

# CLUSTER ANALYSIS OF WIRELESS CHANNEL IMPULSE RESPONSES WITH HIDDEN MARKOV MODELS

*Dmitriy Shutin, Gernot Kubin*

Institute of Communication and Wave Propagation  
Graz University of Technology  
Inffeldgasse 12, A-8010 Graz, Austria  
dshutin@inw.tugraz.at, kubin@inw.tugraz.at

## ABSTRACT

This paper introduces a novel wireless channel clustering technique based on the Saleh-Valenzuela channel model. The channel impulse response is regarded as a realization of the probabilistic channel model, based on which the prior density functions of cluster arrival times are derived. Cluster analysis is done by means of extending the Saleh-Valenzuela model to a non-stationary case and re-interpreting it in terms of mixture models. The parameters of the mixture are then learned with Hidden Markov Models. Once trained, the HMM could be used to optimally cluster the channel taps with Viterbi algorithm. The proposed method has been applied to simulated as well as measured channel impulse responses and showed reasonably good performance.

## 1. INTRODUCTION

Wireless systems are subject to fading - time variations of the receiving conditions caused by multipath propagation and transceiver movements. Due to the fading nature of the wireless channel, it is imperative for the system to follow the variations of the receiving conditions and adapt itself to sustain reasonable communication quality.

Generally, the wireless channel consists of a series of attenuated, time-delayed, phase shifted replicas of the transmitted signal. In the baseband this will be represented as follows:

$$h_t(\tau) = \sum_{i=0}^{N-1} \beta_i(t, \tau) \exp(j\theta_i(t, \tau)) \delta(\tau - \tau_i(t)) \quad (1)$$

where  $\beta_i(t, \tau)$  and  $\tau_i(t)$  are the real amplitudes and excess delays of the  $i$ th multipath component at the time  $t$ , respectively (see, for example, [1], chapter 4). The phase term  $\theta_i(t, \tau)$  lumps together all the mechanisms for phase shifts of a single multipath component within the  $i$ th excess delay bin. Equation (1) is a starting point into further analysis. The straight-forward procedure to follow the channel is to follow each of the contributing reflections. Obviously, it is a hard task to accomplish, since the number of contributions could be significant, but if successful, it will provide a detailed description of the channel dynamics. Although, it is computationally impractical to follow all of the contributing paths, it is still possible to follow at least several strongest [2, 3]. Alternatively, a number of statistical models provide probability density functions (PDFs) for the parameters of interest [4, 5], thus giving probabilistic description of the channel behavior. However, for on-line operation the communication system has to know the instantane-

ous parameter values. It would seem beneficial to join both ideologies and devise a method to extract some information from the instantaneous impulse response, based on the probabilistic model of the channel. This could be accomplished if the channel impulse response is regarded as a sample realization of the corresponding probabilistic model.

Among a number of different statistical models, the Saleh-Valenzuela (S.-V.) channel model [6] has attracted our attention, because it provides a very promising framework for implementing these ideas. The basic idea behind the S.-V. model is based on the assumption that rays arrive in clusters. The clusters arrive with a certain rate at random time instances, as well as the individual rays within each cluster. Such a structure imposed on the impulse response provides, first of all, a basis for hierarchical analysis (i.e., from impulse response to clusters and rays), and physical interpretation of the clusters and model parameters. Although the S.-V. model was initially developed for indoor channels, the same ideas could still be valid for outdoor communications with wideband and directional systems [7]. The clusters present in the impulse response allow one to invoke an ample set of unsupervised clustering techniques based on the known statistical structure of the data. Hidden Markov Models (HMM) is a promising candidate for such a task.

We will show how to re-formulate the channel structure in terms of the HMMs and how to learn the model parameters. The rest of the paper is organized as follows: Section 2 describes the parameter estimation algorithm and clustering procedure, along with some practical considerations; Section 3 shows the application of the proposed clustering algorithm to simulated channels; and Section 4 provides some results of algorithm application to the measured channel impulse responses.

