9 771913 371006    04

# JOURNAL EDITORIAL BOARD

# TABLE OF CONTENTS

## Volume 3　　Number 4　　　　　　　　　　　　　　　　　　　　　April 2010

# International Journal of Communications, Network and System Sciences (IJCNS)

# Journal Information

## SUBSCRIPTIONS

## SERVICES

## COPYRIGHT

## PRODUCTION INFORMATION

# User Selection and Precoding Schemes Based on Partial Channel Information for Broadcast MIMO Systems

**Hamid Reza Bahrami[1], Tho Le-Ngoc[2]**
[1]*Department of ECE, University of Akron, Ohio, USA*
[2]*Department of ECE, McGill University, Montreal, Canada*
*Email*: *hrb@uakron.edu, tho.le-ngoc@mcgill.ca*

## Abstract

Due to the large number of users and the time-varying characteristics of wireless channels, it is very tough to inform the transmitter of full channel information in real multi-user MIMO broadcast systems. On the other hand, the capacity of multi-user systems greatly depends on the knowledge of the channel at the transmitter while this is not always the case in single-user MIMO systems. In this paper, we investigate combined user selection and zero-forcing precoding schemes that use partial channel information, *i.e.*, very low amount of channel information at the base station. We show that while greatly reducing the complexity and channel knowledge feedback load, the proposed schemes preserve the optimality of zero-forcing scheme in term of achievable ergodic sum capacity in limit of large number of active users.

**Keywords:** Broadcast Channel, Multi-User System, Partial Channel Information, Precoding

## 1. Introduction

Linear precoding is an effective tool to achieve better performance and higher throughput in multiple-input-multiple-output (MIMO) communication systems. MIMO precoding techniques for single-user systems have been extensively studied. It was shown that the achievable capacity of a MIMO point-to-point link is limited by the minimum number of transmit and receive antennas [1,2], and linear precoding scheme can be applied to achieve this capacity [3-5].

On the other hand, evaluation of the achievable rate and capacity of multi-user MIMO broadcast (BC) channels is still an open problem. It was the first time in [6] that the achievable region of MIMO BC channels was evaluated using an approach known as dirty paper coding (DPC) [7]. The rate region was obtained for two single-antenna users analytically. Later [8] and [9] suggested a duality between multi-access (MA) and MIMO BC channels. Using this duality, the rate region and capacity of a general MIMO BC system with arbitrary number of (multi-antenna) users can be calculated. It is shown that the DPC approach can achieve the capacity of a general BC channel [8].

The duality between precoder design problem and calculating the capacity of the system, however, does not exist in MIMO BC systems. In other words, there is no lin-

ear scheme that achieves the capacity of MIMO BC channels. Instead, one has to use the DPC approach, which is a nonlinear scheme to achieve the so-called sum-rate capacity of a multi-user BC system. DPC can be implemented using a successive interference pre-subtraction (cancellation) technique at the transmitter [10] based on the structure of generalized decision-feedback equalizer (GDFE) which is a nonlinear approach.

Although the results presented in these papers are interesting from a theoretic point of view, due to the complex nature of optimization processes used in these methods and the assumptions necessary for applying DPC approach, their application in real communication systems is limited. It is mainly because of two important factors; first, the number of users in real multi-user communication systems can be very high and second, the fast changing characteristics of the MIMO wireless channels prohibits time-consuming optimization at the transmitter.

On the other hand, the presence of channel knowledge is more important in multi-user systems [11,12]. One can say that the absence of channel knowledge at the transmit side of multi-user systems will result in higher degradation in terms of performance and capacity as compared to the single-user systems. These facts draw the attention to the use of partial channel information in multi-user communication systems. Compared to the full channel knowledge, partial channel information can be obtained easily

while it is outdated more slowly. Therefore, it is more logical to assume that the transmitter of a multi-user MIMO system has partial channel knowledge. Our emphasis is that the value of partial channel information is far crucial in multi-user as the volume of full channel knowledge in a multi-user system is much larger than that in single-user systems. For example, [12] shows the effect of different levels of channel information over achievable rate in a MIMO broadcast system.

While considering the use of partial channel information in MIMO BC channels of great importance, there are few papers that address this issue in the design of practical transmission schemes. Recently, [13] and [14] suggested a random beamforming scheme based on partial channel information with a rate that can achieve the same scaling factor as the capacity obtained with perfect CSI using dirty paper coding, *i.e.*,

$$\lim_{n \to \infty} \frac{R}{C_{sum}} = \lim_{n \to \infty} \frac{R}{R_{DPC}} = 1 \tag{1}$$

where $n$ is the number of the users, $R$ is the achievable rate of the proposed scheme and $C_{sum}$ is the sum capacity of the system assuming full channel knowledge at the transmitter which is the same as the rate achieved by DPC ($R_{DPC}$).

Our main focus in this paper is to attain the capacity of a multi-user MIMO broadcast channel assuming that the transmitter just has a partial knowledge of users' channels. In other words, we investigate downlink precoding schemes that is capable of achieving the capacity of a MIMO broadcast channel in which a multiple-antenna transmitter communicates with a number of mobile units. This approach is based on the assumption of partial channel side information available at the transmitter while each receiver has perfect channel knowledge.

In [15], using full channel knowledge at the transmit side, a zero-forcing precoding scheme has been introduced and its performance has been evaluated. It has also been proved that when the number of users tends to infinity, the zero-forcing scheme is optimal in term of capacity, *i.e.*, it can achieve the ergodic sum capacity of BC channels when the number of users is large. The problem with this scheme lies, however, in the need for full channel knowledge of the channel at the transmitter which is difficult to obtain in fast-varying wireless channels. Another difficulty with the scheme proposed in [15] is the complexity of the transmit operations. The transmitter has to find a set of best channels in each transmission interval which have the most-orthogonal channel vectors. The authors have recently generalized this scheme to the case of partial channel information by using a quantized feedback scheme in [16]. Note that, the optimality of the above approaches in [13] and [15] is asymptotical, *i.e.* the optimality decreases when the number of users reduces. The problem of zero-forcing in differ-

ent interference scenarios and SNR regimes has been also addressed in [17].

Our study shows that careful selection of channel side information is very important in the sense that it can reduce the feedback cost and transmit complexity while still provides a considerable performance. We propose zero-forcing transmission schemes that use only partial channel information with low feedback load and also facilitate the algorithm of selecting the best users at the transmitter. Although it reduces the feedback load and algorithm complexity, it can be shown that its performance can be comparable to that of the scheme proposed in [15] in term of achievable rate.

Therefore, four important issues in transmission in multi-user systems are addressed in our proposed scheme:

1) Our scheme significantly reduces the feedback load needed to calculate the optimal precoding vectors at the transmitter. The difference in feedback overhead compared to other schemes grows by increasing the number of users and/or transmits antennas.

2) It greatly relaxes the complexity of selecting the best user algorithm at the transmitter by distributing the load of processing among the users communicating in the network. The complexity issue becomes more important when the number of users and/or transmit antennas grows. Therefore, this scheme can be considered as a candidate for distributed network management schemes.

3) Our scheme achieves the same ergodic sum capacity as the scheme in [15] (and hence DPC) and the optimality preserves better especially in the systems with large number of users. However, one important issue in [15] is the fairness of user selection. The original selection scheme in [15] is not fair as the users with high channel gains (users closer to the base station) always have a better chance to be selected. Therefore, a scheduling procedure becomes inevitable to satisfy fairness requirements. We also discuss the issue of fairness for different schemes. We show that using these schemes, the chance of selecting a user does not depend on the channel gain (power) and the distance to the base station and hence all the users have equal chance to be selected as the best (winning) users.

4) We propose four different strategies for user selection, power allocation and precoding. Each of these strategies is suitable for specific propagation scenario and channel condition. However, as the structure and algorithms of all of these schemes are very similar, it is possible to implement all of them at the base station and switch amongst them when necessary. This gives a degree of robustness to the system that can cope with channel impairments and changes.

## 2. System Model

We consider a broadcast system using a transmitter with

$M$ transmit antennas to serve $n$ users, each with $N_i$ antennas ($i = 1, \ldots, n$). The channel for user $i$ can be represented by an $N_i \times M$ matrix $\mathbf{H}_i$ whose entries are assumed to be Gaussian with zero mean and unit variance. At time instant $k$, the $M \times 1$ transmit vector $\mathbf{s}$ to different users has a constrained power, *i.e.*, $E\{\mathbf{ss}^H\} \leq P$ and the $N_i \times 1$ received signal vector for user $i$ is

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{s} + \mathbf{n}_i \qquad i = 1, 2, \ldots, n \qquad (2)$$

where $\mathbf{n}_i$ is the $N_i \times 1$ additive white Gaussian noise sample vector for user $i$. Using precoding, the transmitted vector $\mathbf{s}$ is a linear combination of transmitted symbols for each user $\mathbf{s}_i$,

$$\mathbf{s} = \sum_{i=1}^{n} \sqrt{P_i} \mathbf{W}_i \mathbf{s}_i, \quad \sum_{i=1}^{n} P_i = P$$

where $\mathbf{W}_i$, $\mathbf{s}_i$ and $P_i$ are the $M \times N_i$ precoding matrix, transmitted symbol, and allocated power for user $i$, respectively. For simplicity, in the following discussions, we assume $N_i = 1$, $i = 1, \ldots, n$. The received signal for the $i$th user can be represented as

$$y_i = \sqrt{P_i} \mathbf{h}_i \mathbf{w}_i s_i + \sum_{j \neq i} \sqrt{P_j} \mathbf{h}_i \mathbf{w}_j s_j + n_i \qquad i = 1, 2, \ldots, n$$
$$(3)$$

where the second term is due to the interference from other users. The bold-faced, lower-case symbols in the equation denote vectors corresponding to the case of $N_i = 1$. Our objective is to find the set of precoding vectors to maximize the achievable sum rate, *i.e.*,

$$R = \max_{\mathbf{w}_i, \sum \|\mathbf{w}_i\|^2 \leq P} \sum_{i=1}^{n} \log\left(1 + \frac{P_i |\mathbf{h}_i \mathbf{w}_i|^2}{1 + \sum_{j=1 \neq i}^{n} P_j |\mathbf{h}_i \mathbf{w}_j|^2}\right) \qquad (4)$$

## 3. BC Capacity and Zero-Forcing Precoding

The sum rate capacity of a BC channel achieved by DPC approach has been shown to be [6,8]:

$$C_{sum} = R_{DPC} = \max_{\sigma_i, \sum_{i=1}^{n} \sigma_i \leq P} \log \det\left(\mathbf{I} + \sum_{i=1}^{n} \sigma_i \mathbf{h}_i^H \mathbf{h}_i\right) \qquad (5)$$

which is in fact the capacity of the dual MAC channel.

Note that $\sigma_i$'s are not the same as $P_i$'s in (3). In [13], it has been shown that when the number of users ($n$) is large, the ergodic capacity asymptotically scales like $M \log \log n$. In other words, the ergodic capacity can be written as:

---

$$E\{C_{sum}\} = E\{R_{DPC}\} \underset{n \to \infty}{\sim} M \log\left(1 + \frac{P}{M} \log n\right) \qquad (6)$$

*i.e.*, in the limits, it can provide a diversity order of $\log n$ compare to a single user system. Also, a linear increase in the capacity is achieved by increasing the number of transmit antenna ($M$). It can be seen from (6) that a scheme allocating an average power of ($P/M$) $\log n$ to each of $M$ best users via $M$ independent paths (subchannels), can provide the same ergodic capacity as the DPC and hence, is asymptotically optimal.

On the other hand, although suboptimum, zero-forcing scheme can provide $M$ parallel (independent) subchannels from transmitter to $M$ users by selecting the precoding vectors, $\mathbf{w}_i$'s, such that $\mathbf{h}_i \mathbf{w}_j = 0$ ($i \neq j$) (*i.e.*, no interference from other users). Since the size of $\mathbf{h}_i$'s is $1 \times M$, there will be at most $M$ precoding vectors that can satisfy the above equations. Therefore, at most $M$ users should be selected among $n$ available users. Let $S \subset \{1, \ldots, n\} |S| \leq M$ be a subset of user selected for transmission. By suitable selection of precoding matrices[1], there will be $M$ independent subchannels and hence the achievable rate can be written as:

$$R_{ZF}(S) = \max_{P_i, \sum_{i \in S} P_i \leq P} \sum_{i \in S} \log(1 + \lambda_i P_i) \qquad (7)$$

where $\lambda_i$ is the $i$th subchannel gain. Furthermore, $P_i$'s can be found via a waterpouring process as $P_i = [v - \lambda_i^{-1}]^+$ where $[x]^+ = \max\{0, x\}$ for a scalar $x$ and the constant $v$ is such that the power constraint $P$ is satisfied.

Note that there are different possible selections of user subsets. Achievable sum rate of optimum zero-forcing is defined as [15]:

$$R_{ZF} = \max_{S} R_{ZF}(S) \qquad (8)$$

where $S \subset \{1, \ldots, n\} |S| \leq M$. The optimal solution may need a lengthy and complicated exhaustive search of all possible user subsets for zero-forcing schemes. Furthermore, required knowledge of instantaneous channel vectors (full channel information) at the transmitter may need a large amount of channel feedback load. This limits the application of optimal zero-forcing method to simple cases with small number of users. In [15], it has been shown indirectly that in the limit of large number of users ($n$), the zero-forcing beamforming scheme can provide a sum rate equal to that of DPC in (6),

$$E\{R_{ZF}\} \sim M \log\left(1 + \frac{P}{M} \log n\right) \sim E\{R_{DPC}\} \qquad (9)$$

An appealing sub-optimum zero-forcing scheme with simple selection of $M$ best users was discussed in [5]. However, it still needs full channel knowledge at the transmitter with potentially high complexity and feedback

---

[1]The easiest selection of $\mathbf{w}_i$ is such that $\mathbf{W}(S)$ is the pseudo-inverse of $\mathbf{H}(S)$ where $\mathbf{W}(S)$ is a matrix whose columns are composed of the precoding vectors in $S$ and $\mathbf{H}(S)$ is the channel matrix whose columns are composed of channel vectors of the users in $S$, *i.e.* $\mathbf{W}(S) = \mathbf{H}^\dagger(S)$.

load in real wireless applications.

