An Empirical Investigation of Multi-path Clusters in an Outdoor MIMO Propagation Environment

Lian Chang[†], Jianhua Zhang[*†], Fenghua Zhang[†], Baoling Liu[*†]
[*]Key Laboratory of Universal Wireless Communications, Ministry of Education,[†]Wireless Technology Innovation Institute
Beijing University of Posts and Telecommunications, P.O. box #92, China,100876
Email:{changlian, zhangfenghua}@mail.wtilabs.cn, {jhzhang, blliu}@bupt.edu.cn

***Abstract*—This paper investigates the cluster dynamic behaviors in an outdoor propagation environment to facilitate the outdoor channel modeling and simulation. To identify the relevant clusters from measured channel data in joint spatial-temporal domain, an automatic cluster identification algorithm has been employed. By introducing a closed loop scheme and a multi-dimension filter, significant improvements on the algorithm performance can be obtained. Furthermore, based on analysis on the cluster identification results, this paper proposes two stochastic models for cluster time-variant characteristics including cluster lifespan and cluster spatial-temporal evolution. Empirical examples reveal the applicability and accuracy of the models of which can be made use in related simulations to imitate the continuous evolution of realistic channels.**

I. Introduction

The use of multiple antennas at both link ends in wireless communication systems can promise high spectral efficiency and reliability. An important feature of the multiple-input multiple-output (MIMO) propagation channels is the occurrence of multi-path components (MPCs) in clusters. In order to take such propagation phenomena into account, many geometry-based stochastic models (GBSMs) all involve the concept of clusters.

To specify the cluster, we employ the definition as [1]:

> *The cluster is not only a mathematical but also a physical collection of distinct MPCs those are close to each other from the perspective of parameter distance. In other words, the gathering for MPCs in one cluster always minimizes their total distance from the cluster centroid.*

As well known, to investigate the cluster dynamic behaviors, a cluster identification procedure must be performed to extract MPCs from original measurement data and group them together. Previous results based on single-input multiple-output (SIMO) channel measurements deployed visual inspection to identify clusters [2], [3], [4], which becomes impractical for large amounts of measurement data. Hence, an automatic clustering approach composed of three procedures, i.e., *KPowermeans*, *CombinedValidate*, and *ShapePrune* was proposed in [5] and its performance was assessed by clustering measurement data from an indoor environment. This algorithm used multi-path component distance (MCD) as the distance function [6] and included path power into the cluster identification. In [7], a method to improve the algorithm in terms of convergence rate and accuracy was presented. Nevertheless, current clustering mechanisms all operated in the open loop mode assuming the cluster time-variation is an independent stochastic process. Obviously, this assumption will lead to identification results lack of interrelation of adjoining channel response and over-estimating the link capacity [8]. Moreover, as another problem to be solved, empirical results in existing literatures indicate that sometimes the MPCs with extraordinarily large spatial-temporal parameters, which are termed outliers, may be wrongly assigned to some clusters. In [5], a cluster pruning method based on iteratively calculating the maximum MCD was used. However, the performance was not as ideal as expectation due to failing in distinguishing respectively the distance of delay parameters and the distance of angular parameters. What is more, when investigating the cluster behaviors, current literatures (e.g., [7]) mainly focused on the fading statistics in individual channel snapshot. The time-variant characteristics which are important for modeling continuous channel transitions have not been paid sufficient attention.

In this paper, we propose an improvement scheme, which operates in the closed loop mode, for the automatic cluster identification. That is to say, the current clustering results are fed back to the following clustering process, which can be modeled as an m-step Markov process. Subsequently, to solve the outlier problem, a multi-dimension filter which is self-adaptive to the overall distribution of each MPC parameter is used to filter the clustering results. Furthermore, we statistically analyze the identification results from the perspective of time-variation and get the stochastic models of cluster dynamic behaviors.

The remainder of this paper is organized as follows. Section II describes the outdoor measurement equipments and environment. Section III addresses the improved cluster identification algorithm. Afterwards, the identification results are presented in Section IV and the stochastic models are proposed in Section V. In the end, Section VI summaries the whole work of this paper.

II. Measurement Setup and Environment

A. Measurement Setup

Channel data used in this paper was collected with a prototype wideband channel sounder working at 120.98 Hz of channel sampling frequency. A pseudo-random sequence of length 1023 was continuously generated at the transmitter

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

Fig. 1. Vertical view of measurement scenario. The location of BS and the trajectories of MS are indicated. (Picture from google map)

(Tx) with a chip rate of 100 MHz. The channel snapshot within which all sub-channels are sounded once is referred to as a measurement cycle. To obtain the spatial domain parameters, 56-element omni-directional array (ODA) was utilized at both the base station (BS) and mobile station (MS) side.

