A method of controlling the base station correlation for MIMO-OTA based on Jakes model

Kazuhiro Honda\textsuperscript{a)} and Kun Li

Graduate School of Engineering, Toyama University,
3190 Gofuku, Toyama-shi, Toyama 930–8555, Japan
\textsuperscript{a)} hondak@eng.u-toyama.ac.jp

Abstract: This paper presents a methodology of controlling the spatial correlation of BS (base station) antenna realized by a bilateral fading emulator for MIMO-OTA (Over-The-Air) testing. On the basis of Jakes theory, the base station correlation is controlled by setting the initial phases of scatterers in a two-dimensional fading emulator. The experimental results show that the designed correlation coefficient of an uplink channel using the proposed method agrees well with the theoretical value. Further, it is confirmed that the channel capacity of uplink channel can be controlled when the base station and mobile terminal correlations are changed simultaneously.

Keywords: base station antenna correlation, Jakes model, MIMO-OTA, fading emulator, channel capacity

Classification: Antennas and Propagation

References


1 Introduction

Multiple-input multiple-output (MIMO) is a key technique to the success of forthcoming ultra-high-speed cellular systems [1]. To evaluate the performance of a MIMO device, Over-The-Air (OTA) testing is a widely approved method. In the previous study [2], the channel capacity of a MIMO antenna is evaluated using a two-dimensional fading emulator in downlink channel. However, the evaluation of the channel capacity of uplink channel was not clarified.

In future MIMO systems, a large capacity of uploading data from a handset to base station is anticipated, which means that the channel capacity needs to be evaluated not only in downlink but also in uplink channel. Since the correlation of a base station antenna is known to possess a relatively high value even in the case of the array spacing of several wavelengths in a small cell environment [3], it is important to consider a base station correlation when the evaluation of the channel capacity of a MIMO antenna in uplink channel is conducted.

This paper presents a methodology of controlling the spatial correlation of base station antenna realized by a bilateral fading emulator for MIMO-OTA testing [4]. On the basis of the Jakes theory, a method of achieving the spatial correlation of a base station antenna is introduced by setting the initial phases of scatterers in a two-dimensional channel model. Further, the channel capacity of uplink channel is evaluated when the base station correlation and terminal correlation are varied simultaneously.

2 A method of controlling the uplink correlation

We are developing a bilateral MIMO-OTA evaluation apparatus for measuring the channel capacity of uplink as well as downlink channels [5]. When the channel capacity of uplink channel is measured, RF signals radiated from a DUT MIMO antenna are received by each probe antenna and then the phase of the received signals is controlled by a circuit. Since the base station antenna with a separation of 5\(\lambda\) at 2 GHz is known to have a high correlation [3], it is important to consider a high base station correlation in uplink channel.

We have attempted to control the base station correlation using the initial phase of scatterers. Fig. 1 shows the implementation model for realizing the base station correlation. In Fig. 1, \(h_{ij}\) is the channel response from DUT \#j to Rx \#i. The fading emulator is created based on a two-dimensional channel model [2].

As shown in Fig. 1, RF signals radiated from a DUT MIMO antenna located at the center of the emulator are transmitted to the scatterers and then arrive at the base station (Rx1, Rx2). The different paths are realized by combining different initial phase matrices \(\Phi_1\) and \(\Phi_2\).
In general, the correlation in the uplink channel involves the characteristics of a radio wave propagation represented by multipath waves and a base station antenna. In this paper, the sum of the abovementioned two effects is attributed to the characteristics of the multipath waves. Therefore, the initial phase matrices $\Phi_1$ and $\Phi_2$ have the correlation characteristics that include the combined effects of both the multipath waves and base station antenna in an uplink channel. Based on this concept, the base station correlation characteristics can be controlled by setting the initial phases of scatterers using the Jakes model, as mentioned in the following way.

Although Jakes model is a theory for reception in downlink channel, it can also be applied to the uplink channel for the sake of bilateral nature in the radio wave propagation [6]. When a MIMO antenna placed at the center of the two-dimensional channel model is moved in a distance $d$, the geometrical phase difference $\alpha_i$ between the antenna and each scatterer can be calculated by Eq. (1) using the geometry shown in Fig. 1.

$$\alpha_i = kd \cos \phi_i$$  \hspace{1cm} (1)

where $k = 2\pi/\lambda$ denotes the wave number.

The autocorrelation function between the received signal of Rx1 radiated from the initial position (origin) and that of Rx2 radiated from the position moved in a distance of $d$ is obtained as the Bessel function, $J_0(kd)$, from the Jakes model, as shown by the black curve in Fig. 2. The antenna separation $d$ is determined by a least-mean-square (LMS) function using the designed correlation coefficient $\rho_{BS}$, as shown by the red curve. The LMS function is given by Eq. (2).

$$d(\rho) = -0.46\rho^3 + 0.55\rho^2 - 0.45\rho + 0.39$$  \hspace{1cm} (2)

Using this antenna separation $d$, the geometrical phase difference $\alpha_i$ is calculated by Eq. (1). Using this phase difference, the initial phase matrix $\Phi_2$ of the base station 2 (Rx2) can be calculated as the summation of the initial phase matrix $\Phi_1$ of the base station 1 (Rx1) and the geometrical phase difference matrix $A$, as expressed by Eq. (3).

$$\Phi_2 = \Phi_1 + A$$

$${\begin{bmatrix} \varphi_1 + \alpha_1 & \varphi_2 + \alpha_2 & \ldots & \varphi_n + \alpha_n \end{bmatrix}}$$  \hspace{1cm} (3)
In Eq. (3), the initial phase matrix $\Phi_1$ indicates the uncorrelated initial phase generated by random numbers. Thus, the correlation between the signals of the two scatterering groups forming the base stations 1 and 2 can be controlled in accordance with the designed base station correlation.

### 3 Experimental results

In order to confirm the validity of the proposed method, measurements have been carried out using a bilateral fading emulator for MIMO-OTA [5]. The DUT MIMO antenna is comprised of half-wavelength dipole antennas, which are constructed using a single dipole antenna placed at two different locations forming a quasi-array as shown in Fig. 3(a).

In this paper, the base station correlation $\rho_{BS}$ and terminal correlation $\rho_{MS}$ are defined using the channel responses, as expressed by the following equations.

$$
\rho_{BS} = \frac{1}{2} \left( \frac{|h_{11}h_{21}^*|}{\sqrt{|h_{11}h_{11}^*|\sqrt{|h_{21}h_{21}^*|}}} + \frac{|h_{12}h_{22}^*|}{\sqrt{|h_{12}h_{12}^*|\sqrt{|h_{22}h_{22}^*|}}} \right) 
$$

$$
\rho_{MS} = \frac{1}{2} \left( \frac{|h_{11}h_{12}^*|}{\sqrt{|h_{11}h_{11}^*|\sqrt{|h_{12}h_{12}^*|}}} + \frac{|h_{21}h_{22}^*|}{\sqrt{|h_{21}h_{21}^*|\sqrt{|h_{22}h_{22}^*|}}} \right) 
$$

where the asterisk (*) denotes the complex conjugate.

The correlation of a mobile terminal $\rho_{MS}$ can be controlled by setting the antenna separation using the Jakes model as shown in Fig. 2. The antenna separation of DUT array antenna $L$ is set to 0.29$\lambda$, which means that $\rho_{MS}$ is designed as 0.3. Fig. 3(b) shows a photograph of the DUT antenna at the center of a fading emulator. The measurement frequency is set to 1.95 GHz which is the average frequency of uplink channel at 2 GHz band in Japan.

Fig. 3(c) shows the instantaneous channel response when the designed base station correlation $\rho_{BS}$ is set to 0.9 with a distance $d$ in Fig. 1 of 0.09$\lambda$. The black curve indicates $h_{11}$ whereas the blue curve indicates $h_{21}$, respectively, when the transmitting dipole antenna Tx1 is placed at the center of the fading emulator. In this case, the correlation between the received signals of each base station antenna is found to be 0.91. Moreover, the red curve indicates $h_{12}$ whereas the pink curve indicates $h_{22}$, respectively, when the dipole antenna Tx2 is separated with a distance of 0.29$\lambda$ away from the center of the fading emulator which allows $\rho_{MS}$ to be 0.3. Here, the correlation between the received signals of each base station antenna is
found to be 0.92. Therefore, when the location of the DUT antenna is moved, the correlation between the received signals of base station antennas agrees well with the designed base station correlation $\rho_{BS}$. This means that, using the proposed method, we can realize the same $\rho_{BS}$ values for the DUT antenna placed at different locations in the emulator.

Fig. 3(d) shows the measured results of the correlation coefficient for a 2 × 2 MIMO as a function of the designed base station correlation $\rho_{BS}$ when $XPR$ is 50 dB. In Fig. 3(d), the symbols $\bullet$ show the measured results obtained from the fading emulator while the black curve indicates the analytical outcome calculated.
by the Monte Carlo simulation [7]. It can be seen that the designed base station correlation agrees well with the correlation analyzed by the Monte Carlo simulation. This fact confirms that the base station correlation can be controlled using the proposed method.

In our previous study [4], the impact of the base station correlation on the channel capacity is not investigated. In this paper, the MIMO channel capacity with regard to the different base station correlations is measured. Fig. 3(e) shows the measured and analytical results of $2 \times 2$ MIMO channel capacity with the designed base station correlation $\rho_{BS}$ varied from 0 to 0.95. $\rho_{MS}$ is set to 0.3 and 0.9. $XPR$ is 50 dB while $SNR$ of the incident wave is 30 dB. As can be seen in Fig. 3(e), the measured results agree well with the analytical outcomes. Further, the channel capacity is varied depending on $\rho_{MS}$. Therefore, using the proposed method, the combined evaluation of a handset MIMO antenna in consideration of both uplink and downlink channel properties can be realized.

4 Conclusion

This paper presents a method of controlling the base station correlation for the combined evaluation of a handset MIMO-OTA in consideration of both uplink and downlink channels. The results show that the measured correlation coefficient of uplink channel agrees well with the theoretical values, confirming the effectiveness of the proposed method. Further, the channel capacity of uplink channel is determined by changing not only the base station correlation but also the terminal correlation. Future studies include the derivation of the initial phase of scatterers in accordance with various $XPR$s.

Acknowledgments

This work was entrusted by the MIC-SCOPE 2015.
Success prioritized
distributed coordination
function

Daisuke Umehara\(^{1(a)}\), Hidekazu Murata\(^2\), and Satoshi Denno\(^3\)

\(^1\) Faculty of Information and Human Sciences, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606–8585, Japan
\(^2\) Graduate School of Informatics, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606–8585, Japan
\(^3\) Graduate School of Natural Science and Technology, Okayama University, 3–1–1 Tsushima-naka, Kita-ku, Okayama 700–8530, Japan

\(^a\) umehara@kit.ac.jp

Abstract: In this letter, we present a random access protocol with success priority, success prioritized distributed coordination function (SP-DCF), based on IEEE 802.11 DCF. Success stations (STAs) are prioritized contentionlessly to transmit the next data frame in the transmit queue in a random manner. We develop a performance analysis model of SP-DCF which enables us to estimate a variety of performance indexes such as the throughput or the frame discard rate (FDR) in a high accuracy for saturated wireless networks. The numerical analysis and simulation results show that the proposed SP-DCF achieves higher throughput and lower FDR as compared with the conventional DCF especially for heavily-congested wireless networks.

Keywords: IEEE 802.11 DCF, success priority, performance analysis model, consecutive successes, throughput, frame discard rate

Classification: Fundamental Theories for Communications

References

1 Introduction

IEEE 802.11 distributed coordination function (DCF) is a distributed MAC protocol with the function of clear channel assessment (CCA) established by IEEE 802.11 standards [1, 2]. Bianchi presented a simple and precise performance analysis model by using a two-dimensional Markov chain [3]. A station (STA) sends a frame without contention when the random backoff counter is allocated to zero after it sends the previous frame successfully in the IEEE 802.11 DCF originally developed in the last half of 1990s. The throughput will be higher due to the effect of this priority particularly when the minimum contention window is extremely small. Some of researches have been reported with respect to this phenomenon such as a modification of Bianchi’s performance analysis model [4] and a prioritized access control for bidirectional traffic between an access point (AP) and a group of STAs [5].

In this letter, we propose a success prioritized DCF (SP-DCF), in which a success STA is more actively prioritized among all STAs in the wireless networks [6]. The priority of a success STA becomes high with increasing differential inter-frame space number (DiIFSN) between a success STA and all other deferred STAs. The success STA has an opportunity to contentionlessly transmit the next data frame so that it reduces the frame collision probability. As a result, the success priority enables us to enhance the throughput and reduce the frame discard rate (FDR). We also develop a performance analysis model of SP-DCF with high accuracy extended from Bianchi’s performance analysis model.

2 Success prioritized DCF

Let us consider single-hop wireless networks with an AP and n STAs where \( n \geq 2 \). AP does not have any data frames at STAs while STAs always have data frames at AP. Each STA has one wireless medium interface so that it cannot transmit a data frame and assess the channel simultaneously. AP and STAs are not hidden with each other.