By increasing the number of users in the network, it is more likely that there exists a subset of $M$ users such that a linear precoding scheme (like zero-forcing) can achieve the sum rate capacity of the system. In other words, for a large user population, different suboptimum zero-forcing algorithms have similar asymptotic performance, and hence, complexity becomes the key selection factor. This paper mainly focuses on how to reduce the processing and feedback load in selecting the best $M$ users with as much as possible orthogonal channels to maintain the asymptotic gain of log log $n$ over ergodic capacity in (9).

## 4. Partial Knowledge Zero-Forcing Schemes

Assume that transmitter and all users have a predetermined known orthonormal basis[2] $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_M)$ of size $M$ (e.g., standard basis for $\mathbb{C}^M$). We will discuss the selection of this orthonormal basis in Section 6.

Now, we consider two different strategies: *power* user selection (PUS) and *normalized* user selection (NUS). In the first strategy, at the start of each transmission period, each user calculates the projections of its channel vector on each of the vectors in orthonormal basis ($\mathbf{u}_i$), *i.e.*, for $j^{th}$ user, the projection on $i$th vector is:

$$\gamma_{ij} = \left\langle \mathbf{h}_j, \mathbf{u}_i \right\rangle^2 = \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 \ ; \ i = 1, ..., M; J = 1...n \quad (10)$$

In the second strategy, each user calculates the norm of the projections of its channel vector on each of the vectors in orthonormal basis, *i.e.*,

$$\gamma_{ij} = \left| \mathbf{h}_j \right|^{-2} \left\langle \mathbf{h}_j, \mathbf{u}_i \right\rangle^2 = \left| \mathbf{h}_j \right|^{-2} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 \ ; \ i = 1, ..., M; J = 1...n \quad (11)$$

Next, in both strategies, each user sends its maximum $\gamma_{ij}$ along with its index to the transmitter,

$$\gamma_j = \max_i \gamma_{ij} \ ; \ \alpha_j = \arg \max_i \gamma_{ij} \quad (12)$$

Transmitter then easily selects the best user for each orthonormal basis by finding the maximum $\gamma_j$ over those users. Assume that the indices of users for which $\alpha_j = i$ are saved in a set $S(i)$. Therefore,

$$\gamma_{i,\max} = \max_{j \in S(i)} \gamma_j \quad i = 1...M \quad (13)$$

The transmitter selects these $M$ users as the winning users and asks them to send back their channel vectors (totally $M$ vectors of size $M \times 1$). It then selects the precoding vectors. For selection of precoding vectors, we consider two different precoding schemes: opportunistic precoding (OP) and *channel-aware* precoding (CAP). In *opportun-*

*istic* precoding, as its name implies, the transmitter does not know the channel vector of selected users and therefore transmit in the direction of orthonormal basis, *i.e.*

$$\mathbf{w}_i = \mathbf{u}_i \ ; \ i = 1...M \quad (14)$$

where $\mathbf{w}_i$'s are the $M$ precoding vectors associated to each of the selected users. Sending users' signals in the direction of orthonormal basis clearly minimizes the interference. Therefore, opportunistic precoding can also be referred as *interference-minimizer* precoding. Maximum sum-rate of the users can then be written as:

$$R_{OP} = \sum_{i=1}^{M} (1 + \frac{P}{M} \frac{\left| \mathbf{h}_i \mathbf{u}_i^H \right|^2}{1 + \sum\limits_{j=1 \neq i}^{M} \left| \mathbf{h}_i \mathbf{u}_j^H \right|^2}) \quad (15)$$

In the limit of large number of users, the channel vectors of $M$ selected users are in the direction of orthonormal basis with high probability. Therefore, the sum rate can be rewritten as:

$$R_{OP} = \sum_{i=1}^{M} \log(1 + \frac{P}{M} \left| \mathbf{h}_i \mathbf{u}_i^H \right|^2) = \sum_{i=1}^{M} \log(1 + \frac{P}{M} \left| \mathbf{h}_i \right|^2) \quad (16)$$

(16) is the direct result of (15) and the fact that, in opportunistic precoding, at the limit of large number of users there is no interference between selected users. In channel-aware precoding, transmitter does know the selected users' channel vectors and selects the precoding vectors as the normalized vector of users' channel vectors:

$$\mathbf{w}_i = \left| \mathbf{h}_i \right|^{-1} \mathbf{h}_i \ \ \forall i : 1...M \quad (17)$$

where the best $M$ users in (15) are indexed from 1 to $M$. In this scheme, we target maximizing receive SNR. It is easy to see that by sending in the direction of selected users' channel vectors, we maximize the receive SNR of each user regardless of interference introducing to other users. Therefore, channel-aware precoding can also be called *SNR-maximizer* precoding. Again, the idea behind this scheme is that each channel vector of $M$ selected users is almost in the direction of one of the basis vectors; hence their channel vectors are near-orthogonal to each other. This orthogonality increases with increased number of users in the network. In the best case when the selected channels compose an orthogonal set, the following rate is achievable:

$$R_{CAP} = \max_{P_i} \sum_{i=1}^{M} \log(1 + P_i \left| \mathbf{h}_i \right|^2) \ \text{s.t.} \ \sum_i P_i \leq P \quad (18)$$

[2]A set $\mathbf{U}$ is orthonormal if $\forall i, j \ u_i^H u_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

The set of $M$ power loading values $P_i$'s are selected based on waterpouring as $P_i = [v - |\mathbf{h}_i|^{-2}]^+$. Obviously, other users have zero power.

We can therefore, -distinguish between four different user selection and precoding strategies: PUS-OP, PUS-CAP, NUS-OP, NUS-CAP. The features of each of these schemes are summarized in **Table 1**. Regarding fairness, as PUS selects users based on the maximization of projection of their channel vectors over orthonormal bases, it is more likely that users with strong power (*i.e.* users near base station) are finally selected. Therefore, the issue of fairness is not considered in PUS while it is not the case in NUS scheme. We will elaborate the performance and specifications of each scheme more in Section 7 when we present numerical results.

## 5. Asymptotic Performance of Zero-Forcing Schemes

In this section, we investigate the asymptotic performance of NUS user selection schemes. Since the analysis of PUS scheme is also very similar to CAP, for the sake of brevity we do not discuss it here.

For any two arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$, we define the following orthogonality measure:

$$\varepsilon(\mathbf{a},\mathbf{b}) \triangleq 1 - |\mathbf{a}^H\mathbf{b}|^2 |\mathbf{a}|^{-2} |\mathbf{b}|^{-2} \qquad (19)$$

Clearly, as $\mathbf{a}$ tends to $\mathbf{b}$, $\varepsilon(\mathbf{a},\mathbf{b})$ tends to zero. For the best user channel $\mathbf{h}_i$ in the direction of basis vector $\mathbf{u}_i$, we have:

$$\varepsilon_i = \varepsilon(\mathbf{h}_i,\mathbf{u}_i) = \max_j \varepsilon(\mathbf{h}_j,\mathbf{u}_i) = 1 - \max_j |\mathbf{h}_j\mathbf{u}_i^H|^2 |\mathbf{h}_j|^{-2} \qquad (20)$$

*Lemma* 1: For large number of users, the cumulative distribution function (cdf) function of $\varepsilon_i$ $(i:1..M)$ is linearly increased with $n$, *i.e.*, $\forall i: F_{\varepsilon_i}(x) = \Pr(\varepsilon_i < x) \sim O(nx)$.

*Proof:* Since $\mathbf{U}$ is an orthonormal set, the random vari-

ables $|\mathbf{h}_j\mathbf{u}_i^H|^2$ are i.i.d. over $i$ and $j$ with $\chi^2(2)$ -distribution. As $|\mathbf{h}_j|^2$ is the sum of square of $M$ Gaussian random variables, it is $\chi^2(2M)$ -distributed. It follows that the probability density function (pdf) and cdf of $t_{ij} = |\mathbf{h}_j\mathbf{u}_i^H|^2 |\mathbf{h}_j|^{-2}$ are, respectively,

$$f_t(x) = \int_0^\infty y e^{-yx} \frac{y^{M-1}e^{-y}}{(M-1)!} dy$$
$$= M(1+x)^{-(M+1)} \qquad (21)$$

and

$$F_t(x) = 1 - (1+x)^{-M} \qquad (22)$$

Hence, the cdf of $\varepsilon_i$ is

$$F_{\varepsilon_i}(x) = P(\varepsilon_i < x) = P(\max_j t_{ij} > 1 - x)$$
$$= 1 - F_t^n(1-x) = 1 - \left[1 - (2-x)^{-M}\right]^n \qquad (23)$$

For small $x$, the right hand side of (22) can be approximated as

$$F_{\varepsilon_i}(x) \approx 1 - \left[1 - 2^{-M}(1 + 2^{1-M}Mx)\right]^n$$
$$\approx 2^{1-2M}Mnx \sim O(nx) \qquad (24)$$

for any $\varepsilon_i$ $(i:1..M)$.

*Lemma* 1 indicates that, the probability of $\varepsilon_i$ smaller than a specific small value, increases linearly with the user population, $n$, and the best users' channel vectors become more and more orthogonal to the basis directions (vectors). One can also say that by increasing the number of users, it is more probable that there would be $M$ users with each channel vector very close to a basis. The following Lemma sheds some light on this fact.

*Lemma* 2: For large number of users, the probability that a user has the largest channel vector projection among all

**Table 1. Features of user selection precoding strategies.**

| Scheme | channel knowledge of all users at TX | channel knowledge of selected users at TX | Power allocation and waterpouring | Fairness | Interference between users |
|--------|--------|--------|--------|--------|--------|
| PUS-OP | No | No | No | No | No |
| PUS-CAP | No | Yes | Yes | No | Yes |
| NUS-OP | No | No | No | Yes | No |
| NUS-CAP | No | Yes | Yes | Yes | Yes |

users on two or more directions is linearly proportional to $1/n$.

*Proof*: From Lemma 1, the squared projections, $\left|\mathbf{h}_j\mathbf{u}_i^H\right|^2$, are i.i.d $\chi^2(2)$-distributed variables over $i$ and $j$. Among them, let $Y_n$ and $X_{M-1}$ denote the *largest* of the whole population of $n$ users, and the *second largest* of the selected set of $M$ users, respectively. A user has the best channel for two basis vectors if its *second largest* squared projection is larger than the squared projections of all other users for this specific basis vector. The probability of this event is:

$$P(X_{M-1} > Y_n)$$
$$= 1 - \int_0^\infty (1 - F_X(x)) F_X^{M-1}(x) n f_X(x) F_X^{n-1}(x) dx \quad (25)$$

where $F_X(x) = 1 - e^{-x}$ and $f_X(x) = e^{-x}$ are cdf and pdf of a $\chi^2(2)$ random variable. Therefore,

$$P(X_{M-1} > Y_n)$$
$$= 1 - \int_0^\infty n(1 - e^{-x})^{n+M-2} e^{-2x} dx$$
$$\approx \left[1 + \frac{n}{M}\right]^{-1} \quad (26)$$

The right hand side of (23), behaves like $O(1/n)$ when $n$ is larger than $M$.

The following Lemma guarantees that the proposed NUS-CAP method performs always better than opportunistic zero-forcing precoding (OP) in which $M$ orthonormal vectors are selected randomly at the transmitter and the best users are selected for those random vectors. Transmitter then sends the information to the best $M$ users on those random orthonormal vectors.

*Lemma* 3: The proposed NUS-CAP zero-forcing method always performs better than NUS-OP zero-forcing precoding.

*Proof*: Consider $\varepsilon_i$ $(i:1..M)$ in (20). Assuming equal power allocation, from (16) the average sum rate for the defined NUS-OP zero-forcing precoding can be written as:

$$R_{OP} = E\{\sum_{i=1}^M (1 + \frac{P}{M} \frac{\left|\mathbf{h}_i\mathbf{u}_i^H\right|^2}{1 + \sum_{j=1\neq i}^M \left|\mathbf{h}_i\mathbf{u}_j^H\right|^2})\}$$
$$= E\{\sum_{i=1}^M (1 + \frac{P}{M} \frac{1 - \varepsilon_i}{1 + \varepsilon_i})\} \quad (27)$$

The last equality results from the fact that $\mathbf{u}_i$'s are orthonormal. Our proposed NUS-CAP scheme, however, is

based on the distance of the best users' channel vectors, *i.e.* $\left|\mathbf{h}_j\mathbf{u}_i^H\right|$ $(i = 1...M; j = 1...M)$ and can be found in terms of $\varepsilon_i$ $(i = 1...M)$. It is easy to check that on average:

$$\left|\mathbf{h}_i\mathbf{h}_j\right|^2 \leq M^{-1}\varepsilon_i(1 - \varepsilon_i) \quad i = 1...M \quad (28)$$

Therefore, considering equal power allocation (no waterpouring), the average sum rate of NUS-CAP scheme is:

$$R_{CAP} = E\{\sum_{i=1}^M (1 + \frac{P}{M} \frac{\left|\mathbf{h}_i\mathbf{h}_i^H\right|^2}{1 + \sum_{j=1\neq i}^M \left|\mathbf{h}_j\mathbf{h}_i^H\right|^2})\}$$
$$= E\{\sum_{i=1}^M (1 + \frac{P}{M} \frac{1}{1 + \varepsilon_i(1 - \varepsilon_i)})\} \quad (29)$$

From (27) and (29), as $1 - \varepsilon_i / 1 + \varepsilon_i < [1 + \varepsilon_i(1 - \varepsilon_i)]^{-1}$, the average sum rate provided by the NUS-CAP scheme is always greater than that of NUS-OP zero-forcing.