B. Measurement Environment

Measurement was performed in downtown Beijing. The measurement area is surrounded by buildings ranging from 4 to 8 floors, which is illustrated in Fig. 1. The BS antenna was fixed on the roof of a 22-meter building. The MS antenna was mounted on a trolley. To imitate the height of human users, the height of MS antenna was adjusted to 1.7 m. By moving MS, eight continuous routes were measured. The positions of MS were recorded by the global positioning system (GPS).

III. Improved Cluster Identification Algorithm

A. Methodology of Cluster Identification

Firstly of all, for tractability of the problem, we define the signal model like [5] and [7]:

Considering a number of L MPCs in a set of channel parameters, each of them can be represented by its power $P_l,\ l = 1, \ldots, L$, and a parameter vector $\mathbf{X}_l$ given by

$$\mathbf{X}_l = \left[\tau^l, \varphi^l_{AOA}, \varphi^l_{AOD}, \theta^l_{AOD}, \theta^l_{AOA}\right], \quad l = 1, \ldots, L \quad (1)$$

where τ^l, φ^l_{AOA}, φ^l_{AOD}, θ^l_{AOD}, and θ^l_{AOA} are the spatial-temporal parameters including excess delay, angle of arrival (AOA) azimuth, angle of departure (AOD) azimuth, AOD elevation and AOA elevation of the lth MPC.

Initially, a range $\mathbf{K} = [K_{\min} : K_{\max}]$ for the possible cluster number has to be specified by the user. For each possible $K \in \mathbf{K}$, the *KPowermeans* procedure is performed, and the results are collected in the data sets $\mathbf{R}_K = [\mathbf{c}_1, \ldots, \mathbf{c}_k, \ldots, \mathbf{c}_K]$ containing associated cluster centroids $\mathbf{c}_k \ , \quad k = 1 \ldots K$ which is given by

$$\mathbf{c}_k = \left[\tau^k, \varphi^k_{AOA}, \varphi^k_{AOD}, \theta^k_{AOD}, \theta^k_{AOA}, P^k\right] \quad (2)$$

where τ^k, φ^k_{AOA}, φ^k_{AOD}, θ^k_{AOD}, θ^k_{AOA}, and P^k are the spatial-temporal parameters and normalized power for the kth cluster centroid. Subsequently, each result is validated by the *CombinedValidate* procedure which provides the validation index v_K. The optimum cluster number $K_{opt} \in \mathbf{K}$ with corresponding cluster set $\mathbf{R}_{opt} \subseteq \{\mathbf{R}_{K_{\min}}, \ldots \mathbf{R}_{K_i}, \ldots, \mathbf{R}_{K_{\max}},\ K_i \in \mathbf{K}\}$ is finally determined by the largest v_K .

B. Improvement Solution

In the *CombinedValidate* procedure, the combination of two methods well-known in [9], the Calinski-Harabasz index ($\mathrm{CH}(\cdot)$) and the Davies-Bouldin criterion ($\mathrm{DB}(\cdot)$) are used when clustering L MPCs in K clusters. However, the elements in both the vector

$$\begin{aligned}&\mathbf{CH}\left(\mathbf{K},\mathbf{R_K}\right) = \\ &\left[\mathrm{CH}\left(K_{\min}, \mathbf{R}_{K_{\min}}\right), \ldots, \mathrm{CH}\left(K_{\max}, \mathbf{R}_{K_{\max}}\right)\right]\end{aligned} \quad (3)$$

and vector

$$\begin{aligned}&\mathbf{DB}\left(\mathbf{K},\mathbf{R_K}\right) = \\ &\left[\mathrm{DB}\left(K_{\min}, \mathbf{R}_{K_{\min}}\right), \ldots, \mathrm{DB}\left(K_{\max}, \mathbf{R}_{K_{\max}}\right)\right]\end{aligned} \quad (4)$$

for each channel cycle are merely the function of MPC parameters in the cycle itself, which neglects the interrelation of adjacent cycles. To fill the gap, a novel validation solution specifically involving this property is proposed here.

Considering all the measured channel cycles, we can determine the K_{opt} and $\mathbf{R}_{opt}$ based on two m-step Markov processes which use $\mathrm{CH}\left(K, \mathbf{R}_K\right)$ and $\mathrm{DB}(K, \mathbf{R}_K)$ as probability measure respectively.

