Packet error rate performance of multi-stage cooperative relaying: Outdoor measurement

Hidekazu Murataa) and Makoto Miyagoshi
Graduate School of Informatics, Kyoto University,
Yoshida-hommachi, Sakyo-ku, Kyoto 606–8501, Japan
a) murata@i.kyoto-u.ac.jp

Abstract: The packet error rate (PER) performance of space-time block code based multi-hop cooperative relaying is studied. In the case of multiple relays at each hop, a unique feature of the end-to-end PER performance can be observed up to 3rd hop. In this letter, an outdoor measurement campaign is conducted to verify this PER performance. In addition, a cooperative relaying scheme with received data sharing among relays is also examined. It is experimentally shown that PER at the 3rd hop is better than that of the 1st hop.

Keywords: multi-hop transmission, cooperative relaying, space-time block code, packet error rate, collaborative relaying, measurement campaign

Classification: Wireless Communication Technologies

References

1 Introduction

Multi-hop cooperative relaying has been widely studied [1]. Implementation and experimental results of cooperative relaying are reported in several papers [2, 3, 4]. Space-time code is useful for signal-level (i.e., PHY-layer) cooperation [5], and its implementation issues are investigated in [3]. However, to the best of the authors’ knowledge, the error rate performance of signal-level cooperation is not well verified in actual environments.

A multi-stage concept with multiple intermediate relays at each stage is suggested in [6]. It is pointed out in [7, 8] that the multi-stage concept has a unique advantage over a simple relaying with a single intermediate relay at each hop. The theoretical analysis in [8] reveals that the end-to-end packet error rate (PER) performance can be kept almost constant, or even improved, as the number of hops is increased. This unique feature in the end-to-end PER with the number of hops is demonstrated both theoretically and experimentally [9]. In the case of two relays in each stage, this improvement can be observed up to 3rd hop in most cases [8].

These studies assume independently and identically distributed (i.i.d.) Rayleigh fading for all channels, which is the best condition for signal-level cooperation. Experimental verification of this unique feature by an actual outdoor measurement...
campaign is limited to 2-hop scenarios [10, 11]. Then, a question arises: How about the end-to-end error performance of multi-stage cooperative relaying at the 3rd hop in actual environments (i.e., not i.i.d.), which is exactly the focus of this letter [12].

The PER performance can be further improved when relays at each stage can communicate with each other [13]. Experimental verification using a fading emulator is reported in [13]. This cooperative relaying scheme with data sharing is also examined together with a cooperative relaying scheme without data sharing.

2 Multi-stage cooperative relaying

The system model is shown in Fig. 1(a). The experimental equipment employed in this letter is a multi-stage cooperative relaying system with two relay stations at each stage. In a cooperative relaying scheme without data sharing, two relays (S1-A and S1-B, S2-A and S2-B) independently transmit a space-time block coded packet carrying the received data bits if no error is found by cyclic redundancy check (CRC). In a cooperative relaying scheme with data sharing, the packets received correctly are shared in two relay (S1-B and S1-C, S2-B and S2-C) through a data link. Therefore, the probability of transmitting space-time block coded packets from two relays is improved.

When propagation channels are independently distributed, the performance of multi-stage cooperative relaying is improved by received data sharing among relays. However, the distances between the relays are close so that the relays can communicate together; the diversity gain may be degraded due to correlations between propagation channels. Therefore, employing distant relays without data sharing may be more efficient than employing nearby relays with data sharing. In this letter, the effect of the received data sharing in multi-stage cooperative relaying is also examined via field experiments.

3 Experimental system and setup

The experimental system consists of eight software-defined radio (SDR) based wireless stations. The carrier frequency is 5.11 GHz. Built-in global positioning system (GPS) receivers are employed for timing and frequency synchronization.

3.1 Signaling format

The frame structure is shown in Fig. 1(b). In this signaling format, 100 packets are packed into one block so that the transmit/receive switching frequency is reduced for stable operation of SDR. Two cooperative relaying schemes are switched packet by packet. This enables quasi-simultaneous measurements of PERs of two schemes. The packet consists of 8-symbol training sequence, 50-symbol data symbols, and 8-symbol CRC symbols [12]. Two orthogonal training sequences and Alamouti’s space-time block codes are assigned uniquely for two relay stations of each scheme. The modulation scheme is 100 k symbols/s QPSK except for the training sequence. The data bits are drawn from a pseudo-random sequence.

For pulse shaping, eight times oversampling and a root roll-off Nyquist filter with a roll-off factor of 0.4 is employed at the transmitter. At the receiver side, oversampled signals are utilized only for finding the symbol timing. The symbol
timing is estimated by a simple correlation technique using the training sequence, and the channel state information is also acquired. Note that these processes are performed for each packet independently.

### 3.2 Measurement setup

The locations of eight wireless stations are shown in Fig. 2. As can be seen, distances between a pair (transmit/receive) of wireless stations are 19 m except for two cross links (S1-A to S2-B,C and S1-B,C to S2-A). Antenna separation (i.e., between S1-B and S1-C, and also between S2-B and S2-C) for the cooperative relaying scheme with data sharing is 16 cm. For all wireless stations, omni-direc-
tional antennas with 3 dBi gain are fixed at 0.9 m above the ground. The transmit power is −1 dBm.

4 Experimental results

Fig. 3 shows the measured PER performance and the received signal power at the destination. The PER performance is averaged over 50 packets in this figure. As can be seen in Fig. 3(a) and (b), the received signal amplitude from two corresponding transmitters apparently follows different distributions in both the cases of

![Diagram showing PER and received power at Destination without data sharing.](image)

![Diagram showing PER and received power at Destination with data sharing.](image)

![Diagram showing PER versus hop count.](image)

**Fig. 3.** Measured PER of two cooperative relaying schemes.
(a) without data sharing and (b) with data sharing. This is due to multipath fading and shadowing by trees and plants.

Fig. 3(c) shows the average PER versus the hop count. This PER is averaged over 90,000 (30 minutes observation) packets for each scheme. The average PERs improve on the whole as the number of hops increases. It can be seen that PERs at the 3rd hop are better than those of the 1st hop, and almost comparable to those of the 2nd hop. This is the major feature of multi-stage cooperative relaying. However, we cannot give a detailed discussion of the measured PERs due to the nature of outdoor field experiments.

5 Conclusion

This letter presented the field experimental results of two cooperative relaying schemes. Due to the nature of outdoor field experiments, the difference of the PER performance between two schemes cannot be observed clearly. However, it is shown that multi-stage cooperative relaying can keep the end-to-end PER at the 3rd hop below that of the 1st hop and comparable to that of the 2nd hop. Therefore, the major feature of multi-stage cooperative relaying is observed in an actual environment, where the propagation channels are not i.i.d. but follow different distributions.

Acknowledgments

This work was partly supported by Strategic Information and Communications R&D Promotion Programme (SCOPE) of the Ministry of Internal Affairs and Communications, Japan.
Radio wave propagation simulation considering Doppler shift caused by other vehicles

Naozumi Ando\textsuperscript{1a)}, Kazuma Tomimoto\textsuperscript{2}, Ryo Yamaguchi\textsuperscript{2}, and Mitoshi Fujimoto\textsuperscript{1}

\textsuperscript{1} Graduate School of Engineering, University of Fukui, 3–9–1 Bunkyo, Fukui 910–8507, Japan
\textsuperscript{2} Advanced Technology Research Department Advanced Technology Office, Softbank Corp., 2–5–10 Aomi, Koto-ku, Tokyo 135–0064, Japan
\textsuperscript{a)} ando@wireless.fuis.u-fukui.ac.jp

Abstract: Generally, maximum Doppler frequency can be derived from velocity of moving vehicle. However, frequency shifts beyond the maximum Doppler frequency are observed in actual measurement. In this paper, Doppler shift is calculated in consideration with the shift caused by other moving vehicles. Numerical results show that the actual measured Doppler shift can be reproduced with high accuracy by considering the frequency shift at the other moving vehicles.

Keywords: radio wave propagation, Doppler shift, Doppler spread, ray tracing

Classification: Antennas and Propagation

References


1 Introduction

In mobile communication, frequency shift occurs due to Doppler effect, and transmission quality will be degraded. Generally, maximum Doppler frequency can be derived from velocity of the moving vehicle. And Doppler shift of each path are calculated based on maximum Doppler frequency and arriving angles. Various methods for modeling propagation and obtaining Doppler spectra have been proposed [1, 2, 3, 4]. The formula for calculating the Doppler frequency never exceeds the maximum Doppler frequency. However, in case of actual measurement, frequency shifts beyond the maximum Doppler frequency are observed [5].

This paper aims to show the mechanism of the Doppler shift that exceed the maximum Doppler frequency. Doppler shift is calculated in consideration with the shift caused by other moving vehicles using the angles of arrivals and the number of reflections. It is shown that the calculated result is well agree with the measured one.

2 Analysis model

Fig. 1(a) shows a simulation model for ray-tracing. This is the reproduced model of the area where the actual measurements ware conducted. The receiver and vehicles are arranged at the hatched area (on the road) in Fig. 1. Frequency and velocity of the vehicle is 4.64 GHz and 30 km/h, respectively. Also, the transmitting antenna and receiving antenna are omni-directional. Other specifications of the simulation and measurement are shown in Table I.

