case of FIFO, SP, and EDF scheduling. In Proceedings of the IEEE INFOCOM 2014—IEEE Conference on Computer Communications, Toronto, ON, Canada, 27 April–2 May 2014; pp. 1896–1904.
10. Chen, S.C.; Thomas, J.A. Effective bandwidth in high-speed digital networks. *IEEE J. Sel. Areas Commun.* **1995**, *13*, 1091–1100. [[CrossRef](#)]
11. Wu, D.P.; Negi, R. Effective capacity: A wireless link model for support of quality of service. *IEEE Trans. Wirel. Commun.* **2003**, *2*, 630–644.
12. Poloczek, F.; Ciucu, F. Service-martingales: Theory and applications to the delay analysis of random access protocols. In Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM), Kowloon, Hong Kong, 26 April–1 May 2015; pp. 945–953.
13. Zhao, L.; Chi, X.; Zhu, Y. Martingales-Based Energy-Efficient D-ALOHA Algorithms for MTC Networks with Delay-Insensitive/URLLC Terminals Co-Existence. *IEEE Internet Things J.* **2018**, *5*, 1285–1298. [[CrossRef](#)]
14. Qi, R.; Chi, X.; Zhao, L.; Yang, W. Martingales-Based ALOHA-Type Grant-Free Access Algorithms for Multi-Channel Networks with mMTC/URLLC Terminals Co-Existence. *IEEE Access* **2020**, *8*, 37608–37620. [[CrossRef](#)]
15. Yun, H.; Li, H.; Zheng, C.; Zhu, H. End-to-End Backlog and Delay Bound Analysis for Multi-Hop Vehicular Ad Hoc Networks. *IEEE Trans. Wireless Commun.* **2017**, *16*, 6808–6821.
16. Fantacci, R.; Picano, B. End-to-End Delay Bound for Wireless uVR Services over 6G Terahertz Communications. *IEEE Internet Things J.* **2021**, *8*, 17090–17099. [[CrossRef](#)]
17. Yu, B.; Chi, X.; Sun, H. Delay analysis for aggregate traffic based on martingales theory. *IET Commun.* **2020**, *14*, 760–767. [[CrossRef](#)]
18. Seferoglu, H.; Modiano, E. TCP-aware Backpressure Routing and Scheduling. *IEEE Trans. Mob. Comput.* **2016**, *15*, 1783–1796. [[CrossRef](#)]
19. Yang, S.; Adeel, U.; Tahir, Y.; McCann, J.A. Practical Opportunistic Data Collection in Wireless Sensor Networks with Mobile Sinks. *IEEE Trans. Mob. Comput.* **2017**, *16*, 1420–1433. [[CrossRef](#)]
20. Ozmen, M.; Gursoy, M.C. Wireless Throughput and Energy Efficiency with Random Arrivals and Statistical Queueing Constraints. *IEEE Trans. Inf. Theory* **2016**, *62*, 1375–1395. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.