Let X_n be the cluster number in the nth channel cycle and let $\mathbf{K}$ be the state space. Then the discrete state process $\{X_n, n = 0, 1, \ldots\}$ can be treated as an m-step Markov process based on the assumption that for any n

$$\begin{aligned}&P\left\{X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_n = x_n\right\} = \\ &P\left\{X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_{n-m+1} = x_{n-m+1}\right\}\end{aligned} \quad (5)$$

where x_0 is the initial state of $\{X_n\}$. Generally, equation (5) is termed the transition probability of the Markov process, which can be equivalently expressed as

$$\begin{aligned}&p_n(j|i_0, \ldots, i_{m-1}) = \\ &P\left(X_{n+1} = j | X_n = i_{m-1}, \ldots, X_{n-m+1} = i_0\right).\end{aligned} \quad (6)$$

Furthermore, the k steps transition probability is given by

$$\begin{aligned}&p^{(k)}_n(j|i_0, \ldots, i_{m-1}) = \\ &P\left(X_{n+k} = j | X_n = i_{m-1}, \ldots, X_{n-m+1} = i_0\right) = \\ &\sum_{r \in \mathbf{K}} p_n(r|i_0, \ldots, i_{m-1}) p^{(k-1)}_{n+1}(j|r, i_1, \ldots, i_{m-1}).\end{aligned} \quad (7)$$

In order to get the state distribution of $\{X_n\}$, we construct an m-dimension stochastic vector as

$$\mathbf{Z}_n = [X_n, X_{n+1}, \ldots, X_{n+m-1}], \quad n = 0, 1, \ldots \quad (8)$$

It is easy to prove that $\mathbf{Z}_n$ is an m-dimension Markov vector when X_n is an m-step Markov random variable (RV) [10]. Let $\mathbf{i}^m = [i_1, \ldots, i_m]$ and $\mathbf{j}^m = [j_1, \ldots, j_m]$, the transition probability of $\{\mathbf{Z}_n\}$ can be expressed as

$$\tilde{p}_n(\mathbf{j}^m | \mathbf{i}^m) = P(\mathbf{Z}_{n+1} = \mathbf{j}^m | \mathbf{Z}_n = \mathbf{i}^m), \quad (9)$$

and be further derived from $p(\cdot)$ (defined in (6)) as follows [11]:

$$\tilde{p}_n(\mathbf{j}^m|\mathbf{i}^m) = \begin{cases} p_{n+m-1}\left(j_m|\mathbf{i}^m\right) & j_v = i_{v+1} \\ & v = 1, 2, \ldots, m-1 \\ 0 & others. \end{cases} \tag{10}$$

Moreover, the Chapman-Kolmogorov identity of $\{\mathbf{Z}_n\}$ is given by

$$\begin{aligned} \tilde{p}_n^{(k)}(\mathbf{j}^m|\mathbf{i}^m) = & \sum_{\mathbf{j}_1^m \subseteq \mathbf{S}^m} \sum_{\mathbf{j}_2^m \subseteq \mathbf{S}^m} \cdots \sum_{\mathbf{j}_{k-1}^m \subseteq \mathbf{S}^m} \\ & \tilde{p}_n(\mathbf{j}_1^m|\mathbf{i}^m)\tilde{p}_{n+1}(\mathbf{j}_2^m|\mathbf{j}_1^m)\cdots \\ & \tilde{p}_{n+k-1}(\mathbf{j}^m|\mathbf{j}_{k-1}^m) \end{aligned} \tag{11}$$

where

$$\mathbf{S}^m = [\mathbf{K};\ \mathbf{K}; \ldots; \mathbf{K}]_{m\times 1} \tag{12}$$

is the state space of $\{\mathbf{Z}_n\}$. Based on (10) and (11), we can derive that [11]

$$p_n^{(k+l)}(j|\mathbf{i}^m) = \sum_{\mathbf{r}^m \subseteq \mathbf{S}^m} \tilde{p}_{n-m+1}^{(l)}(\mathbf{r}^m|\mathbf{i}^m) p_{n+l}^{(k)}(j|\mathbf{r}^m). \tag{13}$$

For each possible $\mathbf{i}^m$ and $\mathbf{j}^m$ pair, we can define the k step transition probability matrix $\tilde{\mathbf{P}}_n^{(k)} = \left[\tilde{p}_n^{(k)}(\mathbf{j}^m|\mathbf{i}^m)\right]_{L\times L}$, where $L = (K_{\max} - K_{\min} + 1)^m$. Taken as an instance, we set $K_{\max} = 20$, $K_{\min} = 2$, and $m = 6$, then $L = 19^6$. Obviously, it is too complex to extract such large a matrix from the measurement data. To simplify the solving process, following hypotheses are proposed first.