![Fig. 1. Analysis model](image-url)
3 Doppler shift due to reflection on other vehicle

3.1 Measured Doppler spectrum

When the Doppler shift occurs at the mobile receiver, maximum value of the Doppler shift (maximum Doppler frequency: $f_{D\ max}$) can be calculated by eq. (1) in generally.

$$f_{D\ max} = \frac{v}{\lambda}$$

(1)

Here, $v$ and $\lambda$ is moving velocity and wavelength. And each incoming wave Doppler shift $f_D$ is (2).

$$f_D = f_{D\ max} \cdot \cos \phi$$

(2)

$\phi$ is angle of arrivals in the horizontal plane based on direction of travel. The Doppler shift at the other vehicle is calculated from (1,2) as same as the Doppler shift at the mobile receiver [1, 2, 3, 4]. The Doppler shift on the propagation path is shown in Fig. 1(b) and explained as follows:

Case A) Transmitted radio wave is reflected at the mobile receiver’s body first. Then, the wave is also reflected at the other vehicle that travels the same direction and velocity with the mobile receiver, and received by the mobile station antenna. Doppler shift occurs when reflection and receiving. When reflecting at the mobile receiver and other vehicle, Doppler shifts are negative and positive, respectively. Assuming these two vehicles’ velocity and direction of travel is the same and regular reflection, Doppler shifts are canceled each other that occurs at these reflections. As a result, only the Doppler shift at receiving remains. So, it will not exceed $f_{D\ max}$ in this path. Doppler shift is determined by direction of arrival and $f_{D\ max}$.

<table>
<thead>
<tr>
<th>Table I. Simulation and measurement specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Ray-tracing simulation</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>Moving velocity</td>
</tr>
<tr>
<td>Transmitting antenna</td>
</tr>
<tr>
<td>Height (Transmitter)</td>
</tr>
<tr>
<td>Receiving antenna</td>
</tr>
<tr>
<td>Height (Receiver)</td>
</tr>
<tr>
<td>Maximum reflections</td>
</tr>
<tr>
<td>Physical property value of buildings (concrete)</td>
</tr>
<tr>
<td>Relative permittivity</td>
</tr>
<tr>
<td>Electrical conductivity</td>
</tr>
<tr>
<td>Physical property value of road surface (asphalt)</td>
</tr>
<tr>
<td>Relative permittivity</td>
</tr>
<tr>
<td>Electrical conductivity</td>
</tr>
<tr>
<td>(b) Measurement</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>Moving velocity</td>
</tr>
<tr>
<td>Power</td>
</tr>
<tr>
<td>Transmitting antenna</td>
</tr>
<tr>
<td>Height (Transmitter)</td>
</tr>
<tr>
<td>Receiving antenna</td>
</tr>
<tr>
<td>Height (Receiver)</td>
</tr>
</tbody>
</table>

© IEICE 2020
DOI: 10.1587/comex.2019XBL0142
Received October 9, 2019
Accepted October 30, 2019
Publicized November 12, 2019
Copyedited February 1, 2020
Case B) First, radio wave is reflected at other vehicle’s body moving in the opposite direction to the mobile receiver. And it is received at the mobile station. This reflection causes positive Doppler shift. Doppler shift occurs twice on incidence and reflection at the other vehicle. Doppler shift in reception is also positive. Therefore, triple $f_{D_{\text{max}}}$ can occur at most in this path. Actual Doppler shift is $3f_{D_{\text{max}}}$ or less, because it varies with angle of incidence, reflection and arrival.

Case C) In case of A and B, reflection at other vehicles were considered. However even the path reflected at stationary objects, Doppler shift may exceed $f_{D_{\text{max}}}$. When reflection at the receiver’s body and reception, positive Doppler shift occur. On the other hand, Doppler shift doesn’t occur when reflected by stationary object. Therefore, up to $3f_{D_{\text{max}}}$ Doppler shift is observed.

Doppler shift varies with relationship between mobile receiver and other vehicle velocity and moving direction. As an example, Doppler spectrum in the actual environment is shown in Fig. 2(a) [5]. In this measurement environment, $f_{D_{\text{max}}}$ is approximately 128 Hz, and it can be seen that the Doppler shift of 2 to 4 times of $f_{D_{\text{max}}}$ is observed in the measurement.

3.2 Simulated Doppler spectrum

In the simulation, following two cases is examined to calculate the Doppler shift.

Case I. Doppler shift caused at other vehicles is not considered.

Case II. Doppler shift caused at other vehicles is considered.

The result is shown in Fig. 2(b) and (c). In Case I, there are no frequency shift beyond the maximum Doppler frequency because of using only velocity of receiver. On the other hand, in Case II, considering Doppler shift at other vehicles causes a shift that exceeds the maximum Doppler frequency.

4 Cumulative distribution function of Doppler spread

Cumulative distribution function of Doppler spread in the analysis area of Fig. 1(a) is shown in Fig. 2. If a frequency shift at the other vehicle is not considered, the Doppler spread of simulated result is lower than that of measured result. By considering the frequency shift at the other vehicles, the actual measured Doppler spread can be reproduced with high accuracy.
5 Conclusion

The frequency shift beyond the maximum Doppler frequency was studied. When Doppler shift caused by other vehicle was considered, Doppler spread in actual environment could be reproduced by radio wave propagation simulation.
Novel frequency domain equalization with threshold for molecular communication

Ahsan Sohail\textsuperscript{a)} and Yukitoshi Sanada\textsuperscript{b)}

Department of Electronics & Electrical Engineering, Keio University,
Yagami Campus, 3–14–1 Hiyoshi, Kohoku-Ku, Yokohama, Japan
\textsuperscript{a)} ahsan.sohail@snd.elec.keio.ac.jp
\textsuperscript{b)} sanada@elec.keio.ac.jp

Abstract: This letter proposes a frequency domain equalization (FDE) scheme with a threshold in molecular communication (MC). The responses of a MC channel are tendency to have low pass characteristics and intersymbol interference limits the data rates of MC. However, because of the characteristics of the MC channel, equalization coefficients for high frequency components may increase. Therefore, noise enhancement occurs and it deteriorates communication performance. As large channel responses concentrate at around a zero-frequency component, in the proposed equalization scheme, the high frequency component of the received signal is eliminated if the equalization coefficient exceeds a threshold in the proposed scheme. Numerical results obtained through computer simulation show that the threshold improves the communication performance with the proposed FDE. It is also shown that the trade-off between the noise enhancement and the pulse waveform effects on the performance.

Keywords: molecular communication, frequency selective channel, frequency domain equalization

Classification: Wireless Communication Technologies

References

1 Introduction

Molecular Communication (MC) is a bio-impelled diffusion based system where the trading of data is acknowledged through the transmission, dissemination, and reception of particles such as bacteria’s, pheromones etc. This standard model was initially examined in biological science since it is effectively incorporated in nature by cells for intra-cellular and inter-cellular correspondence [1]. Diffusion-based communication arises as one of the most promising solutions for the communication mechanism between nano-scale devices for its inherent compatibility with living organism and biochemical devices, e.g., pheromone propagation in the air between insects [2] or calcium signaling among living cells [3]. The interest in MC is for applications in a diverse number of fields, including biological engineering, healthcare, manufacturing, and environmental monitoring [1].

The assumed MC system model is presented in Fig. 1 [2]. The transmitter generates and emits molecules into the propagation medium. The transmitter takes the data bits and modulates them to a chemical signal by a device which then emits the chemical molecules into the channel. At the receiver side, the receiver counts the molecules through the sensor and accumulates the number of absorbed
molecules during each time slot [3]. The number of absorbed molecules during each time slot is compared with a predetermined threshold for demodulation. The demodulator and decoder block is programmed in a microcontroller [2].

The characteristics of a diffusion channel play an important role in propagating the molecules. Some of these molecules hit the receiver while few remain in the medium which may reach later or lost [2]. Inter-symbol interference (ISI) is a major problem that is caused by diffusion. Some of the carrier molecules may be unsuccessful to reach the receiver in their anticipated time period and may interfere with data molecules in a consequent symbol period [3]. For molecular communication systems a higher data rate and better efficiency is being demanded though it is limited by noise and ISI.

Currently research for ISI in MC has been focused on time domain equalization (TDE) [3]. As it is known that, in nano communication systems, a simple and miniature-scale equipment is needed [1]. While, in frequency domain equalization (FDE), multiplication operations instead of convolution are applied and they are realized with simpler hardware implementation. The usage of FDE centred on a minimum mean square error (MMSE) criterion based on channel estimation can take the advantage of channel frequency selectivity and hence improves its bit error rate (BER) performance [4].

In [5] and [6] the conventional MMSE equalization scheme incorporating decision feedback channel estimation faces the same noise enhancement issue. The channel estimation could be implemented using decision feedback as well as pilot insertion at the receiver. The conventional scheme in [5] adopts the decision feedback channel estimation only for the subcarriers with a large estimated channel response. If the estimated channel response is small, it applies the response estimated only by the pilot symbols. The latter case is equivalent to the proposed scheme with a threshold of zero in this letter.

MC channels are tendency to have low pass characteristics and large channel responses concentrates at around a zero-frequency component. High frequency components should be eliminated rather than being equalized to prevent noise enhancement. Therefore, this letter proposes a FDE scheme that eliminates frequency component of signal in order not to emphasize noise power if a corresponding equalization coefficient is less than a threshold.