Note that in the proof of Lemma 3, we considered equal power allocation on basis vectors, while if we apply the waterpouring in (18), NUS-CAP scheme performs much better than random zero-forcing scheme. Now, we prove that our scheme is optimum for large number of users.

*Theorem* 1: In the limit of large number of users, the NUS-CAP partial knowledge zero-forcing precoding approach proposed in Section 4 can achieve an average sum rate equal to that of DPC strategy in (6), *i.e.*,

$$E\{R_P\} \sim M\log(1 + \frac{P}{M}\log n) \quad (30)$$

*Proof*: Based on Lemma 3, it is sufficient to show that the discussed NUS-OP zero-forcing precoding scheme can achieve the rate in (30). As the proposed NUS-CAP precoding scheme always outperforms NUS-OP zero-forcing scheme, it turns out that it is capable of achieving the rate in (30) in the limit of large $n$.

Based on the results from extreme value theory in [10],

$$\max_i t_{ij} = \max_i \left|\mathbf{h}_j\mathbf{u}_i^H\right|^2 \left|\mathbf{h}_j\right|^{-2}$$

behaves like $\log n + O(\log \log n)$. Therefore, from (19) and (26), the average sum rate of the NUS-OP zero-forcing precoding for large $n$

$$\lim_{n\to\infty} R_{\text{rand}}$$
$$= \lim_{n\to\infty} E\{\sum_{i=1}^M (1 + \frac{P}{M}(\frac{1}{\varepsilon_i} - 1)\} \sim M\log\frac{P}{M}(1 + \log n) \quad (31)$$

As the NUS-CAP scheme always outperforms NUS-OP zero-forcing precoding, from Lemma 3, it can achieve the sum rate in (31).

Note that in the proof of Theorem 1, we assumed that the available power is distributed equally among $M$ precoder vectors (sometimes called directions). In practice, as mentioned in Section 4, we apply a standard water-pouring to calculate the optimum power allocation. However, as the proof shows, using equal powers on all precoder vectors is also asymptotically optimal (when the number of users is large).

# 6. Feedback Load, Complexity and Design Issues

In this section, we briefly discuss the practical considerations and design issues for our proposed schemes. We show that the proposed schemes can be applied with very low amount of feedback load and transmitter complexity as compared to that of full channel knowledge schemes. At the end of this section we briefly point out the selection of basis vectors set $\mathbf{U}$.

## 6.1. Feedback Load

Consider a BC system using a transmitter equipped with $M$ antennas to serve $n$ single-antenna active users. For schemes based on full channel knowledge such as DPC, optimal zero-forcing described in Section 3, and the zero-forcing scheme proposed in [5], at any transmission period, $2Mn$ real values should be fed back to the transmitter. On the other hand, the proposed CAP sch- eme based on partial knowledge scheme initially needs only $n$ integer indices of the best basis vectors and $n$ real best projections of the users' channel vectors and only $2M^2$ for the best $M$ users channel vectors after initialization. Therefore, in total, it needs $n + 2M^2$ real and $n$ inte- ger values. Clearly, the proposed scheme requires much lower feedback load, especially at large number of users.

It is also possible to exclude bad users (those are not near-orthogonal to any of the basis vectors) from channel information feedback. Similar to the idea in [10], we can define a threshold value and if a user has its maximum channel vector projection below this threshold, it is considered as a *bad* user.

*Lemma* 4: At the limit of large number of users, for each of the basis vectors, there exists at least one user $j$ with its channel vector $\mathbf{h}_j$ satisfying

$$\max_{i=1..M} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 > \log n \cdot$$

*Proof*: Based on Lemma 2, for large number of users, $n$, it is not likely that a particular user can be the best one for two or more basis vectors. Hence, maximization over a set of users that have the maximum projection on a specific basis vector is equivalent to maximization over all users for that particular basis vector, *i.e.*,

$$\max_{j=1..n} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 = \max_{j\in S(i)} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 \quad i=1\ldots M \qquad (32)$$

As $\left| \mathbf{h}_j \mathbf{u}_i^H \right|^2$ are $\chi^2(2)$ -distributed random variables, by applying the extreme value theory [13], for any $i$, $\max_{j\in S(i)} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2$ and its average behave like $\log n$ at large $n$. Therefore, there is always at least one user with channel satisfying

$$\max_{j\in S(i)} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 > \log n = \lambda_{\mathrm{tr}} \quad i=1\ldots M \qquad (33)$$

for all basis vectors.

Lemma 4 indicates that a good threshold value is $\lambda_{\mathrm{tr}} = \log n$. Using this threshold, the feedback load can be expressed as $n\Pr\{\lambda_i > \lambda_{\mathrm{tr}}\}$ ($i:1..M$) real and integer numbers. Note that, in practice $\Pr\{\lambda_i > \lambda_{\mathrm{tr}}\}$ is much smaller than 1 and therefore, inserting this threshold value decrease feedback overhead dramatically.

## 6.2. Complexity

The proposed NUS-CAP scheme which is the most complex amongst four schemes just needs $n$ comparisons at the transmitter and $M$ projections and $M$ comparisons at each user while DPC, optimal and suboptimal zero-forcing schemes are much more complex. For example, the zero-forcing algorithm proposed in [15], needs almost $\sum_{i=1}^{M}(n-1)(i+1) > nM(M+1)/2$ projections and $nM$ comparisons at the transmitter to find a near-ortho-gonal channel vector set. In other words, the proposed NUS-CAP method has a much lower complexity than the other optimal and sub-optimal schemes, and also distrib-utes low processing among users in the system.

## 6.3. Selection of Basis

As mentioned at the beginning of Section 4, the set of orthonormal basis $\mathbf{U}$, can be generally any arbitrary orth-onormal basis in $\mathbb{C}^M$ and as shown in Section 5, the pr-oposed schemes can still work fine. However, orthonor-mal basis can be optimized by a very minor increase in the feedback load from active users to the transmitter.

Let assume that instead of just one predetermined or-thonormal basis, we have a number of $k$ bases, $\mathbf{U}^1$, $\mathbf{U}^2$, …, $\mathbf{U}^k$. Both the transmitter and users know these bases. Each user calculates its channel vector projections onto each of these bases vectors and reports the best value along with the indices of this basis and its own basis. This requires one real and *two* integer numbers (instead of one real and one integer). From the reported informa-tion, the transmitter selects the best orthonormal basis, corresponding to the best $M$ users with the best orthogo-nality measures. As an illustrative example, consider $M = 2$ and two orthonormal bases to select. If we select the

first one as standard basis for $\mathbb{C}^2$, *i.e.*, $\mathbf{U}^1 = \{(1,0),(0,1)\}$, the other one should be the one whose vectors are as much as non-orthogonal to the standard basis, *i.e.*, $\mathbf{U}^2 = \{1/\sqrt{2}\,(1,1), 1/\sqrt{2}\,(1,-1)\}$. Similarly, if we want to select three orthonormal bases, a reasonable choice is $\mathbf{U}^1 = \{(1,0),(0,1)\}$, $\mathbf{U}^2 = \{(a,b),(b,-a)\}$ and $\mathbf{U}^3 = \{(b,a),(a,-b)\}$ where $a = 0.5, b = \sqrt{3/4}$.

The same approach can be applied for larger *M* and larger number of orthonormal bases. Increasing the number of orthonormal bases can greatly improve the performance of the proposed schemes at the cost of a slight increase in the feedback overhead. More precisely, using *k* different orthonormal bases, the probability in (24) is linearly increasing with almost *nk* (instead of *n*) and the probability in (26) is linearly proportional to $1/(nk)$ (instead of $1/n$). It means faster orthogonality rate by increasing the number of users and hence, the proposed schemes become optimal at lower number of users.
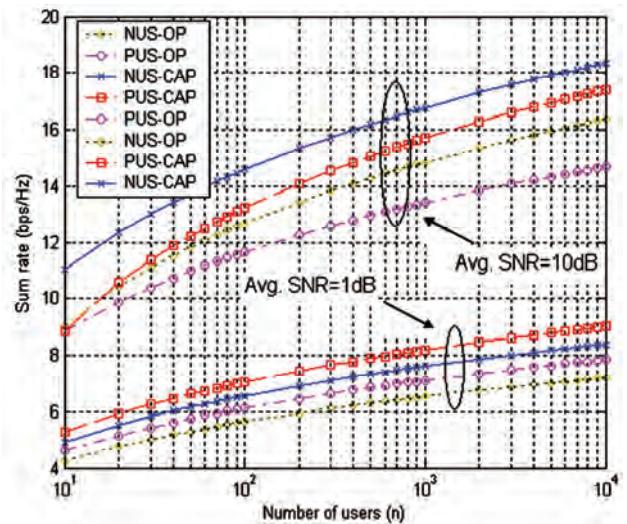
## 7. Numerical Results

In this section we examine the ability of zero-forcing precoder schemes in term of ergodic capacity by means of simulation. We also compare the complexity and feedback overhead of the proposed precoding schemes to that of DPC and full knowledge based zero-forcing precodings.
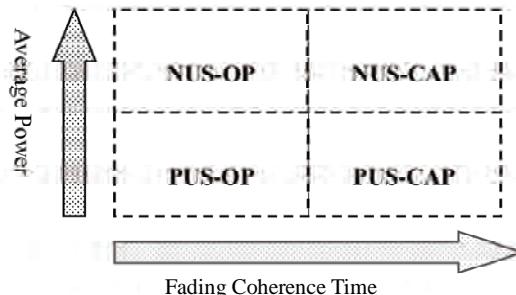
We consider a four antenna transmitter ($M = 4$), for all simulations. Besides, for our schemes we just consider one basis vector set. First, let compare four different schemes in two different scenarios, *i.e.* in low and high average users' channel power. **Figure 1** shows the sum rates of different user selection and precoding schemes in the low and high SNR regimes. By SNR regimes we mean the channel SNRs averaged over all available users' channels. As shown, while PUS selection strategy works better at low SNRs, NUS strategy outperforms it at high SNR regime. Therefore, one can conclude that when the average users' SNR is high, it is more reasonable to use NUS selection scheme while in low average SNR, PUS is more beneficial. Also, consider that at very fast fading

environment, it would be more desirable to use OP rather than CAP schemes. Based on this discussion, the diagram in **Figure 2**, gives us an idea how to select between different schemes. Note that because of similarity of structure and algorithms, it is possible to switch between them in a real time system. In other words, when fading is very fast the system uses OP schemes and in slow fading environment switches into CAP schemes. On the other hand, in high quality channel, it works in NUS regime while in poor-conditioned channels it switches into PUS regimes where users' channel power becomes a very important and significant factor.

Next, we simulate the feedback load, complexity and performance of our scheme and compare it with that of optimum and suboptimum schemes. In **Figure 3**, the feedback load of the schemes for different numbers of users (*n*) is illustrated. As shown, the CAP scheme needs much lower amount of feedback than the full channel knowledge based scheme and this difference in feedback overhead increases with the number of users in the system. At large *n*, the CAP scheme needs the same amount of channel feedback the scheme in [15].
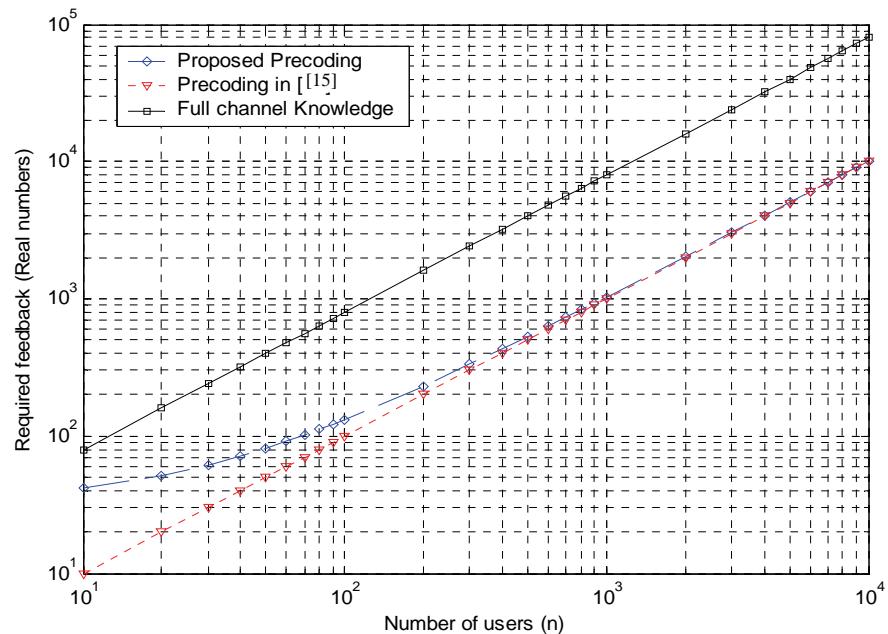