Hypothesis 1. *Using the* $\mathrm{CH}(K, \mathbf{R}_K)$ *and* $\mathrm{DB}(K, \mathbf{R}_K)$ *of the nth channel cycle as the transition probability which is defined in* (9)*.*

Hypothesis 2. *The kronecker tensor product of vector* $\mathbf{CH}(\mathbf{K}, \mathbf{R_K})$ *and vector* $\mathbf{DB}(\mathbf{K}, \mathbf{R_K})$ *in the first m channel cycles are used as distinct initial distribution of* $\{\mathbf{Z}_n\}$ *in the forms of:*

$$\begin{aligned} \tilde{\mathbf{P}}_{\mathrm{CH}}(\mathbf{Z}_0) = & \mathbf{CH}(\mathbf{K},\mathbf{R_K})_1 \otimes \mathbf{CH}(\mathbf{K},\mathbf{R_K})_2 \otimes \cdots \\ & \otimes \mathbf{CH}(\mathbf{K},\mathbf{R_K})_m \end{aligned} \tag{14}$$

and

$$\begin{aligned} \tilde{\mathbf{P}}_{\mathrm{DB}}(\mathbf{Z}_0) = & \mathbf{DB}(\mathbf{K},\mathbf{R_K})_1 \otimes \mathbf{DB}(\mathbf{K},\mathbf{R_K})_2 \otimes \cdots \\ & \otimes \mathbf{DB}(\mathbf{K},\mathbf{R_K})_m \end{aligned} \tag{15}$$

According to **Hypothesis 1** and (10), the transition probability of $\{\mathbf{Z}_n\}$ can be rewritten as

$$\tilde{p}_n(\mathbf{j}^m|\mathbf{i}^m) = \begin{cases} p_{n+m-1}(j_m|\mathbf{i}^m) = \mathrm{DB}(j_m, \mathbf{R}_{j_m})_{n+m-1} \\ j_v = i_{v+1},\ v = 1, \ldots, m-1 \\ 0 \qquad\qquad others \end{cases} \tag{16}$$

and

$$\tilde{p}_n(\mathbf{j}^m|\mathbf{i}^m) = \begin{cases} p_{n+m-1}(j_m|\mathbf{i}^m) = \mathrm{CH}(j_m, \mathbf{R}_{j_m})_{n+m-1} \\ j_v = i_{v+1}, v = 1, \ldots, m-1 \\ 0 \qquad\qquad others \end{cases} \tag{17}$$

for Markov process $\{(\mathbf{Z}_n,\ \mathrm{DB}(K, \mathbf{R}_K)_n)\}$ and $\{(\mathbf{Z}_n,\ \mathrm{CH}(K, \mathbf{R}_K)_n)\}$ respectively. It can be proved that the transition probability matrices of these two Markov processes turn to be sparse and have many identical elements under above conclusions. Substituting (11), (13) turns to

$$\begin{aligned} p_n^{(k+l)}(j|\mathbf{i}^m) = & \sum_{\mathbf{r}^m} \sum_{\mathbf{r}_1{}^m} \sum_{\mathbf{r}_2{}^m} \cdots \sum_{\mathbf{r}_{l-1}{}^m} \tilde{p}_{n-m+1}(\mathbf{r}_1{}^m|\mathbf{i}^m) \\ & \tilde{p}_{n-m+2}(\mathbf{r}_2{}^m|\mathbf{r}_1{}^m)\cdots\tilde{p}_{n-m+l}(\mathbf{r}^m|\mathbf{r}_{l-1}{}^m) \\ & p_{n+l}{}^{(k)}(j|\mathbf{r}^m). \end{aligned} \tag{18}$$

Being equivalent to (18), the k step transition probability matrix is expressed as

$$\mathbf{P}_n^{(k)} = \tilde{\mathbf{P}}_{n-m+1}^{(k-1)} \mathbf{P}_{n+k-1}^{(1)} \tag{19}$$

where $\mathbf{P}_n^{(k)} = \left[p_n^{(k)}(j|\mathbf{i}^m)\right]_{L\times(K_{\max}-K_{\min}+1)}$. Then, the distribution of $\{X_n\}$ in any channel cycle can be calculated by

$$\mathbf{P}(X_{n'}) = \tilde{\mathbf{P}}(\mathbf{Z}_0)\, \mathbf{P}_n^{(k)}|_{n=0,\ k=n'} \tag{20}$$

where $\tilde{\mathbf{P}}(\mathbf{Z}_0)$ is the initial distribution defined as **Hypothesis 2**. For the Markov process $\{(\mathbf{Z}_n,\ \mathrm{DB}(K, \mathbf{R}_K)_n)\}$, substituting (16) and (19) into (20), the distribution of cluster number in any channel cycle can be further written as

$$\mathbf{P}_{\mathrm{DB}}(X_{n'}) = \left\langle \mathbf{P}_{\mathrm{DB}}(X_{m-1}) * \prod_{i=0}^{n'-1} \mathbf{DB}(\mathbf{K}, \mathbf{R_K})_i \right\rangle \tag{21}$$

where $\langle * \rangle$ and $\prod \cdot$ are specific to the product of corresponding elements in two vectors, and

$$\begin{aligned} & \mathbf{P}_{\mathrm{DB}}(X_{m-1}) = \\ & \sum_{X_0 \in \mathbf{K}} \cdots \sum_{X_{m-2} \in \mathbf{K}} \tilde{\mathbf{P}}_{\mathrm{DB}}(\mathbf{Z}_0 = [X_0, \ldots, X_{m-2}, X_{m-1}])\ \cdot \end{aligned} \tag{22}$$