## 2. CLUSTER ANALYSIS ALGORITHM

### 2.1. General description of the Saleh-Valenzuela model

The basic idea of the S.-V. model is easily understood from considering the impulse response in Fig. 1. The model assumes rays arriving in clusters. The time between cluster arrivals is random and distributed exponentially with a parameter  $\Lambda$ . Likewise, within the cluster the time between successive ray arrivals is also exponentially distributed with a parameter  $\lambda$ ,  $\lambda \gg \Lambda$ :

$$\begin{aligned} p(T_i|T_{i-1}) &= \Lambda \exp[-\Lambda(T_i - T_{i-1})] \\ p(\tau_{k,l}|\tau_{k-1,l}) &= \lambda \exp[-\lambda(\tau_{k,l} - \tau_{k-1,l})] \end{aligned} \quad (2)$$

Here,  $T_l$  is the arrival time of the  $l$ th cluster, given the previous arrival at  $T_{l-1}$ ,  $T_l \geq T_{l-1}$ . Likewise,  $\tau_{k,l}$  is the arrival time of the  $k$ th ray within the  $l$ th cluster, given the preceding ray at  $\tau_{k-1,l}$ ,  $\tau_{k,l} \geq \tau_{k-1,l}$ .

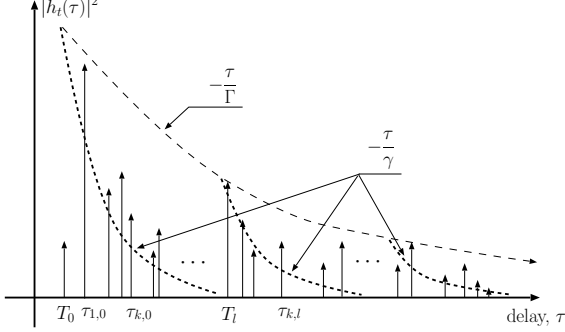


Fig. 1. An example of a wireless channel impulse response.

The power gain  $\beta_{k,l}^2$  is assumed by the model to be independent of the associated delays and distributed exponentially as follows:

$$p(\beta_{k,l}^2) = \left(\overline{\beta_{k,l}^2}\right)^{-1} \exp\left(-\beta_{k,l}^2/\overline{\beta_{k,l}^2}\right)$$

where the expected power gain  $\overline{\beta_{k,l}^2}$  is a function of the delay  $\tau$

$$\overline{\beta_{k,l}^2} = \overline{\beta_{0,0}^2} \exp\left(-\frac{T_l}{\Gamma}\right) \exp\left(-\frac{\tau_{k,l} - \tau_{0,l}}{\gamma}\right)$$

Here,  $\overline{\beta_{0,0}^2}$  is the expected power gain of the first ray in the first cluster.  $\Gamma$  and  $\gamma$  are power-delay time constants for the clusters and the rays, respectively (see Fig. 1). Next, we formulate the clustering algorithm.

## 2.2. PDFs of the cluster arrival times

Let us consider a snapshot of the time-varying channel impulse response  $h_t(\tau)$ , for example, the one shown in Fig. 1. We will drop explicitly the time dependency for notation simplicity and write  $h(\tau)$  only. We can think of it as of a single realization of the S.-V. model.