2.1 IEEE 802.11 DCF

Let us describe the binary exponential backoff algorithm of IEEE 802.11 DCF. The backoff algorithm is conducted before the data frame transmission when the transmit queue is non-empty or the channel is detected busy during the DCF IFS (DIFS) time \( T_{\text{difs}} \). A backoff counter \( k \) is uniform-randomly selected from an integer set of \([0, \text{CW}]\), where the contention window (CW) is initialized to \( \text{CW}_{\text{min}} = W - 1 = W_0 - 1 \) if it is the first time to transmit the data frame. The positive \( k \) decrements by one when the channel is idle during the slot time \( \sigma \). The decrease of \( k \) is deferred when the channel is detected busy and it decrements by one when it
observes an idle channel during $T_{\text{difs}}$. The STA starts transmitting the data frame after the decremented $k$ becomes 0. AP correctly receives the data frame from a STA and then it transmits the ACK frame after elapsing the short IFS (SIFS) time $T_{\text{sifs}}$. The data frame transmission is successful when the ACK frame is received at the source STA and CW is initialized to $CW_{\text{min}}$. It fails otherwise and the STA attempts to retransmit the same data frame with the updated $CW = \min\{2(CW + 1) - 1, CW_{\text{max}}\}$ when the number of retransmissions is less than an integer of $r$, where $CW_{\text{max}} = 2^m(CW_{\text{min}} + 1) - 1$, and $r + 1$ stand for maximum contention window, maximum binary exponent, and retry limit, respectively. The number of retransmissions reaches $r$ and then the data frame is discarded when it fails and CW is initialized to $CW_{\text{min}}$. CW in the $i$-th transmissions of the same data frame for $i \in [0, r]$ is calculated as $CW_i = W_i - 1 = 2^u(CW_{\text{min}} + 1) - 1$ where $u_i = \min\{m, i\}$.

### 2.2 SP-DCF and its performance analysis model

The success priority is established by alternating the IFS with the current STA state in the proposed SP-DCF. The IFS after transmitting the data frame successfully is called success IFS ($SuIFS$), $T_{\text{suifs}}$, whereas the IFS after detecting the channel busy is called busy IFS ($BuIFS$), $T_{\text{buifs}}$. The DiIFSN is defined as an integer of $D = \left(\frac{T_{\text{buifs}} - T_{\text{sufs}}}{\sigma}\right) \in [0, W - 1]$. The $D$ is positive in the proposed SP-DCF whereas it is equal to 0 in the conventional DCF. The next data frame after a success has a backoff counter $k$ uniform-randomly distributed over $[0, CW_{\text{min}}]$ so that it is consecutively transmitted without contention when $k$ is less than $D$. The probability of $m$ consecutive successes is expressed as $(1 - D/W) \cdot (D/W)^{m-1}$ so that the average number of consecutive successes is derived as

$$N_S = \left(1 - \frac{D}{W}\right) \sum_{m=1}^{\infty} m \cdot \left(\frac{D}{W}\right)^{m-1} = \frac{W}{W - D}. \quad (1)$$

The average number of idle slots consumed during consecutive successes without contention is derived as

$$N_{\sigma} = D + \frac{D - 1}{2} \cdot \left(1 - \frac{D}{W}\right) \sum_{m=1}^{\infty} m \cdot \left(\frac{D}{W}\right)^{m} = D \cdot \left(1 + \frac{D - 1}{2(W - D)}\right). \quad (2)$$

Let us define three variable-length slots which have the same backoff counter for channel-assessing STAs and they are classified into idle, success, and collision slots. Let us consider two-dimensional states of backoff stage $i$ and backoff counter $k$, and their state transitions. The collision slot probability $p$ by a frame transmission is assumed to be independent of its current backoff stage. The steady-state probability of a backoff state $(i, k)$ is denoted as $b_{i,k}$. The transmit probability $\tau$ by a slot is expressed as the summation of $b_{i,0}$ over $i \in [0, r]$ and is derived as

$$\tau(p) = \frac{2}{1 - D \cdot (1 - p) + W \cdot \left[\sum_{i=0}^{u-1} (2p)^i + (2p)^n \sum_{i=0}^{r-u} p^i\right]} \bigg/ \sum_{i=0}^{r} p^i. \quad (3)$$
where \( u = u_r \). The \( p \) is expressed as \( p = 1 - (1 - r)^{n-1} \), i.e.

\[
\tau(p) = 1 - (1 - p)^{\frac{1}{1-r}}.
\]

The values of \( r \) and \( p \) are derived by solving the non-linear system, composed of monotonically decreasing Eq. (3) with \( \tau(0) = 2/(1 - D + W) \) and strictly monotonically increasing Eq. (4) with \( \tau(0) = 0 \) and \( \tau(1) = 1 \). Since one or more successes are included in a success slot, the collision probability \( p_c \) by a frame transmission is expressed as

\[
p_c = \frac{p}{N_S \cdot (1 - p) + p} = \frac{W - D}{W - p \cdot D} \cdot p,
\]

by compensating with \( N_S \).

The idle slot probability \( P_\sigma \), the success slot probability \( P_S \), and the collision slot probability \( P_C \) by a slot are expressed as \( P_\sigma = (1 - \tau)^n \), \( P_S = n \tau(1 - \tau)^{n-1} \), and \( P_C = 1 - P_\sigma - P_S \), respectively. Let us denote the average data frame time as \( T_{data} \) and the ACK frame time as \( T_{ack} \). The propagation delay \( \delta \) between any two STAs is assumed to be constant. The time consumed with a success of a single data frame is estimated as

\[
T_S = T_{data} + T_{ack} + T_{sifs} + T_{suifs} + 2\delta.
\]

The time consumed with a collision of two or more data frames for channel-assessing STAs is estimated as

\[
T_C = T_{data} + T_{buifs} + \delta.
\]

The throughput is expressed as

\[
S = \frac{P_S \cdot N_S \cdot M \cdot B}{P_\sigma \cdot \sigma + P_S \cdot (N_\sigma \cdot \sigma + N_S \cdot T_S) + P_C \cdot T_C},
\]

where \( B \) and \( M \) stand for the average data MAC service data unit (MSDU) size and the average number of aggregated MAC protocol data unit (A-MPDU) subframes. Since the data frame is discarded when the ACK frame is not received even in the \( r \)-th retransmission, FDR is expressed as

\[
q_{fdr} = \frac{p^{r+1}}{N_S \cdot (1 - p^{r+1}) + p^{r+1}} = \frac{W - D}{W - p^{r+1} \cdot D} \cdot p^{r+1}.
\]

3 Numerical analysis and simulation results

The parameters in the MAC layers of DCF and SP-DCF are illustrated in Table I. STAs do not detect any other data frames in the air during \( T_{cns} \) before their own frame transmissions. The propagation delay is assumed to be \( \delta = 0.1 \mu s \). Let us employ \( 2 \times 2 \) MIMO, 108 data subcarriers, 16-QAM, rate-1/2 LDPC code, and short guard interval (GI) of 0.4 \mu s. As a result, the PHY rate becomes 120 Mbps.
Simulations including the effect of $T_{\text{cns}}$ and $T_{\text{ato}}$ are conducted to estimate throughput and FDR by using GNU Octave. They start on condition that the contention window of all STAs is set at $CW_{\text{min}}$ and terminate when $10^6$ data frames are successfully transmitted.

Fig. 1 illustrates the numerical analysis and simulation results with respect to throughput and FDR versus the number of STAs for DCF and SP-DCF with $D = 4$, $8$, and $12$. The numerical analysis results are drawn as lines whereas the simulation results are plotted as points, and both the results are agreed with each other in high accuracies. The throughput and FDR of SP-DCF is better than those of DCF and the differences become large as increasing $D$ and $n$. The simulation result for $n = 100$ shows that the throughput of DCF is 47.8 Mbps while that of SP-DCF with $D = 12$ is 76.8 Mbps so that SP-DCF enables to enhance the throughput of

### Table 1. The parameters in the MAC layers.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data MSDU size, $B$</td>
<td>1,500 bytes</td>
</tr>
<tr>
<td># of A-MPDU subframes, $M$</td>
<td>10</td>
</tr>
<tr>
<td>ACK MPDU size</td>
<td>32 bytes</td>
</tr>
<tr>
<td>Slot time, $\sigma$</td>
<td>9 $\mu$s</td>
</tr>
<tr>
<td>SIFS time, $T_{\text{sifs}}$</td>
<td>16 $\mu$s</td>
</tr>
<tr>
<td>DIFS time, $T_{\text{difs}}$</td>
<td>34 $\mu$s</td>
</tr>
<tr>
<td>PHY-RX-START delay, $T_{\text{prsd}}$</td>
<td>25 $\mu$s</td>
</tr>
<tr>
<td>Carrier non-sensing time, $T_{\text{cns}}$</td>
<td>4 $\mu$s</td>
</tr>
<tr>
<td>Minimum contention window, $CW_{\text{min}}$</td>
<td>15</td>
</tr>
<tr>
<td>Maximum binary exponent, $m$</td>
<td>6</td>
</tr>
<tr>
<td>Retry limit, $r + 1$</td>
<td>7</td>
</tr>
<tr>
<td>SUIFS time, $T_{\text{suifs}}$</td>
<td>16 $\mu$s</td>
</tr>
</tbody>
</table>

Fig. 1. Throughput and FDR versus the number of STAs.
29.0 Mbps as compared with DCF. It also shows that the FDR of DCF is 12.6% while that of SP-DCF with $D = 12$ is 3.6% so that SP-DCF is extremely effective for the reduction of FDR and it will be good matched with the behavior in the upper layers.

4 Conclusion

In this letter, we have developed SP-DCF and its performance analysis model. The numerical analysis and simulation results have agreed with each other in high accuracies and have shown that the proposed SP-DCF achieves higher throughput and lower FDR as compared with the conventional DCF especially for heavily-congested wireless networks.

Acknowledgment

This research and development work was supported by JSPS KAKENHI Grant Numbers JP15H04007, JP16K14268, and the MIC/SCOPE #165007006.
MS-SSIA: multiple snapshot spatial smoothing with improved effective array aperture for high-resolution DOA estimation

Yuki Iizuka and Koichi Ichige
Department of Electrical and Computer Engineering, Yokohama National University, 79–5 Tokiwadai, Hodogaya-ku, Yokohama 240–8501, Japan
a) koichi@ynu.ac.jp

Abstract: This paper presents multiple snapshot spatial smoothing with improved aperture (MS-SSIA) for high-resolution direction of arrival (DOA) estimation by uniform linear array (ULA). Spatial smoothing preprocessing (SSP) is often used for DOA estimation of correlated sources, but it reduces the effective array aperture and leads to low DOA estimation accuracy. SSIA and SSOA improve the problem of SSP but they have some other problems. This paper tries to improve SSIA so that it can correspond to the case of multiple snapshots, but we take a different approach with SSOA. The performance of the proposed method is evaluated through computer simulation.

Keywords: direction of arrival estimation, array signal processing, spatial smoothing

Classification: Antennas and Propagation

References


1 Introduction

Direction-of-arrival (DOA) estimation plays an important role in radar, sonar, and indoor and outdoor wireless communications. High resolution DOA estimation methods using sensor arrays have been studied in the last three decades and have attracted much attention. The most well-known methods, MUSIC, Root-MUSIC, ESPRIT, and Unitary-ESPRIT [1], are based on eigenvalue decomposition of a sample covariance matrix of an array input signal. Those algorithms generally require spatial smoothing preprocessing (SSP) [2] or forward-backward SSP (FB-SSP) techniques in estimating DOAs of coherent sources to suppress the correlation between signals. Such techniques are effective to reduce signal correlation but often lead to low DOA estimation accuracy due to the small array aperture caused by spatial averaging.

Spatial smoothing with improved aperture (SSIA) [3] can suppress signal correlation while not reducing the array aperture; however, it can only be applied to the case of a single snapshot because the array steering matrix in SSIA becomes time-dependent and therefore cannot support the case of multiple snapshots. On the other hand, spatial smoothing with overlapped and augmented array (SSOA) has also been proposed [4]. It can correspond to the case of multiple snapshots and can solve the SSP and SSIA problem, but its DOA estimation accuracy does not become superior to that of the SSP method.

In this paper, we try to improve the SSIA algorithm so that it can correspond to the case of multiple snapshots but we take a different approach with SSOA. Starting from the initial DOAs estimated by the MODE method, we create the virtual signal vector that corresponds to the complex conjugates of source signals. Then, the DOAs are recursively updated by augmented array processing. Hereafter, the proposed method is called multiple snapshot SSIA (MS-SSIA) to distinguish it from the original single snapshot SSIA (SS-SSIA). The performance of the proposed method is evaluated through computer simulation and compared with the performances of some conventional methods.

2 Signal model

Assume that $L$ far-field incident signals are received by an $M$-element ULA in an additive white Gaussian noise (AWGN) environment, where the signals and noises are statistically independent. The array input vector $x(t)$ received by the $M$-element ULA can be written as

$$x(t) = As(t) + n(t), \quad (1)$$

where $A$, $s(t)$, and $n(t)$ denote the array steering matrix, incident signal vector, and noise vector, respectively. The array steering matrix $A$ in (1) is given by
\[
A = \begin{bmatrix}
o^T \\
o^T \Phi \\
o^T \Phi^2 \\
\vdots \\
o^T \Phi^{M-1}
\end{bmatrix},
\]

where

\[
\Phi = \text{diag}[e^{i\phi_1}, e^{i\phi_2}, \ldots, e^{i\phi_L}], \quad o = [1, 1, \ldots, 1]^T.
\]

The phase term \(\phi_L\) in (3) is given by \(\phi_L = -2\pi d \sin \theta_L/\lambda\), where \(\theta_L\), \(\lambda\), and \(d\) denote the spatial angle to be estimated, wavelength, and inter-element spacing, respectively.

### 3 Conventional methods

Let \(N\) and \(K(=M-N+1)\) denote the number of subarrays and the number of array elements in a subarray, respectively. SSP [2] is first applied to the array input vector \(x(t)\) to obtain the measurement matrix \(X_{SS}(t)\) of which the rank becomes greater than or equal to \(L\). Note that \(L\) is assumed to be known in advance. Then, the measurement matrix \(X_{SS}(t)\) in SSP can be written as

\[
X_{SS}(t) = A_S \Lambda(t) B_S + N_{SS}(t),
\]

where

\[
A_S = [I_K, 0_{K \times (M-K)}], \quad B_S = [o, \Phi o, \Phi^2 o, \ldots, \Phi^{L-1} o],
\]

\[
\Lambda(t) = \text{diag}(s(t)).
\]

\[
N_{SS}(t) = [J_1^{(M)} n(t), J_2^{(M)} n(t), \ldots, J_N^{(M)} n(t)],
\]

\[
J_n^{(M)} = [0_{K \times (n-1)}, I_K, 0_{K \times (N-n)}], \quad n = 1, \ldots, N.
\]

Note that \(A_S\) and \(N_{SS}(t)\) in (3) denote the modified array steering matrix and the noise matrix, respectively. Each column of \(N_{SS}(t)\) is white but the columns of \(N_{SS}(t)\) are mutually correlated. Besides, both \(A_S\) and \(B_S^T\) become steering matrices of centro-symmetric arrays.