2 System model

A point source transmitter releases $N$ molecules in all directions into an unlimited single dimension (1D) environment with constant temperature since a point segment reception in the receiver is assumed [2, 7]. The released molecules are initially distributed uniformly throughout the surroundings. Local molecule absorptions are low enough to assume that the molecules are diffused into the surroundings with a diffusion coefficient, $D$. Here molecules in the form of small droplets of alcohol are assumed. These molecules are released to the surrounding medium, where they diffuse according to Brownian motion and arrive at the receiver [2]. These particles are detected by a receiver that has a radius of $r_{RX}$ and is placed at a distance of $d$ from the center of the transmitter [7].
A channel impulse response, $h_X(\tau)$, is defined as the number of molecules expected at the receiver after a delay of $\tau$ while assuming that the $N$ molecules are instantaneously released by the transmitter at time $t = 0$ (i.e. considering time period of molecule release is negligibly small). The channel impulse response for a point transmitter is presented below [7].

$$h_X(\tau) = N \ast \text{erfc} \left( \frac{d - r_{RX}}{\sqrt{4Dt}} \right)$$  \hspace{1cm} (1)

In Eq. (1), it is presumed that the channel stays constant throughout the communication, i.e. it is assumed to be a static channel. The diffusion coefficient, $D$, and the radius of receiver, $r_{RX}$, have the constant values which are predetermined. According to the constraint on the magnitude and the power of particles, a straightforward modulation mechanism is presumed to be implemented for nano molecules. The transmitter discharges a predetermined number of particles into the aqueous medium during every duration of the symbol. The $n$th on-off modulation symbol is denoted with $s(t)$ and

$$s(nT_s) = \begin{cases} 
1 & \text{for information bit} = 1 \\
0 & \text{for information bit} = 0 
\end{cases}$$

where $T_s$ is the symbol interval. The expected signal that reaches the receiver is expressed as

$$r(t) = \int (s(\tau) \otimes h_X(\tau))d\tau + n(t)$$ \hspace{1cm} (2)

where $n(t)$ is known as the additive white Gaussian distributed noise (AWGN) [8].

### 3 Proposed scheme

The FDE MMSE eliminates ISI through equalization. At the receiver side, the analog-digital convertor samples the received signal and transforms it to the frequency-domain signal with the help of $M$-point FFT. The $k$th frequency component of the received signal is given as

$$R_k = H_{sk}S_k + N_k$$ \hspace{1cm} (3)

where $S_k$ is the $k$th element of the transmit signal, $H_{sk}$ is the $k$th element of the channel response, and $N_k$ is the noise element in the $k$th frequency component. The FDE MMSE coefficient for the $k$th frequency component of the received signal is given as [4]

$$W_k = \frac{H_{sk}}{H_{sk}^2 + \frac{N_0}{E_S}}$$ \hspace{1cm} (4)

where $N_0$ is the noise spectrum density and $E_S$ is the symbol energy.

The problem with the equalization of the received signal is that equalization also emphasizes the noise power as well. In order to minimize the effect of the noise enhancement through equalization, in the proposed scheme, the channel response has to be compared with a pre-selected threshold and the signal component after equalization is set to zero if the corresponding channel response is lower than the threshold. This process is represented as
\[ A_k = \begin{cases} 0 & \text{for } H_{xk} < T_h \\ W_k R_k = W_k H_{xk} S_k + W_k N_k & \text{for } H_{xk} > T_h \end{cases} \] (5)

where \( T_h \) is the preselected threshold value. If the channel response is too small, the term, \( W_k N_k \), increases as the denominator of the right term in Eq. (4) approaches to \( \frac{N_0}{E_S} \) that may be a small value. Therefore, to prevent noise enhancement, \( W_k \) is set to zero for \( H_{xk} < T_h \).

4 Numerical results

4.1 Simulation conditions

For the purpose of numerical evaluation through computer simulation the following simulation conditions are assumed: \( 10^5 \) symbols are transmitted for each plot, a diffusion coefficient of alcohol is \( 9.59 \times 10^{-6} \text{ m}^2/\text{s} \). The radius of a receiver is 1 \( \mu \text{m} \), and the distance between the transmitter and the receiver is 5 \( \mu \text{m} \). The total number of molecules transmitted are \( N = 10^3 \). The symbol interval \( T_s = 0.5 \) second (100 samples) and the size of the FFT is \( M = 200 \). The normalized threshold for FDE MMSE is set between 0.00 and 0.20.

The frequency response of the MC channel for the FDE MMSE with the threshold (\( \{ R_k \} \) in Eq. (5)) and the normalized pulse waveform at the receiver as well as those after FDE are presented in Fig. 2(a) and 2(b). The channel responses are concentrated at around a zero frequency component. It is clearly depicted that the responses of the channel in higher frequency are very small so that they can be neglected to suppress noise enhancement.

![Normalized channel response](image1)

![Equalized signal before detection](image2)

Fig. 2. Molecular communication channel response and signal waveform before detection.

It is clear from Fig. 2(b) that the pulse amplitude at the point of 100 samples causes ISI as the symbol interval is 100 samples. The ratio between the maximum amplitude to the amplitude at a point of 100 samples varies depending on the threshold values as shown in Fig. 2(a). For example, if the threshold value is 0.10 or 0.14, the interference to the following pulse is relatively small while the interference to the following pulse is large if it is 0.12. Thus, the system performance is closely related to the amount of ISI.
4.2 Performance results

4.2.1 Relation between BER and threshold

As explained in Fig. 3(a), as threshold increases, BER decreases until it reaches 0.10. However, the BER increases at the thresholds of 0.11 and 0.12. This is due to the fact that the ISI affects the system performance; Fig. 2(b) and Fig. 3(a) correspond to each other. The rise in the BER at the particular values of the threshold implies that a specific threshold value has to be selected based not only on the amount of noise enhancement, but also on the pulse waveform. Therefore, the pulse waveform as well as the amount of noise after equalization decides selection of threshold values.

4.2.2 Relation between BER and $E_b/No$

The BER performance versus bit-energy-to-noise-spectrum-density ($E_b/No$) for different threshold values is presented in Fig. 3(b). The conventional FDE-MMSE with feedback channel estimation at a threshold value of zero shows the worst performance because of noise enhancement. On the other hand, it is clear that the proposed system achieves the best performance for a threshold value of 0.10. Although the threshold value varies from 0.06 to 0.14, BERs with the other threshold values show worse BERs because of the ISI.

5 Conclusion

In this letter FDE MMSE with the threshold has been proposed to achieve better system performance for the diffusion based molecular communication. The proposed technique is used to suppress the amount of noise enhancement caused by FDE MMSE equalization with the selective threshold. As seen in the above sections that large channel responses are concentrated at around a zero frequency component. According to the threshold value the amount of ISI after FDE varies. Therefore, the tradeoff between the noise enhancement and the pulse waveform after FDE effects on the BER performance.
Joint computation and bandwidth resources allocation for generalized computing model in mobile edge computing

Qiaobin Kuang¹, Jie Gong², Xiang Chen¹[a], and Xiao Ma²
¹ School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China
² School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China
a) chenxiang@mail.sysu.edu.cn

Abstract: Mobile edge computing (MEC) has restricted computation resource and limited bandwidth resource when serving multiple users simultaneously. In general, a user’s task cannot be always parallelized. In this paper, we study a joint computation and bandwidth resource allocation problem, considering a generalized task computing model. The computation capacity is modeled as a non-decreasing concave function about allocated computation resource, and the joint computation and bandwidth resources allocation problem is solved by cvx opt. framework. Simulation results show that our proposed scheme is greatly superior to the serial processing benchmark scheme for Non-linear model. With non-linear model, higher computational multiplexing gain can be achieved, however, for linear model, there is no gain that can be obtained from the proposed scheme.

Keywords: mobile edge computing, computation offloading, resource allocation

Classification: Wireless Communication Technologies

References

1 Introduction

With the emergence of novel mobile applications such as face recognition, augmented reality technology and autonomous driving, the requirement for high-intensity computation and low latency is becoming more stringent. However, mobile users are often resource constrained due to their physical limitation. Mobile edge computing (MEC) which deploys computation resources at access points (APs) to offload mobile users’ computing tasks, is a promising paradigm to solve the problem, as it not only can provide sufficient computing capability at the mobile edge, but also shortens the response time compared with centralized clouds [1, 2]. A large number of previous papers in mobile edge computing have investigated computation offloading problems. For a single-user MEC system, [3, 4] presented task offloading schedule approaches. In multi-user case, it is crucial on how to jointly allocate radio and computation resources [5, 6, 7].

As seen in the existing works, computation time is not always inversely proportional to the allocated computation resource, as task may not be fully parallelized. In particular, some parts of task are so highly integrated or relatively that cannot be partitioned and some components’ outputs are the inputs of others [1]. Even if a large amount of parallel computation resources are allocated to one user, due to the partition limitation of the user’s task, some parallel computation resources are not utilize, that is, as the allocated computation resources increase, it does not mean that the computation time will decrease. To the best of our knowledge, there is no generalized model on the relation between computation time and computation resource.

In this work, we propose a new joint computation and bandwidth resources allocation approach for generalized computing model in an MEC system. Each user requests system bandwidth resources to transfer their respective tasks, and the edge server allocates certain computation resources to each user. Specifically, we model the generalized computing model as a non-decreasing concave function to represent the relation between computation capacity and computation resource. We have jointly optimized the allocation of both computation and bandwidth resources
to minimize the overall latency of the system. The simulation results show that the proposed method greatly reduces the latency of the MEC system compared to random serial processing (RSP) method for non-linear model. However, for linear computing model, the RSP method outperforms the joint computation and bandwidth resources allocation approach.

2 System model

Consider an MEC system consisting of an AP and \( K \) mobile users. The AP contains an MEC server and the set of mobile users is denoted by \( \mathcal{K} \), with \( |\mathcal{K}| = K \). All these users are equipped with a single antenna. The total available frequency band of the MEC system is \( B \), and is orthogonally allocated to the mobile users.

Each mobile user \( i \in \mathcal{K} \), would like to offload its computation-intensive task to the MEC server to reduce its own energy consumption and latency. Let a pair \((l_i, w_i)\) represents the corresponding information of the mobile user \( i \in \mathcal{K} \), in which \( l_i \) is the number of task input-bits at the mobile user \( i \), and the execution of the task needs \( w_i \) CPU cycles. In our setting, the AP serves as a central controller to decide how to assign the orthogonal frequency band and the CPU of the server for mobile users. The control signalling over-head is ignored for simplicity. The whole offloading process can be divided into the following steps.