Similarly, for the Markov process $\{(\mathbf{Z}_n,\ \mathrm{CH}(K, \mathbf{R}_K)_n)\}$, substituting (17) and (19) into (20), we can get the distribution expressed as

$$\mathbf{P}_{\mathrm{CH}}(X_{n'}) = \left\langle \mathbf{P}_{\mathrm{CH}}(X_{m-1}) * \prod_{i=0}^{n'-1} \mathbf{CH}(\mathbf{K}, \mathbf{R_K})_i \right\rangle \tag{23}$$

where $\mathbf{P}_{\mathrm{CH}}(X_{m-1})$ is defined just like $\mathbf{P}_{\mathrm{DB}}(X_{m-1})$. Considering (21) and (23), the *CombinedValidate* algorithm can be rewritten as

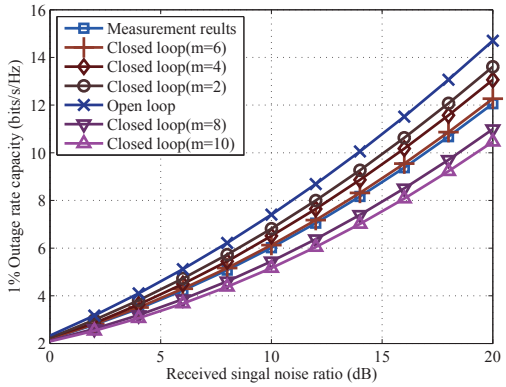
$$p_{\mathrm{DB}}(x_{n'} = K_i) \le t \cdot \min\left\{\mathbf{P}_{\mathrm{DB}}(X_{n'})\right\}, \tag{24}$$

where t is an empirical coefficient defined in [5], and assuming the set $\mathbf{F} \subseteq \mathbf{K}$ consists of all the possible K_i which satisfies (24), the K_{opt} and corresponding $\mathbf{R}_{opt}$ for the n'th cycle can be determined by

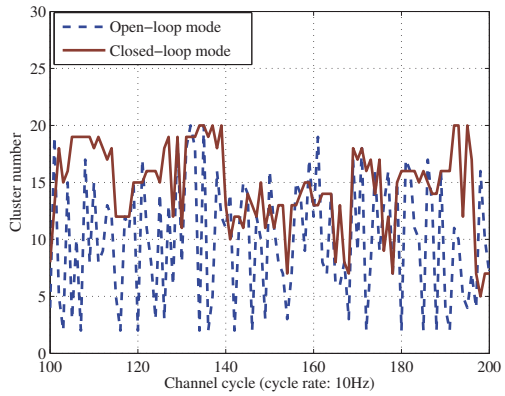
$$K_{opt}^{(n')} = \arg\max\left\{p_{\mathrm{CH}}(x_{n'} = K_i)\right\}, \quad \forall K_i \in \mathbf{F}. \tag{25}$$

At the last step of the cluster identification algorithm, the results obtained from the *ShapePrune* procedure are filtered by a multi-dimension filer in delay and angular domain. The filter is modeled as a five-dimension vector expressed as

$$\begin{aligned} \mathcal{F} = & \mathcal{F}_\tau \vec{i_\tau} + \mathcal{F}_{\varphi_{AOA}} \vec{i_{\varphi AOA}} + \mathcal{F}_{\varphi_{AOD}} \vec{i_{\varphi_{AOD}}} \\ & + \mathcal{F}_{\theta_{AOA}} i_{\theta_{AOA}} + \mathcal{F}_{\theta_{AOD}} i_{\theta_{AOD}} \end{aligned} \tag{26}$$

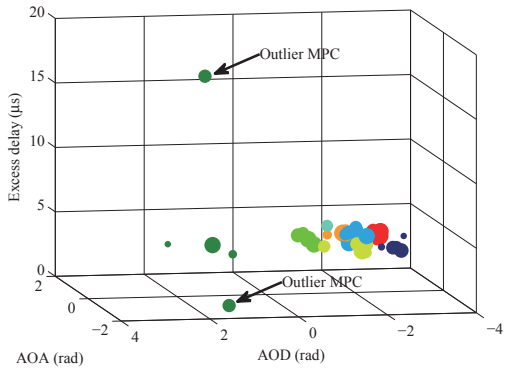


(a)