**Figure 1. Comparsion of sum rates of different precoder schemes in different SNR regimes.**
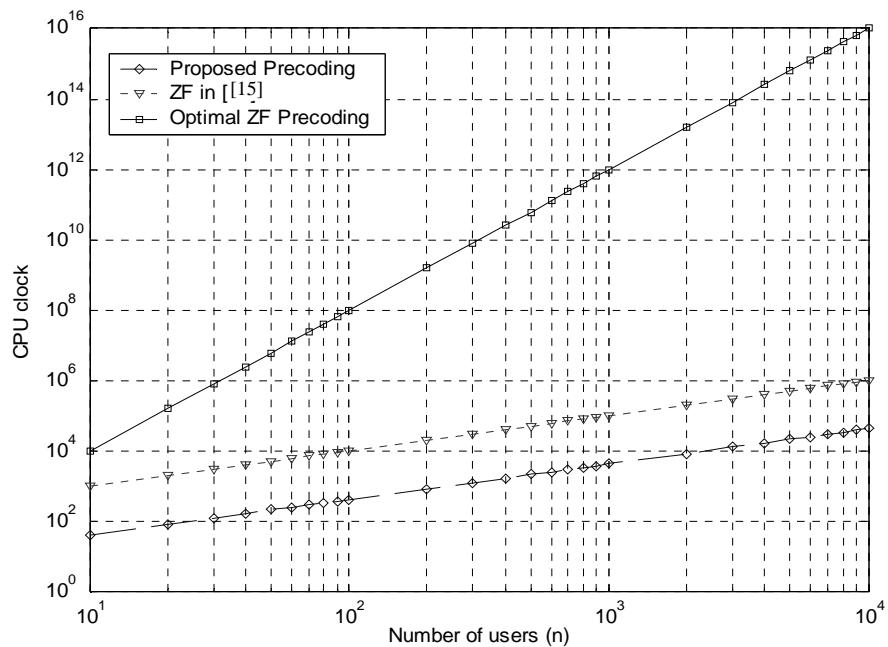


**Figure 2. Selection of different schemes in various scenarios.**

**Figure 4** shows the complexity of different schemes in term of the number of CPU operation needed to finish the optimization tasks. Complexity of NUS-CAP scheme grows linearly with the number of users while the complexity of the optimal zero-forcing and DPC is exponentially increasing with the number of users. Besides, NUS-CAP scheme needs less CPU clock than the full channel knowledge based zero-forcing in [15].
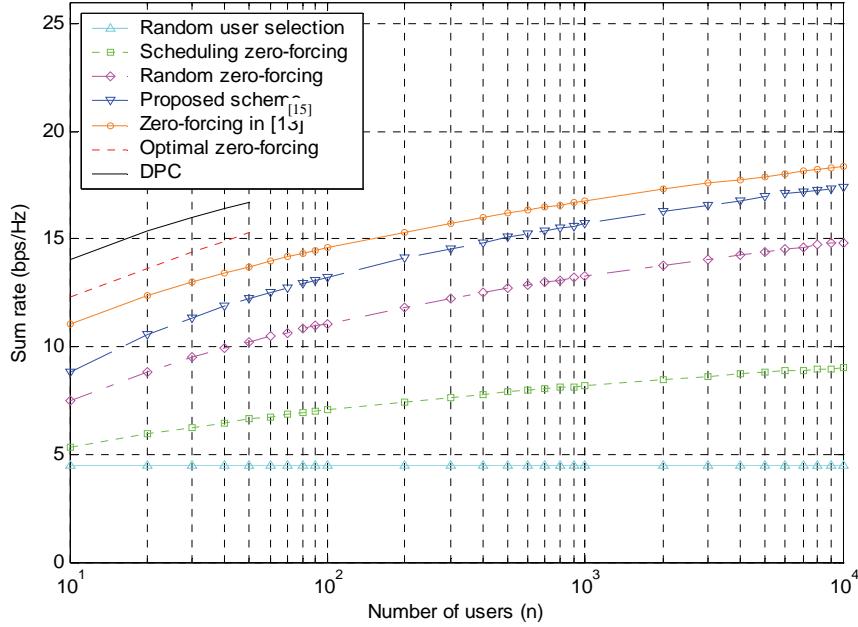
In **Figure 5**, achievable sum rates of different precoding schemes are compared for different number of users. We assume a fixed available power of $P = 10$ dB at the transmitter. Different schemes are considered, including random user selection ($M$ users are selected randomly each time), scheduling zero-forcing discussed in [15], random zero-forcing (OP) scheme discussed in Section 5, proposed NUS-CAP scheme, full channel zero-forcing



**Figure 3. Feedback overhead comparison of different schemes with $M = 4$.**



**Figure 4. Complexity comparison of different schemes with $M = 4$.**

**Figure 5. Sum rates for different schemes with *M* = 4 and *P* = 10 dB.**

scheme in [15], optimal zero forcing described in Section 3 and dirty paper coding. Due to high level of complexity, the sum rate of DPC and optimal zero-forcing has been just depicted for small number of users. As shown, NUS-CAP outperforms the schemes based on partial channel knowledge. Furthermore, its achievable sum rate is close to that of full channel knowledge based schemes and the rate of growth is the same as that of optimal schemes at the limit of large number of users. Note that, NUS-CAP works better in high number of users; however, by increasing the number of basis sets it is possible to compensate this drawback in low number of users. As mentioned we have just considered one basis vector set for the proposed schemes. By increasing the number of basis sets, the proposed schemes perform better even when the number of users is small.

## 8. Concluding Remarks

In this paper, we investigated the design of a partial knowledge based precoding scheme for broadcast MIMO systems. By partial knowledge we mean that the transmitter does not have the instantaneous channel realizations of all active users in the network. This assumption arises because the fast changing nature of wireless channels makes full channel knowledge difficult to obtain at the transmitter.

We proposed various combined user selection and precoding schemes with no or small amount of channel feedback and low complexity, which can achieve the performance of full knowledge schemes in the limit of large number of users. More precisely, it was shown that the ergodic capacity offered by the proposed schemes can

achieve the same growing rate with the number of users (log log *n*) as that of DPC and optimal zero-forcing precoding.

The ability of the proposed schemes in achieving optimality in term of ergodic capacity growth rate was demonstrated by analysis and numerical results. It was shown that as the number of users increases, the proposed schemes need lower amount of feedback and processing load than other schemes. Moreover, the proposed schemes are superior to each other in different propagation scenarios and SNR regimes. Nevertheless, as their structures are very similar, it is possible for the base station to switch between them in order to adapt to the propagation environment. This fearure gives flexibility in system configuration as well as more robustness in compensating for channel changes.

## 9. References

[1]  E. Telatar, "Capacity of Multi-Antenna Gaussian Channel," *European Transactions on Telecommunications*, Vol. 10, November 1999, pp. 585-595.

[2]  G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas," *Wireless Personal Communication*, Vol. 6, March 1998, pp. 311-335.

[3]  Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis and H. Sampath, "Optimal Design for Space-Time Linear Precoders and Decoders," *IEEE Transactions on Signal Processing*, Vol. 50, No. 5, May 2002, pp. 1051-1063.

[4]  J. Akhtar and D. Gesbert, "Spatial Multiplexing over

Correlated MIMO Channels with a Closed-Form Precoder," *IEEE Transactions on Wireless Communications*, No. 4, Vol. 5, September 2005, pp. 2400-2409.

[5] H. R. Bahrami and T. Le-Ngoc, "Precoder Design Based on Correlation Matrices for MIMO Systems," *IEEE Transactions on Wireless Communications*, Vol. 5, No. 12, December 2006, pp. 3579-3587.

[6] G. Caire and S. Shamai (Shitz), "On the Achievable throughput of a Multiantenna Gaussian Broadcast Channel," *IEEE Transactions on Inforation Theory*, Vol. 49, No. 7, July 2003, pp. 1691-1706.

[7] M. Costa, "Writing on Dirty Paper," *IEEE Transactions on Information Theory*, Vol. 29, No. 3, May 1983, pp. 439-441.

[8] S. Vishwanath, N. Jindal and A. Goldsmith, "Duality, Achievable Rates and Sum Rate Capacity of Gaussian MIMO Broadcast Channel," *IEEE Transactions on Information Theory*, Vol. 49, No. 10, October 2003, pp. 2658-2668.

[9] P. Viswanath and D. N. C. Tse, "Sum Capacity of the Vector Gaussian Broadcast Channel and Downlink-Uplink Duality," *IEEE Transactions on Information Theory*, Vol. 49, No. 8, August 2003, pp. 1912-1921.

[10] W. Yu and J. M. Cioffi, "Sum Capacity of Vector Gaussian Broadcast Channels, *IEEE Transactions on Information Theory*, Vol. 50, No. 9, September 2004, pp. 1875-1892.

[11] A. Goldsmith, S. A. Jafar, N. Jindal and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE Journal on Selected Areas in Communications*, Vol. 21, No. 5, June 2003, pp. 684-702.

[12] G. Caire, N. Jindal, M. Kobayashi and N. Ravindran, "Achievable Rates with Simple Channel State Estimation and Feedback Schemes," Submitted to *IEEE Transactions on Information Theory*, November 2007.

[13] M. Sharif and B. Hassibi, "On the Capacity of MIMO Broadcast Channels with Partial Side Information," *IEEE Transactions on Information Theory*, Vol. 51, No. 2, February 2005, pp. 506-522.

[14] M. Sharif and B. Hassibi, "On the Capacity of MIMO Broadcast Channel with Partial Side Information," *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Vol. 1, November 2003, pp. 958-962.

[15] T. Yoo and A. Goldsmith, "On the Optimality of Multiantenna Broadcast Scheduling Using Zero-Forcing Beam-Forming," *IEEE Journal on Selected Areas in Communication*, Vol. 24, No. 3, March 2006, pp. 528-541.

[16] T. Yoo, N. Jindal and A. Goldsmith, "Multi-Antenna Broadcast Channels with Limited Feedback and User Selection," *IEEE Journal on Selected Areas in Communication*, Vol. 25, No. 6, August 2007, pp. 1069-1078.

[17] K. Huang, J. G. Andrews and R. W. Heath, "Performance of Orthogonal Beamforming for SDMA with Limited Feedback," *IEEE Transactions on Vehicular Technology*, January 2009, pp. 152-164.

Scientific
Research

# Envelope Correlation Parameter Measurements in a MIMO Antenna Array Configuration

**Constantinos Votis, George Tatsis, Panos Kostarakis**
*Physics Department, University of Ioannina, Ioannina, Greece*
*Email: kvotis@grads.uoi.gr, gtatsis@grads.uoi.gr, kostarakis@uoi.gr*
*Received January* 11, 2010; *revised February* 12, 2010; *accepted March* 15, 2010

## Abstract

In a $2 \times 2$ MIMO antenna array system envelope correlation coefficient "$\rho$" shows the influence of different propagation paths of the RF signals that reach the antenna elements. The approximated value of this coefficient is based on a simple closed-form equation and also varies from 0 to 1. Quite perfect performance for MIMO applications is achieved when this parameter approximates to zero. In this paper, we evaluate an antenna diversity MIMO system by measuring the envelope correlation coefficient. The corresponding results in our antenna array configurations show that the measured "$\rho$" has very small values and approximates to zero. This observation indicates quite perfect behavior and performance of our MIMO antenna array system.

**Keywords:** Scattering Parameters, Envelope Correlation, Printed Dipole Antenna

## 1. Introduction

Multiple Input Multiple Output (MIMO) systems have received a great attention, recently. This architecture uses more than one antenna elements in transmitter and receiver ends and is able to overcome the limit of channel capacity in a rich multipath environment [1]. The theoretical capacity of the system increases linearly with the number of elements in MIMO antenna arrays. However, practical considerations indicate that the corresponding capacity of the system may be reduced if the received signals in any of the different antenna elements are correlated [2]. This effect proposes that diversity gain is obtained in the antenna system when the value of "$\rho$" is less than 0.5 [3]. It is obvious that correlation affects MIMO performance and represents a crucial parameter for modern wireless applications [4].

Moreover, MIMO design considerations include these antenna diversity techniques that also increase spectrum efficiency. It is also recognized that mutual coupling of the antenna degrades the performance of these systems.

These observations and an amount of corresponding research activities indicate that MIMO system performance is a crucial topic and for this investigation some parameters need to be considered. The envelope correlation between antenna elements is one of most important because it is related with the spectral efficiency and may provide degradation on performance of these applications.

Antenna correlation calculation procedure is provided by appropriate methods of analysis. Basically, three methods are used for these envelope correlation coefficient calculations. One of them is based on the far-filed radiation pattern. However, it is a time-consuming process [5,6]. This requires the corresponding numerical or experimental analysis and therefore is a cumbersome process. The second method is based on Clarke's formula [7] and has recently been used [8,9]. The third method is suitable for experimental measurements and requires the knowledge of scattering parameters obtained on the antenna elements. This last method is the one we adopted throughout this paper. The procedure of calculating the correlation between antennas in a two - antenna system using the scattering parameters is proposed in [10]. In our study, envelope correlation of eight antenna array types are presented and investigated for two indoor environments.

The present paper is structured as follows: in Section 2, the basic theoretical background is presented; the proposed antenna array implementation aspects are introduced in Section 3. Antenna array configurations are investigated in terms of envelope correlation and the corresponding results are discussed in Section 4. The experimental observations are summarized in Section 5.

