IEICE Communications Express, Vol.13, No.4, 101–105

ComEX

LETTER

Novel micro-burst detection method in access/edge networks using periodic IoT messages

Keita Nishimoto[1, a)], **Kota Asaka**[1], **Tatsuya Shimada**[1], **and Tomoaki Yoshida**[1]

Abstract Real-time access edge computing demands access networks to connect a lot of IoT devices to edge servers with ultra-low latency. A lot of connected devices will cause the data bursts and sudden increase of delay (i.e. micro-burst) in the upstream queues in gateways. This study proposes a novel method for detecting such bursts to ensure low-latency communication. The proposed method leverages the periodic messaging behavior of certain IoT devices. It detects bursts based on the variation in the arrival times of messages transmitted from already deployed multiple IoT devices. This approach enables monitoring of delays and detecting bursts without the need for additional probes or modifications to gateways. The study validates the detection accuracy and feasibility through numerical analysis and simulations.

Keywords: micro-burst, delay monitoring, edge computing
Classification: Network management/operation

1. Introduction

Recently, researchers have proposed fixed access network architectures for access edge computing, where applications on edge servers control IoT devices (e.g. industrial robots and sensors) in a real-time manner [1] (Fig. 1).

In such an architecture, the access network must guarantee ultra-low latency between edge servers and devices connected via IoT Gateways (GW). In particular, we focus on micro-bursts as a technical obstacle to provide stable low-latency communication in the network. A micro-burst is a spike of traffic that flows into a router/switch in a very short period of time (a few tens of microseconds to a few milliseconds) and causes increased latency and packet drops [2]. Although this phenomenon is currently observed in Data Center networks [2], we anticipate it will also arise in the upstream flow in the future access network due to the increase of connected devices.

A general method for detecting micro-bursts is to monitor the queue lengths in switches/GWs and periodically report them to the detection system. However, this method may cause these devices to be expensive due to the additional features for monitoring and may consume the bandwidth by the report messages. Several methods [3, 4] have been proposed to estimate the one-way queueing delay by measuring the variation in arrival timing of these periodic messages (Fig. 2 (a)). For example, the authors in a previous study [4] proposed the passive measurement method for one-way queueing delay, which utilizes the periodic VoIP messages. These methods will be useful for detecting the micro-bursts because it does not need additional features on each device (i.e. switch/GW) and not consume bandwidth.

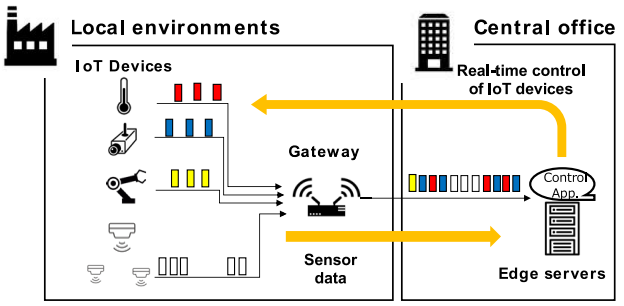

Fig. 1 Real-time control for IoT devices from edge applications across the fixed access network.

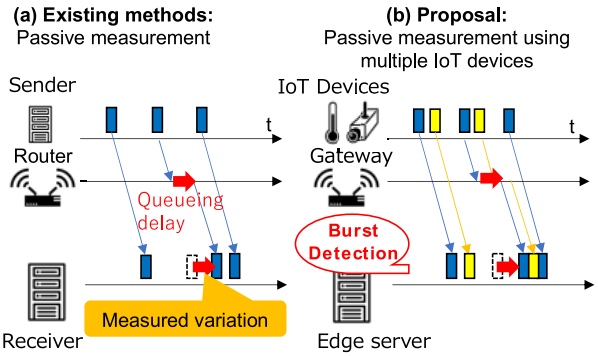

Fig. 2 Comparison of our proposal with the existing studies.