Exploiting the structures of \(A_S\) and \(B_S\), equation (3) is modified into the following augmented matrix representation:

\[
X_{SS-SSIA}(t) = \begin{bmatrix}
X_{SS}(t) \\
\Pi_K X_{SS}(t) \Pi_N
\end{bmatrix} = \begin{bmatrix}
A_S \Lambda(t) \\
A_S \Lambda(t) A^*(t) A_2
\end{bmatrix} B_S + \begin{bmatrix}
N_{SS}(t) \\
\Pi_K N_{SS}(t) \Pi_N
\end{bmatrix} = A_{SS-SSIA}(t) B_S + N_{SS-SSIA}(t),
\]

(4)
where $\Lambda_1$ and $\Lambda_2$ are unitary diagonal matrices of size $L \times L$, and $\Pi_p$ is a $p \times p$ exchange matrix with ones on its anti-diagonal and zeros elsewhere [3].

The matrix $A_{SS-SSIA}(t)$ in (4) becomes a centro-symmetric array of size $2K$ and satisfies the shift invariance property. Therefore, we can apply the Unitary ESPRIT method to (4) and obtain the estimated DOAs $\{\hat{\theta}_i\}_{i=1}^L$ for the case of a single snapshot. However, SS-SSIA does not work well for the case of multiple snapshots because the matrix $A_{SS-SSIA}(t)$ depends on time and varies for each snapshot.

SSOA [4] is an improved technique using the augmented array steering matrix in a similar manner to SS-SSIA and can work well for the case of multiple snapshots. However, with our confirmation through numerical simulation, the value of $K$ in [4] for the existing methods is not optimized. The performance of SSOA becomes the same as that of SSP if we choose the optimum value of $K$ for the existing methods.

4 Proposed method: MS-SSIA

As discussed in the previous section, the matrix $A_{SS-SSIA}(t)$ in SS-SSIA varies in accordance with the snapshot because of two time-dependent matrices $\Lambda(t)$ and $\Lambda^*(t)$. This section presents how the SS-SSIA is modified to obtain the proposed method MS-SSIA. We construct an augmented matrix, which has only one time-dependent matrix $\Lambda(t)$, and this leads to a method that can work well for the case of multiple snapshots.

4.1 Virtual signal vector

The proposed MS-SSIA is an iterative algorithm that is based on updating the array steering matrix $A$ and DOAs $\theta_i$. Let $\{\hat{\theta}_{i,j}\}_{j=1}^L$ denote the estimated DOAs for the $i$-th iteration. We first apply the MODE method to (3) and obtain the initial DOA estimates $\hat{\theta}_{i,0}$.

For the case of the $i$-th iteration ($i = 0, 1, 2, \ldots$), the Unitary ESPRIT method can estimate the DOAs $\hat{\theta}_{i,j}$ as well as the corresponding steering matrix $\hat{A}_i$ and source signal vector $\hat{s}_i(t)$. Then, the virtual input vector $\tilde{x}(t)$, which corresponds to the complex-conjugate source signal matrix $\Lambda^*(t) = \text{diag}\{s^*(t)\}$, can be written as

$$\tilde{x}(t) = As^*(t) + \tilde{n}(t) \approx \hat{A}_i\hat{s}_i^*(t),$$

where

$$\hat{A}_i = [a(\hat{\theta}_{1,j}), a(\hat{\theta}_{2,j}), \ldots, a(\hat{\theta}_{L,j})],$$

$$\hat{s}_i(t) = \hat{A}_i^+x(t),$$

where $\hat{A}_i^+ = (\hat{A}_i^H\hat{A}_i)^{-1}\hat{A}_i^H$ is the pseudo-inverse matrix of $\hat{A}_i$. Then, its measurement matrix $\tilde{X}_{SS}(t)$ is given by

$$\tilde{X}_{SS}(t) = A_S\Lambda^*(t)B_S + \tilde{N}_{SS}(t),$$

where

$$\tilde{N}_{SS}(t) = [J_1^{(M)}\tilde{n}(t), J_2^{(M)}\tilde{n}(t), \ldots, J_N^{(M)}\tilde{n}(t)],$$
4.2 DOA estimation using augmented matrix

In a similar manner to SS-SSIA, we construct the following augmented matrix:

\[
X_{\text{MS-SSIA}}(t) = \begin{bmatrix}
X_{\text{SS}}(t) \\
\Pi_k \hat{X}_{\text{SS}}^*(t) \Pi_N
\end{bmatrix} = \begin{bmatrix}
A_S A(t) B_S \\
A_S A_1 A(t) A_2 B_S
\end{bmatrix} + \begin{bmatrix}
N_{\text{SS}}(t) \\
\Pi_k \hat{N}_{\text{SS}}(t) \Pi_N
\end{bmatrix} = A_{\text{MS-SSIA}} A(t) B_S + N_{\text{SS-SSIA}}(t), \tag{7}
\]

where

\[
A_{\text{MS-SSIA}} = \begin{bmatrix}
A_S \\
A_S A_1 A_2
\end{bmatrix},
\]

\[
N_{\text{MS-SSIA}}(t) = \begin{bmatrix}
N_{\text{SS}}(t) \\
\Pi_k \hat{N}_{\text{SS}}(t) \Pi_N
\end{bmatrix}.
\]

Now, we see that the matrix \(A_{\text{MS-SSIA}}\) in (7) becomes a time-independent steering matrix of a centro-symmetric array of size \(2K\) and therefore works for the case of multiple snapshots. We again apply the Unitary ESPRIT method to (7) and obtain the updated DOA estimates \(\hat{\theta}_{l,j+1}\). We apply the multiple augmentations (the iterative application of the augmented matrix) and repeat the procedures (5)–(7) until the estimated DOAs \(\hat{\theta}_{l,j+1}\) converge.

5 Simulation

The DOA estimation accuracy of the proposed MS-SSIA method is evaluated through computer simulation and compared with the accuracies of (i) Unitary ESPRIT with SSP, (ii) Unitary ESPRIT with SSIA [3], (iii) Unitary ESPRIT with SSOA [4] and (iv) MODE [5]. Note that the MODE is also adopted as one of the representative DOA estimation methods for coherent sources. The estimation accuracy was evaluated using the Root Mean Square Error (RMSE) of the estimated DOAs, which is calculated as the average of 500 Monte-Carlo simulation results. The stochastic Cramer-Rao lower bound (CRLB) is also plotted in all the figures [6]. We consider a ULA with \(M = 28\) array elements with an array interval \(d = \lambda/2\) and \(L = 9\) coherent source signals. The nine DOAs are given by \(\theta = [-30, -24, -18, -12, -6, 0, 6, 12, 18, 30]\) (deg), which is the same as Thakle et al.’s scenario [3]. The number of snapshots is normally set to 3, but it is set to 1 for the case of SS-SSIA because it only works for a single snapshot.

Fig. 1 shows the behavior of RMSE as a function of \(K\) for the case of 10 dB SNR, where the optimum values \(R = 2\) and \(\Delta = 5\) of the parameters in SSOA [4] are used. We can see from Fig. 1 that the optimum value of \(K\) is 11 for SSIA and MS-SSIA but 14 for SSOA and 19 for spatial smoothing preprocessing. Note that the value of \(K\) used by Sekine et al. [4] is not appropriate; therefore, we hereafter use those optimum values of \(K\).
Fig. 2 shows the behavior of RMSE as a function of SNR. We see from Fig. 2 that the proposed MS-SSIA method achieves better performance than the other methods for the case of low SNRs and as good performance as the MODE method for the case of high SNRs. The reason for this result is that applying the augmented matrix in [3] can improve the DOA estimation accuracy but the application of only once merely results a limited improvement. We see from Fig. 2 that the multiple augmentations in the proposed approach can further improve the DOA estimation accuracy at low SNRs.

Fig. 3 shows the behavior of RMSE as a function of the number of snapshots for the case of (a) SNR is $-4$ dB and (b) SNR is 10 dB. We can see from Fig. 3(a) that the proposed MS-SSIA method gives smaller RMSEs than those of the other methods in the case of a small number of snapshots, and as good performance as the MODE method for the case of a large number of snapshots. We also observed the similar behavior in Fig. 3(b) in the case of high SNR. The proposed MS-SSIA method can be regarded as a powerful tool that works in a severe environment like low SNRs and a small number of snapshots.
6 Concluding remarks

This paper introduced the MS-SSIA method, which can be regarded as an extension of the SS-SSIA method, to correspond to the case of multiple snapshots. The performance of the proposed method was evaluated through computer simulation, and we confirmed that the proposed MS-SSIA method worked effectively in severe environments like when there was a low SNR and a small number of snapshots. The performance of RMSE for low SNR levels should further be improved as one of future studies.

Acknowledgment

This work was supported in part by KDDI foundation and JSPS Grant-in-Aid for Scientific Research #25420401. The authors are very grateful for their support.
Computational complexity reduction of NCSP-OFDM receiver for sidelobe suppression

Hikaru Kawasaki, Jun Sasaki, Masaya Ohta\textsuperscript{a)}, and Katsumi Yamashita

Graduate School of Engineering, Osaka Prefecture University, 1–1 Gakuen-cho, Nakaku, Sakai, Osaka 599–8531, Japan
\textsuperscript{a)} ota@eis.osakafu-u.ac.jp

Abstract: \(N\)-continuous symbol padding orthogonal frequency division multiplexing (NCSP-OFDM) is a modulation technique for sidelobe suppression, which adds the correction symbol only into the guard interval to enable the seamless connection of OFDM symbols. NCSP-OFDM requires the calculation of an inverse matrix per channel estimation, and so its computational complexity is large. This paper proposes a simple method for the calculation of the inverse matrix. Numerical experiments demonstrate that the proposed method does not affect the error rate performance and effectively reduces the computational complexity.

Keywords: OFDM, \(N\)-continuous symbol padding OFDM, sidelobe suppression, computational complexity reduction

Classification: Wireless Communication Technologies

References


1 Introduction

Orthogonal frequency division multiplexing (OFDM) is adopted in several telecommunications systems because of its high spectral efficiency and robustness against multipath fading. However, there is a problem in that high sidelobes arise from the discontinuity of adjacent OFDM symbols. Various methods of sidelobe suppression have been proposed [1, 2, 3, 4].

$N$-continuous OFDM [5] is a precoding method to seamlessly connect OFDM symbols up to the high order derivative for sidelobe suppression, which is suitable for suppressing out-of-band radiation. This precoding method modifies the whole of OFDM symbol by inserting the correction symbol into data symbols and it degrades the error rate severely as increasing the continuous derivative order. Although Ref. [5] has proposed an iterative algorithm to remove the correction symbol, the receiver must use it many times per received symbol to achieve practical error rate performance, so this leads to increases the computational complexity. Orthogonal precoding [6, 7] does not degrade the error rate, however, it sacrifices the data rate so as to consume some subcarriers, and has the enormous computational complexity.

We have proposed $N$-continuous symbol padding OFDM (NCSP-OFDM) [8] to improve the error rate of $N$-continuous OFDM without data rate reduction, in which the correction symbol is added only into the guard interval to enable the seamless connection of OFDM symbols up to high order derivative. In the NCSP-OFDM, the sidelobe suppression performance is identical to that of $N$-continuous OFDM, and the receiver does not require the iterative algorithm because the correction symbol does not leak to the body of OFDM symbol following the guard interval. The computational complexity of the demodulation is certainly reduced, compared with the conventional $N$-continuous OFDM and orthogonal precoding. However, NCSP-OFDM requires the calculation of an inverse matrix per channel estimation for its demodulation. Therefore, this computational complexity should be reduced.

In this paper, we propose a simple method for calculating the inverse matrix: the inverse matrix is expanded in a Neumann series and approximated by neglecting higher-power terms. In section 2, we explain NCSP-OFDM. In section 3, we propose a method to reduce the computational complexity in the receiver of NCSP-OFDM. In section 4, we evaluate the effectiveness of the proposed method by numerical experiments, and in section 5 we conclude this paper.
2 NCSP-OFDM

2.1 Transmission

The precoding of NCSP-OFDM [8] corrects the symbols in the guard interval to be \(N_c\)-continuous to the preceding and following symbols.

The OFDM signal is written as

\[
s(t) = \sum_{i=0}^{\infty} s_i(t - i(T_s + T_g)),
\]

where \(T_s\) is the OFDM symbol duration and \(T_g\) is the guard interval length. The \(i\)th NCSP-OFDM symbol \(s_i(t)\) is expressed as

\[
s_i(t) = \begin{cases} 
\sum_{k \in K} \tilde{d}_{i,k} e^{j2\pi k t / T_s}, & -T_g \leq t < 0 \\
\sum_{k \in K} d_{i,k} e^{j2\pi k t / T_s}, & 0 \leq t < T_s 
\end{cases},
\]

where \(\tilde{d}_{i,k}\) are the results of precoding information symbol \(d_{i,k} \in \mathcal{C}\) (\(\mathcal{C}\) is a symbol constellation), \(K = \{k_0, \ldots, k_{K-1}\}\) is a set of data subcarrier indices, and \(K\) is the number of subcarriers. The symbol \(s_i(t)\) satisfies the following constraints so that the symbols in the guard interval are continuous with both the preceding and following symbols at the connecting boundaries until the \(N_c\)-th order derivative:

\[
\begin{align*}
\left. \frac{d^n}{dt^n} s_i(t) \right|_{t = -T_g} &= \left. \frac{d^n}{dt^n} s_{i-1}(t) \right|_{t = T_s}, \\
\left. \frac{d^n}{dt^n} s_i(t) \right|_{t = 0} &= \left. \frac{d^n}{dt^n} s_i(t) \right|_{t = T_s},
\end{align*}
\]

for \(n = 0, 1, \ldots, N_c\).