## 2. Theory

The method of calculating envelope correlation of ele-

ments in each antenna array configuration is based on a fundamental Equation (1) that requires 3-dimensional radiation pattern considerations.

$$\rho_e = \frac{\left|\iint\limits_{4\pi}\left[\overrightarrow{F_1}(\theta,\phi)\bullet\overrightarrow{F_2}^*(\theta,\phi)\right]d\Omega\right|^2}{\iint\limits_{4\pi}\left|\overrightarrow{F_1}(\theta,\phi)\right|^2 d\Omega \iint\limits_{4\pi}\left|\overrightarrow{F_2}(\theta,\phi)\right|^2 d\Omega} \quad (1)$$

The parameter $\overrightarrow{F_i}(\theta,\phi)$ is the field radiation pattern of the antenna system when only the port *i* is excited and all other ports are terminated to 50 Ω load. The symbol • also denotes the Hermitian product [4,10].
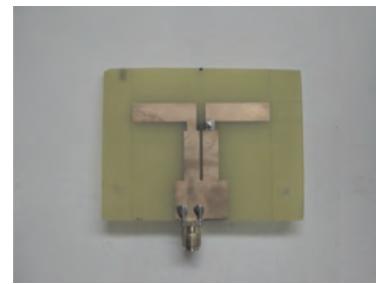
Recent research activities have shown that the envelope correlation can be well defined by a simple closed-form equation that relates the scattering parameters of the elements in an antenna array configuration. Especially, in case of a multipath indoor environment with a uniform distribution of Equation (2) is proved to be a good approximation [4]. For two antenna elements this equation using the scattering parameters becomes:

$$\rho_e = \frac{\left|S_{11}^*S_{12} + S_{21}^*S_{22}\right|^2}{(1-\left|S_{11}\right|^2-\left|S_{21}\right|^2)(1-\left|S_{22}\right|^2-\left|S_{12}\right|^2)} \quad (2)$$
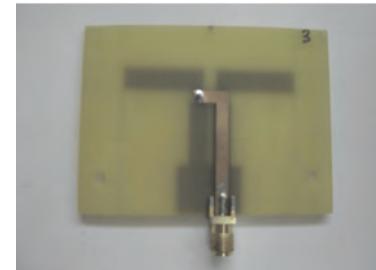
It is obvious that radiation pattern in Equation (1) makes the calculation more complicated than the envelope correlation calculations based on in Equation (2). The practical advantage of the third method that is based on second equation is that not only is quite simple to use it experimentally, but also provides sufficiently accurate results in many experimental environments such as indoor environments with rich multipath propagation performance.

## 3. Antenna Array Aspects

The mathematical consideration given by in Equation (2) is related to the corresponding antenna array structure that is comprised by two identical printed dipole antennas with integrated balun and a plane reflector of aluminum. **Figure 1** presents the layout of antenna dipole. Geometry parameters of the printed dipole have been further studied and investigated [11-13]. Its characteristics are summarized in **Table 1**. Each of the two identical printed dipoles has a resonance point close to the frequency range of 2.4 GHz and the corresponding resonance bandwidth is quite 500 MHz. In addition, the reflector backplane (**Figure 2**) is designed and implemented to allow the positioning of the antenna elements in various configurations. From these considerations it is obvious that this antenna array structure supports wireless applications in frequency range of 2.4 GHz. A typical antenna array configuration is shown in **Figure 3**.


(a) Bottom Layer
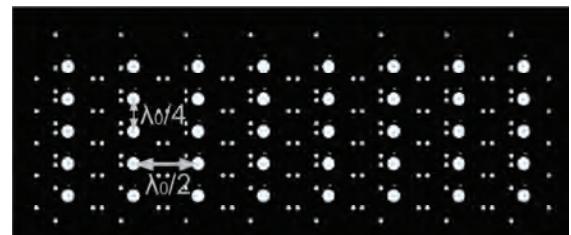

(b) Top Layer
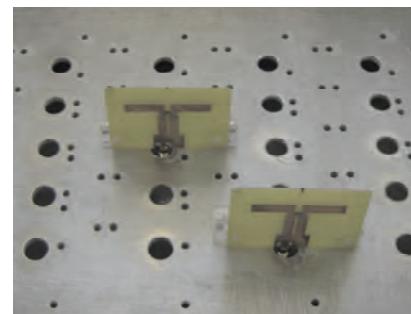
**Figure 1. Printed dipole antenna.**

**Table 1. Results for printed dipole (simulated/measured).**

| Definition | Symbol | Simulated | Measured |
|---|---|---|---|
| Resonance Center Frequency | $f_0$ (GHz) | 2.3 | 2.4 |
| Resonance Bandwidth | *BW* (GHz) | 0.5 | 0.5 |
| Return Loss | *RL* (dB) | -58 | -42 |



**Figure 2. Top site of plane reflector.**



**Figure 3. Typical antenna array configuration.**
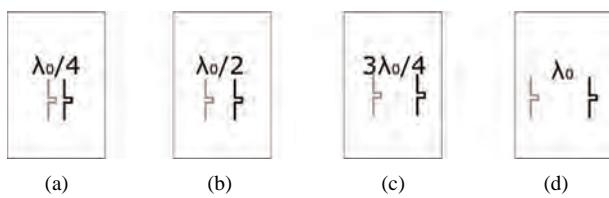
# 4. Analysis and Discussion

The antenna system that has been introduced is studied and investigated in terms of envelope coefficient measurements. The corresponding results have been obtained in each case of eight antenna array configurations. The first four of them corresponds to side by side placement of identical printed dipoles and are grouped to the first part of the present investigation. The last four represent collinear form and provides the second part. **Figure 4** and **Figure 5** show the corresponding groups of antenna array configurations for each part of the proposed investigation.
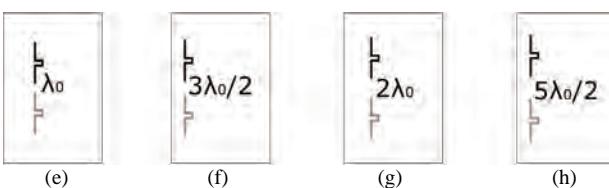
The corresponding *S*-parameters that are provided by experimental procedure have been concentrated and used to the mathematical formula proposed by in Equation (2). These experimental measurements of scattering parameters have been provided by a Network Analyzer and have taken place in two Laboratory's Environments A and B. In each of the two indoor environments and for Group A and Group B antenna array configurations these parameters have been obtained. With these measured results we calculate the envelope correlation coefficient. As mentioned above, this procedure is based on in Equation (2).

The corresponding calculations provide envelope correlation values in a frequency range of 1 GHz with center frequency of 2.4 GHz. The experimental results for Group A antenna array configurations are shown in **Figure 6** and **Figure 7** for two different Laboratory Environments A and B, respectively.
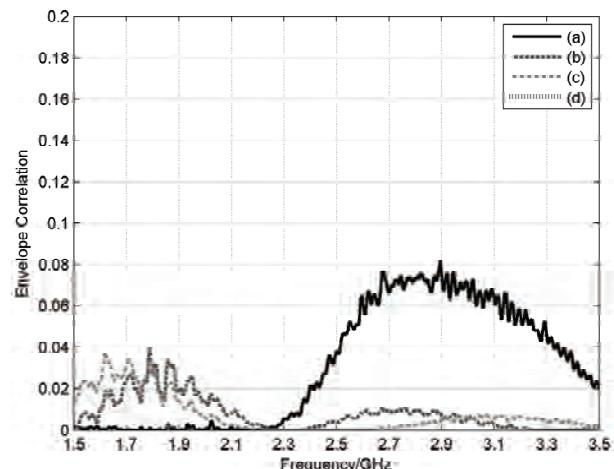
In each case of Group A configurations the corresponding envelope correlation curve has been plotted. From **Figure 6** and **Figure 7**, it seems that the "$\rho$" parameter has almost small values in this frequency range. In particular, at frequency range of 1.5 to 2.1 GHz the



**Figure 6. Measured results of envelope correlation for Group A in Laboratory A environment.**



**Figure 7. Measured results of envelope correlation for Group A in Laboratory B environment.**

envelope correlation coefficient has a value close to 0.015 for all antenna array configurations in Group A. This observation indicates that the parameter "$\rho$" remains stable as the distance between the dipoles increases in side by side antenna array configuration of Group A. Instead, at frequency range from 2.1 to 3.5 GHz the value of envelope correlation coefficient has variations in a range of 0 to 0.05 for the smallest distance between the dipoles in Group A antenna array configuration. In other cases these variations are neglected.

For Group B the corresponding results for parameter "$\rho$" are depicted in **Figure 8** and **Figure 9** for Laboratory's environment A and B, respectively.

From these experimental results, it seems that the parameter of envelope correlation has also quite small val-



**Figure 4. Relating antenna elements positions in the reflector backplane (Group A).**



**Figure 5. Relating antenna elements positions in the reflector backplane (Group B).**

ues in the frequency range of 1.5 to 3.5 GHz. Small variations of this parameter is introduced at the frequency range from 1.5 to 2 GHz only for the smallest distance between the dipoles in antenna array configuration of Group B. These variations are not provided by other values of distance between the elements bigger than one wavelength for collinear antenna array configuration. Moreover, in the frequency range from 2 to 3.5 GHz the parameter of envelope correlation coefficient remains stable and close to zero for each of the collinear antenna array configuration in Group B.

It is also important that these considerations for group A and B antenna array configurations are valid for laboratory's environments A and B. The corresponding declines between the two different propagation envi-

ronments are neglected. In each of these indoor environments it is the rich multipath propagation procedure that provides this important agreement in the corresponding measured results. It is convenient that the indoor environments offer propagation processes with a big number of independent signal paths. This observation indicates that in rich multipath environments the corresponding signal paths are uncorrelated and are impossible to be in deep fade, simultaneously. This feature improves the performance of MIMO communication applications and systems.

Another important consideration is that in Group B antenna array configurations the envelope correlation is quite low relative to Group A. Different antenna array geometry may be an important issue for this effect. In any case of them, the value of envelope correlation is quite small and less of value 0.5.

To sum up, these antenna array configurations provide low envelope correlation between the signals that have received or transmitted from the corresponding elements. In particular, Group B configurations have quite better correlation performance and offer more efficiently in MIMO applications' performance.



**Figure 8. Measured results of envelope correlation for Group B in Laboratory A environment.**

## 5. Conclusions
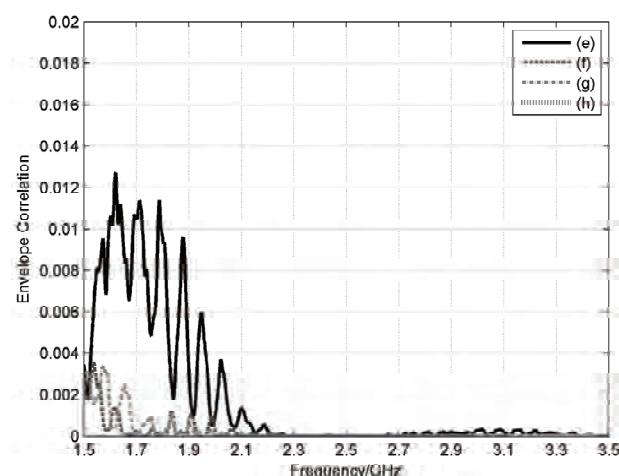
Antenna array configurations have been studied and investigated in terms of envelope correlation coefficient from the scattering parameters of two dipole antenna elements. The corresponding closed-form expression has been introduced and used to calculate this crucial parameter. This mathematical formula requires less laborious calculations and provides knowledge on antenna diversity optimization. Measured results indicate that rich multipath environment yields to low correlation coefficients for the proposed antenna array configurations. Several distances between antenna elements do not affect this parameter, dramatically. The corresponding considerations are very crucial for MIMO system's performance and are supported by an accurate method for investigation on modern wireless applications' design.

## 6. Acknowledgment
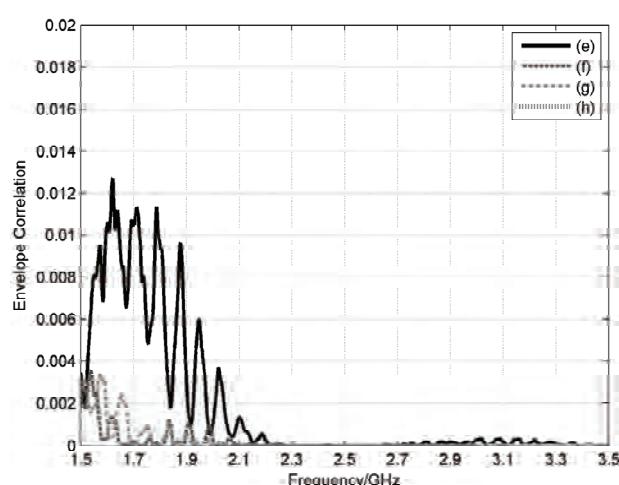
## 7. References

[1]	G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas," *Wireless Personal Communications*, Vol. 6, 1998, pp. 311-335.



**Figure 9. Measured results of envelope correlation for Group B in Laboratory B environment.**

[2] R. Janaswamy, "Effects of Mutual Coupling on the Capacity of Fixed Length Linear Arrays," *IEEE Antennas and Wireless Propagation Letters*, Vol. 1, 2002, pp. 157-160.

[3] R. G. Vaughan and J. B. Andersen, "Antenna Diversity in Mobile Communications," *IEEE Transactions on Vehicular Technology*, Vol. 36, 1987, pp. 149-172.

[4] R. G. Vaughan, "Signals in Mobile Communications," *IEEE Transactions on Vehicular Technology*, Vol. 35, 1986, pp. 133-145.

[5] G. Lebrun, S. Spiteri and M. Faulkner, "MIMO Complexity Reduction through Antenna Selection," *Proceedings on Australian Telecommun Cooperative Research Center*, ANNAC'03, Vol. 5, 2003.

[6] S. Jacobs and C. P. Bean, "Fine Particles, Thin Films and Exchange Anisotropy," *In Magnetism,* G. T. Rado and H. Suhl, Eds., Academic, New York, Vol. 3, 1963, pp. 271-350.