Let us also assume, that this particular realization has  $L$  clusters, arriving at the unknown times  $T_l$ ,  $l = 0, 1, \dots, L-1$ . Without any loss of generality, let us also assume that  $T_0 \equiv 0$ , i.e., the first cluster arrives at the moment  $\tau = 0$ . Defining the time between consecutive arrivals as  $\Delta T_l = T_l - T_{l-1}$ , the positions of the clusters are given as:

$$T_l = \sum_{m=1}^l \Delta T_m \quad (3)$$

From the original model it follows that the inter-arrival times  $\Delta T_m$ 's are statistically independent exponentially distributed random variables. This leads to the derivation of the PDFs  $p_l(T)$  of the cluster arrival times. Indeed, from (3) it follows that  $p_l(T)$  is a convolution of  $l$  exponential distributions. It can be shown that

$p_l(T)$  is a chi-square distribution with  $2l$  degrees of freedom:

$$p_l(T) = \frac{\Lambda^l}{(l-1)!} T^{l-1} \exp(-\Lambda T) \quad (4)$$

Expression (4) could be interpreted differently. For instance, the case  $l = 1$  gives the distribution of the arrival time of the second cluster. But, the same function could also be used as a likelihood of an arbitrary time instance  $\tau$  belonging to the first cluster. Indeed,  $p_1(\tau)$  decays exponentially with its maximum at zero, since by definition the first cluster starts at  $\tau = 0$ .

Unlike the original formulation (2), expression (4) is not conditioned on any of the previous arrival instances and, in this sense, it is a prior distribution of the cluster arrival times. This property makes it useful in the proposed clustering algorithm.

## 2.3. Parameter estimation and clustering

Now, when the PDFs of the cluster arrival times are known, we proceed with the clustering algorithm. We start with the following assumptions:

1. The samples come from a known number of  $L$  clusters.
2. The channel impulse response  $h(\tau)$  is defined on a finite ordered set  $\mathcal{D} = \{\tau_0, \tau_1, \dots, \tau_{N-1}\}$  of time instances where the pulses were registered in the impulse response. Thus,  $T_l \in \mathcal{D}$  and  $\tau_{k,l} \in \mathcal{D}, \forall k, l$
3. We extend the original model by assuming different  $\Lambda_l$  for each cluster. Thus, each  $\Lambda_l$  uniquely specify the position of the corresponding cluster allowing clusters to move in a non-stationary environment.

For a given impulse response the time instance  $\tau_j$  could possibly belong to any of the  $L$  clusters. Mathematically speaking, it belongs to a mixture of  $L$  density functions:

$$p(\tau_j|\mathbf{\Lambda}) = \sum_{l=1}^L p(\tau_j|\Lambda_l, \omega_l) p(\omega_l) \quad (5)$$

$$p(\tau_j|\Lambda_l, \omega_l) = \frac{\Lambda_l^l}{(l-1)!} \tau_j^{l-1} \exp(-\Lambda_l \tau_j)$$

Here,  $p(\omega_l)$ 's are the prior probabilities for each of the clusters,  $\omega_l \equiv l$  is a cluster index, and  $p(\tau_j|\Lambda_l, \omega_l)$  is a likelihood of the  $\tau_j$  belonging to the  $l$ th cluster. The parameters of the mixture (5) could be effectively learned with HMMs. HMMs are extensively used in many applications involving the modeling of the sequential data. To determine the HMM we have to specify its basic elements: the number of states  $L$ , the state transition matrix  $\mathbf{A}$ , the initial state distribution  $\pi_l$ , and the observation PDFs  $p_l(\tau)$ ,  $l = 1 \dots L$ . The most suiting to our case would be the classical *left-right* model [8]. The choice of this type of the HMM is based on the following reasoning: each state in the HMM is associated with the corresponding cluster. Assuming sequential cluster arrivals, the jumps backward or over several states are now allowed. This dictates the special form of the transition matrix  $\mathbf{A}$ , characteristic for the left-right models. The initial state distribution is also fixed, since by definition we start in the first cluster. In our formulation the observation sequence is formed from the delays  $\tau_j$ , drawn from the mixture (5). Each component in the mixture is the corresponding observation PDF. Thus,  $p_l(\tau_j) \equiv p(\tau_j|\Lambda_l, \omega_l)$ .

