HMM-based Modelling of Roadside-to-Vehicle WLAN Communications

Syed Faraz Hasan[§], Nazmul H. Siddique[‡], *Senior Member IEEE,* Shyam Chakraborty[‡], *Member IEEE,*
and Swades De[⋆], *Member IEEE*

***Abstract*—Roadside-to-vehicle communications has recently gained significant research attention. The idea is to exploit opportunistically encountered public WLAN APs from vehicles moving on their normal routes. Different aspects of this challenging communication scenario has been explored in several previous works, however, nothing notable has been done on mathematically modelling this kind of connection. In this article, we characterize a typical roadside-to-vehicle connection using Hidden Markov Models after showing that previously known models are not applicable for vehicular environments. Such a model can help analyze the performance of WLANs in delivering network services to mobile nodes travelling at vehicular speeds. We base our model on the data collected from the drive runs conducted in domestic and commercial areas of the city.**

***Index Terms*—802.11 WLANs, Baum-Welch Algorithm, Hidden Markov Model, Roadside-Vehicle Communication.**

I. INTRODUCTION

IEEE 802.11 wireless LANs have been conventionally used to offer high speed internet services to users with restricted mobility. These networks are inherently designed to cater for indoor applications such as in a house, a coffee shop, an airport, etc. A few recent works have extended the application domain of WLANs to vehicular communications. In one such application, the performance of WLAN access point (AP) has been studied for providing continuous network services while moving at vehicular speeds. In principle, a mobile terminal (MT) placed in a vehicle moves at normal vehicular speeds on roads, connects to an available WLAN AP and uses its resources until it gets out of the footprint of the AP. The MT remains disconnected until it finds another AP footprint, when it is able to re-establish a connection. This cycle continues till the end of the journey. Figure 1 shows the frequent loss of WLAN connectivity as the MT attempts to exploit the coverage of public-deployed WLAN APs while moving at vehicular speeds. As can be seen from the figure, an MT faces connection and disruption periods as it moves along its normal path. The vehicle spends different amounts of times inside and out of the AP footprint while moving on roads. In the figure, these durations are 'x' and 'y' seconds long respectively. This irregularity in service is mainly because of random deployment of WLAN APs across cities and towns. This intermittent nature of outdoor WLAN connection is often referred to as *disruption* [1], which must be minimized for better performance to support convergent communication services. In this article, we address the issue of mathematically modelling a MT-AP connection in vehicular mobility context. There have been some prior works on mathematical modelling of AP performance as presented in Section II, however, modelling this unique kind of connection is a new research direction and very few works on this issue are known to authors.

The modelling reported here is based on authors' ongoing *Net-on-Roads* project [2], which addresses the possibility of using this intermittent connection to get broadband services at vehicular speeds. We believe that the first step towards effectively utilizing such a network is to mathematically model it using the available techniques. While some works have been done on tolerating disruption [3] [4], we argue that there must be a mathematical interpretation of disruption so that the effectiveness of any proposed disruption tolerant scheme may be tested. To this end, this paper is an attempt toward disruption modelling using Hidden Markov Models. This paper not only provides a Hidden Markov Model for representing this disruptive behaviour, it also introduces one single mathematical parameter to reflect on the amount of disruption available in a typical Net-on-Roads trace. We collect data from test drives in the commercial and domestic areas of the city and calculate the transition probabilities, as discussed in Section V.

Rest of paper is organized as follows. Section II outlines the previous works on mathematical modelling of different communication setups. Section III highglights the model de-

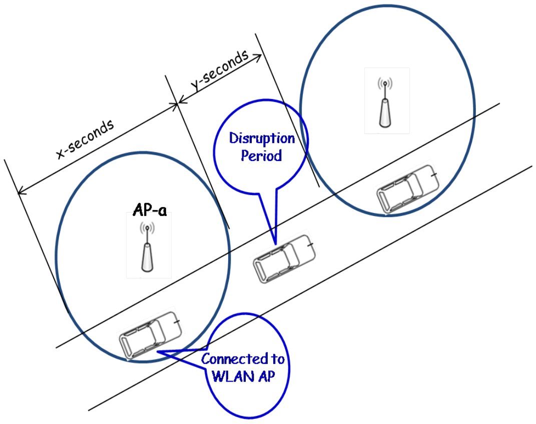

Fig. 1. An MT faces connection and disruption periods as it attempts to get continuous services from WLAN APs at vehicular speeds.