[7] R. H. Clarke, "A Statistical Theory of Mibile Reception," Bell System Technology Journal, 1968, pp. 957-1000. R. Nicole, "Title of Paper with Only First Word Capitalized," *Journal Name Standard Abbreviations*, in Press.

[8] K. Boyle, "Radiation Pattern and Correlation of Closely Spaces Linear Antennas," *IEEE Transactions on Antennas Propagation*, Vol. 50, 2002, pp.1162-1165.

[9] H. T. Hui, W. T. OwYong and K. B. Toh, "Signals Correlation between Two Normal-Mode Helical Antennas for Diversity Reception in a Multipath Environment," *IEEE Transactions on Antennas Propagation*, 2004, pp. 572-577.

[10] J. Blanch, J. Romeu and I. Cordella, "Exact Representation of Antenna System Diversity Performance from Input Parameter Description," *Electronics Letters*, Vol. 39, 2003, pp. 705-707.

[11] H.-R. Chuang and L.-C. Kuo, "3-D FDTD Design Analysis of a 2.4 GHz Polarization – Diversity Printed Dipole Antenna with Integrated Balun and Polarization – Switching Circuit for Wlan and Wireless Communication Application," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 51, No. 2, 2003.

[12] Z. Fan, L. Ran and K. Chen, "Printed Dipole Antenna Designed with Microstrip Balun on V-Shaped Ground Plane," *Progress in Electromagnetics Research Symposium Hangzhou*, 2005, pp. 23-26.

[13] D. M. Pozar, "Microwave Engineering," Wiley, 1998.

        

# QoS-Guaranteed Secure Multicast Routing Protocol for Satellite IP Networks Using Hierarchical Architecture

**Zhizhong Yin[1,2], Long Zhang[1], Xianwei Zhou[1], Peng Xu[1], Yu Deng[3]**

[1]*Department of Communication Engineering, School of Information Engineering, University of Science and Technology Beijing, Beijing, China*
[2]*The Academy of Equipment Command & Technology, Beijing, China*
[3]*Communications Research Group, Department of Electronics, University of York, Heslington, UK*
*Email*: *iceberg206@163.com*
*Received January 30, 2010; revised February 28, 2010; accepted March 17, 2010*

## Abstract

Most recent satellite network research has focused on providing routing services without considering security. In this paper, for the sake of better global coverage, we introduce a novel triple-layered satellite network architecture including Geostationary Earth Orbit (GEO), Highly Elliptical Orbit (HEO), and Low Earth Orbit (LEO) satellite layers, which provides the near-global coverage with 24 hour uninterrupted over the areas varying from 75° S to 90° N. On the basis of the hierarchical architecture, we propose a QoS-guaranteed secure multicast routing protocol (QGSMRP) for satellite IP networks using the logical location concept to isolate the mobility of LEO and HEO satellites. In QGSMRP, we employ the asymmetric cryptography to secure the control messages via the pairwise key pre-distribution, and present a least cost tree (LCT) strategy to construct the multicast tree under the condition that the QoS constraints are guaranteed, aiming to minimize the tree cost. Simulation results show that the performance benefits of the proposed QGSMRP in terms of the end-to-end tree delay, the tree cost, and the failure ratio of multicasting connections by comparison with the conventional shortest path tree (SPT) strategy.

**Keywords:** Satellite Networks, Security, Multicast Routing, Quality of Service, Hierarchical Architecture

## 1. Introduction

Satellite networks are characterized by global coverage, cost-effective broadcast and multipoint capabilities, flexible network configuration and capacity allocation, bandwidth-on-demand flexibility, etc. There is no doubt that satellite networks will be an integral part of the newly emerging Next Generation Networks (NGN) and the evolution of Future Networks (FN), and also play an increasing critical role in providing broadband Internet access, personal communications, broadcast and multi-cast of digital content to geographically diverse user groups, and so on. Although satellite networks offer great potential, they also present significant security challenges that need to be addressed. These security-related challenges can be summarized as follows [1].

• Satellite channels are wireless broadcast media, which makes it possible for an unauthorized user to receive the signal and eavesdrop on the communications, if it is not encrypted.

• Without the proper security mechanisms, any sufficiently well-equipped adversary can send spurious commands to the satellite and jam or disrupt the communications.

• Security systems should add minimal delays to the communications and have mechanisms to recover from loss in security information.

The above challenges may incur more security threats for satellite networks, including *insider attacks*, *packet modification*, *sending forged commands*, *denial of service*, etc [2].

With the explosive growth of the Internet-based multimedia applications, such as push media, file distribution and caching, multimedia conferencing, chat groups, multi-player games, etc, multicasting constitutes an important service to perform the simultaneous distribution of the same multicast packets from a single source node to a group of destinations in satellite networks. Multicast

routing is one of the key technologies in multicast service for satellite networks. In recent years, many conventional multicast routing protocols for terrestrial networks have been proposed [3,4]. However, these existing multicast routing protocols can not be very well suited for satellite IP networks. At present, only a few multicast routing schemes in the literature have been developed for satellite IP networks. In [5], using the datagram routing algorithm (DRA) [6] to create the multicast trees, a multicast routing algorithm for LEO satellite IP networks is introduced, which minimizes the end-to-end delay for real time multimedia services. The bandwidth-efficient multicast routing mechanism [7] based on *rectilinear Steiner trees* (RSTs) for LEO satellite IP networks minimizes the total bandwidth and gains the limited overhead. Two multicast routing algorithms based on the dynamic approximate center (DAC) core selection method, *i.e.*, the core-cluster combination-based shared tree (CCST) algorithm and the weighted CCST algorithm (*w*-CCST), are presented in [8]. The former significantly decreases the average tree cost, and the latter reduces the average end-to-end propagation delay. The distributed multicast routing protocol in [9] aims to minimize the total cost of the multicast trees in multi-layered satellite IP networks, including Geostationary Earth Orbit (GEO), Medium Earth Orbit (MEO), and Low Earth Orbit (LEO) layers.

In addition, the future media rich applications such as media streaming, content delivery distribution and real time broadband access require satellite networks that inherently offer user level quality of service (QoS) guarantees. In this regard, one of the challenges for multicasting communications in satellite IP networks is to design the QoS multicast routing protocols. To our knowledge, there is no QoS multicast routing protocol so far specifically developed for satellite IP networks. In general, a combination of different layers of satellite constellations, such as GEO, LEO, MEO, and Highly Elliptical Orbit (HEO) satellite constellations, to build up a solid satellite network with multiple layers, can yield a much better performance than these layers individually, e.g., higher efficiency in the spectrum usage, flexible user's access and route selection, larger transmission capacity. In this paper, for the sake of better "global coverage", we take into account the demand of satellite communications over the high-latitude areas, and present a triple-layered LEO/HEO/GEO satellite network architecture. On the basis of the novel hierarchical architecture, we adopt the concept of logical locations [6] to isolate the mobility of LEO and HEO satellites and propose a QoS-guaranteed secure multicast routing protocol (QGSMRP) for satellite IP networks. In QGSMRP, we employ the asymmetric cryptography to protect the data integrity of the control messages via the pairwise key pre-distribution and propose a least cost tree (LCT) strategy to construct the multicast tree under the condi-

tion that the QoS constraints are guaranteed, aiming to minimize the tree cost. Simulation results demonstrate that the proposed QGSMRP owns the better performance over the end-to-end tree delay, the tree cost, and the failure ratio of multicasting connections in contrast with the traditional shortest path tree (SPT) strategy.
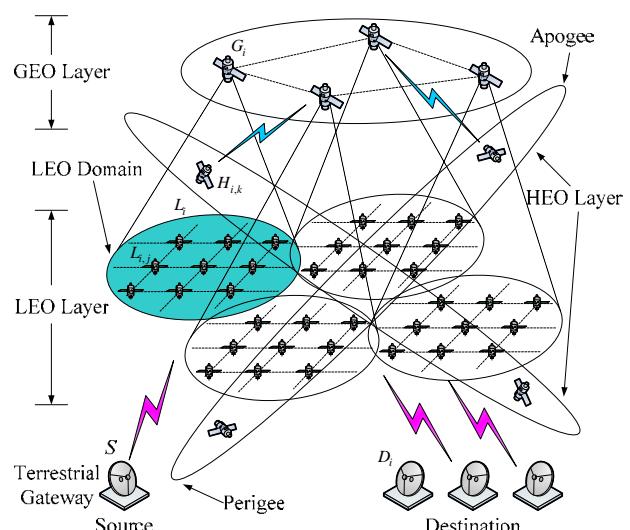
The rest of the paper is organized as follows. Section 2 describes the satellite network architecture. In Section 3, the problem formulation is introduced. Section 4 presents the QGSMRP. Simulation results are given in Section 5. Section 6 concludes this paper.

## 2. Satellite Network Architecture

As illustrated in **Figure 1**, the triple-layered satellite network architecture consists of a GEO constellation, several LEO and HEO constellations, and some fixed terrestrial gateways. The hierarchical architecture is divided into three satellite layers, *i.e.*, GEO, LEO, and HEO layers. The total number of GEO satellites is $N_G$ in GEO layer and a GEO satellite is denoted by $G_i$, $i = 1, \ldots, N_G$. In LEO layer, the total number of LEO satellites is $N_L$ and a LEO satellite is denoted by $L_{i,j}$, which is in the coverage area of the GEO satellite $G_i$. Assume that *Walker star* pattern constellation is applied in LEO layer to organize the LEO satellites. We introduce the HEO constellations to provide coverage to selected areas of the Earth, e.g. the Polar Regions, over which most GEO satellites lack. In HEO layer, the total number of HEO satellites is $N_H$ and a HEO satellite is denoted by $H_{i,k}$, which is in the coverage area of the GEO satellite $G_i$.
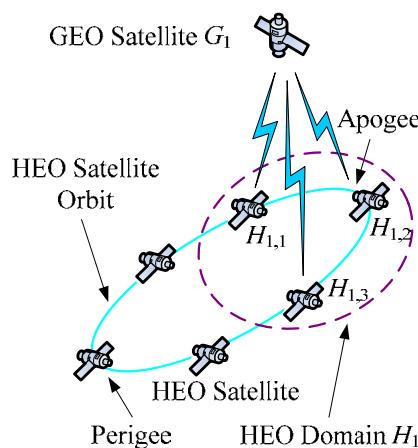
Three types of full duplex links are maintained in the architecture. Satellites are connected to each other within the same layer via Inter-Satellite Links (ISLs), while the



**Figure 1. The triple-layered satellite network architecture.**

communications between satellites (e.g. GEO and LEO satellites) in different layers is completed via Inter-Orbital Links (IOLs). Note that communication capabilities from HEO satellites is only provided when HEO satellites are moving very slowly relative to the globe while in the vicinity of apogee. For that reason, assume that the communications between GEO satellites and HEO satellites is accomplished via IOLs when HEO satellites are moving near apogee, while the communications among HEO satellites cannot be maintained through ISLs in the architecture. The terrestrial gateways are directly connected to LEO and GEO satellites via User Data Links (UDLs) and assumed to be the sources and destinations of multicasting communications.

Considering the logical locations of satellites, we introduce the satellite domains to organize satellites in a hierarchical manner in order to isolate the mobility of satellites from upper layer, *i.e.*, GEO layer. A LEO satellite domain $L_i$ is the set of logical locations of LEO satellites that are within the coverage of a GEO satellite $G_i$. ISLs in LEO layer can be categorized into two types, *i.e.*, intra-domain ISLs and inter-domain ISLs. Note that the LEO satellites are connected to their adjacent neighbors over the grid points in the same layer via intra-domain ISLs. Here, $L_i = \{L_{i,j} \big| j = 1, \ldots, \mathcal{S}(L_i)\}$, where $\mathcal{S}(\cdot)$ is a size function that generates the total number of satellites in a satellite domain. Similarly, a HEO satellite domain $H_i$ is the set of logical locations of HEO satellites that are within the coverage of a GEO satellite $G_i$, and $H_i = \{H_{i,k} \big| k = 1, \ldots, \mathcal{S}(H_i)\}$. Assume that half of the HEO satellites within a HEO constellation in the same orbit have IOLs with a GEO satellite at time instant and different GEO satellites may own the same HEO satellite domains. We give an example to illustrate the HEO domain in the architecture. As shown in **Figure 2**, a HEO constellation owns 6 satellites and a HEO domain contains 3 satellites.



**Figure 2. A HEO domain for example.**

## 3. Problem Formulation

The topology of satellite network based on our architecture is modeled as a connected directed graph $G = (V, E)$, where $V$ is the set of nodes representing satellites and terrestrial gateways in our architecture and $E \subseteq V \times V$ is the set of links connecting the nodes, *i.e.*, ISLs, IOLs and UDLs.

**Definition 1**. Let a terrestrial gateway $S \in V$ denote a source, and other terrestrial gateways constitute a set of destinations, *i.e.*, $D \subseteq V - \{S\}$, called a multicast group. A multicast tree $T = (V_T, E_T)$, for $V_T \subseteq V$ and $E_T \subseteq E$, is a subtree of the graph $G = (V, E)$ rooted from *S*, which contains all of the nodes of *D* and an arbitrary subset of $V - D$.

**Definition 2**. The link state of a link *l* consists of delay $\mathcal{D}(l)$ and available bandwidth $\mathcal{B}(l)$, for $l \in E$, where $\mathcal{D}(l) : E \rightarrow \boldsymbol{R}^+$ and $\mathcal{B}(l) : E \rightarrow \boldsymbol{R}^+$ are delay function and available bandwidth function, respectively. Note that the delay $\mathcal{D}(l)$ of a link *l* contains three delay components: 1) radio propagation delay, 2) queuing delay, and 3) protocol processing delay.