Ref. [8] has proposed the precoding which minimizes the EVM (error-vector magnitude) between the data symbol \(d_i = [d_{i,k_0}, \ldots, d_{i,k_{K-1}}]^T\) and the modified symbol \(\tilde{d}_i = [\tilde{d}_{i,k_0}, \ldots, \tilde{d}_{i,k_{K-1}}]^T\) as the solution of (2) and (3), such that

\[
\tilde{d}_i = (I_K - P)d_i + P\Phi^H d_{i-1},
\]

where

\[
P = \left[ \begin{array}{cc} \Phi \Phi^T & \Phi \\ \Phi^T & I_{(N_c+1) \times K} \end{array} \right],
\]

\(X^\dagger = X^H (XX^H)^{-1} X\) represents the Moore–Penrose pseudoinverse of \(X\), \(A\) is an \((N_c + 1) \times K\) matrix with elements \([A]_{i,n} = (k_i)^{n-1}\), \(\Phi = \text{diag}(e^{j\phi k_0}, \ldots, e^{j\phi k_{K-1}})\), and \(\phi = -2\pi T_g / T_s\).

The discrete expression of \(s_i(t)\) can be expressed in matrix form as

\[
s_i = \begin{bmatrix} O_{L \times N} & I_N \end{bmatrix} u_i + \begin{bmatrix} O_{L \times (N-L)} & I_L \end{bmatrix} \tilde{u}_i,
\]

where \(u_i = D^{-1} \tilde{d}_i\), \(u_i = D^{-1} \tilde{d}_i\), \(D^{-1}\) is the \(N \times K\) inverse discrete Fourier transform matrix

\[
[D^{-1}]_{mn} = \frac{1}{N} e^{j2\pi mk/N}, \quad k_m \in \mathcal{K},
\]
$N$ is the number of samples for the body of one OFDM symbol (FFT points), and $L = NT_s/T$ is the guard interval sample length assumed to be longer than maximum delay spread in this study.

### 2.2 Reception

Assuming that the channel is time-invariant and can be estimated perfectly, the received symbol in time-domain can be expressed as $r_i = Hs_i + n_i$, where $H$ is an $N \times (N + L)$ Toeplitz matrix that expresses the channel characteristics in the time-domain, $n_i$ is a complex-valued zero-mean Gaussian noise vector. Then the receiver performs FFT-demodulation to generate the frequency-domain symbol $\tilde{r}_i$. Assuming that the channel is time-invariant and can be estimated perfectly, the reception maximum delay spread in this study.

In Ref. [8], the detected data symbols $\hat{d}_i = [\hat{d}_{1,k_0}, \ldots, \hat{d}_{1,k_{k-1}}]^T$ are obtained from the hard decision $\hat{d}_{1,k} = \arg \min_{d \in C} \{ |\hat{d}_{1,k} - d|^2 \}$ against the demodulated symbol $\tilde{d}_i = [\tilde{d}_{1,k_0}, \ldots, \tilde{d}_{1,k_{k-1}}]^T$ calculated by

$$\tilde{d}_i = (\Lambda_0 - \Lambda_2 P)^{-1}(\bar{r}_i - \Lambda_2 P \Phi^H \hat{d}_{i-1}),$$

where $\Lambda_0 = DH\begin{bmatrix} O_{L \times (N-L)} & I_L \\ I_N & 0 \end{bmatrix}D^{-1}$ is a diagonal matrix which can be estimated by zero forcing algorithm and $\Lambda_2 = DH\begin{bmatrix} O_{L \times (N-L)} & I_L \\ O_{N \times N} & 0 \end{bmatrix}D^{-1}$.

### 3 Proposed method

In the receiver of the NCSP-OFDM, $(\Lambda_0 - \Lambda_2 P)^{-1}$ and $\Lambda_2 P \Phi^H$ must be recalculated per estimation of channel characteristics $\Lambda_0$ and $\Lambda_2$. The computational complexity of $(\Lambda_0 - \Lambda_2 P)^{-1}$ is $O(L^3)$ as described in [8], and thus the computational load on the receiver is very large and must be reduced.

In order to reduce this computational complexity on the receiver, we firstly rewrite the demodulated symbol (8) as

$$\tilde{d}_i = Q^{-1} \Lambda_0^{-1} (\bar{r}_i - \Lambda_2 P \Phi^H \hat{d}_{i-1}),$$

where

$$Q = I_K - \Lambda_0^{-1} \Lambda_2 P.$$  

Here we consider an approximation of the inverse matrix $Q^{-1}$. This approximation is neglecting higher-power terms in a Neumann series expansion of $Q^{-1}$ such as

$$Q^{-1} = I_K + T + T^2 + \cdots + \sum_{m=1}^{\infty} T^{m-1},$$

where $T = \Lambda_0^{-1} \Lambda_2 P$. If the higher-power terms are neglected such that only the $m = 1$ term remains, then $Q^{-1}$ is approximated by the unit matrix $I_K$ which requires no computation. The computational complexities of the approximated $Q^{-1}$ are $O(K^2)$ if $m = 2$ due to $T$ and $O(K^3)$ if $m \geq 3$ due to $T^2$. Therefore, $m$ must be 1 or 2 to reduce the computational complexity of the receiver.

Especially only if $m = 1$, $Q^{-1}$ can be removed and thus the demodulation (8) itself can save half of both multiplications and additions compared with those of $m \geq 2$, as ignoring those for $\Lambda_0^{-1}$ same as the plane OFDM.
4 Numerical experiments

To evaluate the proposed method, we conducted numerical experiments. We firstly estimated the computational complexity for calculating matrices $Q^{-1}$ and $A_P\Phi^H$ per channel estimation. Fig. 1 shows the comparisons in multiplications and additions. As shown in Fig. 1(a), the proposed method of $m = 2$ is worthless at $K < 1200$ although it is superior to the conventional method at $K > 1200$. In contrast, the proposed method of $m = 1$ can drastically reduce the numbers of multiplications and additions. For example, for $K = 1200$ the proposed method of $m = 1$ requires only 1% of the multiplications and 2% of the additions of the conventional method.

Next, we investigated error rate performance of the proposed method. Fig. 2 show the bit error rate (BER) in two types of multipath fading channel models [9]: typical urban channel model (TUx) supposes many paths and a low attenuation rate, while rural area channel model (RAx) supposes few paths and a high attenuation rate. In this experiment, subcarriers are modulated with 16-QAM, $(N, K, L) = (512, 300, 36)$ and $(128, 72, 9)$, $N_c = 3$, $T_s = 1/15$ ms, $T_g = 9T_s/128$, and over-sampling rate is 4. These results indicate that the BER performances of the proposed method of $m = 1$ is identical to that of the conventional NCSP-OFDM. In other words, the proposed method does not affect the error rate performance.

Lastly, we analyzed the validity of $Q^{-1} \approx I_K$ of (11) at $m = 1$ to confirm the performance of the proposed method. Fig. 3 shows the distribution of the absolute values of each element of $Q^{-1} - I_K$ from randomly generating 1,000 channels of TUx and RAx, respectively. In Fig. 3, when $K = 72$, for example, the mean value is equal to $1.10 \times 10^{-3}$ and the number of elements of $Q^{-1} - I_K$ with larger absolute values than $10^{-1}$ is only $2.37 \times 10^{-2}$% and most of the rest are much smaller than 1. From this analysis, the results eventually show that $Q^{-1}$ is very close to $I_K$ both in these practical fading channels. Therefore the proposed method of $m = 1$ has the performance identical to that of the conventional method.

Fig. 1. Comparison of computational complexity.

(a) Multiplications  (b) Additions
5 Conclusion

In this paper, we have proposed a simple method for calculating the inverse matrix per channel estimation, which involves a Neumann series expansion and an approximation that neglects higher-power terms. The experimental results indicated that the inverse matrix can be approximated by the unit matrix and that the proposed method does not degrade the error rate performance and effectively reduces the computational complexity.

Fig. 2. BER performance.

Fig. 3. Distribution of absolute values of each element of $Q^{-1} - I_K$ in TUx/RAx.
A low power and high speed data transmission system based on 2D communication

Yuichi Masuda\(^a\), Akihito Noda, and Hiroyuki Shinoda
Department of Complexity Science and Engineering, The University of Tokyo, Chiba, Japan
\(^a\) masuda@hapis.k.u-tokyo.ac.jp

Abstract: Low power communication systems such as ZigBee enable battery-driven sensor nodes in wireless sensor networks (WSNs). Recent growing demand for transferring large amount of data across a WSN requires more energy-efficient communication schemes. Energy efficiency, evaluated in terms of energy-per-bit rate (EBR), determines the upper bound of information quantity that can be transferred while consuming all the energy stored in a battery embedded in each sensor node. TransferJet is one of the most energy efficient wireless communication schemes. It achieves 2–3 orders of magnitude lower EBR than ZigBee, at the expense of very short communication range up to a few centimeters. In this paper, we report on a preliminary experiment of energy efficient signal transmission using TransferJet devices on two-dimensional communication (2DC) system. It worked at reasonably high transmission rate of 71 Mbps with 1.7-nJ/bit EBR. We also explain the feasibility of a room-scale high-speed communication system with such a low EBR based on 2DC technology.

Keywords: two-dimensional communication, wireless sensor network, Internet of Things (IoT)

Classification: Wireless Communication Technologies

References


1 Introduction

A developing CMOS technology and increasing variety of low power communication systems enable battery-driven sensor nodes in wireless sensor networks (WSNs) and the Internet of Things (IoT). ZigBee and Bluetooth are representative wireless communication standards used in battery-driven WSN. On the other hand, their energy-per-bit rate (EBR), which is the energy required per a single bit transfer, is higher than that of conventional Wi-Fi [1]. EBR determines the maximum amount of information that can be transferred within their battery life. For example, a ZigBee device driven by an AA-sized battery, in which roughly 10,000 J of energy is stored, can transmit at most 1 GB of information [2].

In other words, the EBR and the battery capacity determine the battery life. Suppose a temperature sensing application that operates at 10-Hz sampling-rate with 10-bit resolution of digital to analog converter (DAC). Its data rate, 100 bps, enables 28,000-hours operation of ZigBee transmitter with a 10,000-J AA-battery. A sound sensor, operating at 40-kHz sampling-rate and 10-bit DAC, generates 400-kbps data stream. In this case, the battery life is reduced to only 7 hours. To extend the battery life, the EBR has to be reduced.

At the expense of very short communication range up to a few centimeters, TransferJet [3] can operate at 2–3 orders of magnitude lower EBR than ZigBee. Although its remarkably high energy efficiency is attractive, its very short transmission range is not acceptable in room-scale WSNs.

In this paper, we report on a preliminary experiment of energy efficient signal transmission using TransferJet devices on a two-dimensional communication
(2DC) system, as a feasibility study of low-EBR room-scale communication system. It achieved a data rate higher than 70 Mbps with an EBR of 1.7 nJ/bit on a 50-cm square 2DC sheet. The EBR is significantly lower than that of ZigBee. We also explain the feasibility of a room-scale communication system at a higher data rate with such a low EBR of the mobile nodes based on 2DC technology. The possibility of 2DC of TransferJet was once demonstrated at CEATEC [4] by a Japanese company, but the sheet size was small and the technical report has not been published yet.

The rest of this paper is organized as follows. In Section 2, our motivation will be described. Section 3 will present a brief review of 2DC tile systems and Section 4 will explain the possibility of energy-efficient signal transmission on 2DC. Experiment will be shown in Section 5. Finally, Section 6 concludes this paper.

2 Low power communication systems

To figure out the motivation of this work, related technologies are compared in Fig. 1, in terms of EBR, transmission rate, and communication range. Those data are based on evaluation of commercially available modules [3, 5, 6]. Related technologies mapped in the figure are: ZigBee, Bluetooth, WiFi (IEEE 802.11b), Passive Wi-Fi [6] and TransferJet [3]. The experimental result presented in Section 5 is also shown as specification of 2DC.

Fig. 1 shows that transmission system based on 2DC technology can reduce EBR by reducing the power consumption of analog radio frequency (RF) components as well as TransferJet and Passive Wi-Fi, while achieving a room-scale communication range.

The major factors of power consumption are classified into the following two components: the analog RF component to generate RF signal transmitted from the antenna and the digital component to process baseband signals. The EBR can be drastically reduced as the analog RF power consumption is reduced. As the
evidence, the EBR of TransferJet and Passive Wi-Fi are at least 1–2 orders of magnitude lower than other schemes in the dashed box shown in Fig. 1.

Recently, we have proposed a 2DC tile system [7]. By connecting tiles side by side, the communication range is extended to a room-scale. It enables signal transmission/reception by all the devices laid on the floor, and even on any furniture covered with 2DC sheets. The devices can communicate with low emission power density of $-41.3 \text{ dBm/MHz}$ [8]. This is significantly lower than the antenna power of conventional ZigBee devices, which is $0–+20 \text{ dBm}$ in a 2-MHz channel bandwidth. Hence, the EBR can be reduced in 2DC.

As shown in Fig. 1(a), 2DC tile systems can provide longer communication range than TransferJet. We have recently analyzed the relationship between the number of tiles connected and signal-to-noise ratio (SNR). Assuming $-41.3$-dBm/MHz power density transmitted, up to 17-m communication range can be achieved with 50-cm square 2DC tiles, while 20-dB SNR is required [9].

Higher transmission rate of 2DC, compared with Passive Wi-Fi, is due to wide frequency range it occupies. It can operate as ultra-wideband (UWB) radio system [8]. In this paper, we show that the transmission rate higher than 70 Mbps can be achieved with the EBR of 1.7 nJ/bit.