2.1 Computation offloading from mobile users to the AP

In this case each mobile user \( i \in \mathcal{K} \) will occupy bandwidth \( b_i \), to offload its task to the AP. The data rate of user \( i \) can be denoted as

\[
R_i(b_i) = b_i \log_2 \left( 1 + \frac{h_ip_i}{N} \right),
\]

where \( p_i \) is the transmission power of user \( i \) which is determined by the base station according to some power control algorithm, \( h_i \) is the channel power gain between the user \( i \) and the AP, and \( N \) denotes the power of the additive white Gaussian noise. Therefore, the time for offloading the task of user \( i \) can be calculated as

\[
T_{\text{tran}}^{i}(b_i) = \frac{l_i}{R_i(b_i)}, \quad \forall i \in \mathcal{K}.
\]

Furthermore, the sum bandwidth allocated to the users can not exceed the total budget, i.e.

\[
\sum_{i \in \mathcal{K}} b_i \leq B.
\]

2.2 Computing at AP

In the MEC server in AP, a virtual machine is associated to user \( i \) for its task computation. The total computation resource of all the virtual machines will not exceed the maximum computation resource of the system \( F \). Let \( f_i \) denote the computation resource assigned to user \( i \), due to the nature of computation tasks, the computation capacity for user \( i \) is not always linear to the computation resource. Thus, we define a function \( g(f_i) \) to represent the relation between computation resource and capacity, where \( g(\cdot) \) can be any non-decreasing and concave function.
Accordingly, we can obtain the time of computing at the MEC server for user \( i \), which can be denoted as

\[
T_{i}^{\text{comp}}(f_i) = \frac{w_i}{g(f_i)}, \quad \forall i \in K. \tag{4}
\]

The total limitation of the computation resources of the AP is

\[
\sum_{i \in K} f_i \leq F. \tag{5}
\]

### 2.3 Sending back computing result

After completing the task computing, the result is sent back to each user. Assume the result of the computation is much smaller than the input-bits. Therefore, the time to send back the result can be ignored.

According to the offloading process, the total latency for user \( i \) to complete its task can be expressed as

\[
T_i(b_i, f_i) = T_i^{\text{trans}}(b_i) + T_i^{\text{comp}}(f_i), \quad \forall i \in K. \tag{6}
\]

Since AP computes the tasks of users in parallel, the process latency of the system is given by

\[
T = \max_{i \in K} a_i T_i(b_i, f_i), \tag{7}
\]

where \( a_i \) is the positive weight for user \( i \) to complete its task. Our target is to minimize the weighted latency of the system for users’ tasks execution, by efficiently optimizing the computation and bandwidth resources allocation of the MEC system. The decision variables include the bandwidth allocation \( b_i \) and the computation resource allocation \( f_i \). The problem is formulated as:

\[
(P1) : \min_{b_i, f_i} \max_{i \in K} a_i T_i(b_i, f_i)
\]

s.t. \( b_i \geq 0, \ f_i \geq 0, \ \forall i \in K, \tag{8} \)

\[
(3), (5),
\]

where (3), (5) are the constraint of bandwidth and computation resource of the MEC system, respectively.

Note that problem (P1) is convex, and we solve it by standard convex optimization techniques such as the interior point method [8].

### 3 Numerical results

In this part, the numerical results are presented to validate the performance of the proposed joint computation and bandwidth resources allocation design for generalized computing model. The linear model \( g_l(f) = \beta_l f \) denotes the tasks of users can be parallelized completely. Therefore, we use a natural logarithmic function \( g_{\log}(f) = \beta_{\log} \log(1 + rf) \) as the chosen non-linear function to present the limited parallel processing, where \( \beta_{\log} \) and \( r \) are the parameters of the function. Compared to the following benchmark scheme without such a joint design under two computation models.

- **random serial processing**: The scheme is to arbitrarily choose a user to transmit at first, and then transmit one of the remaining users’s tasks when
the transmission of the previous user’s task is completed, and at the same time, the server is serving the previous user. In short, the channel and the server are available for only one user simultaneously. In this paper, we don’t consider the energy consumption. Therefore, in order to reduce the latency of the system, we choose the maximum data rate and computation capacity for those users. The maximum data rate for user $i$ is the same in both computing models, i.e., $R_{i,\max} = B \log_2(1 + \frac{h_{i,p_i}}{N})$. The maximum computation capacity for both linear computing resource (LCR) models and non-linear computing resource (nLCR) models are $g_{i,\max} = \beta_i F$, $g_{\log,\max} = \beta_{\log} \log(1 + rF)$, respectively. We assume that the finish time of the first user for both transmission and computation for LCR model is $T'_{1,1} = \frac{a_i l_i}{R_{i,\max}} + \frac{a_i w_i}{g_{i,\max}}$. Accordingly, the finish time for the second user is $T'_{1,2} = \max\{\frac{a_1 l_1}{R_{1,\max}} + \frac{a_2 l_2}{R_{2,\max}}, T'_{1,1}\} + \frac{a_2 w_2}{g_{2,\max}}$. In the end, we can get a recursion formula as $T'_{l,k} = \max\{\sum_{i=1}^{k} \frac{a_i l_i}{R_{i,\max}}, T'_{l,k-1}\} + \frac{a_k w_k}{g_{k,\max}}$, which is the $k$ users’ system latency. Similarly, for nLCR models, the recursion formula can be obtained as $T'_{\log,k} = \max\{\sum_{i=1}^{k} \frac{a_i l_i}{R_{i,\max}}, T'_{\log,k-1}\} + \frac{a_k w_k}{g_{k,\max}}$.

In this simulation, we assume the distance between the users and the server satisfies the discrete uniform distribution of interval $[120, 150]$ meters. The path-loss between users and the server is denoted as $\beta_0(d/d_0)^{-\zeta}$, where $\beta_0 = -60$ dB corresponds to the path-loss at the reference distance of $d_0 = 10$ m, $d$ denotes the distance from the transmitter to the receiver, and $\zeta = 3$ is the path-loss exponent. Furthermore, we set $a_i = 1$, $p_i = 26$ dBm, $\forall i \in K$, $\beta_i = 0.98$, $\beta_{\log} = 20$, $r = 10^3$, $B = 40$ MHz, $N_0 = -100$ dBm, $F = 5$ GHz. The input-bits of the tasks $l_i$ is yield the uniform distribution of interval $(0,1)$ Mbits, the required CPU cycles of users $w_i$ satisfy the discrete uniform distribution of interval $[1,3] \times 10^9$ cycles. The reasonability of the selected parameters above is supported by the references [4] and [7]. Different values of the above parameters will influence the performance of the proposed model. Therefore, we choose the commonly used values in the literature of computation offloading in MEC to guarantee the well behaving of the model and the appropriateness to model an actual CPU.

Fig. 1 shows the performance gains of the linear computing resource (LCR) models and non-linear computing resource (nLCR) models for different numbers of users.
users, where the performance gain is expressed as $\eta_{\text{gain}} = 1 - \frac{T_l}{T_0}$ and $\eta_{\log,\text{gain}} = 1 - \frac{T_{\log}}{T_{0,\log}}$ for linear and non-linear models, respectively. Denote $T_l$ and $T_{\log}$ as the system latency of proposed scheme for LCR and nLCR models, respectively. $T'_l$ and $T'_{\log}$ are the system latency of RSP. As shown in Fig. 1, the nLCR model has a gain under any number of users, and with the increasing of user number, the amplitude of the gain is increasing up to 1. This means that under the nLCR model, the user’s parallel processing capacity is limited, even if a user is allocated a lot of parallel computation resources, these computation resources can not be fully utilized due to the limitation of the user’s task partition. When the system serves users in parallel, these parallel computation resources can be more fully utilized, therefore, the proposed joint computation and bandwidth resources allocation scheme can achieve superior performance to RSP. Notice that the relation between the two times is essentially determined by the practical applications. For example, when focusing on the computation offloading problem as in our paper, the computation time is dominant. But when the communication time is dominant, our proposal still has equal performance compared to RSP.

As the number of users increases, however, the LCR model has no gain, meaning that the user’s task can be processed completely in parallel. In this case, simultaneously serving these users does not bring gain, and also wastes channel resources, because all users’ tasks must have been transmitted before they are computed, thus, the channels allocated are idle. For the LCR model, the channels are always occupied for RSP, thus, serial processing is more resource efficient than the proposed scheme. The gain of serial processing by transmission and computation over time multiplexing tends to be stable with the number of users increasing.

4 Conclusions

In this paper, we investigated a computation and bandwidth resource allocation scheme. Numerical results presented that, compared to the benchmark scheme, the proposed scheme can achieve remarkable gain for non-linear computing resource model. For linear computing resource model, the joint computation and bandwidth resource allocation scheme is inferior to random serial processing. How to combine resource optimization with more MEC servers and more users is worth further investigation in the future work.