[§] PhD student at the School of Computing and Intelligent Systems, Ulster University, UK. Email: faraz_hasan-s@email.ulster.ac.uk.

[‡] School of Computing and Intelligent Systems, Ulster University, UK. Email: nh.siddique@ulster.ac.uk, ss.chakraborty@ulster.ac.uk.

[⋆] Department of Electrical Engineering, IIT Delhi. Email: swadesd@ee.iitd.ac.in.

scription while Section IV breifly describes the experiments conducted to collect the data. Section V gives the proposed model and also discusses why previous models are not applicable in case of Net-on-Roads. Conclusions are given in Section VI.

II. PREVIOUS WORKS ON MATHEMATICAL MODELLING OF COMMUNICATIONS

Some works on modelling communication processes and systems are known in literature. Here we summarize the ones that are pertinent to our current work. Frame error losses in 802.11 [5] and packet losses in 802.11g networks [6] have already been modelled using Markov Models. The work done in [5] considers 802.11a and 802.11b networks whereas the popular GE model has been modified to facilitate multimedia applications in [6] for 802.11g networks. While both these models consider the wireless channel and its characteristics, our model is based on the factors like vehicle speed, population of APs, and their supported authentication schemes etc. Our model is based on transitioning of an MT between the two states (as will be highlighted later) and applies to all WLAN APs of all radio types. Packet Loss and Frame Loss are also modelled in [7] and [8] respectively, but for indoor WLAN applications. Our model, on the other hand, is meant for depicting the performance of WLANs in outdoor environments over successive APs at high vehicular speeds.

A similar approach to the idea, in which an MT transits between *connected* and *disconnected* states, is used earlier in the Brady Speech Model [9], 6-state Brady Model, and modified Brady Model [10]. However, all these works use the relations of exponential distribution to calculate transition probabilities. An important difference between our case and theirs lies in the fact that time spent in each state for a Net-on-Roads connection is not exponentially distributed and hence exponential relations cannot be used. We give a detailed description of this fact while discussing the proposed model in Section V.

Breadcrumbs proposed in [11] and the Markov model presented in [12] are most closely related to our work. Markov model presented in [11] grows as the vehicle moves from one place to another and the transition probabilities are dynamically evaluated. The model is afterwards used to predict the connectivity certain steps ahead of time. It can, therefore, be seen in the domain of disruption tolerance i.e. tolerating the disruption periods during the drives. In our model, the transition probabilities are calculated offline and a mathematical parameter is derived therefrom to represent the amount of disruption available in a typical Net-on-Roads trace. These works are certainly inter-related but with distinct difference that one focuses on improving user experience [11] and ours focusing on evaluating disruption. Mathematical representation of a streaming proxy for vehicular networks has been given in [12] but the backhaul networks are cellular or satellite. The disruption modelled in this work occurs when the fast moving vehicle momentarily loses connection (e.g. train passing through a tunnel etc). Our work deals with WLAN networks and represents a much larger disruption. Following section gives a brief description of the proposed model.

III. MODEL DESCRIPTION

As noted previously, while attempting to use randomly deployed WLAN APs, an MT experiences periods of connectivity and disruption as it traverses along the roads and motorways. In this sense, an MT can be seen to transit between *connected* (when MT is within AP footprint) and *disconnected* (when MT is away from the AP) states in vehicular environments. Once an MT establishes an association with an AP, it is said to be in connected state because it is now connected to the external network and hence can use the broadband services. As soon as it gets out of the footprint of an AP, it enters the disconnected state because it can no longer access the external network. Therefore, we represent the entire transition phenomenon with a 2-state Markov Model in which an MT transits between *connected* and *disconnected* states with certain transition probabilities.

According to the model described so far, an MT transits from one state to another upon encounter with a WLAN AP while moving at vehicular speeds on roads. Based on the authentication scheme, APs are classified as *Open* and *Closed*. Open APs can make connections with any device with a suitable wireless card without requiring authorization. Such APs are commonly deployed in restaurants and shopping malls to attract customers by allowing free internet access. Closed APs, on the other hand, restrict unauthorized use of their network by employing authentication schemes such as WPA, WEP, etc, mainly preferred by commercial entities. Therefore, if an MT encounters a closed AP, it will be denied network services and hence cannot make a transition to the connected state. This infers that a transition to connected state not only requires an encounter with an AP, but also that the encountered AP follows *open* authentication. Since an MT encounters both open and closed APs on its way, it may still not enter connected state if the encountered AP is closed. In our model, we incorporate the information of authentication scheme of an AP as *observable symbols* hidden from the MT at the time of attempting a connection. After the introduction of the hidden states, the entire model is represented by a set of transition and observation probabilities as a Hidden Markov Model [13]. Figure 2 gives an schematic diagram of the proposed model; the transition and observation probabilities are calculated later in section V. The model shown in the figure will be modified in Section V based on the idea that the closed APs may be further classified as *WPA-personal*, *WPA2-Personal* and *WPA-*