The training of HMM consists in estimating the model parameters given the observation sequence. This task is achieved by

EM-like algorithms and has been treated in a number of sources (see [9, 10], for example). There is a slight modification to the parameter re-estimation formulas that accounts for of the non-normal form of the observation distribution  $p(\tau_j|\Lambda_l, \omega_l)$ . Once the parameters are learned the HMM can be used to optimally cluster the taps of the impulse response. Since each state corresponds to the cluster, the most probable state sequence given the observations and the model parameters will tell us which taps have been emitted at which state, i.e., from which cluster they come. This is solved by the Viterbi procedure[8].

One could also be interested in the expected time of the cluster arrival, which in this case will be given as

$$\bar{\tau}_{\cdot,l} = E\{\tau|l\} = \int_0^\infty \tau p(\tau|\Lambda_l, \omega_l) d\tau = l/\Lambda_l \quad (6)$$

## 2.4. Practical considerations

The practical implementation of the algorithm is quite straightforward except for some particularities.

The distributions in the form (4) are very loose, meaning there is a significant uncertainty in classifying the points belonging to the late clusters. To amend this we regularize the classification algorithm by introducing an additional weighting function  $w(\tau|\Lambda_l, \omega_l)$  used as a window to truncate the original density  $p(\tau_j|\tilde{\Lambda}_l^{(k)}, \omega_l)$ :

$$w(\tau_j|\mu_l, \alpha_l) = \frac{1}{\alpha_l \sqrt{2\pi}} \exp(-0.5(\tau_j - \mu_l)^2/\alpha_l^2)$$

Assuming functional independence of the window parameters  $\mu_l$  and  $\alpha_l$  from  $\Lambda_l$ , the former can be easily estimated from the data within the same HMM framework by setting the derivatives of the corresponding likelihood function with respect to  $\mu_l$  and  $\alpha_l$  to zero.

It is crucial to operate on a sparse impulse response. Otherwise no clusters can be defined, since a uniform sequence of instances  $\tau_j$  does not form clusters. Sparseness can be enforced by thresholding channel coefficients with power gains smaller than the noise level and then selecting local maxima.

Due to the iterative nature of the algorithm, the choice of the initial values is an important step. In our implementation the initial values were set as follows:

$$a_{ll}^{(0)} = 0.95, \quad \tilde{\Lambda}_l^{(0)} = l \left/ \frac{\max_j(\tau_j)}{L+1} \right., \text{ for } l = 1 \dots L$$

## 3. SIMULATION RESULTS

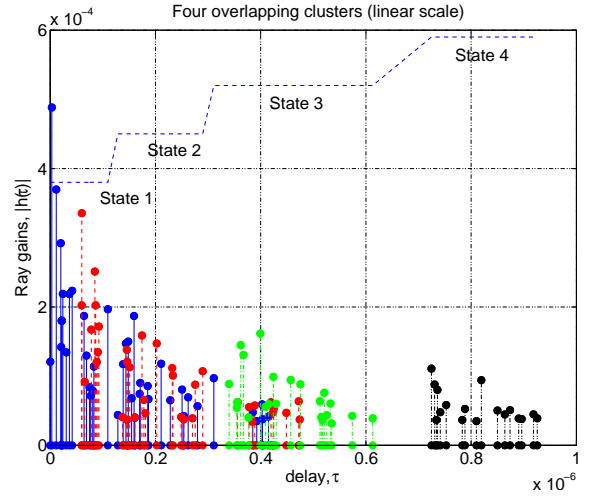
To test the algorithm we simulate an impulse response according to the original S.-V. model (Fig. 2). The parameters used in the simulation are summarized in Table 1.

$$\begin{aligned} L &= 4; \Lambda = 5.7 \cdot 10^6, [\text{sec}^{-1}]; \quad \lambda = 1.0 \cdot 10^8, [\text{sec}^{-1}] \\ 1/\Gamma &= 3.0 \cdot 10^6, [\text{sec}^{-1}]; \quad 1/\gamma = 1.0 \cdot 10^7, [\text{sec}^{-1}] \\ \overline{\beta_{0,0}^2} &= 5.0 \cdot 10^{-8}, [\text{W}]; \quad \text{noise floor} = 1.0 \cdot 10^{-9}, [\text{W}] \end{aligned}$$

**Table 1.** Parameters used in the simulation.

This example shows the case when the 1st and 2nd clusters

are strongly overlapping, 3rd cluster only partially overlaps with the two previous ones, and cluster 4 is clearly distant from the preceding neighbors.