Acknowledgments

The work is supported in part by NSFC (No. 61771495, No. U1734209), the State’s Key Project of Research and Development Plan under Grant 2017YFE0121300-6, the Science and Technology Program of Guangzhou under Grant201707010166, and Guangdong Provincial Special Fund For Modern Agriculture Industry Technology Innovation Teams (No. 2019KJ122).
Feasibility verification of direct spectrum division transmission over multiple satellite transponder

Fumihiro Yamashita\(^1\), Daisuke Goto\(^1\), Yasuyoshi Kojima\(^1\), Hiroki Shibayama\(^1\), Hiroyuki Kobashi\(^2\), and Daiki Haraguchi\(^2\)

\(^1\) NTT Access Network Service Systems Laboratories, NTT Corporation, 1–1 Hikari-no-oka, Yokosuka, Kanagawa 239–0847, Japan
\(^2\) SKY Perfect JSAT Corporation, 1–8–1 Akasaka, Minato, Tokyo 107–0052, Japan

\(a)\) fumihiro.yamashita vp@hco.ntt.co.jp

Abstract: We have been proposing a direct spectrum division transmission (DSDT) technique that can divide a single carrier signal into multiple sub-spectra and assign them to dispersed frequency resources of a satellite transponder to improve spectrum utilizing efficiency. This letter overviews the concept of DSDT and shows the results of feasibility verification on DSDT via satellite experiments. In a past study, we carried out fundamental satellite experiments on DSDT over a single transponder. This time, we conduct satellite experiments on DSDT over multiple transponders by using the latest DVB-S2X format signal. The BER performances of signals divided by the DSDT technique were almost the same as that of a single-carrier. We confirmed that DSDT was practical.

Keywords: direct spectrum division transmission (DSDT), satellite experiments

Classification: Satellite Communications

References


1 Introduction

The rapid adoption of network services requires more efficient wireless network infrastructures. Broadband satellite links are needed to establish IP access networks or mobile backhaul systems because satellite communications are the sole access network for aircrafts, vessels, and disaster-struck areas [1, 2]. In typical satellite communications systems, Demand Assigned Multiple Access (DAMA) assigns the satellite transponder’s frequency resources to user earth stations (UESs) independently, so the total usage of the satellite transponder changes over time [3]. The repeated acquisition and release of frequency resources among various services makes the unused frequency resources on the satellite transponder scatter and become individually insufficient to accommodate new broadband users, degrading the frequency utilization efficiency. Several transmission schemes to minimize the discontinuity in unused frequency resources have been studied. Multi-Frequency TDMA (MF-TDMA) [4] utilizes multiple frequency and timing slots. However, a MF-TDMA user cannot use multiple discontinuous frequency slots simultaneously. OFDMA [5] and Single-Carrier FDMA (SC-FDMA) [6] can utilize multiple frequency slots concurrently and improve frequency utilization efficiency when the frequency and timing of all the UESs are synchronized. However, precise synchronization control is difficult and complicated because satellite channels have long transmission delays.

To tackle these problems, we have been proposing “Direct Spectrum Division Transmission (DSDT),” a new transmission method based on a spectrum editing technique that better utilizes unused frequency resources on the satellite transponder. The transmitter divides the single carrier modulated signal into multiple “sub-spectra” in the frequency domain and arranges the sub-spectra to match the unused frequency slots. Conversely, the divided sub-spectra in the receiver are combined in the frequency domain and demodulated. Unlike OFDMA and SC-FDMA, the guard band between the DSDT user’s sub-spectrum and the primary user’s signal can be reduced because the bandwidth of each sub-spectrum on the Tx side is shaped by a root roll-off filter and its side lobes are sufficiently suppressed. Basically, DSDT signals do not have to be synchronized with adjacent signals. In a previous paper, we introduced feasibility verification of DSDT over a single satellite transponder [7]. We expect using DSDT over multiple transponders to...
provide a new type of satellite communication service by raking unused frequency resources dispersed over multiple transponders. Since the satellite on-board analog equipment performance differed transponder to transponder, practical satellite experiments can effectively verify DSDT over multiple transponders. Therefore, this letter reports the DSDT performance over multiple transponders.

2 Overview of direct spectrum division transmission

Fig. 1 shows the overall system concept using DSDT. On the Tx side, the DSDT adapter is inserted between the existing RF equipment and the existing modem. On the Rx side, the other DSDT adapter is inserted between the existing RF equipment and the existing modem. In the DSDT adapter on the Tx side, a single carrier modulated signal output from the existing modem is converted into the frequency domain by using a fast Fourier transform (FFT), then divided into multiple sub-spectra with arbitrary bandwidth by a spectrum dividing filter bank. The dividing filters have a special frequency response that reduces the roll-off factor of each sub-spectrum to generate sub-spectra with root roll-off characteristics and decrease the occupied bandwidth after spectrum division [8]. The divided sub-spectra are set on unused frequency resources and converted into the time domain by using inverse FFT (IFFT).

On the Rx side, received sub-spectra are converted into the frequency domain by FFT, and all sub-spectra are re-shifted and recombined by the spectrum combining filter bank. In the combining filter bank, root roll-off filters are the combining filters. Since both sub-spectra and combining filters have root roll-off characteristics and the cutoff frequencies between combining filters are the same, the signal after the spectrum combination has a flattened pass band and exhibits the
full roll-off characteristics of the original single carrier modulated signal. After phase compensation, the combined signal is converted into the time domain by IFFT.

The phase characteristics of the combined signals become discontinuous because each Rx sub-spectrum undergoes independent phase offset due to the unsynchronized clock timing between Tx and Rx side. The DSDT adapter estimates the phase differences between Rx sub-spectra from the transition band of adjacent sub-spectra and compensates them [9]. Thus, the phase characteristics become continuous. Existing modems can realize the phase synchronization of the combined signals.

3 Feasibility verification of DSDT performance via satellite experiments

For feasibility verification, we developed DSDT adapters. Fig. 2(i) shows the experimental setup for Ku-band satellite experiments. We used typical DVB-S2X formatted signals to check if the DSDT adapter could divide/combine the latest commercial signal. Fig. 2(ii) shows the representative spectrum arrangement patterns for satellite experiments. First, we measured the BER performance without spectrum division for reference, as shown in Fig. 2(ii)-(a). Next, we measured the BER performance of signals evenly divided into 2 sub-spectra over a single transponder, as shown in Fig. 2(ii)-(b). Finally, we measured the BER performance of signals divided and arranged over multiple transponders, as shown in Fig. 2(ii)-(c). As Fig. 2(ii)-(c) shows, the guard band between adjacent transponders was 4 MHz.
(i) Experimental Setup for satellite experiments

(a) Spectrum arrangement without spectrum division

(b) Spectrum arrangement over a single transponder

(c) Spectrum arrangement over multiple transponders

(ii) Spectrum arrangement over satellite transponders

Fig. 2. Configuration of satellite experiments
Detailed signal conditions are as follows. In Fig. 2(ii)-(a), the symbol rate was set at 20M-baud. Its roll-off ratio is 0.05. On the other hand, DSDT divides 20M-baud single carrier signals into two 10M-baud sub-spectra, as shown in Figs. 2(ii)-(b) and (c). QPSK, 8PSK, 16APSK, 32APSK, and 64APSK modulation types changed. LDPC \( (R = \frac{3}{4}) \) was applied for FEC. Fig. 3 shows the measured BER performances in the satellite experiments. As Fig. 3 shows, BER performances over satellite transponders were almost the same regardless of spectrum division. This demonstrates that spectrum division by DSDT is feasible and practical in satellite communications.

4 Conclusion

We have been proposing a direct spectrum division transmission (DSDT) technique that can divide a single carrier signal into multiple sub-spectra and assign them to dispersed frequency resources of a satellite transponder to improve the spectrum efficiency of a whole system. For its feasibility verification, we introduced the results of satellite experiments with DSDT over multiple satellite transponders. Our experiments confirmed that DSDE was practical because BER performances of divided signals were almost the same as that of a single carrier.
Color-based registration of point cloud data by video camera for electromagnetic simulation

Zhihang Chen1), Kentaro Saito1, Wataru Okamura2, Yukiko Kishiki2, and Jun-ichi Takada1

1 School of Environment and Society, Tokyo Institute of Technology, O-okayama, Meguro-ku, Tokyo 152–8550, Japan
2 Kozo Keikaku Engineering Inc., 4–38–13 Hon-cho, Nakano-ku, Tokyo 164–0012, Japan

a) czh@ap.idc.titech.ac.jp

Abstract: Radio propagation simulation is a valuable tool to predict the propagation channel characteristics for system design and service-cell planning. In the simulation, the accurate environment model is needed to improve the prediction accuracy of the site-specific channel. Recently, the ability of smartphones makes remarkable progress, and the point cloud data of the environment can be obtained by the video camera output of the smartphone. However, because of the memory size limitation of those devices, it is necessary to repeat the measurement and combine the several point cloud datasets to model the large-volume environment. In this paper, we proposed the color-based registration method by extending the fast descriptor algorithm. We applied our proposed algorithm to the point cloud data taken by the video camera of the smartphone, and proved that the calculation time was reduced by 70% in the office scenario. The work is expected to be utilized for the radio propagation simulation in actual environments.

Keywords: color moment, gray scale variation, indoor environment, point cloud registration, radio propagation simulation

Classification: Antennas and Propagation

References


1 Introduction

To assist with the cell planning of the wireless communication system, the site-specific radio propagation simulation and its prediction accuracy improvement have been actively researched throughout the years. One vital input is the environment model, which should be as close to the actual environment as possible to achieve good prediction performance. It can be manually drawn by commercial software like computer-aided design (CAD) or SketchUp. However, it is very cumbersome and the small details could be occasionally missed, which degrade the prediction performance.

Therefore, to reflect the environment accurately, obtaining the point cloud by the laser scanner, which is a set of points in 3D space presenting the geometry of objects, is an effective way. Recently, the ability of smartphones makes remarkable progress, and using video camera of smartphone to generate the point cloud data becomes a practical solution. However, because of the memory size limitation of those devices, they are not suitable for obtaining the large volume environment at once. Therefore, it is necessary to scan the environment multiple times and register several point clouds to a common coordinate system. The registration is normally done by finding the points from two different sets of point clouds that correspond to the same physical location. These points are called correspondences. After that, these point cloud sets can be combined by using those correspondences as references. Thus, the accuracy of correspondences detection is vital for obtaining a decent registration result.