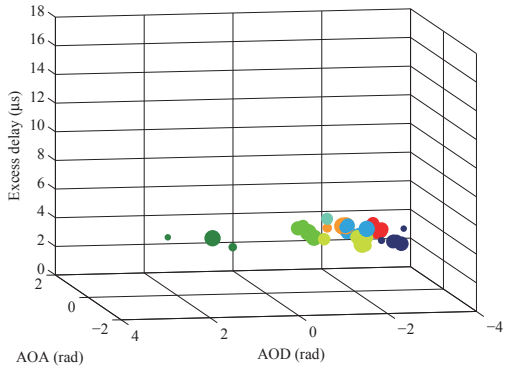


(b)

Fig. 2. Performance validation based on the closed loop mode: (a) outage capacity of the reconstructed channels for different m compared with that of the measurement data, and (b) time-variant cluster numbers extracted from the open and closed loop mode respectively.



(a)



(b)

Fig. 3. Cluster identification results: (a) without multi-dimension filter, (b) with multi-dimension filter.

where $\mathcal{F}_{\tau}$, $\mathcal{F}_{\varphi_{AOA}}$, $\mathcal{F}_{\varphi_{AOD}}$, $\mathcal{F}_{\theta_{AOA}}$, and $\mathcal{F}_{\theta_{AOD}}$ are transfer function (TF) in spatial-temporal domains, and $\vec{i}$ denotes the directional vector for corresponding domain. The passband width of the TF in these five domains is determined by the probability distribution function (pdf) of corresponding parameters. The filtered results are fed back to the *ShapePrune* procedure, until the pruned clusters do not change any more.

IV. Cluster Identification Results

Varying the step of the Markov process, series of K_{opt} and $\mathbf{R}_{opt}$ can be obtained. Then, we input these results into the generic model framework in [12], and reconstruct the radio channels. Based on the channel impulse responses (CIRs) collected from the channel model and outdoor measurements described in Section II, we evaluate the outage capacity of corresponding cases. Fig. 2(a) validates the conclusion in [8] that uncorrelated clustering results may over-estimate the link capacity. Besides, we can see in Fig. 2(a) that as the step of the Markov process increases, the capacity of reconstructed channels gets closer to the realistic situation. But, when the step is over 6, the approximation gets worse. Fig. 2(b) shows the time-variant cluster number after applying the m-step Markov process to cluster identification. The dashed line stands for the cluster number obtained from the open loop mode. For comparison, the cluster number obtained from the closed loop mode is plotted by the solid line. Obviously, thanks to the improvement scheme, variation of cluster number between adjacent channel cycles turns to be slower, which is closer to actual situation.

Fig. 3 shows clustering results after applying the multi-dimension filter in the *ShapePrune* procedure. Obviously, in Fig. 3(b), the MPCs with relatively larger and smaller delay which appear in Fig. 3(a) are removed from their own cluster. The filter improves the visibility without changing cluster behavior significantly, and leads to a more intrinsic and intuitive definition of a "cluster" itself.

V. Time-Variant Properties Modeling

A. Cluster Lifespan

In this paper, the cluster lifespan is referred to the time interval across which the cluster exists from its first appearance until it finally disappears as the MT moves along its trajectory. Here, we employ the MCD-based cluster tracking method in [1] to discriminate the birth and death of clusters. In Fig. 4, the cluster lifespan is showed by the statistic histogram. The lognormal distribution,

$$f\left(x;\ u,\ \sigma\right) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - u)^2/2\sigma^2} \tag{27}$$

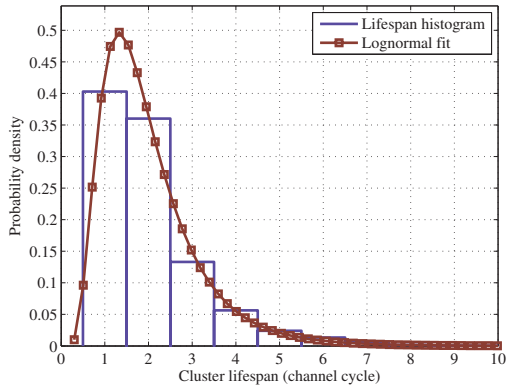


Fig. 4. Distribution of cluster lifespan and its lognormal fit. (Using channel cycle as the measure of lifespan.)