**Fig. 2.** Impulse response with overlapping clusters.

Cluster index, $l$	1	2	3	4
True $\bar{\tau}_{\cdot,l} \times 10^{-6}, [\text{sec}]$	0.16	0.23	0.46	0.81
Estimated $\bar{\tau}_{\cdot,l} \times 10^{-6}, [\text{sec}]$	0.07	0.20	0.43	0.81
Estimated $\Lambda_l \times 10^6, [\text{sec}^{-1}]$	17.7	10.0	6.8	4.9

**Table 2.** Comparison of true and estimated expectations of cluster arrival times for the channel shown in Fig.2.

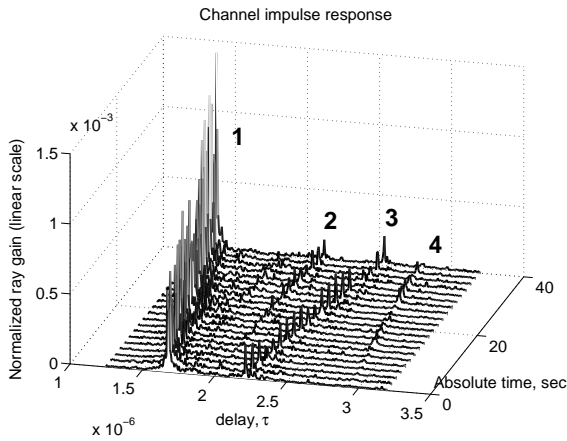
The clustering results are shown in Fig.2 by the corresponding state transition curve superimposed on the plot. By comparing the expected cluster arrival times computed according to (6) with the true ones, known from the simulations, we can judge the performance of the clustering algorithm. The corresponding values are summarized in Table 2. We can see that the algorithm fails to properly distinguish the strongly overlapping, however the estimated arrival time of the 4th cluster is very close to the true value. We can also compute an empirical cluster arrival rate  $\Lambda^{emp}$  as an average of the  $\tilde{\Lambda}_l$  (excluding the  $\tilde{\Lambda}_1$  since the arrival rate is measured with respect to the first cluster). For this simulation,  $\Lambda^{emp} = 7.2 \cdot 10^6 \text{ sec}^{-1}$  which is overestimated in this case.

## 4. APPLICATION OF THE CLUSTERING ALGORITHM TO THE MEASURED CHANNEL IMPULSE RESPONSES

This section presents some results of the application of the clustering algorithm to channel impulse responses obtained from the field-trial Multiple-Input-Multiple-Output (MIMO) channel measurements, performed by Forschungszentrum Telekommunikation Wien, FTW, Vienna, Austria, under the supervision of Helmut Hofstetter<sup>1</sup>[11]. For our purposes we select only a Single-Input-Single-Output (SISO) subset by taking one transmitting and one receiving antenna from the array.

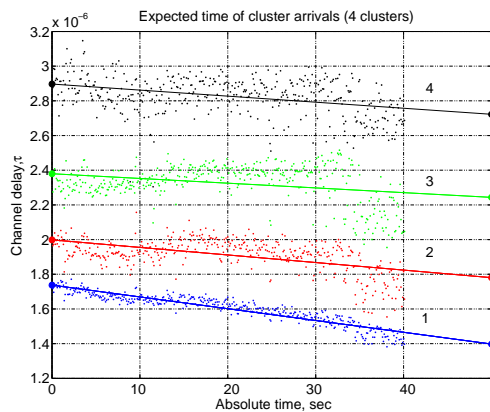
<sup>1</sup>The authors wish to thank Forschungszentrum Telekommunikation Wien for providing MIMO channel measurements data.