The fast descriptor algorithm [1] is one of the most popular existing techniques of point cloud registration. In this algorithm, the seed correspondence is found in the initial matching step. Subsequently, more correspondences are estimated based on that seed correspondence. This algorithm has proven to have a good accuracy. However, many seed correspondences derived from the initial matching step are incorrect. This leads to an increase in computation time due to the useless propagation matching of those incorrect correspondences. Thus, this paper aims to improve computation time by introducing the color-based filter to remove improper correspondences.

2 The proposed registration method

2.1 Overview

The goal of point cloud registration is to estimate the transformation matrix between the two sets of point clouds data from a set of seed/point correspondences. Here the correspondence represents a set of points that can be identified or projected in both point cloud data by using transformation matrix. The proposed modification of the fast descriptor algorithm is shown in Fig. 1. In a nutshell, while the conventional descriptor relies only on the geometry of point clouds to initialize the point correspondences, our modification introduces the use of point cloud’s color properties to refine the point correspondences after initial matching. As a result, the number of points is drastically decreased as well as the computation time in the propagation matching and latter processes where the detailed procedures are provided in [1].
2.2 Color moment descriptor

Because the video camera outputs the color properties of the point in addition to its geometry information, utilizing the color property is an effective way to find good correspondences. According to [2], the color moment is an index describing a distribution of color intensity among three color channels; red (R), green (G), and blue (B). Intuitively, the true correspondences between two sets of point clouds, though taken from a different perspective, should have similar color distributions because both are corresponding to the same physical object. In other words, a set of potential incorrect correspondences between two point clouds could be detected and easily filtered out by considering the statistical difference from a color moment.

Suppose the color moment of a particular point cloud $x$ at $k$-th color channel is characterized as the center of a sphere surrounded by a bunch of neighboring points $x_i$ and spanned with radius $r(l)$, three statistical measurements can be described as follows

$$
\begin{align*}
\mu^{l,k} &= \frac{1}{|S(l)|} \sum_{i \in \{j|x_j \in S(l)\} \cap \{k\}} c_{ik}, \\
\sigma^{l,k} &= \left( \frac{1}{|S(l)|} \sum_{i \in \{j|x_j \in S(l)\} \cap \{k\}} (c_{ik} - \mu^{l,k})^2 \right)^{1/2}, \\
\gamma^{l,k} &= \left( \frac{1}{|S(l)|} \sum_{i \in \{j|x_j \in S(l)\} \cap \{k\}} (c_{ik} - \mu^{l,k})^3 \right)^{1/3},
\end{align*}
$$

where $S(l) = \{x_i \mid \|x_i - x\| \leq r(l)\}$, $k \in \{R,G,B\}$. $c_{ik}$, $\mu$, $\sigma$ and $\gamma$ represent color index, mean, standard deviation and skewness, respectively. For the $l$-th layer of radius, the 9-dimensional vector of color moment corresponding to the point $x$ is defined as

$$
c_l = [\mu^{l,R}, \mu^{l,G}, \mu^{l,B}, \sigma^{l,R}, \sigma^{l,G}, \sigma^{l,B}, \gamma^{l,R}, \gamma^{l,G}, \gamma^{l,B}]^T.
$$

Thus, the color moment matrix $C$ with $L$ layers can be represented by

$$
C = [c_1 \ c_2 \ c_3 \ \ldots \ c_L].
$$
2.3 Gray scale variation descriptor
The color moment descriptor alone is not enough to select the correct correspondences because there still exist multiple objects having a similar color distribution. Therefore, in this paper, we introduced the novel descriptor called the gray scale variation to assist the registration process. This descriptor assumes that correspondences between two sets of point clouds should share a dominant direction of gray scale variation \( \mathbf{o}^* \) because they are exposed to the same light intensity. Assuming a unit vector from a particular point \( x \) to the \( i \)-th neighboring point is \( x_i = x - x \), the optimal dominant direction of gray scale variation is determined by the following

\[
\mathbf{o}^* = \arg \max_{\mathbf{o}} \sum_{x \in S(l)} w_i |x_i^T \mathbf{o}|^2.
\] (4)

Where \( w_i = |I[x_i] - I[x]| \), and \( I[\cdot] \) indicate an amplitude variation of gray scale and the function to derive a gray scale value. The problem in Eq. 4 can be easily solved via Singular Value Decomposition (SVD). Let \( o_l^* \) be the gray scale variation at \( l \)-th layer, the gray scale variation matrix \( O \) of \( x \) can be written as,

\[
O = [o_{x_l}^*, o_{y_l}^*, o_{z_l}^*, \ldots, o_{l_l}^*].
\] (5)

2.4 Filtering rule
In fast descriptor algorithm, the correct correspondences between a two set of point clouds \( x \) and \( y \) are determined based on the similarity of normal vector \( n_l \) and variance within the area of \( l \)-th layer. In our modified descriptor algorithm, however, \( O \) and \( C \) are additionally included as additional parameters to effectively filter out the non-correspondences before the propagation matching step as shown in Fig. 1. Specifically, only a pair of \( x \) and \( y \) whose parameters meet all following conditions at all \( L \) layers will be selected as correspondences

\[
\begin{cases}
|\mu_{x_l}^{l,k} - \mu_{y_l}^{l,k}| < \chi_{\mu} \\
|\sigma_{x_l}^{l,k} - \sigma_{y_l}^{l,k}| < \chi_{\sigma} \\
\sigma_{x_l}^{l,k} \cdot \sigma_{y_l}^{l,k} > 0, & l = 1, \ldots, L, \quad k \in \{ R, G, B \}.
\end{cases}
\] (6)

Where \( \theta_l = \cos^{-1}(n_l, o_l) \). \( \chi_{\mu} \), \( \chi_{\sigma} \), and \( \chi_{\theta} \) are the filter rule thresholds of three statistical measurements of certain point.

3 Experiment procedure
In this measurement, point clouds data from two scenarios, office desk, and poster, as shown in Fig. 2 were captured using iPad with Structure Sensor device [3]. Then, these data were processed offline by MATLAB installed on local machine with Intel i7-7700 and 16GB of memory. The office desk scenario is more complicated than the poster because it contains plenty of complicated components such as corners and edges.

In the experiment, point clouds data were captured at the same position twice, and one point cloud dataset was intentionally rotated by random angle. We applied the fast descriptor algorithm and our proposed algorithm to register those two
datasets and evaluated the rotation error of the estimated transformation matrix and calculation time. The thresholds in the filtering rule were determined where $\chi_{11}$, $\chi_{22}$, and $\chi_{0}$ were set to 10, 10, and 15, respectively.

4 Experiment result

The result is shown in Fig. 3. In both scenarios, the proposed algorithm achieved similar accuracy to the conventional algorithm. However, the computation time was drastically decreased, especially in the office scenario. The proposed method was 70% faster than the conventional algorithm. The reason why the improvement was significant in the office scenario is that there are numerous correspondences such as corners and edges of structures. But the proposed algorithm could filter the non-correspondence effectively by combining the color properties. In contrast, the improvement of computation time was rather moderate in the poster scenario because the structures were relatively simple, and the correct correspondences were obtained without the color information.

The results are interpreted that the proposed method is effective in improving the calculation time, and the improvement is expected to be significant in the large-volume environment with the presence of complex structures.
5 Conclusion

This paper aimed to improve the computation time of the conventional fast descriptor algorithm during point cloud registration. The proposed algorithm introduced the color filter to reject incorrect correspondences. The filter contained two types of color descriptors namely color moments and gray scale variation. The result illustrated that, with the effect of color descriptors, the complexity of point clouds registration has been reduced by around 70% in office scenario while maintaining the same level of accuracy. On the contrary, the computation time was marginally decreased in the case of poster scenario due to the simple environmental structure.

Acknowledgments

This work was supported by the MIC SCOPE 185103006. The authors would like to thank Dr. Panawit Hanpinititsak and Mr. Nopphon Keerativoranan for improving the quality of this paper.
Evaluation of a TEM horn antenna for radiated immunity tests in close proximity

Katsushige Harima1(a), Takayuki Kubo2, and Takeshi Ishida2

1 National Institute of Information and Communications Technology,
4–2–1 Nukui-kitamachi, Koganei, Tokyo 184–8795, Japan
2 Noise Laboratory Co., Ltd.,
1–4–4 Chiyoda, Chuo-ku, Sagamihara, Kanagawa 252–0237, Japan
a) harima@nict.go.jp

Abstract: We have proposed a transverse electromagnetic (TEM) horn antenna for radiated immunity tests in close proximity as specified in IEC 61000-4-39 and evaluated the main characteristics of the antenna required in the immunity tests. The TEM horn antenna has broadband radiation characteristics with a low VSWR and generates a homogeneous field over the entire test frequency range from 380 MHz to 6 GHz.

Keywords: immunity, TEM horn antenna, shortened exponential taper, radiated immunity test in close proximity

Classification: Electromagnetic Compatibility (EMC)

References


1 Introduction

Recently, with the widespread availability of portable wireless services such as mobile phones, the situations in which these radio frequency (RF)-transmitting devices are used in close proximity to electronic equipment have become increasingly common. For this reason, immunity requirements for electronic equipment with the aim of ensuring protection against RF sources used in close proximity have been specified in the basic standard IEC 61000-4-39 [1].