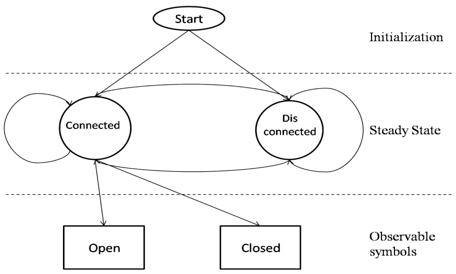

Fig. 2. Schematic of the proposed Hidden Markov Model for roadside-to-vehicle communications.

TABLE I

POPULATION OF WLAN APS IN THE AREA WHERE TESTS WERE CONDUCTED

Speed	25-30km/hr	45-50km/hr
Total APs	65	52
Open	43	32

TABLE II

STATISTICS REPRESENTING ENCOUNTER TIMES BETWEEN A MOBILE NODE AND AN AP

Speed	25-30km/hr	45-50km/hr
Mean encounter time (sec)	57.3	55.76
Median encounter time (sec)	57	57
Standard deviation (sec)	1.41	9.21

Enterprise. We comment on this when we give the final model in Section V.

IV. ON ROAD EXPERIMENTATION

In order to develop the model described in the previous section, required information is recorded using Vistumbler [14] for two test drives in a commercial area. The MT (laptop with Realtek 802.11g card with Windows Vista) placed in a car is driven at two different speeds; each representing a particular traffic pattern. By collecting data at two different speeds, we are interested in finding out any differences in the transition probabilities. Thus, we propose two models, one each for dense and normal traffic scenarios and comment on the differences between them in the next section.

We identify a commercial area as one dominated by entities such as businesses, shops, banks, petrol stations, etc. Observed AP population at both ranges of vehicle speed is given in Table I. We also observed that a significant proportion of the encountered APs was 802.11g networks, offering 54Mbps data rates. The availability of 802.11g networks alongside roads is not too relevant for this work, however, their presence is very encouraging for the overall Net-on-Roads project. As can be seen from Table I, many encountered APs follow *open* authentication i.e. they do not require user credentials to connect to the ISM bands. More precisely, an MT transits to connected state in over 60% of the total encounters in accordance with our observations for both speed ranges. We further note from Table II that encounter time statistics are not significantly affected by the changes in speed. Since our model is based on these observations, it also does not fluctuate very much with the changing vehicular speed, as shown in section V. The mean encounter time is almost same and also considerably high at both vehicle speeds. An obvious difference between the two traffic scenarios, as can be seen from Table II, is the standard deviation in the encounter times. Under the light of this observation, a typical roadside-to-vehicle connection can experience smoother connection services at slower speeds[1]. Encounter times and the frequency of encounters are the basic building blocks of the model, which are later highlighted in the next section.

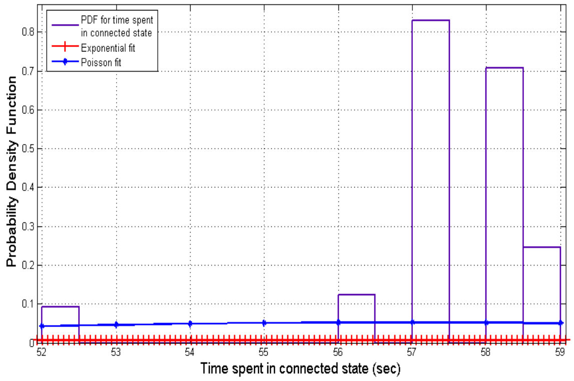

Fig. 3. PDF for data observed at 25-30km/hr vehicle speed.

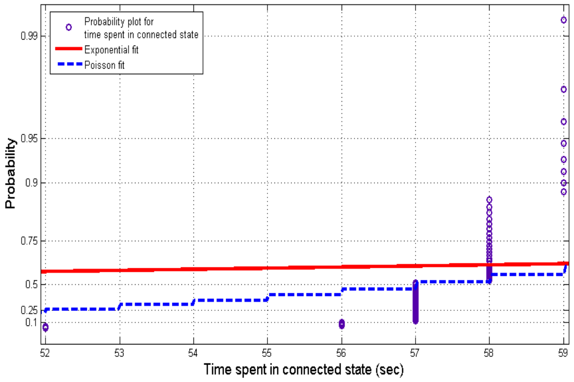

Fig. 4. Probability plot for data observed at 25-30km/hr vehicle speed.