**Definition 3**. The available path bandwidth $B(P)$ of a path *P* is the minimum bandwidth of the links along the path, *i.e.*,

$$B(P) = \arg \min\{\mathcal{B}(l_i) \big| l_i \in P, i = 1, \ldots, \mathcal{PL}(P)\} \quad (1)$$

where $\mathcal{PL}(\cdot)$ is a function that returns the number of links in the path *P*.

**Definition 4**. The available tree bandwidth $B(T)$ of a multicast tree *T* is the minimum bandwidth of the links in the multicast tree, *i.e.*,

$$B(T) = \arg \min\{\mathcal{B}(l_i) \big| l_i \in T, i = 1, \ldots, \mathcal{TL}(T)\} \quad (2)$$

where $\mathcal{TL}(\cdot)$ is a function that returns the number of links in the tree *T*.

**Definition 5**. The path delay $D(P)$ of a path *P* is the sum of the delay of the links on the path, *i.e.*,

$$D(P) = \sum_{i=1}^{\mathcal{PL}(P)} \mathcal{D}(l_i) \quad (3)$$

**Definition 6**. The tree delay $D(T)$ of a multicast tree *T* is the end-to-end maximum delay of the paths from source *S* to the destinations of *D* on the multicast tree, *i.e.*,

$$D(T) = \arg \max\{D(P_{S \rightarrow D_i}) \big| i = 1, \ldots, |D|\} \quad (4)$$

where $P_{S \rightarrow D_i}$ denotes a path from source *S* to destination $D_i$, and $|D|$ denotes the number of destinations.

**Definition 7**. The path cost $C(P)$ of a path $P$ is defined as the product of available path bandwidth and path delay of path $P$, *i.e.*,

$$C(P) = B(P) \times D(P) \qquad (5)$$

**Definition 8**. The least cost path $P^*_{A \to B}$ from node $A$ to node $B$ is defined as a path that satisfies

$$C(P^*_{A \to B}) = \arg \min\{C(P^i_{A \to B}) \mid i = 1, \ldots, \mathcal{PN}(P_{A \to B})\} \qquad (6)$$

where $P_{A \to B}$ denotes a feasible path from node $A$ to node $B$, and $\mathcal{PN}(\cdot)$ is a function that returns the number of feasible paths from node $A$ to node $B$.

**Definition 9**. The tree cost $C(T)$ of a multicast tree $T$ is defined as the product of the available tree bandwidth and the tree delay of the multicast tree $T$, *i.e.*,

$$C(T) = B(T) \times D(T) \qquad (7)$$

Our problem is: given a satellite network $G = (V, E)$, a source $S$, a multicast group $D$, a delay bound $\Delta$, and a bandwidth bound $\Omega$, to construct a multicast tree $T = (V_T, E_T)$ which spans $S$ and $D$ such that the tree cost defined in (7) is minimized under the condition that the available tree bandwidth and tree delay of the multicast tree $T$ satisfy the required QoS constraints, namely, 1) delay constraint: $D(T) \le \Delta$, and 2) bandwidth constraint: $B(T) \ge \Omega$.

## 4. The Proposed Protocol

Our proposed QoS-guaranteed secure multicast routing protocol (QGSMRP) is composed of four processes, *i.e.*, link state report, route discovery and reply, route maintenance, and multicast tree creation. In QGSMRP, we assume that every node and its neighbor node in the satellite network $G = (V, E)$ mutually have the pairwise keys, *i.e.*, public keys and private keys, preloaded by key pre-distribution mechanism. We also assume that every node in the satellite network has unique node address. **Table 1** provides the notation description which will be used for cryptographic operations in the paper.

### 4.1. Link State Report

The link state report process is initiated when a source receives a QoS request for setting up a multicasting connection with a multicast group $D$ and the delay bound $\Delta$ and bandwidth bound $\Omega$. The source initially creates a report request (REPORT_REQ) message and then transmits it to a GEO satellite via a UDL. When receiving the REPORT_REQ, the GEO satellite follows the steps below to complete the link state report process.

**Table 1. Notation description.**

| Notations | Descriptions |
|---|---|
| $\mathcal{K}_{A+}$ | The public key held by node $A$ |
| $\mathcal{K}_{A-}$ | The private key held by node $A$ |
| $\mathcal{T}$ | Timestamp |
| $\{M\}\mathcal{K}_{A-}$ | Encryption of message $M$ with key $\mathcal{K}_{A+}$ |
| $[M]\mathcal{K}_{A+}$ | Decryption of message $M$ with key $\mathcal{K}_{A-}$ |
| *SRid* | STATE_REPORT ID |
| *RRQid* | RREQ ID |
| *RRPid* | RREP ID |
| *JRQid* | Join_REQ ID |
| *JRPid* | Join_REP ID |
| *JNid* | Join_NAK ID |

1) Link State Report Request

**Step 1:** The GEO satellite sends the REPORT_REQ to other adjacent GEO satellites via ISLs.

**Step 2:** When the REPORT_REQ are received by all the GEO satellites in GEO layer, each GEO satellite sends the REPORT_REQ to the LEO and HEO satellites within its covered LEO and HEO domain through IOLs.

**Step 3:** In LEO layer, LEO satellite floods the REPORT_REQ to other LEO satellites within the same domain via intra-domain ISLs and across different domains via inter-domain ISLs.

**Step 4:** The members of the multicast group $D$ receive the REPORT_REQ from the GEO and LEO satellites via UDLs.

After all the nodes in the satellite network acquire the REPORT_REQ, the link state interaction process is initiated.

2) Link State Interaction

**Step 1:** The member of the multicast group $D$, $D_i \in D$, sends a state report (STATE_REPORT) message to the GEO and LEO satellites via the reverse UDLs. The STATE_REPORT is constructed as follows:

<{*SRid*, *Node Address*, *Downstream Node Address*, $\mathcal{T}$ }$\mathcal{K}_{D_i-}$, *Available Bandwidth*, *Delay*>

The pair <*Node Address*, *SRid*> uniquely identifies the STATE_REPORT. The *SRid* is monotonically increasing when a node issues a new STATE_REPORT to its downstream node and can be used to check the duplicate copies of an old STATE_REPORT for the downstream node. The destination $D_i \in D$ encrypts the message {*SRid*, *Node Address*, *Downstream Node Address*, $\mathcal{T}$ } with its private key $K_{D_i-}$. The *Available Bandwidth* and *Delay* record the available bandwidth and the delay between a node and its downstream node over the corresponding link. When a GEO satellite (or a LEO satellite)

receives the STATE_REPORT, it will decrypt the ciphertext in the STATE_REPORT with its public key to verify whether $\mathcal{T}$ expires:

$$[\{SRid,\ Node\ Address,\ Downstream\ Node\ Address,\ \mathcal{T}\}\mathcal{K}_{D_i-}]\,K_{G_i+}$$

If so, the GEO satellite (or LEO satellite) drops the STATE_REPORT. When receiving STATE_REPORT from all the destinations, the GEO and LEO satellites acquire the link state information, *i.e.*, the available bandwidth and delay from the GEO and LEO satellites to destinations.

**Step 2:** In LEO layer, LEO satellite receives the STATE_REPORT from the upstream node (*i.e.*, the destination) and performs the same cryptographic operations, and then issues its STATE_REPORT to other LEO satellites within the same domain via intra-domain ISLs and across different domains via inter-domain ISLs.

**Step 3:** The LEO satellites in the same domain transmit their STATE_REPORT via IOLs to their manager, the GEO satellite.

**Step 4:** In GEO layer, GEO satellite also transmits the STATE_REPORT to other adjacent GEO satellites via ISLs. When exchanging their link state information, GEO satellites delivers the STATE_REPORT to HEO satellites within its covered HEO domain via IOLs and also to the source via UDLs.

When the link state report process is completed, the available bandwidth and the delay of each link $l \in E$ in the satellite network $G(V, E)$ are established and the related link state information is recorded in each node.

## 4.2. Route Discovery and Reply

When the link state report process is completed, the source *S* initiates the route discovery by flooding a route request (RREQ) message to its neighbor nodes. The RREQ is constructed as follows:

$$<\{RRQid,\ Source\ Address,\ Multicast\ Group\ Address\ List, \mathcal{T}\}\mathcal{K}_{S-},\ \Omega,\ \Delta,\ Accumulated\ Delay,\ Path>$$

The pair *<Source Address, RRQid >* uniquely identifies the RREQ. The *Path* records the routing information during route discovery and the *Accumulated Delay* records the sum of delay along the path. When an intermediate node *A* receives a RREQ, it decrypts the ciphertext in the RREQ with its public key $\mathcal{K}_{A+}$, and checks three items to decide whether to send the received RREQ:

**Item 1:** Whether $\mathcal{T}$ has expired. If so, the RREQ will be dropped.

**Item 2:** Whether there is enough available bandwidth $\mathcal{B}(l')$ over link $l'$ between the last hop node and itself according to link state information, *i.e.*, whether

$\mathcal{B}(l') \geq \Omega$. If $\mathcal{B}(l') < \Omega$, it means that there is no available bandwidth to meet the QoS requirements over that link and the RREQ will be dropped.

**Item 3:** Whether the sum of *Accumulated Delay* and the delay $\mathcal{D}(l')$ over link $l'$ meets the delay constraint, *i.e.*, $AD + \mathcal{D}(l') \leq \Delta$, where $AD$ denotes the value in *Accumulated Delay*. If $AD + \mathcal{D}(l') > \Delta$, it means that delay bound cannot be guaranteed and the RREQ will be dropped.

Otherwise, if the RREQ is received by the intermediate node for the first time, the intermediate node enters its own address to the *Path*, and inputs the value of $(AD + \mathcal{D}(l'))$ into *Accumulated Delay*, encrypts the corresponding message with its private key $\mathcal{K}_{A-}$ and then sends the RREQ out. If the newly received RREQ was received before, it means that there exists another path from the source to the node. The node records this path information and discards that RREQ.

This operation in route discovery will be repeated node by node until the delay bound or available bandwidth bound cannot be guaranteed. Eventually, a RREQ message will arrive at a destination $D_i \in D$ and this destination will also check the three items to determine whether the QoS constraints are satisfied. If so, this destination will wait for a certain time to receive multiple RREQ, and then will create a route reply (RREP) message and send the RREP back to the source *S*. The RREP is constructed as follows:

$$<\{RRPid,\ Destination\ Address,\ Source\ Address,\ \mathcal{T}\}\,K_{D_i-},\ \Omega,\ \Delta,\ Path\ Set,\ Lifetime>$$

The pair *<Destination Address, RRPid>* uniquely identifies the RREP and the *Lifetime* denotes a value of a pre-defined time for which the nodes receiving the RREP consider the route to be valid. The *Path Set* is the set of multiple possible parallel paths from the source to the destination and each path is marked with the information of accumulated delay and available bandwidth from the source to the destination. Meanwhile, the destination can also act as an intermediate node and continues to forward the RREQ until the QoS constraints are not guaranteed. When the source receives the RREP, it first decrypts the ciphertext in the RREP with its public key $\mathcal{K}_{S+}$ to verify whether $\mathcal{T}$ expires, and also records the *Path Set* into routing table. When the source receives all the RREP, it gets a partial topology from it to the multicast group in satellite network $G = (V, E)$.

## 4.3. Route Maintenance

1) Joining Multicast Group

The joining multicast group process is initiated when a

new terrestrial gateway wants to join a multicast group $D$. The new terrestrial gateway $D_{new}$ firstly creates a STATE_REPORT and then sends the STATE_REPORT to the GEO and LEO satellites via UDLs. Through link state interaction, the nodes with the corresponding links acquire the link state information. Then the new terrestrial gateway broadcasts a join request (JOIN_REQ) message which is constructed as follows:

$<\{JRQid,$ *Terrestrial Gateway Address,* $\mathcal{T} \} \mathcal{K}_{D_{new-}},$

*Path, Accumulated Delay>*

The pair $<$*Terrestrial Gateway Address, JRQid>* uniquely identifies the JOIN_REQ. When an intermediate node receives a JOIN_REQ, it checks the three items to decide whether to reflood the received JOIN_REQ. The new gateway will send multiple JOIN_REQ out within a pre-defined time until they arrive at a desitiation $D_i \in D$. The destination firstly waits for a certain time to receive multiple JOIN_REQ. Secondly, according to the *Path*, it will check the three items to determine whether the QoS constraints are satisfied. If so, the destination creates a join reply (JOIN_REP) message and sends it back to the new gateway. Note that this destination serves as a forwarding node and the source does not know the information about this new gateway. The JOIN_REP is constructed as follows:

$<\{JRPid,$ *Destination Address, Terrestrial Gateway Address,* $\mathcal{T} \} K_{D_i-},$ $\Omega$, $\Delta >$

The pair $<$*Destination Address, JRPid>* uniquely identifies the JOIN_REP. After a pre-defined time, the new gateway does not receive any more JOIN_REP and the joining multicast group process terminates.