Fig.3 shows consecutive measurements of the channel impulse response as the transmitter moves. The channel snapshot was taken every 20msec while the transmitter was moving at  $\approx 1\text{m/s}$ . This particular example shows a subset from 2000 consecutive channel snapshots which is equivalent to 40sec of measurements.



**Fig. 3.** Measured channel impulse response (non-stationary behavior).

The visual inspection of the impulse responses reveals several prominent stripes. It is reasonable to assume the existence of clusters in a neighborhood of these peaks. Assuming four cluster, we perform clustering on every 5th channel snapshot from the measurement data. At each step the algorithm was pre-initialized with the parameters estimated on the preceding snapshot. Fig.4 shows the estimates of expected time of arrival for four clusters (scattered traces) with superimposed linear regression lines.



**Fig. 4.** Expected time of cluster arrivals with superimposed linear regressions.

The variance of the estimates is explained by the uncertainty due to the noise present in the measurements, and final number of taps in the channel impulse response.

## 5. CONCLUSION

Clustering the channel impulse responses could be an interesting approach to extract important channel parameters. Based initially on the S.-V. channel model, the expressions for the PDFs of the cluster arrival times are derived. The latter turn out to be chi-square distributions with  $2l$  degrees of freedom, where  $l$  is the cluster index. These could be thought of as a prior distributions over cluster arrival times. The HMM framework is used to learn the parameters of the distribution and classify the taps in the impulse response. By exploiting the property of the chi-square distributions the expected time of cluster arrivals can be effectively computed. This value can then used as the cluster center.

To fully exploit this approach, the stationarity restriction has been relaxed, i.e., an independent parameter for each distribution has been used. This allows a better adaptation of the clustering algorithm to the non-stationary scenarios.

The algorithm has been applied to simulated, as well as to real data and, in both cases, it has shown reasonable performance.

## 6. REFERENCES

- [1] G.L. Turin et al., "A statistical model of urban multipath propagation," *IEEE Transactions on Vehicular Technology*, vol. Vol. VT-21, pp. 1–9, February 1972.
- [2] T. Ekman, *Prediction of Mobile Radio Channels, Modeling and Design*, Ph.D. thesis, Uppsala University, Nov. 2002.
- [3] T. Eyceoz, A. Duel-Hallen, and H.; Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *Communications Letters, IEEE*, vol. Volume:2, Issue:9, pp. 254–256, September 1998.
- [4] H. Hashemi, "Impulse response modeling of indoor radio propagation channels," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 11:, pp. 967–978, September 1993.
- [5] J.B. Andersen, T.S. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *Communications Magazine, IEEE*, vol. Volume: 33 Issue: 1, January 1995.
- [6] A.A.M Saleh and R.A. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. Volume:SAC-5, no. 2, pp. 128–137, February 1987.
- [7] "Mission report – modelling unification workshop," Tech. Rep. COST 259 TD(99) 061, Vienna, 22-23 April 1999.
- [8] Lawrence R Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. Vol. 77, no. 2, pp. 257–286, February 1989.
- [9] J. Bilmes, "A gentle tutorial on the EM algorithm and its application to parameter estimation for gaussian mixture and Hidden Markov Models," Tech. Rep. ICSI-TR-97-021, University of Berkeley, 1997.
- [10] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern classification*, John Wiley & Sons, Inc., second edition, 2000.
- [11] E. Bonek et al., "Double-directional superresolution radio channel measurements," in *Proc. Conf. on Comm., Control, and Computing*, Tokyo, Japan, Oct. 2001, vol. 3.