V. PROPOSED MODEL

A. On Probability Distribution of Observed Data

In this section, we comment on the distribution of data set observed during the drive tests. We are particularly interested in evaluating the distribution of amount of time spent in the connected state by an MT in vehicular set up. Motivation of analyzing the said distribution comes from the fact that previous models [5]-[10], which represent the mobile device to transit between various states, are based on the assumption that the amount of time mobile device spends in one state is exponentially distributed. Here, we show that the data obtained from our test runs is not exponentially distributed and hence the previously known models are unusable in this case.

Figure 3 and Figure 4 show the exponential and poisson fits to probability distribution and probability plot, respectively, for the time spent in the connected state. It is clear that the

---

[1] This work is concerned with mathematical modelling of this type of connection; we therefore do not comment on the impact of vehicle speed on the user experience.

observed data deviates significantly from these distributions and hence the transition probabilities can not be calculated using the relations of exponential distribution. While these figures represent the data recorded for tests conducted at speed 25-30km/hr, similar plots can be shown for the second trace as well. Another approach in modelling this type of connection could be through Poisson process. We considered poisson distribution along with exponential distribution in Figure 3 and 4 because, intuitively, the AP hits experienced by an MT may be viewed as events and entire process may be modelled as a Poisson Counting Process. We argue, on the basis of following, that such a consideration may still be invalid.

- Poisson Processes have stationary increments, i.e. arrivals depend only on the length of the time interval [15]. In case of a typical roadside-to-vehicle connection, observing an AP hit depends on other factors such as population of APs, their presence near the roadside, vehicle speed, etc. Therefore, AP hits (arrivals in Poisson process) do not necessarily depend on the time duration.
- Events in a Poissons process do not occur in batches [16]. For the concerned connection, this property would mean that a mobile node can only observe one AP hit at one time instant. However, according to our observations, several AP hits were recorded to occur at the same time (in batches).
- If a process proceeds as Poisson process, the amount of time between successive arrivals is exponentially distributed [17]. From Figure 5 and Figure 6, showing the probability plot of inter-arrival times for observed AP hits at both vehicular speeds, it is obvious that inter-arrival times do not follow the exponential distribution.

Based on above arguments, we claim that Net-on-Roads connection neither follows an exponential distribution nor it can be represented as a Poisson process. However, a *Renewal Process*, which is another counting process, may be used to model Net-on-Roads connection. Our interest in studying Net-on-Roads as renewal process comes from the fact that, unlike Poisson process, its data set follows any arbitrary distribution and that the distribution function for Net-on-Roads data may be calculated using the Renewal eq. (1). We do not address renewal processes here but their detailed account in this context may be useful.

$$m(t) = F(t) + \int_0^t m(t-x)f(x)dx \tag{1}$$

Where *m(t)* is the mean-value function, *F(t)* is inter-arrival distribution, and *f(x)* is the density function at *x*, where *x* is the time of first renewal.

B. Use of Baum-Welch (BW) Algorithm

We have already shown that the connection we are trying to model does not follow exponential or poisson distributions. Investigation on determining the exact distribution of such a roadside-to-AP connection or modelling it as a Renewal function are recognized as future works. Instead of calculating transition probabilities using relations belonging to specific distributions, we use BW algorithm which is one of the most popular methods for calculating transition probabilities using an iterative procedure. The algorithm starts with an initial model based on parameters given in eq. (2) and eq. (3), and is re-estimated using the previous values until it converges to a certain maxima [18].

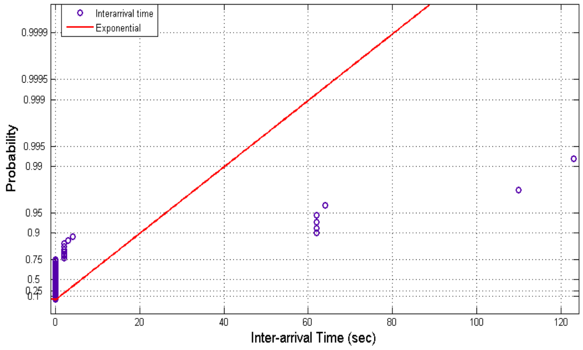

Fig. 5. Probability plot for Inter-arrival times for the drive done at 25-30km/hr.