Here, we focus on the characteristic of IoT devices that periodically send messages [5]. Specifically, our method detects micro-bursts by analyzing the time-series of one-way queueing delays estimated by measuring the variation in arrival times of these periodic messages from multiple IoT devices already deployed (Fig. 2 (b)). The existing studies [3, 4] have already established queueing delay measurement method which uses the pair of the periodic packet sender/receiver. However, they do not satisfy high sampling granularity required for detecting a micro-burst that occurs in a few tens of microseconds to a few milliseconds. The novelty of our work is to expand the existing method that focused a single sender, to asynchronous and multiple periodic senders (i.e. IoT devices) for improving the sampling granularity, and to further propose the detection method considering the above characteristics. We evaluate the detection

[1] NTT Corporation, 1–1 Hikarinooka, Yokosuka-shi, Kanagawa 239–0847, Japan
[a)] keita.nishimoto@ntt.com

DOI: 10.23919/comex.2023XBL0174
Received December 6, 2023
Accepted December 26, 2023
Publicized February 16, 2024
Copyedited April 1, 2024

accuracy by numerical analysis, varying the number of devices sending periodic messages, which is crucial for the feasibility.

2. Micro-burst detection system

2.1 Network architecture

Figure 3 shows the architecture of the fixed access network and a GW, focusing on the upstream traffic flow. The GW connects numerous IoT devices in a local environment (e.g. factory) to edge servers in a central office via the fixed access network. For simplicity, we assumed the GW can be modeled as M/M/1 [6] in the upstream traffic.

A control application on the edge server monitors these devices and controls them in a real-time manner, responding to the sensor data periodically sent from them. Our aim is to detect a sudden increase of the upstream queueing delay inside the GW. We assume some of IoT devices each of which sends a message every T second, while other devices send messages at random. We further define the number of these *periodic devices* as a variable N, and assume T is a few tens of milliseconds, according to the previous study [1], for realizing the real-time control.

2.2 One-way queueing delay estimation

Our proposed method utilizes the existing technology [4] to estimate the queueing delay experienced by messages sent from each periodic device. The estimation relies on measuring the time gap between the actual and expected (i.e., periodic) arrival times, achieved by capturing the messages. Figure 4 shows the time-series of actual queueing delays and estimated samples. The estimated samples from a single periodic device, as indicated by each color in Fig. 4, provides delay samples at a coarse granularity with interval T. However, by aggregating delay estimation values from multiple periodic devices (all in colors in Fig. 4), it becomes possible to sample delays at a finer garrulity (with sample interval Δt). In this proposed method, messages from asynchronously transmitting devices are captured at the central office's capture point. Please note that the timings when the system estimates the delays are when each of periodic messages arrives and are not distributed at the same intervals since each device sends messages asynchronously every T second. The feature of our proposed method is to detect a micro-burst considering the variable sampling intervals.

2.3 Micro-burst detection

Our method estimates the queue length inside GW based on the time-series of estimated delays and judges a micro-burst occurs when the estimated queue length exceeds pre-defined data size L. Figure 5 depicts a part of actual delay time-series (solid line) and estimated samples (colored circles).

Here, we assume two messages sent from two different periodic devices arrive on the capture point at t_i and t_{i+1}, and estimated queuing delays at these timings are $d(t_i)$ and $d(t_{i+1})$, respectively. If a burst occurs at t_{burst} in the interval period $\Delta t_i (= t_{i+1} - t_i)$, and the queue length exceeds data size L, then the queue length at t_{i+1} must be more than $L - B\Delta t_i$ (B: outflow rate of the queue) and thus $d(t_{i+1}) > L/B - \Delta t_i$ as

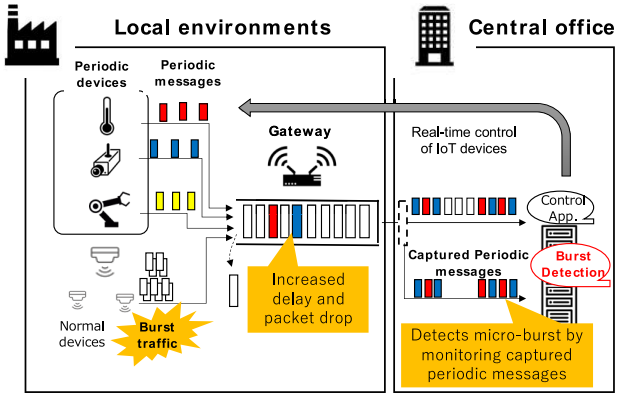

Fig. 3 Architecture for detecting micro-bursts by utilizing periodic messages (colored with red, blue and yellow) sent from IoT devices.