The use of TEM horn antennas as radiating antennas is explicitly defined in the standard, and several types of antennas have been proposed for immunity tests [2, 3, 4]. For a linearly tapered TEM horn antenna, resistive tapers are often used for matching the characteristic impedance at the feeding point and to the aperture [2]. A TEM horn antenna with an exponentially tapered structure based on a tapered transmission line does not require resistive loading for impedance matching [5]. This type of antenna has the most suitable characteristics as a broadband antenna among typical tapered transmission lines [6]. However, it does not maintain a single main lobe in the radiation pattern over part of its frequency range. Modified exponentially tapered TEM horn antennas with improved directivity have been proposed [6, 7]. The TEM horn antenna is a balanced antenna such as a dipole and is usually connected to a balun circuit. For radiated immunity testing, the balun should have a broadband, low-loss, and high-power handling capability. For this purpose, a tapered coaxial balun [8] or a balanced feeding mechanism [9] is applicable.

In this work, we proposed a TEM horn antenna with a balanced feeding mechanism for use in close-proximity radiated immunity tests. The antenna characteristics and field uniformity of the exposure area required by the test standard were evaluated experimentally and numerically.

2 Design of the TEM horn antenna

A TEM horn antenna was designed for EMC measurements, especially for a radiated immunity test in close proximity. The close-proximity test is carried out in the frequency range of 380 MHz to 6 GHz, which is used in mobile wireless services [1]. A TEM horn antenna is used as the field-generating antenna; the use of a double-ridged guide horn, LPDA, and other antennas is not allowed in the standard. The TEM horn antenna consists of two metal plates tapered from the
feeding section to the aperture. The tapered structure has to be designed considering the impedance matching at the feeding section and to the aperture. For a linearly tapered TEM horn antenna, which is easy to construct, a resistively loaded taper is used to match the impedances [2]. On the other hand, by applying tapered transmission lines to the antenna shape, the characteristic impedance can be transformed continuously and smoothly from the feeding section to the aperture; that is, resistive loading for matching the impedance is not necessary [5]. The exponential taper has the most suitable radiation characteristics as a broadband antenna among typical tapered transmission lines [6].

Fig. 1 shows a proposed TEM horn antenna with an exponentially tapered shape. The plate of the antenna was made of brass of 0.6 mm thickness, and the spacing between the two plates was maintained using an expanded hard foam with a relative dielectric constant of 1.07. The configuration of these dielectric supporting objects was adjusted while taking into account its influence on the radiation characteristics by computer aided design (CAD) modeling and numerical simulations using CST Studio Suite based on the finite integration technique (FIT). As the antenna parameters for the exponentially tapered shape that we used, the antenna length \( L \) and aperture dimension \( h_L = w_L \) are set to a half wavelength (60 cm) at the lowest frequency (250 MHz). The characteristic impedance \( Z(z) \), the separation \( h(z) \) between the two plates, and the width \( w(z) \) of the plate at the location \( z, 0 \leq z \leq L \), are expressed respectively as [5]

![Diagram](image)

**Fig. 1.** (a) External view and (b) internal structure of proposed TEM horn antenna. Calculated surface current distribution on antenna plate of (c) 870 MHz and (d) 2.45 GHz.
\[ Z(z) = Z_0 e^{z/L \ln(Z_L/Z_0)}, \]  
\[ h(z) = h_0 e^{z/L \ln(h_L/h_0)}, \]

and
\[ w(z) = \frac{h(z)}{Z(z)} 120\pi, \]

where \( Z_0 (= Z(0)) \) is the input impedance at the coaxial feeding point of 50\,\Omega and \( Z_L (= Z(L)) \) is the characteristic impedance at the square aperture of 377\,\Omega. The TEM horn antenna with the exponentially tapered shape does not maintain a single main lobe in the radiation pattern over part of its frequency range. This directivity problem can be improved by shortening the plate, that is, the antenna plate was cut off at 10% \((dL = 0.1L)\) of the antenna length from the aperture [6], as shown in Fig. 1(b).

Since the TEM horn antenna is a balanced antenna such as a dipole, it is usually connected to a balun circuit when feeding by an unbalanced coaxial cable. In particular, for radiated immunity testing, the balun should have a broadband, low-loss, and high-power handling capability. However, by considering the current distribution on the antenna, the common-mode leakage current flowing outside the outer conductor of the coaxial cable can be suppressed without using the balun circuit [9]. The surface current distribution on the tapered plate of the TEM horn antenna is calculated by the FIT solver. Some examples of the current distribution are shown in Figs. 1(c) and (d). It shows that there are high-current parts at the narrow portion and edge of the plate, and low-current parts at the center of the aperture side. A semirigid cable was attached to the plate from the antenna feeding point to a point in the low-current area.

3 Measured results

In a proximity radiated immunity test [1], the equipment under test (EUT) is placed at a distance of 100 mm from the antenna aperture. The field uniformity in the area illuminated by the TEM horn antenna is validated by measurements, and the usable window size of the uniform area of the E-field is defined to be in the range of 0 to 4 dB below the maximum field strength. The antenna is moved according to the window size to fully illuminate all surfaces of the EUT. Therefore, TEM horn antennas that generate larger uniform areas improve test efficiency by reducing the test time.

The experimental setup for the measurement of the field uniformity is shown in Fig. 2(a). The frequency-selective E-field probe was scanned on a test area in 2.5 cm steps using an XY-positioner, and the E-field distribution generated by the TEM horn antenna was measured. The measured field uniformities at frequencies for different mobile wireless services, which are 380 MHz, 870 MHz, 1.845 MHz, 2.45 GHz, and 5.85 GHz, are shown in Figs. 2(b) to (f), respectively. The maximum field strength was located on the antenna axis and at the center of the uniform area over the entire range of test frequencies. The uniform area tends to decrease as the frequency increases, which results in the minimum uniformity window size of 20 cm x 20 cm. The measured \( E_y, E_x, \) and \( E_z \) field distributions at 5.85 GHz are
shown in Figs. 2(f) to (h), respectively. Although the cross-polarization level is not required in the close proximity test, its level is less than $-20 \text{ dB}$ in the on-axis direction. The $E_z$ component appeared to increase at the upper and lower sides owing to the direction of the $E$-field vector. The measurement results are in good agreement with the results of numerical simulations using the antenna CAD model, as shown in Fig. 2(f).

The VSWR of the antenna should not exceed 3:1 as required in a proximity radiated immunity test. The measured VSWR of the TEM horn antenna is shown in Fig. 3(a). Considering that the reflection characteristic rapidly deteriorates at the lower frequency, the antenna geometry parameter in the design was adjusted to 250 MHz so that the VSWR satisfies the IEC requirement at the test frequency of 380 MHz. A good reflection characteristic within a VSWR of 2:1 is obtained owing to the effect of the exponentially tapered transmission line. The reflected wave from the EUT placed close to the antenna causes the deterioration of the VSWR. Assuming the worst-condition, i.e., an EUT with a huge metal surface, the aperture
of the TEM horn antenna was directed at a distance of 100 mm from the metal floor in a semi-anechoic chamber. The VSWR worsens owing to the reflection wave from the metal floor, yet it still satisfied the VSWR limit specified in all test frequency bands.

The test levels for RF fields are classified as 10, 30, 100, and 300 V/m according to the wireless communication equipment [1]. The forward power required to generate 300 V/m at a 100 mm distance from the aperture of the TEM horn antenna is shown in Fig. 3(b). The proposed antenna is highly efficient because it does not use resistive loading or a balun circuit. That is, a forward power of 100 W is required at low frequencies, and that of 40 W is required above 2 GHz.

4 Conclusion

We have presented a shortened exponentially tapered TEM horn antenna used for radiated immunity tests in close proximity. A TEM horn antenna was designed and fabricated using a numerical simulation based on the finite integration technique. The experimental results indicate that the proposed TEM horn antenna has broadband radiation characteristics covering a test frequency range from 380 MHz to 6 GHz and a low VSWR of less than 2:1, which satisfies the requirement of the test standard, and that it generates a homogeneous RF field area with less deformation over the entire test frequency range.
Imbalance state resolving considering flow types

Takuna Kaiwa\textsuperscript{a)} and Nattapon Kitsuwan
Department of Computer and Network Engineering,
The University of Electro-Communications,
1–5–1 Chofugaoka, Chofu-shi, Tokyo 182–8585, Japan
\textsuperscript{a)} tkaiwa@uec.ac.jp

Abstract: This paper proposes a scheme to resolve the imbalance state in a network in which flow types, mice and elephant flows, are considered. A combination of link utilized rate and transmission delay of each link are considered as a link cost. In the proposed scheme, the load imbalance state is resolved by dividing the elephant flow into several subflows and injecting each subflow into multiple paths. The maximum utilization rate of the proposed scheme decreases 38.9%, compared to a conventional scheme.

Keywords: elephant flow, mice flow, routing, scalability

Classification: Network Management/Operation

References


1 Introduction

In computer network, flow type is classified into an elephant flow and mice flow. The elephant flow is an enormously large continuous flow, while the mice flow is a small flow. The number of elephant flows in the network is less than 10% but the elephant flow occupies 40% of the amount of traffic in the network [1]. A load balancing-based routing method must be adopted for each flow type to avoid link congestion and to protect packet processing delay for the mice flow.
Load balancing considering flow types in a network has been studied. An algorithm that combines both a static and dynamic load balancing was introduced [2]. The static load balancing computes the routing in case of no imbalance state. Once the imbalance state occurs, the dynamic load balancing put the elephant flow to a low priority queue. The elephant flow is transmitted when there is no remaining flow with high priority queue in a network node. Head-of-line blocking occurs for the mice flow at the output port when the elephant is transferring. It results in long tail latency for the mice flow. In [1], an elephant flow is split into tiny flows according to the ratio which is inversely proportional to the link load. Each tiny flow is transmitted to its destination with different paths. Both elephant and mice flows are routed with the same routing. As a result, the head-of-line blocking still occurs.