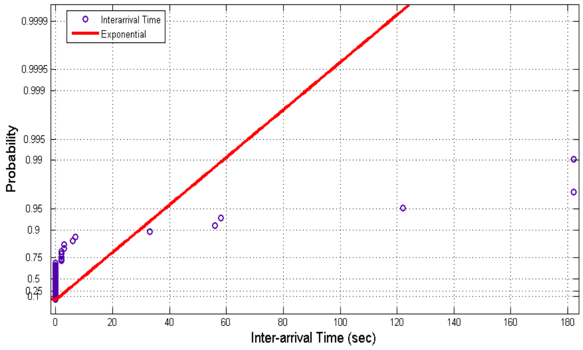

Fig. 6. Probability plot for Inter-arrival times for the drive done at 45-50km/hr.

$$a_{ij} = P(q_{n+1} = j | q_n = i) \tag{2}$$

$$b_j(k) = P(O_n = k | q_n = j) \tag{3}$$

where, a_{ij} is the probability of transiting from state i to j, $b_j(k)$ is the observation probability while in state j, q_{n+1} and q_n are the future and present states respectively and O_n represents the symbol observed while in state n.

In this article however, we focus on developing an initial model which may be used for proposing more accurate models. As mentioned in the Section III, the closed APs may follow different authentication schemes such as Wireless Personal Access (WPA). WPA itself has several varieties; for instance, commercial users would use WPA-Enterprise which uses an authentication server to authenticate clients. Domestic users may choose to use WPA-Personal to avoid complexity associated with an external server based authentication. WPA2 has also been introduced with differences that it uses separate encryption methods and allows ad hoc mode [19]. Using the state and observation information given in eqs. (4) and (5) we

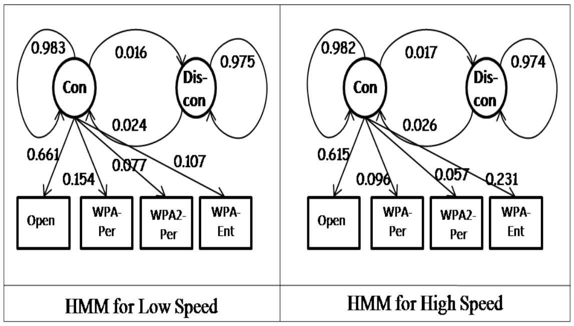

Fig. 7. Proposed HMMs for Net-on-Roads connections.

evaluate the transition and observation matrices from eqs. (2) and (3). The results are shown in Figure 7.

$$i, j \in \{connected, disconnected\} \tag{4}$$

$$k \in \{Open, WPA_{Per}, WPA2_{Per}, WPA_{Ent}\} \tag{5}$$

It is interesting to note from the transition and observation probabilities shown in Figure 7 that cross-over probabilities depicted in both models are very small. Both states in the proposed model may be treated as *recurrent states*, more precisely, *positive recurrent states,* as the probability of re-entering a state starting from the same state is sufficiently high. Also, the steady state transition probabilities for models proposed for both speed ranges are very similar. This small difference in transition probabilities at two different vehicle speeds needs to be investigated in a greater detail. Authors maintain that the difference may become considerable if the two speed ranges have a larger separation between them. The main reason for this is that the model is based on the encounter time characteristics, which are observed to be the same at both these speeds. The similarity between the models had also been highlighted previously in Section IV. From the transition probabilities of the low speed model shown in Figure 5, Long Term Error Rate (LTER) [20] is calculated to be 0.3902 using eq. (6).

$$LTER = \frac{P_{cd}}{1 - P_{dd} + P_{cd}} \tag{6}$$

where, P_{cd} is the probability of transiting from *connected* to *disconnected* state and P_{dd} is the probability of re-entering *disconnected* state starting from the *disconnected* state. LTER evaluated suggest that 39.02% of the transmissions done in this vehicular set up will be received in error. Put another way, the overall connection would be almost 40% disruptive. Authors believe that identifying such a parameter which could reflect on the overall disruption in a roadside-to-AP communications is necessary to judge the disruption tolerance algorithms. Although LTER indicates the performance of a roadside-to-vehicle communication, it does not take into account the effect of authentication scheme (hidden states of HMM). Identification of a more representative mathematical expression for the HMM can be an area of future research.