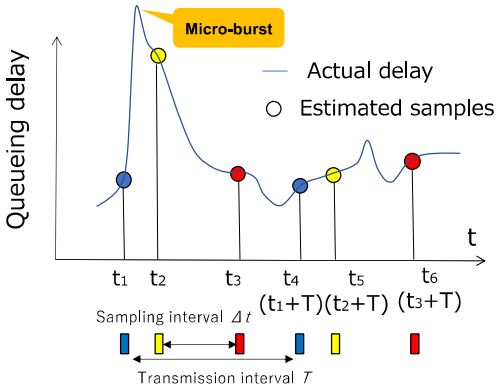

Fig. 4 Time-series of estimated queuing delays inferred from periodic messages (blue, yellow, red) transmitted from three devices.

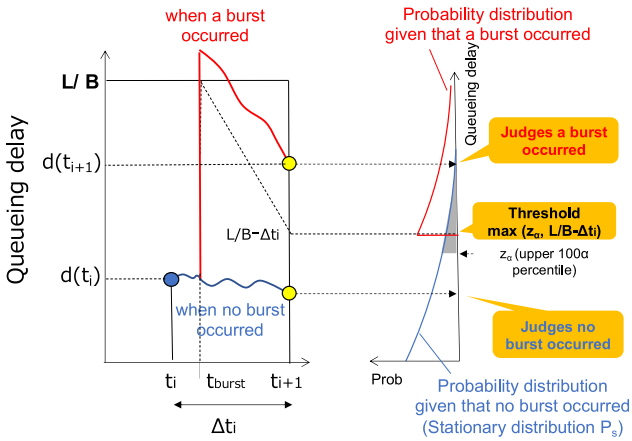

Fig. 5 Conceptual diagram of the proposed burst detection method.

Fig. 5 shows. On the other hand, if it does not occur, $d(t_{i+1})$ is considered to follow the stationary distribution defined by the background traffic and the outflow rate B. Following the above discussion about $d(t_{i+1})$, the system judges a micro-burst occurs during $[t_i, t_{i+1})$ when $d(t_{i+1})$ exceeds the threshold $L/B - \Delta t_i$ and deviates from the stationary distribution P_s. The latter condition can be defined by the threshold z_α: the upper α-percentile determined by the predefined acceptable upper limit of false positive rate (FPR: filled with gray in Fig. 5). Please note $\alpha = 100 \times FPR$. The pseudo-code is shown in Fig. 6.

input: $d(t_{i+1})$, t_i, t_{i+1}, L, B, D, FPR
output: is_burst_happened
P(X): Probability distribution of queue length X in stationary state
$F^{-1}(X)$: Inverse cumulative distribution function of P(X)
$\Delta t_i = t_{i+1} - t_i$
$\alpha = 100 \times FPR$
$z_\alpha = F^{-1}(\alpha)$
if $d_{ti+1} > \max(L/B - \Delta t_i, z_\alpha)$:
    is_burst_happened = True
else:
    is_burst_happened = False

Fig. 6 Burst detection algorithm (pseudo-code).

3. Evaluation

3.1 Basic analysis when $N = 100$

To verify the basic performance of our proposal, we conducted simulations using the M/M/1 queue model [6] as GW (thus stationary distribution follows the geometric distribution) using the parameters shown in Table I. We generated burst traffic (1MB) at the 5 second and injected it into the queue while generating background traffic for ten seconds. We measured the detection accuracy by whether it can detect the burst at the 5 second with three metrics (true positive/false positive/false negative rates).