The EBR of 2DC is measured by using TransferJet devices on 2DC system. The symbol duration of TransferJet is shorter than the delay spread of 2DC system, which causes the inter-symbol interference (ISI). By optimizing the symbol rate, the ISI can be reduced and transmission rate can be increased.

3 Previous works of two dimensional communication

2DC is a short-range communication scheme using a sheet-like waveguide [10]. The 2DC sheet guides electromagnetic waves and generates evanescent (non-radiative) waves above its surface. It enables low emission and wide frequency range communication between transceivers laid on the sheet surface. The system can be constructed as “an extremely low power radio station (ELPRS)”, which is defined as a radio station that generates an electric field intensity less than $35 \mu/m$ at 3 m distant from the radio equipment [11]. While the requirement on the intensity of the radiated electromagnetic field is satisfied, the frequency range that an ELPRS occupies is not restricted.

![Fig. 2. 2DC tile system installed on the floor of a room.](image-url)
A 2DC tile system is shown in Fig. 2. Two-dimensional area (floor) can be covered with multiple parallel 1-D chains or with a long meander chain of 2DC tiles [11]. To keep the available signal power almost constant across the entire tile system, amplifier is embedded in each base layer and compensates signal loss [12].

Each waveguide sheet is electromagnetically isolated from the other sheets, i.e., the guided modes of them do not interact with each other. This means that the transmitted signal reverberates inside the 50-cm square waveguide sheet. The reverberation in such a lossy small area results in a small delay spread about 8–10 ns [7]. The delay spread of ordinary indoor wireless communication, where radio signals reverberate in a room-scale three-dimensional space, is 100–150 ns on an average and generally varies an order of magnitude due to scattering and absorption by objects/people in the environment. The symbol rate of ZigBee and Bluetooth is less than 1 Mbps. Its symbol duration is greater than 1 µs which is 7–10 times longer than the average delay spread. On the other hands, since the signal propagation paths in tiles are electromagnetically almost isolated from the outside space, the delay spread is little affected by objects/people in the room.

4 Energy-efficient signal transmission on 2DC

While the symbol duration is longer than the average delay spread, signal can be transferred with acceptable ISI. Thus, the devices can transmit signal at high symbol rate, which enables high data transmission rate with a single or a few carriers. The digital logic component of single-carrier transceivers can be much simplified compared with a modulation circuit for orthogonal frequency-division multiplexing (OFDM).

To avoid significant ISI, the symbol duration should generally be about 10 times greater than the delay spread of a channel. In a waveguide sheet with an 8–10-ns delay spread, a 100-ns symbol duration signal, which corresponds to 10-MHz symbol rate, can be transferred without significant ISI. Assuming complementary code keying (CCK) with 8 bit per symbol, 80 Mbps will be enabled.

As a result, the transceiver which optimizes the symbol rate for room-scale 2DC environments can communicate with low emission power density of −41.3 dBm/MHz while providing 80 Mbps transmission rate. Its communication range is at most 17 meter.

5 Experimental results

In this section, we present experiments to demonstrate the feasibility of low emission and high transmission rate 2DC. We measured the transmission rate of a single-carrier high-symbol-rate 2DC system, by using a commercially available TransfeJet adapter, SANWA SUPPLY ADR-TJMUBK.

TransferJet is one of ELPRSs. Its symbol rate of 280 MHz, which corresponds to a symbol duration of 3.57 ns, will undergo significant ISI in a 50-cm square sheet.

Fig. 3(a) shows the measurement environment. One of two TransferJet adapters, attached to a PC, was connected to the feeding point of 2DC sheet with a coaxial cable. The other one, attached to another tablet PC, is connected to a
proximity coupler laid on the sheet. To connect the coaxial cables, an SMA connector was soldered on each TransferJet circuit board.

We measured the transmission rate at three different coupler positions. The delay spread of the 2DC channel, from the feeding point fixed on a sheet edge to the proximity coupler, depends on the coupler position, because a standing wave is generated due to the open-edges of the sheet [7]. The delay spread was measured with a vector network analyzer (VNA) at each coupler position. While keeping the coupler position unchanged, VNA ports were connected to the feeding point and the coupler, instead of the TransferJet devices. VNA measured a scattering parameter (S-parameter) from the feeding point to the coupler in the frequency domain. It was converted into the time-domain impulse response by the inverse Fourier transform [7].

The three measurement results are shown in the transmission-rate versus delay-spread plot, Fig. 3(b). The EBR also shown in the same graph was calculated from its transmission rate and power consumption of 118 mW, that was published in [2]. At a coupler position where the delay spread was 3.9 ns, the transmission rate

![Graph showing transmission rate and EBR versus delay spread](image)

**Fig. 3.** Transmission rate and EBR were evaluated on a 2DC sheet by using a pair of TransferJet adapters. (a) The measurement environment and (b) measured results.
achieved 71.1 Mbps and the corresponding EBR was 1.7 nJ/bit. This result is also plotted in Fig. 1. The transmission rate slower than the original TransferJet is due to the bit error more frequently caused by a significant ISI. By optimizing the symbol duration, the transmission rate will be maximized at any arbitrary coupler position on the entire 2DC area.

6 Conclusion

In this paper, we reported a low power consumption and high transmission rate communication system based on 2DC technology. We evaluated the feasibility of our proposed communication system by using TransferJet adapter. The experimental system achieved 71.1 Mbps transmission rate while its EBR was 1.7 nJ/bit. Thus, the feasibility of low EBR and high transmission rate 2DC was demonstrated. It will enable to exchange large amount of data across a WSN with a practical battery life of distributed nodes.

Acknowledgments

This work was supported in part by the Strategic Information and Communications R&D Promotion Programme (SOPCE) 155103003.
Theoretical examination of channel estimation method for TS-OFDM signal under symbol timing offset

Pongsathorn Reangsuntea¹a), Pisit Boonsrimuang², Kousuke Sanada¹, Kazuo Mori¹, and Hideo Kobayashi¹

¹ Department of Electrical and Electronic Engineering, Graduate School of Engineering, Mie University, Mie 514–8507, Japan
² Department of Telecommunication Engineering, Faculty of Engineering, King Mongkut’s Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand

a) pongsathorn@com.elec.mie-u.ac.jp

Abstract: Recently the channel estimation method of using a time domain Training Signal (TS) for the Orthogonal Frequency Division Multiplexing (TS-OFDM) has been considered actively by many researchers as an alternative method of using pilot signals in the frequency domain. However most of them were investigated assuming the ideal symbol timing. This paper conducts the theoretical examinations for the effect of channel estimation method of using the TS method when the detected symbol timing has an offset from the ideal symbol timing. This paper also verifies the theoretical analysis by computer simulation results.

Keywords: OFDM, training signal, channel estimation, symbol timing

Classification: Wireless Communication Technologies

References

1 Introduction

In the employment of Cycle Prefix-Orthogonal Frequency Division Multiplexing (CP-OFDM) in multipath fading channels, the channel frequency response (CFR) is usually estimated by using pilot signals in which the pilot signals are inserted into data symbols periodically in the frequency domain. As an alternative method of using pilot signals, a CFR estimation method of using a time domain Training Signal (TS) was proposed for the TS-OFDM [1, 2, 3, 4]. In [4], the authors demonstrated that the TS aided method can achieve higher CFR estimation accuracy than the pilot aided method especially under higher mobile environments. However these papers [1, 2, 3, 4] assume the ideal symbol timing (ST) and as far as we know there was no detailed investigation under the symbol timing offset (STO). In this paper, we perform theoretical examinations for the effect of CFR estimation method for the TS-OFDM under the STO. This paper also verifies the theoretical analysis by computer simulation results.

2 Design of TS-OFDM under symbol timing offset

At the CP-OFDM receiver, a ST must be firstly established to discard the CP from the received signal and to decide the FFT window for the data symbol. The ST is usually detected by taking an auto-correlation between the transmitted and received signals [5]. However, this method would establish the ST at the delay path with the maximum amplitude among the multiple delay paths which leads the inter-symbol interference (ISI) in earlier received delay paths than the detected ST after removing the CP. To solve this problem, [5] proposed the early gate FFT window method for the CP-OFDM in which the ST for FFT window is established by adding a pre-advanced offset to the detected ST to avoid the ISI.

Based on the early gate FFT window method for the CP-OFDM, this paper employs a frame format for the TS-OFDM as shown in Fig. 1. In the figure, the ideal ST is defined at the start of TS1, \( L \) is the number of delay paths in the real channel, \( S \) is the length of TS1 and TS2 with the same data pattern which are added at the both ends of data symbol, \( \eta \) is the firstly detected ST and \( \beta \) is the pre-advanced offset. When assuming \( \eta \) is detected at the delay path with the maximum amplitude, \( \eta \) is existed within \( 0 \leq \eta \leq L - 1 \) and the ST for the FFT window is established at \( \eta_o = -\beta + \eta \). The required length of \( S \) can be decided by considering two cases \( \eta = 0 \) and \( \eta = L - 1 \) as shown in Fig. 1 which correspond to the established ST \( \eta_o = -\beta \) and \( \eta_o = -\beta + L - 1 \), respectively. Hence the following conditions are required to satisfy the ISI free for both cases.

\[
\begin{align*}
\beta + L & \leq S \quad \text{when } \eta = 0 \\
-\beta + L - 1 & \leq 0 \quad \text{when } \eta = L - 1
\end{align*}
\]

From Eq. (1), the required length of \( S \) is given by \( S \geq 2L - 1 \) which satisfies the ISI free if \( \eta \) is detected within \( 0 \leq \eta \leq L - 1 \). Here the length of \( S \) could be reduced from the fact that the amplitude of longer delay path is stochastically smaller than the shorter delay path in the real channel. However since the purpose of this paper is to evaluate the effect of estimated CFR under the STO, the conditions of \( 0 \leq \eta \leq L - 1 \), \( S = 2L - 1 \) and \( \beta = L - 1 \) under the ISI free are assumed in the following theoretical analysis. In the following, the sub-indexes \( Id \), \( So \), \( TS \) and \( D \) represent the ideal ST, ST with offset, training and data signals, respectively.
From Fig. 1, the received TS1 mod where

By using Eqs. (3) and (4), the unknown parameters

The simple solution for Eq. (5) can be given by assuming

L2e L1 − 1

Where $mod[n, S]$ represents $n$ modulo $S$, $w_{TS}^{ld}(m, n)$ is an additive white Gaussian noise and $h_l(m)$ is the channel impulse response (CIR) for the $l$-th delay path at the $m$-th symbol in the real channel. Since the number of delay paths $L$ in the real channel is unknown at the receiver, the length of TS ($= S$) which be taken by larger than $L$ is employed in Eq. (3). Assuming that $h_l^0(m)$ is the unknown parameters to be estimated, the expected received TS1 $\hat{y}_{TS}^{ld}(m, n)$ can be given by,

By using Eqs. (3) and (4), the unknown parameters $\hat{h}_l^{ld}(m)$ under the ideal ST can be estimated by solving the following maximum likelihood (ML) equation.

The simple solution for Eq. (5) can be given by assuming $y_{TS}^{ld}(m, n) = \hat{y}_{TS}^{ld}(m, n)$ and it is expressed by the following simultaneous equations.

where $[b_{mod[n-l,S]}]_{Sx1}$ is the circulant matrix with the size of $(S \times S)$. By using the property of circulant matrix, its inverse matrix is also the circulant matrix [6]. Let $[c_{mod[n-l,S]}]$ is the inverse matrix of $[b_{mod[n-l,S]}]$, the unknown parameter $\hat{h}_l^{ld}(m)$ in Eq. (6) can be estimated by,

Here it should be noted that $[c_{mod[n-l,S]}]$ can be calculated in advance because all elements of $[b_{mod[n-l,S]}]$ are known at the receiver which leads the considerable reduction of computation complexity in the estimation of CIR at every symbol under mobile environments. From Eq. (7), $\hat{h}_l^{ld}(m)$ can be given by,
\[
\hat{h}_l^d(m) = \sum_{n=0}^{S-1} y_{TS}^{ld}(m, n) \cdot c_{\text{mod}(n-l, S)} = \sum_{n=0}^{S-1} \left\{ \sum_{l=0}^{S-1} h_l(m) \cdot b_{\text{mod}(n-l, S)} + w_{TS}^{ld}(m, n) \right\} \cdot c_{\text{mod}(n-l, S)}
\]

(8)

\[
h_l(m) + \sum_{n=0}^{S-1} w_{TS}^{ld}(m, n) \cdot c_{\text{mod}(n-l, S)} = h_l(m) + z_{TS}^{ld}(m, l)
\]

where \(z_{TS}^{ld}(m, l)\) is the noise component at the \(l\)-th delay path under the ideal ST. From Eq. (8), it can be observed that the estimated \(\hat{h}_l^d(m)\) under the ideal ST can be expressed by \(h_l(m)\) in the real channel although it includes the noise component.