VI. CONCLUSION

This article has proposed a Hidden Markov Model based on Baum-Welch Algorithm for a typical Roadside-to-Vehicle connection. The main concept behind this modelling is representing an MT as transitioning between *connected* and *disconnected* states as it moves along its normal commute. The hidden part of the model is based on the idea that encounter with all APs cannot make MT transit to a connected state, where the information of authentication scheme is assumed embedded in the model as the observable symbols. Although the proposed model can be improved using iterative re-estimation, it gives a starting point in this area which has not been addressed before. All results are based on experimental findings and reflect an actual drive through environment. The current outcome can be further extended to a single mathematical parameter to represent the performance of an entire roadside-to-vehicle trace.

REFERENCES

[1] J. Eriksson, H. Balakrishnan and S. Madden, “Cabernet: Vehicular Content Delivery using WiFi”, *in Proc. 14th ACM Mobicom*, pp. 199-210, 2008.

[2] S. F. Hasan, N. H. Siddique and S. Chakraborty, “WLAN Data Rates achievable from roads in Low and High Mobility Environments”, Workshop on Vehicular Connectivity, *IEEE International Communications Conference*, pp. xx-xx, 2010 (accepted for publication).

[3] J. Ott and D. Kutscher, “The Drive-Thru Architecture: WLAN-based Internet Access on the Road”, *IEEE Semiannual Vehicular Technology Conference*, 2004.

[4] N. Frangiadakis, D. Kuklov and N. Roussopoulos, “Pegasus: 802.11 connectivity at high speed”, *IEEE Globecom Workshops*, pp. 1-8, 2007.

[5] J. Arauz and P. Krishnamurthy, “Markov Modelling of 802.11 Channels”, *in Proc. 58th IEEE Vehicular Tech Conf*, vol. 2, pp. 771-775, 2003.

[6] L. Carvalho, J Angeja and A. Navarro, “A New Packet Loss Model for IEEE 802.11g Wireless Network for Multimedia Communications”, *IEEE Transactions on Consumer Electronics*, vol. 51, no. 3, pp. 809-814, 2005.

[7] K. Cardoso, and J. Rezende, “Accurate Hidden Markov Modelling of Packet Loss in Indoor 802.11 Networks”, *IEEE Communications Letters*, vol. 13, no. 6, pp. 417-419, 2009.

[8] J. Hartwell and A. Fapojuwo, “Modelling and Characterization of Frame Loss Process in IEEE 802.11 Wireless Local Area Networks”, *in Proc. of 60th IEEE Vehicular Tech Conf*, vol. 6, pp. 4481-4485, 2004.

[9] D. Goodman and S. Wei, “Effeciency of Packet Reservation Multiple Access”, *IEEE Transactions on Vehicular Technology*, vol. 40, no. 1, pp. 170-176, 1991.

[10] B. Tsankov, R. Pachamanov and D. Pachamanov, “Modified Brady Voice Traffic Model for WLAN and WMAN”, *IEEE Electronics Letters*, vol. 43, no. 23, pp. 1295-1297, 2007.

[11] A. J. Nicholson and B. D. Noble, “BreadCrumbs: Forecasting Mobile Connectivity”, *MobiCom'08*, pp. 46-57, 2008.

[12] M. B. Hassan and M. Hassan, “A Markov Chain Model of Streaming Proxy for Disconnecting Vehicular Networks”, *check*.

[13] L. R. Rabiner, “A tutorial on Hidden Markov Models and selected applications in Speech Processing”, *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257-286, 1989.

[14] Vistumbler, http://www.vistumbler.net. (as on August 13, 2009).

[15] S. Ross, “Introduction to Probability Models”, 8th Ed, *Academic Press*, 2003.

[16] Y. Viniotis, “Probability and Random Processes for Electrical Engineers”, *McGraw-Hill International Edition*, 1998.

[17] R. Bernstein, “Schaum's Outline of Theory and Problems of Elements of Statistics II: Inferential Statistics”, *McGraw-Hill*, 1999.

[18] National Research Council (U.S.), “Spatial Statistics and Digital Image Analysis”, *National Academies Press*, 1991.

[19] S. M. Rahman and M. R. Syed, “Multimedia technologies: concepts, methodologies, tools, and applications”, *Idea Group Inc (IGI)*, 2008.

[20] E. Vergetis, E. Pierce, M. Blanco and R. Guerin, “Packet-level diversity - from theory to practice: an 802.11-based experimental investigation”, *Proceedings of the 12th annual international conference on Mobile Computing and Networking*, pp. 62-73, 2006.