To demonstrate the superiority of our proposal, Fig. 7 compares it with other detection methods. The first method (colored with green) for comparison is a naïve approach, which detects micro-bursts when an estimated delay takes an outlier from the steady-state distribution (i.e., when $d(t_{i+1}) > z_\alpha$). The second method (colored with blue) is based on a fixed interval T/N, without considering variability of $\Delta t_i = t_{i+1} - t_i$. In Fig. 7, the gray line represents the estimated delay samples, and the star marks indicate the timings when each method detected micro-bursts, and the dotted lines represent the detection thresholds.

With the method using the steady-state distribution (green), a peak value was falsely detected as a burst (false positive) at a point (4.985 sec) other than the actual burst occurrence, as the estimated delay exceeded the fixed threshold z_α. On the other hand, when using the fixed sampling interval (blue), the threshold is set higher than the actual burst, leading to missed detections (false negative). This misdetection happens when a burst occurs in a large ($> T/N$) sampling interval period. In contrast, the proposed method successfully detected the burst indicated by the magenta star mark by dynamically adjusting the threshold (magenta dotted line) according to Δt_i.

To further compare the performance of our proposal with the fixed sampling interval, Fig. 8 shows the true positive rates (1 − false negative rates) as background traffic rate is varied. It shows the proposed method could achieve almost 100% true positive rate, while the other missed approximately 40% of detections when the background traffic rate was 1Gbps. We also measured their false positive rates and observed the both methods achieved 0%.

3.2 Minimum required number of periodic devices

To assess the feasibility of our proposal, it is essential to determine how many periodic devices are required. A small N (number of periodic devices) expands the sampling interval, reducing granularity and increasing false negatives. Conversely, requiring thousands of devices makes our proposal impractical. Here, we assess the feasibility through formulating false negative rate as a function of N.

In the proposed method, if a burst occurs at t_{burst}, and Δt_{burst} (the interval between the time t_{burst} and the nearest sampling timing t_{i+1}) satisfies the condition $L - B\Delta t_{burst} > z_\alpha \iff \Delta t_{burst} < (L - z_\alpha)/B$, the system judges a micro-burst occurs. That is, if we define X as $(L - z_\alpha)/B$, the lower bound of the true positive rate can be expressed as P ($\Delta t_{burst} < X$), thus the upper bound of the false negative rate we want to derive is 1 − P ($\Delta t_{burst} < X$). We derived the theoretical values of 1 − P ($\Delta t_{burst} < X$) as follows (see Appendix for details):

$$1 - \mathrm{P}\ (\Delta t_{burst} < X)$$

Table I Settings and parameters.

B (Outflow-rate of GW)	10Gbps
In-flow background traffic rate	1Gbps
L (Maximum acceptable queue length)	1MByte
Acceptable FPR (False Positive Rate)	0.001
N (Number of periodic devices)	100
T (Transmission interval)	10 ms (refer to [1])

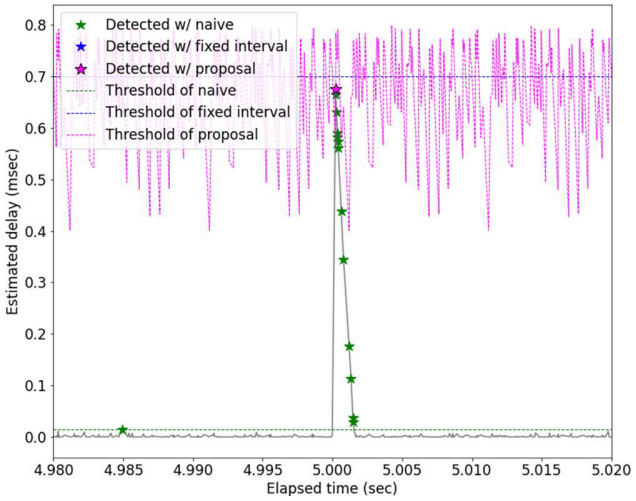

Fig. 7 Detection of a micro-burst with several methods and their thresholds.