When the ST is established at \(\eta_c \neq 0\) (\(\beta > \eta\)) which includes \((-\eta_c, \eta_c)\) offset from the ideal ST as shown in Fig. 1, the received TS \(y_{TS}^{So}(m, n)\) can be expressed by,

\[
y_{TS}^{So}(m, n) = y_{TS}^{ld}(m, n), \quad \text{for } (-\beta - \eta) \leq n \leq S - (\beta - \eta) - 1
\]

\[
y_{TS}^{So}(m, n) = y_{TS}^{ld}(m, n) - (\beta - \eta)), \quad \text{for } 0 \leq n \leq S - 1
\]

(9)

The CIR \(\hat{h}_l^{So}(m)\) under the STO must be estimated by using the inverse matrix \([c_{\text{mod}(n-l, S)}]\) given in Eq. (7) because the Ideal ST is unknown at the receiver. From Eqs. (3) and (9), \(\hat{h}_l^{So}(m)\) under the STO can be estimated by,

\[
\hat{h}_l^{So}(m) = \sum_{n=0}^{S-1} y_{TS}^{ld}(m, n - (\beta - \eta)) \cdot c_{\text{mod}(n-l, S)}
\]

\[
= \sum_{n=0}^{S-1} \left\{ \sum_{l=0}^{S-1} h_l(m) \cdot b_{\text{mod}(n-(\beta-\eta)-l, S)} + w_{TS}^{ld}(m, n - (\beta - \eta)) \right\} \cdot c_{\text{mod}(n-l, S)}
\]

(10)

\[
h_{\text{mod}(-\beta-\eta, S)}(m) + z_{TS}^{So}(m, l)
\]

where \(z_{TS}^{So}(m, l)\) is the noise component under the STO. From Eq. (10), it can be concluded that \(\hat{h}_l^{So}(m)\) under the STO can be given by \(h_{\text{mod}(-\beta-\eta, S)}(m)\) in the real channel although it includes the noise component \(z_{TS}^{So}(m, l)\).

In this paper, we employ the MMSE-FDE method for the TS-OFDM which can obtain the frequency diversity gain by decomposing the received time domain signal by \((N + S)\)-point FFT/IFFT [2]. Then CFR under the STO can be given by,

\[
\hat{H}^{So}(m, k) = \sum_{l=0}^{L+S-1} \hat{h}_l^{So}(m) \cdot e^{-j\frac{2\pi kl}{L+S}} = \sum_{l=0}^{S-1} \left\{ h_{\text{mod}(-\beta-\eta, S)}(m) + z_{TS}^{So}(m, l) \right\} \cdot e^{-j\frac{2\pi kl}{L+S}}
\]

(11)

When \(\eta_c \neq \beta - \eta\), the CIR in Eq. (11) can be split into three terms as given by,

\[
\hat{H}^{So}(m, k) = \sum_{l=0}^{(-\beta-\eta)-1} h_{\text{mod}(-\beta-\eta, S)}(m) \cdot e^{-j\frac{2\pi kl}{L+S}} + \sum_{l=\beta-\eta}^{L+(\beta-\eta)-1} h_{\text{mod}(-\beta-\eta, S)}(m) \cdot e^{-j\frac{2\pi kl}{L+S}}
\]

\[
+ \sum_{l=L+(\beta-\eta)}^{S-1} h_{\text{mod}(-\beta-\eta, S)}(m) \cdot e^{-j\frac{2\pi kl}{L+S}} + Z_{TS}^{So}(m, k), \quad 0 \leq k \leq N + S - 1
\]

(12)

As for the 1st and 3rd terms in Eq. (12), \(h_{\text{mod}(-\beta-\eta, S)}(m)\) for \(0 \leq l \leq (\beta - \eta) - 1\) and \(L + (\beta - \eta) \leq l \leq S - 1\) are given by \(h_l(m)\) for \(S - (\beta - \eta) \leq l \leq S - 1\) and \(L \leq l \leq S - 1 - (\beta - \eta)\), respectively. Since \(h_l(m)\) is the CFR in the real channel which exists only within \(0 \leq l \leq L - 1\). From this fact, the 1st and 3rd terms can be assumed to be zero and Eq. (12) can be expressed by,

\[
\hat{H}^{So}(m, k) = \sum_{l=-\beta-\eta}^{L+(\beta-\eta)-1} h_{\text{mod}(-\beta-\eta, S)}(m) e^{-j\frac{2\pi kl}{L+S}} + Z_{TS}^{So}(m, k)
\]

\[
e^{-j\frac{2\pi kl}{L+S}} \cdot \sum_{l=0}^{L-1} h_l(m) e^{-j\frac{2\pi kl}{L+S}} + Z_{TS}^{So}(m, k)
\]

\[
e^{-j\frac{2\pi kl}{L+S}} \cdot H(m, k) + Z_{TS}^{So}(m, k), \quad 0 \leq k \leq N + S - 1
\]

(13)
where $Z^S_{D_0}(m,k)$ is the noise component in the frequency domain and $H(m,k)$ is the CFR in the real channel. From Eq. (13), it can be concluded that the estimated $\hat{H}^S_{D_0}(m,k)$ under the STO is affected by the phase rotation as the function of detected ST $\eta_0(=\beta + \eta)$ from the ideal ST to the $H(m,k)$ in the real channel. However it should be noted that the estimated $\hat{H}^S_{D_0}(m,k)$ includes the noise components $Z^S_{D_0}(m,k)$ over $(N+S)$ samples which are caused from the noise component added to the estimated CIR $\hat{h}_l^S(m)$ within $S$ samples even when $\hat{h}_l^S(m)$ is only existed at $L$ samples within $S$ samples.

In the data demodulation with the MMSE-FDE method [2], the received signal with the length of $(N+S)$ samples after discarding the length of $S$ from the established ST $\eta_0$ as shown in Fig. 1 can be given by,

$$
y^S_D(m,n) = y^D_D(m,n), \text{ for } S - (\beta - \eta) \leq n \leq N + 2S - (\beta - \eta) - 1$$

$$= \sum_{i=0}^{S-1} h_i(m) \cdot x^T(m, n - (\beta - \eta) - l) + w^D_D(m, n - (\beta - \eta)), \text{ for } S \leq n \leq N + 2S - 1$$

(14)

where $x^T(m, n - (\beta - \eta) - l)$ is the transmitted time domain signal in Eq. (2) including the data with the length of $N$ and some parts of TS1 and TS2 with the total length of $S$ which are added at the start and end of data symbol. Let $X^T(m, i)$ is the frequency domain signal converted from the time domain signal $x^T(m, n)$ given in Eq. (2), $x^T(m, n - (\beta - \eta) - l)$ in Eq. (14) is given by,

$$x^T(m, n - (\beta - \eta)) = \frac{1}{\sqrt{N+S}} \sum_{i=0}^{N+S-1} X^T(m, i) \cdot e^{j2\pi n (i - S)/N}$$

(15)

By using Eq. (15), the received frequency domain signal which is converted from the time domain signal in Eq. (14) by $(N+S)$-point FFT, can be given by,

$$Y^S_D(m,k) = \frac{1}{\sqrt{N+S}} \sum_{n=S}^{N+S-1} y^D_D(m,n) \cdot e^{-j2\pi n (k - S)/N}$$

$$= \frac{1}{(N+S)} \sum_{i=0}^{N+S-1} X^T(m,i) \cdot e^{-j2\pi n (k - S)/N} \sum_{l=0}^{S-1} h_l(m) \cdot e^{-j2\pi n l/N} \sum_{n=S}^{N+S-1} e^{j2\pi n (k - S)/N}$$

$$+ \frac{1}{\sqrt{N+S}} \sum_{n=S}^{N+S-1} w^D_D(m, n - (\beta - \eta)) \cdot e^{-j2\pi n (k - S)/N}$$

$$= e^{-j2\pi n (k - S)/N} X^T(m,k) \cdot H(m,k) + Z^S_{D_0}(m,k), \text{ for } 0 \leq k \leq N + S - 1$$

(16)

In the derivation of Eq. (16), the following relationship is used.

$$\sum_{n=S}^{N+S-1} e^{j2\pi n (k - S)/N} = \begin{cases} N+S & \text{when } i = k \\ 0 & \text{when } i \neq k \end{cases}$$

(17)

From Eq. (16), it can be observed that frequency domain received data signal $Y^S_D(m,k)$ under the STO is affected by the phase rotation to the transmitted frequency domain data symbol $X^T(m,k)$ which is completely the same phase rotation as the estimated CFR $\hat{H}^S_{D_0}(m,k)$ in Eq. (13). From this fact, the phase rotation of $Y^S_D(m,k)$ in Eq. (16) due to the STO can be compensated by the MMSE-FDE of using the estimated $\hat{H}^S_{D_0}(m,k)$ in Eq. (13).

In the demodulation of frequency domain data information $X(m,k)$, the estimated $\hat{X}^T(m,k)$ after the MMSE-FDE is converted to the time domain signal $\hat{x}^T(m,n)$ by $(N+S)$-points IFFT. Here it should be noted that the frequency domain $\hat{X}^T(m,k)$ corresponds to the time domain $\hat{x}^T(m,n)$ from $n = S$ to $N + 2S - 1$ as given in Eq. (2). From this fact, the data information $\hat{X}(m,k)$ in the frequency domain can be demodulated by converting $\hat{x}^T(m,n)$ from $n = S$ to $S + N - 1$ which
corresponds to the time domain data information \(x(m,n)\) in Eq. (2) by \(N\)-points FFT [2].

4 Proof of theoretical analysis by computer simulation results

To verify the theoretical analysis derived in Section 3, this paper conducts the computer simulations. In the simulation, the modulation method is quadrature phase shift keying (QPSK), TS-OFDM bandwidth is 5 MHz, radio frequency is 2 GHz, the number of data subcarriers is \(K = 96\) and the number of IFFT points at the transmitter is \(N = 128\). The channel is the time-varying multipath fading channel based on the WSSUS (wide sense stationary uncorrelated scattering) model with \(L = 10\) delay paths which have the exponential power delay profile with \(-1\) dB decay constant. The normalized Doppler frequency is \(f_d T_s = 0.006\) where \(f_d\) is maximum Doppler frequency and \(T_s\) is the symbol duration. To satisfy the ISI free conditions given in Eq. (1), the length of TS is \(S = 19\) and the pre-advance offset is \(\beta = 9\). In the demodulation of data symbol, the MMSE-FDE method with the observation period of \((N + S)\) samples is employed.

Fig. 2 shows an example of simulation result for the estimated CIR to verify Eqs. (12) when the ST is detected at the delay path \(l = 2\) \((\eta = 2)\) with the maximum amplitude. From the figure, it can be observed that the estimated \(\hat{h}_l^{So}(m)\) is existed from \(l = \beta - \eta = 7\) to \(l = L + \beta - \eta - 1 = 16\) which correspond to the real CIR \(h_{mod(l-(\beta-\eta),S)}(m)\) within \(0 \leq l \leq L - 1\) \((0 \leq l \leq 9)\) as given in Eq. (12). From the results, the theoretical analysis in Eq. (12) under the STO is proved by the computer simulation results. Fig. 3 shows the bit error rate (BER) performances with the MMSE-FDE method when changing the detected symbol timing \(\eta\) from 0 to \(L - 1\). From the figure, it can be observed that the TS method can keep better BER performance when the ST is detected within \(0 \leq \eta \leq L - 1\) \((= 9)\). From these results, it is concluded that the theoretical analysis under the STO presented in this paper is proved by the computer simulation results.

5 Conclusions

This paper presented the theoretical examinations for the CFR estimation method of using TS under the symbol timing offset. From the theoretical analysis, it was confirmed that the symbol timing offset causes the same phase rotation both to the estimated CFR and the frequency domain received data and it can be canceled completely by the MMSE-FDE. This paper also verified the theoretical analysis presented in this paper by computer simulation results.
Identifying anomalous traffic using dynamic programing based differential analysis

Keisuke Ishibashi, Tsuyoshi Kondoh, Mashiro Maruyoshi, and Takeshi Kuwahara
Nippon Telephone and Telegraph Corporation,
3–9–11 Midori-cho, Musashino-shi, Tokyo 182–0012, Japan
a) ishibashi.keisuke@lab.ntt.co.jp

Abstract: This paper proposes an identification method of anomalous traffic such as DDoS attacks. Identification results are represented as a set of aggregated flows; such as source/destination IP address ranges(prefixes), source/destination port numbers and protocols and can be used as ACL (Access Control List) rules at routers. We set requirements for the identification can be summarized as the following three conditions; 1) covering the anomalous traffic, 2) avoiding to cover normal traffic, 3) with small number of aggregated flows. To accomplish these requirements, we propose a method to generate a set of aggregate flow that achieves the highest score representing the requirements by comparing before and after attacks and searching a optimal set with dynamic programming to avoid exponential computation explosion.

Keywords: DDoS, anomaly identification, dynamic programing

Classification: Internet

References


1 Introduction

With threats against Internet security increasing, protecting network and/or server resources from anomalous traffic, such as DDoS (distributed denial of service) attacks, has become a critical task in network operations. We have mature technologies on detecting anomalous traffic by statistical traffic analysis [1]. However, only detecting anomaly is not enough. To take an action for the detected anomaly, it should be identified; such as source/destination IP address ranges (prefixes), source/destination port numbers and protocols. Once anomaly is identified, then we can mitigate the anomaly such as configuring ACL (Access Control Lists) to block or rate-limit the anomaly at routers.

This paper proposes a method for anomaly identification. From the viewpoint of identification, requirements for the identification can be summarized as the following three conditions; 1) covering the anomalous traffic (maximizing the filter coverage), 2) avoiding to cover the normal traffic (minimizing collateral filtering), 3) with small number of flows (minimizing the number of ACLs and avoiding overfit).

However, there are conflicts among the above three requirements. For example, if we specify DDoS traffic by its destination (victim) addresses, we can satisfy conditions 1) and 3), but it leads collateral filtering for normal user accesses to the victim and condition 2) is not satisfied. Thus we should also identify the source IP addresses of such DDoS to filter only anomalous (attack) traffic. In identifying the source IP addresses, it arises the other conflict between condition 2) and 3). Today, DDoS attack are caused by botnets, composed with a huge number of compromised hosts. Identifying such number of IP addresses is not only difficult but also useless, because current router cannot be configured for such number of ACLs. Here, it is observed that origin of DDoS traffic is not uniformly distributed in the IP address space, but concentrated on specific Autonomous Systems (ASes), or address ranges [2]. Thus, to reduce the number of ACLs, in spite of identifying each individual source IP address, we should identify the compromised hosts as IP address range (prefixes). By specifying the address with more larger prefix, the number of ACLs decreases, but the collateral filtering may increase. Thus these three conditions are interchangeably in conflicts relationships.