This paper proposes a scheme to resolve the imbalance state by reducing the network congestion and head-of-line blocking for the mice flow by using different path calculations for both the elephant and mice flows in a TCP/IP network. The proposed scheme calculates a path cost based on a transmission delay and link utilization rate. The elephant flow is divided into multiple subflows. Each subflow is transmitted in a different path. Computer simulation shows that 38.9% reduction of the maximum utilization in the proposed scheme is achieved.

2 Related works

2.1 Scalable and fair forwarding of elephant and mice traffic in network

The elephant flow is divided by the number of routes and processed using a source routing method as same as the mice flow [1]. Therefore, the load on the original route that the elephant flow passes before dividing the elephant flow can be reduced. A source node adds the routing to the destination. The controller only updates the routing for the divided elephant flows at the source node. The routing at the source node is not flexible.

2.2 Disturbance based dynamic load balancing

The load balancing process is split into two parts, static load balancing routing (S-LBR) and disturbance based rerouting (D-LBR) [2]. S-LBR determines a static routing for a request by considering the normalized residual bandwidth \( W_{ij}^o \) and normalized link utilization rate \( \eta_{ij}^o \), which are defined as follows.

\[
\eta_{ij}^o = \frac{(\eta_{ij} - \eta_{\text{min}})}{(\eta_{\text{max}} - \eta_{\text{min}})} \quad \text{(1)}
\]

\[
W_{ij}^o = \frac{(W_{ij} - W_{\text{min}})}{(W_{\text{max}} - W_{\text{min}})} \quad \text{(2)}
\]

where \( W_{ij} \), \( W_{\text{max}} \), and \( W_{\text{min}} \) are a residual bandwidth of link \((i, j)\), the maximum, and minimum residual bandwidth of entire links in the network, respectively. \( \eta_{ij} \), \( \eta_{\text{max}} \), and \( \eta_{\text{min}} \) represent a link utilization rate of link \((i, j)\), the maximum, and minimum link utilization of entire links in the network, respectively.

S-LBR finds paths with the minimum hop count from a source to a destination. \( \eta_{ij} \) is considered as a link cost. A path with the minimum cost is selected as a
transmission path. If there are several paths, the best path is selected by considering the distance and the residual bandwidth.

D-LBR, which makes adjustments to balance the load by rerouting the paths in S-LBR when the following condition is satisfied.

\[ \eta(t) \times e < \eta \] (3)

where \( \eta(t) \) represents the average of all link utilization rates at time \( t \), and \( 1 < e < 2 \). First, a flow is categorized into elephant flow and mice flow. Second, a flow that passes a congested link is detoured to reduce the link utilization rate of the congested link by selecting a path with the minimum hop on the detoured path. If there is no candidate for the detoured path, the mice flow is placed in a high priority queue, and the elephant flow is placed in a low priority queue. The mice flow may be routed on the same path with the elephant flow. The transmission delay for the mice flow is long.

3 Proposed scheme

The proposed scheme performs the different path computations between the elephant and mice flows. The elephant flow is split into multiple subflows. The subflows are forwarded along the computed paths as a max-min fairness policy [3]. The mice flow is forwarded along the shortest path. As a result, the proposed method can suppress the rapid increase of the maximum link utilization rate caused by the elephant flow and a congested link. A packet reordering problem occurs in the proposed scheme since the elephant flow is split. This problem can be solved by state reconciliation or threshold adjustment algorithms [4].

3.1 Rerouting mice flow and elephant flow

The proposed method considers the normalized transmission delay, \( d_{ij}^{\text{nl}} \), and normalized link \((i, j)\) utilization rate, \( \eta_{ij}^{\text{nl}} \) as a cost. \( d_{ij}^{\text{nl}} \) at time \( t \) is defined as follow.

\[ d_{ij}^{\text{nl}}(t) = \frac{(d_{ij}(t) - d_{\text{min}})}{(d_{\text{max}} - d_{\text{min}})}, \quad (4) \]

where \( d_{\text{min}} \) and \( d_{\text{max}} \) are the minimum and maximum transmission delays of the entire network link, respectively. The cost of link \((i, j)\), \( l_{ij}^{\text{nl}} \), is defined as follow.

\[ l_{ij}^{\text{nl}} = \alpha \times d_{ij}^{\text{nl}} + (1 - \alpha) \times \eta_{ij}^{\text{nl}}, \quad (5) \]

where \( \alpha \) is a constant value and \( 0 \leq \alpha \leq 1 \). We use two processes to decide the rerouted flow.

1. Selecting of flow rerouting

The proposed method specifies the forwarding route for each flow passing through a congested link. A variance of link utilization rate for traffic pattern \( q \), \( \sigma_{q}^{2} \), is used as a load splitting value of the entire network. The load splitting value of the entire network is calculated as follows.

\[ \sigma_{q}^{2} = \frac{\sum_{k=1}^{K} (\eta_{k}(q) - \overline{\eta}(q))^{2}}{K - 1}, \quad (6) \]

where \( K \) represents the number of links. A traffic pattern \( q \) that achieves the minimum load splitting value \( \sigma_{q}^{2} \) is selected to be rerouted.
2. Flow rerouting

If the selected flow is the mice flow, forward the flow along the shortest path. Otherwise, split the flow as a max-min fairness policy and forward the split flows along the computed paths.

3.2 Flow splitting algorithm for elephant flow

A flow requests bandwidth $B$ from a source to a destination. Let $P$ be a set of splitting paths for the requested flow, where $|P|$ is given. $B$ is split into $|P|$ paths, each path takes the bandwidth $b_p$, where $p$ is an index of element in $P$. A residual bandwidth for path $p$, which is the minimum $W_{mn}$ for links $(m, n)$ on path $p$, is defined as $W_p = \min_{(m,n)} W_{mn}$, where $W_{mn} = C_{mn} - (\eta_{mn} \times C_{mn})$ and $C_{mn}$ is a capacity of link $(m, n)$. The process of flow allocation with flow splitting is as follows.

- Step 1: Compute $t_{ij}$ for all links in the network.
- Step 2: Find all possible set of paths, $Q$, from source to destination. Compute cost of all paths in $Q$, $s_q = \max_{(i,j)} \{ t_{ij} \}$, where link $(i, j)$ is on path $q$ and $q \in Q$.
- Step 3: Among the possible paths, select $k$ paths that have the minimum value of $s_q$. Put those $k$ paths into $P$, and sort the elements in $P$ in ascending order.
- Step 4: If $\sum_{p \in P} W_p < B$, splitting process fails.
- Step 5: Set $p = 0$ and $a = |P|$.
- Step 6: If $B > 0$, set $b = \lfloor \frac{B}{a} \rfloor$. Otherwise, finish the process.
- Step 7: If $W_p \geq b$, set $b_p = b$. Otherwise, set $b_p = W_p$.
- Step 8: Assign $b_p$ to path $p$.
- Step 9: $B = B - b_p$.
- Step 10: Increase $p$ by one, decrease $a$ by one and repeat from step 6.

Fig. 1 shows an example of flow splitting for the elephant flow of the proposed scheme. $\alpha = 0$ is assumed. And all link’s capacity are 100 Mbps. It is assumed that five paths, paths 1 to 5, are available. Fig. 1(a) shows remain bandwidth of every link. Fig. 1(b) shows the computed link cost of every link. $s_1$ to $s_5$ are 0.87, 0.53, 0.47, 0.47, and 0.33, respectively. Assume that $B = 100$ Mbps is split into three paths. The algorithm selects $P = \{\text{Path 5}, \text{Path 3}, \text{Path 4}\}$ as split paths. $b$ on each path is 33 Mbps. The algorithm assigns 33 Mbps to $s_5$ and $s_3$, and 34 Mbps to $s_4$. 

![Image](delay_bandwidth.png)

(a) Delay and utilized bandwidth for each link

![Image](link_cost.png)

(b) Link cost

**Fig. 1.** Example of rerouting elephant flow of proposed scheme.
4 Performance evaluation

We evaluate the maximum utilization of the proposed scheme comparing to the conventional scheme by simulation. The German17 topology which consists of 17 nodes and 26 undirected links, as shown in Fig. 2(a), and the USIP topology which consists of 24 nodes and 41 undirected links, as shown in Fig. 3(a), are used in the simulation. 30 flows are generated. Three of them are the elephant flows and 27 of them are the mice flows. The size of the elephant flow is assumed between 100 Mbps and 200 Mbps, the flow size that less than 100 Mbps is assumed as the mice flow. Source and destination of each flow are randomly selected. Each link capacity is randomly set between 500 Mbps and 800 Mbps. The number of simulation is repeated for 1,000 times. If the elephant flow is split into subflows, the maximum size of each subflow is assumed to 30 Mbps. The subflows are split as a per-packet manner. We determine the path of the flow using SLBR before network operation, and perform load balancing when the imbalance state occurs. The strict source and record route is used in the simulation. Packet reordering is omitted.

Fig. 2 and 3 compare the maximum utilization rates of the conventional and proposed schemes in the German17 and USIP topologies, respectively. Figs. 2(b) and 3(b) show the number of congestions of each link. A link with the high number of congestions implies that the bottleneck easily occurs. Figs. 2(c) and 3(c) show that the maximum utilization rate of the proposed scheme achieves 21.1% and 38.9% reduction for German17 and USIP, respectively.

![Fig. 2. Simulation with German17](image1)

![Fig. 3. Simulation with USIP](image2)
5 Conclusion

A scheme to resolve the imbalance state in a network was proposed. In this scheme, a flow to be rerouted is determined by a traffic pattern. If the rerouted flow is the mice flow, it is rerouted through the shortest path. If the rerouted flow is the elephant flow, the flow is split as max-min fairness. Each split flow is routed on a different path. The simulation results showed that the proposed scheme reduces 38.9% of the maximum bandwidth for our examined topologies, compared to the conventional scheme.