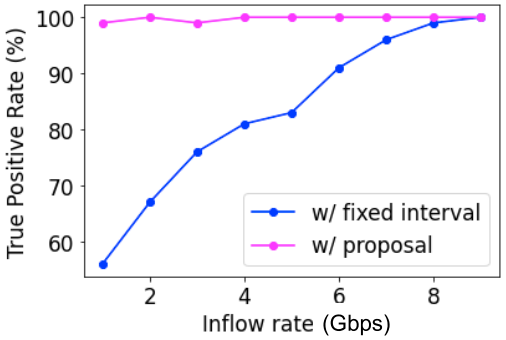

Fig. 8 Comparison of true positive rate between proposed method with adjusted threshold and the other method with the fixed interval.

$$= 1 - N/T \times \left(\int_0^X \Delta t_i f\left(\Delta t_i\right) d\Delta t_i + \int_X^T X f\left(\Delta t_i\right) d\Delta t_i \right) \quad (1)$$

where $f(\Delta t_i) = (N-1)(1-\Delta t_i/T)^{N-2}/T$.

To verify the above formulation, we compared the theoretical values obtained from the formula with the false negative rates obtained through the simulations while varying the number of N under the same conditions as in Section 3.1. As Fig. 9 shows, the theoretical values are consistent with the simulation results.

The minimum necessary number of periodic devices can be calculated by finding the minimum N that satisfies the condition that $1 - \mathrm{P}(\Delta t_{burst} < X)$ is higher than the predefined acceptable false negative rate. Please note that X is determined by parameters L, B and z_α affected by the background traffic. In Fig. 10, we show the minimum periodic devices, when setting the acceptable false negative rate = 1% and varying the transmission interval T with each background traffic rate. The result shown in Fig. 11 indicates that a micro-burst can be detected with an accuracy of less than 1% false negative rate if approximately 50 periodic devices are connected to the GW. Since this number of devices can be accommodated by a commodity-available 64-port switch with fixed lines, we consider our proposal has a reasonable feasibility.

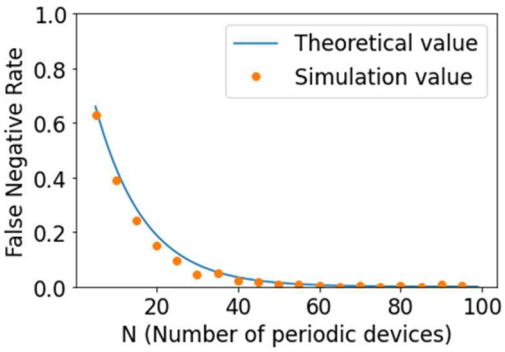

Fig. 9 Comparison between the false negative rates obtained from theoretical values and those from simulation results.

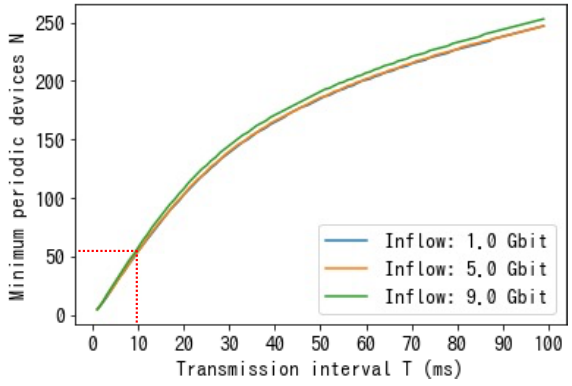

Fig. 10 Number of periodic devices required to achieve a false negative rate of 1% when the transmission interval T is varied.

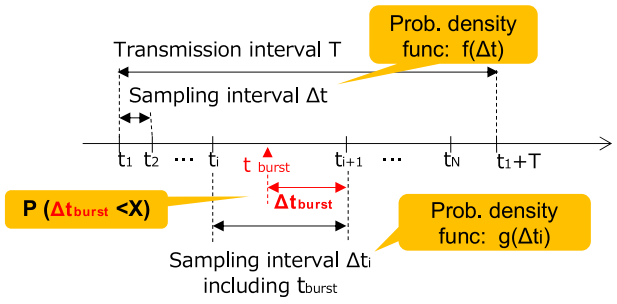

Fig. 11 Transmission interval T and the sampling intervals.