In this paper, we present a method to simultaneously accomplish these requirements, we define a score for a set of address prefixes that represents above three requirements, coverage ratio, collateral ratio, and number of address prefixes. Then, identifying the anomalous traffic results in selecting the IP address prefixes that has highest scores. However, the algorithm that greedy evaluates all possible combination of IP address prefixes to find a set with highest scores requires huge computing resources because the number of possible combination easily increases as the number of prefixes grows. We adopt dynamic programming to decrease the evaluation time. Because the address prefixes can be represented as a tree whose node corresponds to a prefix, we can adopt dynamic programming as evaluating a sub tree and recursively use the result for evaluating larger tree that include the sub tree.

There some works focusing on filtering attack (malicious) traffic by identifying it as source IP address ranges [4, 5]. However, both works do not simultaneously...

© IEICE 2016
DOI: 10.1587/comex.2016XBL0114
Received June 9, 2016
Accepted June 22, 2016
Publicized July 7, 2016
Copyedited September 1, 2016
optimize above three conditions. For example, in [4], number of address ranges is limited (e.g., 20) and in [5], volume of traffic. Thus, there might be the case where we cannot filter sufficient volume of traffic or number of filter exceeds the limitation of routers.

We evaluate our algorithm by comparing the computing time between greedy method and dynamic programing-based method. We then show an example of an anomaly identification with actual DDoS traffic provided by CAIDA [3] and show that we can identify 92% of anomalous traffic with 2.7% collateral.

In this paper, we propose a method that simultaneously accomplish these requirements, coverage ratio, collateral ratio, and number of address prefixes. To do this, we define a score that reflects the above requirements for each the set of IP address prefixes and identifying the anomalous traffic by selecting the set that has highest scores.

2 Anomaly traffic identification

In this section, we first formalize the identification of traffic anomaly, and then propose the adoption of dynamic programming. Due to space limitation, we focus on the identifying source IP addresses of anomaly traffic. Other attributes of flows can also be identified with the same way. We independently identify the anomaly traffic on each attribute, then take a product set of each identified set and then evaluate it to judge whether or not the set is to be aggregate flow sets of anomalous traffic.

2.1 Formalization

We assume traffic before anomaly as normal traffic and identify the anomalous traffic as the difference before and after the anomaly. We build the IPv4 address binary tree where the root corresponds to 0.0.0.0/0, a leaf corresponds to /32 address and each intermediate node has two child nodes.

Each node $n$ has the coverage and collateral values.

- $\text{Cov}(n)$ [bps]: Attack traffic rate to node $n$ that is measured as difference between and after anomaly.
- $\text{Col}(n)$ [bps]: Legitimate traffic rate to node $n$ that is measured before anomaly. $\text{Cov}(n)$ and $\text{Col}(n)$ can be calculated by measuring some periods of traffic (e.g., 5 minutes) before and during the attack and counting bytes for each node (source address) $n$ that appears in the measured traffic. If a node $n$ only appears before (during) the attack, then we set $\text{Cov}(n)$ ($\text{Col}(n)$) as zero. Then, we define the following values for a set of nodes $S$, where each node $n \in S$ is not ancestor of another node $m \in S$.

- $\text{Cov}(S) := \sum_{n \in S} \text{Cov}(n)$
- $\text{Col}(S) := \sum_{n \in S} \text{Col}(n)$
- $N(S) := |S|$ (Number of nodes in the set)

From the view point of identification, requirements for the identification can be summarized as the following three conditions; 1) covering the anomalous traffic (maximizing the filter coverage), 2) avoiding to cover the normal traffic (minimizing collateral filtering), and 3) with small number of flows (minimizing the number
of ACLs). A desirable set is that has higher Cov and smaller Col and n. Thus we define a function that represents desirability of the node set, such as

\[
f(S) := \alpha \text{Cov}(S) - \beta \text{Col}(S) - \gamma N(S),
\]

where \(\alpha\), \(\beta\), and \(\gamma\) are weighting parameters for corresponding indexes. For each node sets in IP address tree, we evaluate the value \(f(S)\), and choose the set of nodes \(\hat{S}\) whose has the highest values as \(\hat{S} = \arg \min_S f(S)\).  

Fig. 1 shows the example. Each node corresponds to an address prefix \(n\), that has \(\text{Cov}(n)\) and \(\text{Col}(n)\). Then we chose a set of nodes \(S\) which has the highest \(f(S)\). In the figure, blue nodes are chosen as the set with \(N(S) = 4\).

### 2.2 Dynamic programming

Though we can obtain address prefix set \(S\) which has highest scores by naively evaluates all possible combination of IP address prefixes, it requires huge computing resources because the number of possible combination easily increases as the number of prefixes grows. Here, we adopt dynamic programming to decrease the evaluation time.

Here, if a node set in a sub-tree is selected as a part of best node set in the whole tree, the set must be the best node set in the sub-tree. Note that this fact holds when the objective function \(f\) is additive in that \(f(S \cup T) = f(S) + f(T)\) for each disjoint node sets \(S\) and \(T\). It is easily shown that Eqn. 1 satisfies the condition. In that case, the following equation holds for disjoint node set \(S, S', \) and \(T^2\):

\[
f(S) > f(S') \rightarrow f(S \cup T) > f(S' \cup T).
\]

Thus we can start with a smallest sub-tree with two leaf nodes and their parent node, and select a best nodes set among them. Then recursively use the result for evaluating larger tree that include the sub tree.

**Step 1:** Chose a node \(n\) whose child nodes are only leaf node. If such a node does not exist, then algorithm stops and generate the existing nodes as the best node set.

---

1. Actually, we implement some constraint on Cov, Col, N for practical use such as \(\text{Cov}(S) < 0.05\). However, we focus on the optimization for the sake of simplicity.

2. In fact, additive property is sufficient condition but not necessary condition. Therefore, more loose condition can be considered, but in this paper, we limit the function as additive one.
Step 2: Make all combinations of the node $n$ and its child nodes (at most two), say, $c$ and $c'$ and possible combinations are $\{n\}$, $\{c\}$, $\{c',c\}$\(^3\). Chose a combination $S$ who has the highest $f(S)$ in Eqn. 1.

Step 3: Replace $n$ with $S$ and consider $S$ as a leaf node. Return to Step 1.

3 Numerical evaluation

We evaluate our algorithm with actual traffic data.

3.1 Performance comparison

Firstly, we compare the computing time for naive method and dynamic programming.

We evaluate our algorithm with the WIDE MAWI traffic data [6]. Number of flow in the data set is 457,716 and the number of unique IP addresses is 61,124. We change the number of nodes and measure the computing time to generate best node set. While the computing time for naive method exponentially increase as the number of node increases, that for dynamic programming remains negligible as in Fig. 2. Actually, when the address tree is perfect binary, then it is easy to show that the number of evaluation for dynamic programing is linear to the number of leaf nodes. However, because the increase for naive method is so rapid, that for dynamic programing cannot be observed.

3.2 Application example

We apply our algorithm to actual DDoS traffic data provided by CAIDA [3]. The data consists of DDoS traffic only, and we combine it with normal traffic data also provided by CAIDA captured at a 10G link. Then, we check how our algorithm extracts the DDoS traffic from combined traffic. In both data, IP addresses are anonymized with prefix preserving way, thus we can evaluate our address prefix based identification even if the addresses are anonymized.

Fig. 3 shows the results (The blue nodes are selected nodes with highest score). As described in the previous section, naive search method cannot handle a tree larger than twenty leaf nodes, but by using dynamic programming, we can

\[^3\text{Actually, we can omit the set which has both child node and parent node, because it means there are two ACLs whose addresses are overlapped each other.}\]
successfully analyze large tree which has 52 leaf nodes, and achieves 92.0% of anomalous traffic with 2.7% collateral by selecting 14 address prefix nodes. On the other hand, if we limit the number of leaf node to twenty, then the coverage of best node sets becomes 67.7% and the collateral is 1.2%.

![Application example (dynamic programming)](image)

4 Conclusion

This paper proposes an identification method of anomalous traffic such as DDoS attacks which meets the requirements for the identification can be summarized as the following three conditions; 1) covering the anomalous traffic, 2) avoiding to cover normal traffic, 3) with small number of aggregated flows. We evaluate our algorithm by using actual DDoS traffic data and show that computation time does not exponentially increase as the conventional method does. In addition, we experimentally apply our algorithm to actual DDoS traffic, and achieves 92.0% of anomalous traffic with 2.7% collateral.

Acknowledgments

The dataset was provided by CAIDA.
Enhancing aperture efficiency of reflectarray by accurately evaluating mutual coupling of reflectarray elements

Keisuke Konno and Qiang Chen
Department of Communications Engineering,
Graduate School of Engineering, Tohoku University,
6–6–05 Aramaki Aza Aoba, Aoba-ku, Sendai-shi, Miyagi 980–8579, Japan
a) konno@ecei.tohoku.ac.jp

Abstract: In this letter, aperture efficiency of reflectarrays is enhanced using an efficient full-wave method. The full-wave method deals with the effect of real mutual coupling between all reflectarray elements exactly during its design process. Resultant long CPU time is greatly reduced by developing an efficient algorithm which is optimized for the vector supercomputer. Numerical simulation shows that the aperture efficiency of the reflectarrays can be enhanced using the proposed method. The proposed method is computationally efficient and applicable for various design targets.

Keywords: reflectarray, vector supercomputer
Classification: Antennas and Propagation

References


1 Introduction

A reflectarray has received much attention as a compact and lightweight reflector antenna [1, 2, 3]. The reflectarray is a so-called semi-periodic structure because the array itself is periodic while every element has a different dimension. In order to design the reflectarray, phase of reflection coefficient is obtained in advance as a function of dimension of reflectarray element. However, it is difficult to obtain the phase of reflection coefficient exactly because the effect of real mutual coupling is unknown before the reflectarray structure is finally designed. As a result, phase of reflection coefficient is obtained by using various techniques which deal with the effect of mutual coupling approximately.

It is very popular to obtain the phase of reflection coefficient using an unit cell analysis under periodic boundary condition (PBC) [4, 5, 6]. It is known that the unit cell analysis under the PBC works well when the dimension of reflectarray element varies continuously in the reflectarray. However, in practice, the dimension of the reflectarray element often varies discontinuously and so-called local periodicity cannot be assumed. In addition, a real reflectarray is a finite structure, but not an infinite structure. As an alternative approach, isolated element approach [7] and surrounded element approach [8, 9] have been proposed respectively. The isolated element approach neglects the effect of mutual coupling while the surrounded element approach assumes mutual coupling in a finite array. Both of these approaches deal with the reflectarray as a finite structure but the effect of real mutual coupling cannot be reflected to the phase of reflection coefficient. To the best of our knowledge, design of reflectarray which deals with the real mutual coupling between reflectarray elements has not been performed yet.

In this letter, the aperture efficiency of reflectarrays is enhanced using an efficient full-wave solver and the real mutual coupling between reflectarray elements is dealt with exactly except for the setup of the initial reflectarrays. Initial dimensions of every reflectarray element in reflectarrays are given using conventional isolated or surrounded element approach. After that, dimensions of every reflectarray element are renewed in order to enhance the gain of reflectarrays. The
performance of the resultant reflectarray is evaluated using the full-wave solver every time dimensions of reflectarray element are renewed.

2 Proposed design method of reflectarray

A structure of reflectarray is shown in Fig. 1. A metal-only reflectarray which consists of a linear element is assumed here and a full-wave analysis was performed using induced electromotive forced (EMF) method [10, 11]. In order to reduce the number of unknowns as much as possible, two types of reflectarray elements, i.e. a linear dipole element w/ and w/o parasitic dipole elements are mixed. In this letter, the linear dipole element w/parasitic dipole elements is deployed only when the linear dipole element cannot compensate the required phase shift. Dimensions of the two kinds of reflectarray elements except for their length were optimized in advance in order to have a large phase shift as much as possible using the isolated or surrounded element approach. Although a horn antenna is often used as a primary source of reflectarrays in practice, a dipole antenna was used as a primary source to simplify the numerical simulation here because this letter focuses on the real mutual coupling effect between reflectarray elements, not primary source. The effect of an infinite ground plane was included in our numerical simulation using the image theory.

Design of reflectarrays which deals with the real mutual coupling is performed as follows.

1. Phase of reflection coefficient of reflectarray elements is calculated approximately using isolated or surrounded element approach.
2. Initial length of every reflectarray element is obtained using the phase of reflection coefficient.
3. Length of \( m \)th reflectarray element is incremented/decremented from \( l_m \) to \( l_m \pm \Delta l \) and impedance matrix of the entire reflectarray is renewed. Here, \( l_m \) is the initial length of \( m \)th reflectarray element and \( \pm \Delta l \) is its increment/decrement.
4. Gain \( G(l_m \pm \Delta l) \) of the renewed reflectarray is obtained using full-wave solver based on a preconditioned conjugate gradient method.
5. If \( G(l_m + \Delta l) > G(l_m) \) and \( G(l_m + \Delta l) > G(l_m - \Delta l) \), the length of \( m \)th reflectarray is incremented until \( G \) reaches local maximum value. If \( G(l_m - \Delta l) > G(l_m) \) and \( G(l_m - \Delta l) > G(l_m + \Delta l) \), the length of \( m \)th reflectarray is decremented until \( G \) reaches local maximum value. If \( G(l_m) > G(l_m + \Delta l) \) and \( G(l_m) > G(l_m - \Delta l) \), the length of the \( m \)th reflectarray element is kept.
6. 3∼5 is repeated until \( m \) reaches to \( M \), where \( M = M_x M_y \) is the number of reflectarray elements.

It should be indicated that the design method only gives the suboptimum design because no global optimization technique is used. However, the real mutual coupling is dealt with exactly and the resultant reflectarray is expected to have an enhanced performance.