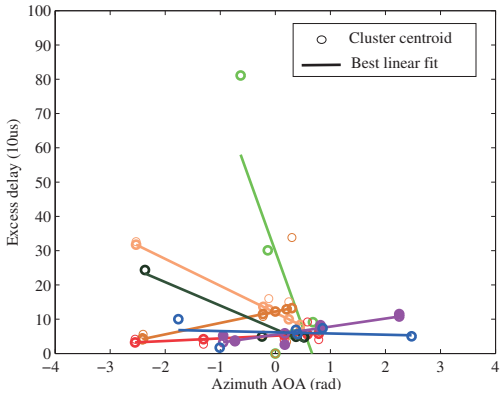


Fig. 5. Example of the spatial-temporal variation of clusters within their lifespan.

fits the measured data as shown in Fig. 4 with $\mu = 2.0$ and $\sigma = 1.30$.

B. Cluster Spatial-Temporal Vector

Due to the variation of radio channels, fluctuation of the cluster spatial-temporal parameters would occur within the cluster lifespan. Results from data analysis reveal that the variation of these parameters can be modeled by four linear polynomial approximations corresponding to $\tau-\varphi_{\mathbf{AOA}}$, $\tau-\varphi_{\mathbf{AOD}}$, $\tau-\theta_{\mathbf{AOD}}$, $\tau-\theta_{\mathbf{AOA}}$ respectively (hereinafter referred to as: $\tau-\varphi_{\mathbf{AOA}}$ vector, $\tau-\varphi_{\mathbf{AOD}}$ vector, $\tau-\theta_{\mathbf{AOD}}$ vector, and $\tau-\theta_{\mathbf{AOA}}$ vector). Thus, linear least squares regression is used to find the best line fits through all data points of each cluster in the form of

$$y = kx + b \tag{28}$$

where k is the gradient of the spatial-temporal vectors. Fig. 5 shows examples of the empirical $\tau-\varphi_{\mathbf{AOA}}$ vector and illustrates the best line fit for cluster transition. It indicates that the direction of change is independent of the location in the spatial-temporal domain. Thus, the gradient k of each vector can be described by an RV. Here, we find that $f(k)$ can be well modeled by a Gaussian pdf as illustrated by the plots in Fig. 6.

If taking the correlation of gradient possessed by distinct vectors into consideration, the cluster time-variant behaviors can be modeled as a four-dimension stochastic vector $(k_{\tau-\varphi_{AOA}}, k_{\tau-\varphi_{AOD}}, k_{\tau-\theta_{AOD}}, k_{\tau-\theta_{AOA}})$ with four-dimension normal distribution:

$$f(x_1, x_2, x_3, x_4) = \frac{1}{(2\pi)^2 |\mathbf{B}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{a}-\mathbf{x})\mathbf{B}^{-1}(\mathbf{a}-\mathbf{x})'} \tag{29}$$

where $\mathbf{a}$ denotes the expectation vector, $\mathbf{B}_{4\times4}$ denotes the covariance matrix, and $(\cdot)'$ denotes matrix transposition.

Calculating the integration of $k_{\tau-\theta_{AOD}}$ and $k_{\tau-\theta_{AOA}}$, the joint pdf of the gradient of delay-azimuth vectors can be derived from

$$f(x_1, x_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{(2\pi)^2 |\mathbf{B}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{a}-\mathbf{x})B^{-1}(\mathbf{a}-\mathbf{x})'} dx_3 dx_4. \tag{30}$$

The joint pdf of the gradient of delay-elevation vectors can be derived from the integration of $k_{\tau-\varphi_{AOA}}$ and $k_{\tau-\varphi_{AOD}}$ similarly.

Fig. 7 shows the joint distribution of $k_{\tau-\varphi_{AOA}}$, $k_{\tau-\varphi_{AOD}}$ and that of $k_{\tau-\theta_{AOD}}$, $k_{\tau-\theta_{AOA}}$. Observably, the gradient of delay-elevation vectors is more clustered to small values than that of delay-azimuth vectors. That is to say, the intra-cluster elevation angles of departure and arrival will take a slower change with MT motion, so that we need not to update the cluster elevation angles such frequently as the azimuth angles when modeling or simulating the time-variant channels.

VI. Conclusion

In this paper, the cluster dynamic behaviors in the outdoor MIMO environment are investigated. We propose an improvement solution for the cluster identification firstly. By introducing the closed loop mode, we significantly improve the validity of identification results making them closer to real cluster behaviors. Thanks to the multi-dimension filter, the outlier paths are eliminated effectively. Subsequently, through analysis on the identification results, the cluster lifespan is found to be well fitted with lognormal distribution and the continuous transition of cluster parameters can be modeled by several spatial-temporal vectors whose gradient obeys to normal distribution. Their joint and partial joint pdf are also obtained in this paper.

Acknowledgment

The research is supported in part by China Key Projects in the National Science & Technology under Grant NO. 2012BAF14B01, by China Important National Science & Technology Specific Projects under Grant NO. 2012ZX03001043-009 and by China Mobile Research Institute.