4. Conclusion

This paper presents a method that leverages the periodic transmission characteristic of IoT devices to detect micro-bursts. Simulation results demonstrate that the proposed method detects the bursts with fewer false positives and negatives compared to other methods. Additionally, we established how the number of periodic devices impacts detection accuracy to validate our proposal's feasibility.

References

[1] Y. Koyasako, T. Suzuki, T. Yamada, T. Shimada, and T. Yoshida, "Demonstration of real-time motion control method for access edge computing in PONs," *IEEE Access*, vol. 10, pp. 168–175, 2022. DOI: 10.1109/ACCESS.2021.3136876

[2] D. Shan, F. Ren, P. Cheng, R. Shu, and C. Guo, "Micro-burst in data centers: Observations, analysis, and mitigations," Proceedings of IEEE ICNP 2018, pp 88–98, 2018. DOI: 10.1109/ICNP.2018.00019

[3] W. Lu, W.-X. Gu, and S.-Z. Yu, "One-way queuing delay measurement and its application on detecting DDoS attack," *Journal of Network and Computer Applications*, vol. 32, no. 2, pp. 367–376, 2009. DOI: 10.1016/j.jnca.2008.02.018

[4] B. Ngamwongwattana and R. Thompson, "Sync & sense: VoIP measurement methodology for assessing one-way delay without clock synchronization," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 5, pp. 1318–1326, 2010. DOI: 10.1109/TIM.2010.2043978

[5] T. Hoßfeld, F. Metzger, and P.E. Heegaard, "Traffic modeling for aggregated periodic IoT data," Proceedings of ICIN 2018, pp. 1–8, 2018. DOI: 10.1109/ICIN.2018.8401624

[6] L. Kleinrock, *Queueing Systems, Volume I (Theory)*, Wiley-Interscience, 1975.

[7] T. Yoshikawa and A. Okabe, "Analysis of the probability distribution of interval length of random points on a line segment," *Comprehensive Urban Studies*, vol. 43, pp. 99–105, 1991 (in Japanese).

Appendix

For deriving $\mathrm{P}(\Delta t_{burst} < X)$, we first derived the probability density function for Δt_i, the length of the sampling interval where a micro-burst occurs. As Fig. 11 shows, each sampling interval Δt can be seen as fractions of [0, T) divided by $(N-1)$ random points since each periodic device asynchronously transmits messages every T second. Therefore, this problem can be seen a *broken-stick problem* (see [7]). The probability density function for sampling interval Δt is derived according to the existing solution: $f(\Delta t) = (N-1)(1-\Delta t/T)^{N-2}/T$ (2). Further, assuming that micro-burst occurs at a random timing t_{burst}, the probability density function for Δt_i can be derived as $g(\Delta t_i) = f(\Delta t_i) \times \Delta t_i N/T$ (3).

Our target condition, i.e., $\Delta t_{burst} < X$, can be classified into two cases: (a) where Δt_i is smaller than X, and (b) where Δt_i is greater than X but Δt_{burst} is smaller than X. Therefore, $\mathrm{P}(\Delta t_{burst} < X)$ can be represented as the sum of

these two cases, by substituting Eq. (2) and Eq. (3).

$$\begin{aligned} &\mathrm{P}\left(\Delta t_{burst} < X\right) \\ &= \mathrm{P}\left(\Delta t_i < X\right) + \mathrm{P}\left(\Delta t_i \geq X \cap \Delta t_{burst} < X\right) \\ &= \int_0^X g\left(\Delta t_i\right) d\Delta t_i + \int_X^T g(\Delta t_i) \times X / \Delta t_i d\Delta t_i \\ &= N/T \times \left(\int_0^X \Delta t_i f\left(\Delta t_i\right) d\Delta t_i + \int_X^T X f\left(\Delta t_i\right) d\Delta t_i \right) \end{aligned} \quad (4)$$