In order to reduce the long CPU time of the numerical simulation, a vector supercomputer SX-ACE at Cyberscience Center, Tohoku University was used. It is well known that the vector supercomputer performs so-called vector operation
which deals with iterative operations simultaneously. Therefore, outstanding computing performance is expected when the vector operation is fully utilized. Our in-house code was optimized for the SX-ACE in order to maximize vector operation ratio which shows ratio of vector operation to total operation. As a result, vector operation ratio over 99.2% was achieved by tuning the program carefully.

3 Numerical results

50 × 50 reflectarrays were designed using the isolated and surrounded element approach combined with the proposed method. Here, phase of reflection coefficient of reflectarray elements in the center of 3 × 3 array of identical elements was used for surrounded element approach. Gain of the reflectarrays is shown in Figs. 2(a)–2(d). It is found that the gain of the reflectarrays designed using conventional isolated and surrounded element approach is enhanced after the proposed method is applied. Enhancement of the gain is caused by accurately evaluating mutual coupling between reflectarray elements. In other words, the proposed method can utilize the effect of real mutual coupling between reflectarray elements in order to enhance the gain. On the other hand, it is demonstrated that the gain of the reflectarrays designed using the isolated or surrounded element is quite low when their array spacing is small. This is because the effect of real mutual coupling on the gain of the reflectarrays is strong when the array spacing is small and should be dealt with exactly when the reflectarrays are designed. As shown in Fig. 2(a) and 2(c), the proposed method works well even when the array spacing of the reflectarrays is small.

In order to clarify the effect of real mutual coupling on the performance of the reflectarrays, the aperture phase distribution of the reflectarrays is shown in Figs. 2(e)–2(h). Figs. 2(e)–2(h) imply that the dimensions of a reflectarray element vary discontinuously and have no local periodicity especially when the array spacing is small. Therefore, it can be said that the performance of the reflectarrays which are designed using any of conventional approaches (i.e. isolated element approach, surrounded element approach, and unit cell analysis with PBC) can degrade due to the effect of the real mutual coupling. On the other hand, the reflectarray designed using the proposed method shows no performance degradation because the effect of real mutual coupling has already been dealt with exactly.
Summary of the design examples is shown in Table I. It is found that the gain and aperture efficiency of the reflectarrays designed using the surrounded element approach are higher than those designed using the isolated element approach when the array spacing is the same. Therefore, the reflectarrays designed using the...
surrounded element approach are assumed to be better initial structures than those
designed using the isolated approach. As a result, relatively high gain and high
aperture efficiency are achieved using the proposed method at the expense of small
computational cost when the initial reflectarrays are designed using the surrounded
element approach, not the isolated element approach.

<table>
<thead>
<tr>
<th>Initial design of reflectarrays</th>
<th>Isolated</th>
<th>Surrounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array spacing $d_x = d_y$ [λ]</td>
<td>0.55</td>
<td>0.75</td>
</tr>
<tr>
<td>Gain [dBi]</td>
<td>30</td>
<td>38.7</td>
</tr>
<tr>
<td>(w/o proposed method)</td>
<td>(19.1)</td>
<td>(37.6)</td>
</tr>
<tr>
<td>Aperture efficiency [%]</td>
<td>11</td>
<td>44.1</td>
</tr>
<tr>
<td>(w/o proposed method)</td>
<td>(0.9)</td>
<td>(34)</td>
</tr>
<tr>
<td>Total number of unknowns</td>
<td>10,858</td>
<td>10,858</td>
</tr>
<tr>
<td>Total CPU time [sec.]</td>
<td>105,508</td>
<td>36,374</td>
</tr>
<tr>
<td>Total number of full-wave simulations</td>
<td>5,029</td>
<td>4,333</td>
</tr>
<tr>
<td>CPU time per full-wave simulation [sec.]</td>
<td>21</td>
<td>8.4</td>
</tr>
</tbody>
</table>

4 Conclusion

In this letter, the aperture efficiency of reflectarrays was enhanced using the
efficient full-wave solver. It was shown that the accurate evaluation of real mutual
coupling in reflectarray design enhance the aperture efficiency of the reflectarray at
the expense of CPU time. In particular, the proposed method is quite helpful to
enhance the gain of reflectarrays when their array spacing is small. The proposed
method can be easily modified corresponding to various design targets because a
full-wave solver and objective function in the proposed method can be replaced by
the other one.

Acknowledgments

We would like to thank staffs in Cyberscience Center, Tohoku University for their
helpful advices. This research used computational resources of the HPCI system
provided by Cyberscience Center, Tohoku University through the HPCI System
Research Project (Project ID:hp140052). This work was financially supported by
JSPS KAKENHI Grant Number 26820137 and JSPS Postdoctoral Fellowships for
Research Abroad.
Design of a dual-band indoor antenna with orthogonal bi-directional radiation pattern

Ho-Yu Lin\textsuperscript{a}) and Masayuki Nakano

KDDI R&D Laboratories Inc.,
2–1–15 Ohara, Fujinomi-shi, Saitama 356–8502, Japan
\textsuperscript{a}) ho-lin@kddilabs.jp

Abstract: In recent years, MIMO (Multi-Input Multi-Output) transmission technique has attracted increasing attention, as it can satisfy the need for exponential increase in data traffic for mobile communications, especially when used for indoor areas. In this paper, an antenna design is proposed that can operate in the 2.6 GHz and 3.5 GHz bands and produce an orthogonal bi-directional radiation pattern. Some characteristics of the proposed antenna, such as reflection coefficient ($S_{11}$), transmission coefficient ($S_{21}$) and far-field radiation pattern are simulated and evaluated by means of FDTD (finite-difference time-domain) methods. The results show that $S_{11}$ and $S_{21}$ are below $-10$ dB and $-15$ dB, and moreover, the MIMO performance can probably be improved because the difference between the two directional gain levels is achieved by approximately 15 dB when four slits are etched into the ground plane.

Keywords: indoor antenna, base station antenna, MIMO, dual band

Classification: Antennas and Propagation

References


1 Introduction

In recent years, the MIMO (Multi-Input Multi-Output) transmission technique has attracted increasing attention as it can satisfy the need for the exponential increase in data traffic required by mobile communications, especially for indoor and densely populated areas. In general, base stations use diversity reception to minimize the effect of Rayleigh fading on the received signal. Polarization spatial diversity techniques are the most common forms of diversity reception [1, 2, 3]. Another form of diversity, angular diversity techniques are also effective at mitigating multipath situations [4], because multiple antennas are oriented in different directions, but multiple antennas are required to form such systems. However, for indoor base stations, it is considered that a device must be compact, and its location should be optimized for maximum antenna coverage.

A number of studies on indoor base station antennas have been published in journals and conference papers. According to [5], in 1997 the main design target for an indoor base station antenna and the effect of the ground plane were introduced. A design for a cylindrical monopole antenna was proposed to reduce the physical antenna size [6]. In order to enhance the bandwidth, a dual-sleeve antenna structure and the use of a shorting wire were proposed [7] and [8].

In this paper, we propose a novel compact indoor base station antenna that is designed to operate in the 2.6 GHz and 3.5 GHz bands. The $S_{11}$ and $S_{21}$ are satisfying the requirement of $S_{11} < -10$ dB and $S_{21} < -15$ dB in MIMO system [9] and [10], and moreover, the proposed antenna has an orthogonal bi-directional radiation pattern for each frequency band, which means that angular diversity gains could be improved to a practically useful extent in a multipath environment. Moreover, the MIMO performance can probably be improved because the difference between the two directional gain level can be achieved by approximately
15 dB when we etched four slits into ground plane. The simulation was performed with CST Microwave STUDIO [11].

2 Antenna structure

2.1 Basic principles of radiation patterns

Fig. 1 illustrates the radiation pattern of two isotropic sources in the same phase and opposite phase, and the difference between the two directional gain levels on an orthogonal bi-directional radiation pattern. Fig. 1(a) shows the two isotropic sources separated by a distance \( d \) and located along the \( y \)-axis. As we know, beam-width and main-beam direction can be controlled by changing the signal amplitude and phase of each antenna [12]. For instance, when two isotropic sources have the same amplitude and phase (Phase difference, \( \delta = 0 \)-degree), the main-beam of the vector sum of the fields is orthogonal to the \( x \)-axis direction. Similarly, when they have the same amplitude but opposite phase (Phase difference, \( \delta = 180 \)-degree), the main-beam is in the \( y \)-axis direction. The vector addition of the fields as mentioned above can be expressed as

\[
E = 2E_0 \cos\left(\frac{d}{2} \cos \theta\right) 
\]

(1)

\[
E = 2jE_0 \sin\left(\frac{d}{2} \cos \theta\right) 
\]

(2)

where \( d = \lambda/2 \) and the amplitude of \( 2E_0 \) and \( 2jE_0 \) = 1 are considered as a special case. We should note that the field pattern in the opposite phase given by (2) is a relatively broad figure-eight with the maximum field.

According to this concept, when we set the phases of two isotropic sources to be the same or opposite by use of 180-degree hybrid couplers, an orthogonal bi-directional radiation pattern can be formed to extend simultaneously in the \( x \)-axis and \( y \)-axis directions. However, it should be noted that the overlapping patterns also cause the undesirable high correlated or correlated MIMO signals to affect MIMO performance. Fig. 1(b) shows an image of the desirable radiation pattern. From the image, the difference between the two directional gain level should be large in order...
to reduce the overlapping on the orthogonal bi-directional radiation pattern, that meaning the MIMO performance can probably be improved. Therefore, we considered that the difference should be at least 15 dB in order to improve MIMO performance. In next section, the structure of the proposed antenna and the detailed simulation results are shown and discussed.

2.2 Antenna structure and simulation result

Fig. 2(a) shows a prototype model of the vertical polarized antenna. Two pairs of monopole antennas, both with a diameter of 2 mm, are mounted on a ground plane and located along the $y$-axis. 2.6 GHz and 3.5 GHz antenna elements are separated by a distance $d_1$ of 0.5λ and $d_2$ of 0.59λ, respectively. The circular ground plane is formed from an indoor base station for practical use, and its diameter and thickness are set to 100 mm and 0.5 mm. In order to improve the input impedance matching, we introduced an L-type matching-pin that is designed to connect the monopole antenna for 3.5 GHz with the ground plane as shown in Fig. 2(b). The height ($h$) and distance ($w$) between matching-pin and monopole antenna are 5.5 mm and 1 mm. Four slits are etched into the ground plane and symmetrically arranged to improve the diversity performance of the proposed antenna and the width and length ($l$) of each slit are 1 mm and 35 mm. Gap feeding is set at the bottom of each antenna. The parameters of $w$ and $h$ can be changed to obtain the optimized result for impedance matching. For the radiation pattern, the optimized result can be obtained by tuning the width and length of slit and $d_2$.

![Antenna structure](image)

**Fig. 2.** Prototype of the proposed antenna structure.

Fig. 3 shows the simulation results for the $S$-parameters and far-field radiation pattern. From the results show in Fig. 3(a), when the matching-pin is not used, it can be seen that $S_{11}$ in the 3.5 GHz band does not satisfy the required value of $-10$ dB, possibly because the antennas are close to each other and the coupling between them becomes strong. $S_{21}$ at 2.6 GHz is $-17$ dB, which satisfies the $-15$ dB requirement, but at 3.5 GHz it is $-13$ dB and needs to be improved. Fig. 3(b) shows the results with the matching-pin, the input impedance in the 3.5 GHz band can be improved ($S_{11} < -10$ dB) and $S_{21}$ at the 2.6 GHz and 3.5 GHz is $-15$ dB and $-24$ dB, which satisfies the required $-15$ dB. The bandwidth of the proposed antenna at 2.6 GHz and 3.5 GHz are 50% and 17%, respectively.
Fig. 3(c) and (d) plot the radiation pattern without four slits. From the result, although the proposed antenna seems to have an orthogonal bi-directional radiation pattern, the difference between the two directional gain levels is 10.8 dB (in the $x$ direction) and 20.6 dB (in the $y$ direction) at the 2.6 GHz, 22.5 dB (in the $x$ direction) and 14.2 dB (in the $y$ direction) at the 3.5 GHz band, so some results need to be improved to satisfy the 15 dB requirement. Fig. 3(e) and (f) show the radiation pattern when the four slits are etched into ground plane. From the results, the difference between the two directional gain levels is 26.1 dB (in the $x$ direction) and 25.6 dB (in the $y$ direction) at the 2.6 GHz band, and 22.8 dB (in the $x$ direction) and 16.9 dB (in the $y$ direction) at the 3.5 GHz band, all results fulfill the 15 dB requirement by the addition of the four slits. For the opposite phase case at 2.6 GHz the difference between two directional gain levels has been improved from 10.8 dB to 26.5 dB (in the $y$ direction), the amount of improvement is approximately 15 dB.

Fig. 3(g) and (h) present the excited surface current distribution, obtained from the simulation of the ground plane of the proposed antenna, for the opposite phase case at 2.6 GHz. Comparison of (g) and (h) shows that when four slits are etched into the ground plane, the current in the $x$ direction is lower than it is without the slits. This is thought to result from currents flowing ($I_1$ and $I_2$) along the slits that mutually cancel, making that the main reason for the improvement in the difference between the two directional gain levels in orthogonal bi-directional radiation patterns.
3 Conclusions

In this paper, we propose a novel compact indoor base station antenna with an orthogonal bi-directional radiation pattern for dual-band operation at 2.6 GHz and 3.5 GHz. The simulated results indicate that input impedance in the 3.5 GHz band can be improved by adding a matching-pin, bringing the $S_{11}$ below the required value of $-10$ dB. Moreover, the MIMO performance can probably be improved because the difference between the two directional gain levels is achieved by approximately 15 dB when four slits are etched into the ground plane. The proposed antenna’s structure is simple and compact, which will be useful when designing high-performance base station antennas. In future work, the estimation of MIMO capacity will be implemented and an experiment will be carried out to test the simulated results presented here for the proposed antenna.

Fig. 3. Simulation results of the proposed antenna