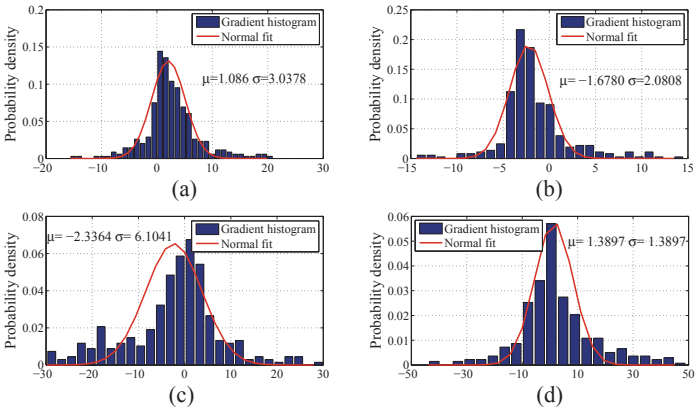


Fig. 6. The distribution of the gradient of (a) $\tau-\varphi_{\mathbf{AOA}}$ vector, (b) $\tau-\varphi_{\mathbf{AOD}}$ vector, (c) $\tau-\theta_{\mathbf{AOD}}$ vector, and (d) $\tau-\theta_{\mathbf{AOA}}$ vector.

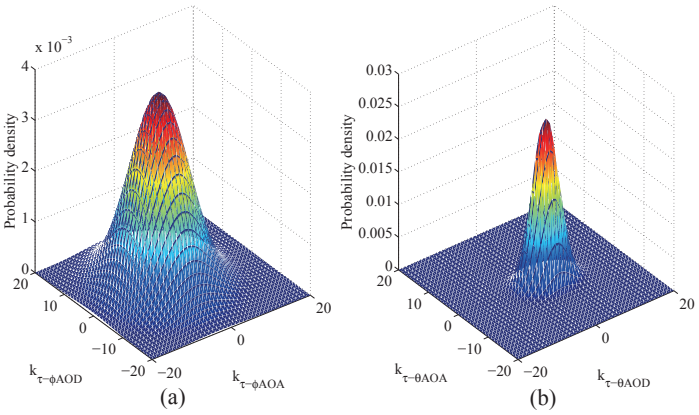


Fig. 7. Joint pdf of the gradient of spatial-temporal vectors: (a) $k_{\tau-\varphi_{AOA}}$, $k_{\tau-\varphi_{AOD}}$, and (b) $k_{\tau-\theta_{AOD}}$, $k_{\tau-\theta_{AOA}}$.

REFERENCES

[1] N. Czink, C. Mecklenbrauker, and G. Del Galdo, "A novel automatic cluster tracking algorithm," in *Proc. IEEE 17th Int Personal, Indoor and Mobile Radio Communications Symp*, 2006, pp. 1–5.

[2] L. Vuokko, P. Vainikainen, and J. Takada, "Clusters extracted from measured propagation channels in macrocellular environments," *IEEE Trans. Antennas Propag.*, vol. 53, no. 12, pp. 4089–4098, 2005.

[3] K. Yu, Q. Li, and M. Ho, "Measurement investigation of tap and cluster angular spreads at 5.2 ghz," *IEEE Trans. Antennas Propag.*, vol. 53, no. 7, pp. 2156–2160, 2005.

[4] C.-C. Chong, C.-M. Tan, D. I. Laurenson, S. McLaughlin, M. A. Beach, and A. R. Nix, "A new statistical wideband spatio-temporal channel model for 5-ghz band wlan systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 139–150, 2003.

[5] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric mimo channel data including path powers," in *Proc. VTC-2006 Fall Vehicular Technology Conf. 2006 IEEE 64th*, 2006, pp. 1–5.

[6] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuuninen, and J. Ylitalo, "Automatic clustering of mimo channel parameters using the multi-path component distance measure," in *Proc. WPMC2005*, 2005, pp. 829–833.

[7] W. Dong, J. Zhang, X. Gao, P. Zhang, and Y. Wu, "Cluster identification and properties of outdoor wideband mimo channel," in *Proc. VTC-2007 Fall Vehicular Technology Conf. 2007 IEEE 66th*, 2007, pp. 829–833.

[8] K.-H. Li, M. A. Ingram, and A. Van Nguyen, "Impact of clustering in statistical indoor propagation models on link capacity," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 521–523, 2002.

[9] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1650–1654, 2002.

[10] J. Menhi, *Stochastic process*, 2nd ed. New Age International Press, 2002.

[11] J. Wang, "Ergodic for multiple nonhomogeneous markov chains and application," M.S. thesis, Inst. Sci., Jiangsu Univ., Zhenjiang, China, Dec. 2006.

[12] "WINNER II channel models," 2007, IST-4-027756 WINNER II, D1.1.2, V1.0.