2) Leaving Multicast Group

The leaving multicast group process is initiated when a destination $D_{leave}$ wants to leave a multicast group $D$. The destination firstly sends out multiple join negative acknowledgement (JOIN_NAK) messages to its neighbors and then deletes all the routing information of the neighbors. The JOIN_NAK is constructed as follows:

$<\{JNid,$ *Terrestrial Gateway Address,* $\mathcal{T} \} K_{D_{leave-}},$ *Path>*

The pair $<$*Terrestrial Gateway Address, JNid>* uniquely identifies the JOIN_NAK. When receiving a JOIN_NAK, a neighbor checks whether it has an upstream node or a downstream node. If so, the neighbor prunes the link from this destination and deletes the routing information of this destination, and then transmits the JOIN_NAK to notify that this destination has been leaving the multicast group. Otherwise, the neighbor will check whether it is a member of multicast group. If so, the neighbor just prunes the link from this destination. Otherwise, the neighbor becomes a non-forwarding node

and withdraws from the QoS multicasting communications.

## 4.4. Multicast Tree Creation

The multicast tree creation process is activated by the source at the end of the route discovery and reply process. As mentioned previously, the source has maintained multiple parallel paths from itself to several destinations in multicast group. Therefore, the main goal of the source is to select one of the parallel paths to set up a multicasting connection, and then proceed to create a multicast tree. In this paper, we present a least cost tree (LCT) strategy to construct the multicast tree under the condition that the QoS requirements are guaranteed. The basic idea of the LCT strategy works as follows. Assume that the *Path Set* from a destination $D_i$ $i=1,\ldots,|D|$, is denoted by $PS_i$. When receiving a RREP from a destination $D_i$, the source gets the information of accumulated delay and available bandwidth from the source to this destination along a path $P_{i,j}$, $j=1,\ldots,|PS_i|$, *i.e.*, path delay $D(P_{i,j})$ and available path bandwidth $B(P_{i,j})$. Therefore, the source can compute the path cost $C(P_{i,j}) = B(P_{i,j}) \times D(P_{i,j})$ for the path $P_{i,j}$, and then compare $C(P_{i,j})$ to select a path $P_{i,j}^*$ with the least path cost, *i.e.*,

$$C(P_{i,j}^*) = \arg \min\{C(P_{i,j})\big| i=1,\ldots,|D|, j=1,\ldots,|PS_i|\} \quad (9)$$

After a pre-defined time, the source gains all the information about the paths with least path cost from itself to each destination in the multicast group $D$. Afterwards, the source follows the steps below to create a multicast tree. When the multicast tree is built up, the multicast session begins.

**Step 1:** Construct two node sets $K = \{S\} \bigcup D$ and $H_0 = \{S\}$.

**Step 2:** Start with a subtree $T_0 = (V_0, E_0)$, where $V_0 = \{S\}$ and $E_0 = \varnothing$.

**Step 3:** For $\alpha = 1,\ldots,|D|$, the source finds a node in $K - H_{\alpha-1}$, *i.e.*, a destination $D_i$, such that the path cost from the source to $D_i$ is minimum among all the paths with the least path cost, *i.e.*,

$$D_i = \arg \min\{C(P_{i,j}^*)\big| i=1,\ldots,|D|, j=1,\ldots,|PS_i|\} \quad (10)$$

Construct the subtree $T_\alpha = (V_\alpha, E_\alpha)$ by adding the path $P_{i,j}^*$ between the source and $D_i$ to $T_\alpha$, *i.e.*, set $V_\alpha = V_{\alpha-1} \bigcup \{\text{nodes in } P_{i,j}^*\}$ and $E_\alpha = E_{\alpha-1} \bigcup \{\text{links in } P_{i,j}^*\}$. Meanwhile, set $H_\alpha = H_{\alpha-1} \bigcup \{D_i\}$.

# 5. Performance Evaluation
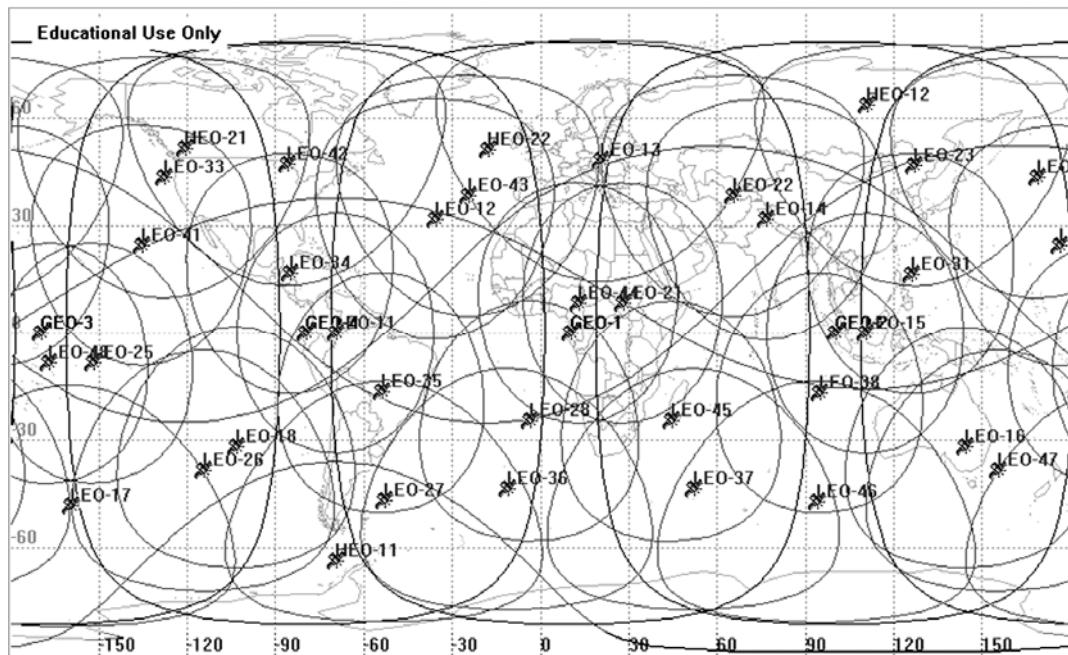
In this section, we evaluate the performance of QGSM-RP by comparing it with the conventional shortest path tree (SPT) strategy [10] via computer simulations. In our empirical study, three metrics are considered in the performance evaluations, *i.e.*, 1) the end-to-end tree delay, 2) the tree cost, and 3) the failure ratio of multicasting connections.

In our simulations, the parameters of the triple-layer satellite network are described in **Table 2**. The performance of coverage from the triple-layer satellite network is illustrated in **Figure 3**. According to **Figure 3**, triple-layered satellite network can offer coverage over the areas varying from 75° S to 90° N with 24 hour uninterrupted. We use the non-uniform distribution [9] to determine the positions of the terrestrial gateways, including the source and the multicast group. Moreover, we assume that the capacity of all ISLs, IOLs, and UDLs are set to 800 Mb/s, each outgoing link has a buffer space of 20 MB, and the delay bound and bandwidth bound are set to $\Delta = 450$ ms and $\Omega = 256$ Mb/s, respectively.
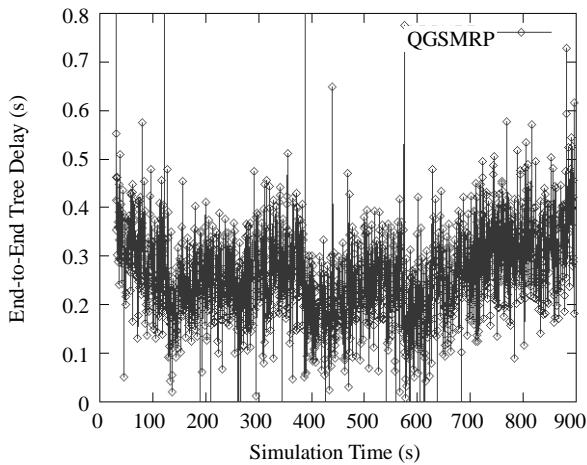
**Figure 4** and **Figure 5** illustrate the performance of the end-to-end tree delay of the proposed QGSMRP and the SPT strategy, respectively. The size of multicast group is set to 50. We can observe that with the growth of the simulation time, the variation of the end-to-end tree delay of the proposed QGSMRP and the SPT strategy is almost uniform with the range of 0.1 s to 0.5 s. Furthermore, the end-to-end tree delay changes quickly and obviously as the simulation time increases.

**Table 2. Parameters for triple-layered satellite networks.**
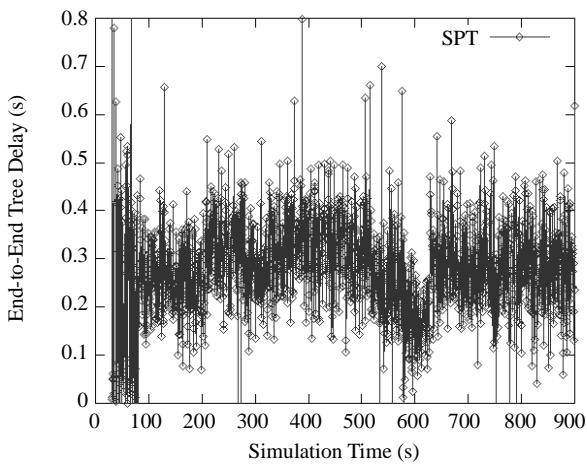
| Parameters | LEO Layer | HEO Layer | GEO Layer |
|---|---|---|---|
| Type of orbit | LEO Recursive orbit | HEO Recursive orbit | GEO |
| Altitude | 1262 km | 27000 km(A) 800 km(P) | 35786 km |
| Orbital period | 6628 s | 8 h | 24 h |
| Number of satellites | 32 | 4 | 4 |
| Number of orbital planes | 4 | 2 | 1 |
| Orbit inclination angle | 48° | 63.4° | — |
| Minimum elevation angle | — | 10° | 5° |
| Constellation type | Walker star | Draim | — |
| Semi-major axis | — | 20278 km | — |
| Eccentricity | — | 0.646 | — |
| Argument of perigee | — | 270° | — |
| Phase factor | 1 | — | — |
| Ascending node longitude | — | 90° E | — |



**Figure 3. Near-global coverage from the satellite network using the proposed triple-layered LEO/HEO/GEO architecture.**
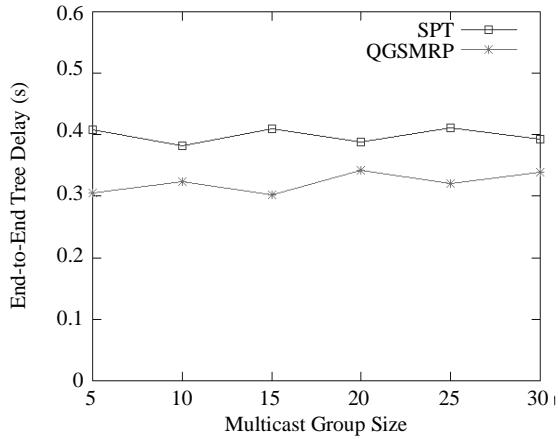
**Figure 4. The end-to-end tree delay of the proposed QGS-MRP.**
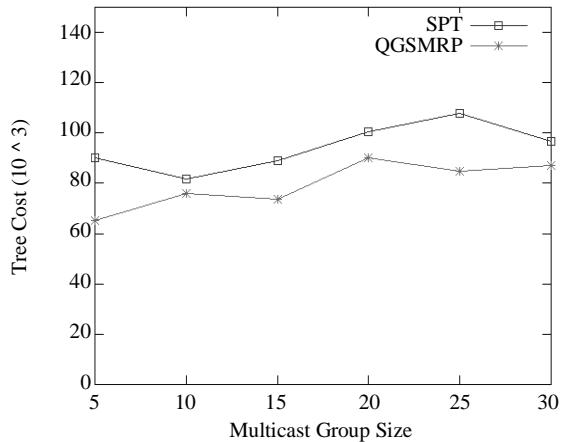


**Figure 5. The end-to-end tree delay of the SPT strategy.**



**Figure 6. Comparison of the end-to-end tree delay between the SPT strategy and the proposed QGSMRP.**



**Figure 7. Comparison of the tree cost between the SPT strategy and the proposed QGSMRP.**

**Figure 6** shows the comparison of the end-to-end tree delay versus the multicast group size between the SPT and the proposed QGSMRP. It can be easily seen that the end-to-end tree delay of the proposed QGSMRP is much lower than that of the SPT strategy with the increase of the size of the multicast group. **Figure 7** compares the performance of the tree cost versus the multicast group size between the SPT and the proposed QGSMRP. We can obviously see that as the multicast group size grows, the tree cost of the proposed QGSMRP is much smaller than that of the SPT strategy. This can be explained by the fact that the proposed QGSMRP focuses on the optimization of the tree cost during the process of the route discovery and reply, whereas the main target of the SPT strategy is to bring the better performance of the path delay in the multicast tree without taking into account the available bandwidth in the process of the multicast tree construction.

**Figure 8** demonstrates the comparison of the failure ratio of multicasting connections between the SPT strategy and the proposed QGSMRP. It can be observed from **Figure 8** that the failure ratio of multicasting connections of the proposed QGSMRP is smaller than that of the SPT strategy with the range of 15 to 30 of the multicast group size, which indicates that the success ratio of the QoS multicasting requests of the proposed QGSMRP is superior to that of the SPT strategy with the growth of the multicast group size. For that reason, the proposed QGSMRP can easily establish the QoS multicasting connections when the scale of the network is large, *i.e.*, a large quantity of the terrestrial gateways. However, the failure ratio of multicasting connections of the SPT strategy is better than that of the proposed QGSMRP in the range of 5 to 10 of the size of the multicast group. We can also observe that the failure ratio of multicasting connections of both of the multicast routing strategies are represented as a gradual increasing trend.

**Figure 8. Comparison of the failure ratio of multicasting connections between the SPT strategy and the proposed QGSMRP.**

## 6. Conclusions

In this paper, considering the difficulty to provide the coverage over the special regions or the areas of high latitudes by the existing hierarchical satellite networks, we introduce a novel triple-layered LEO/HEO/GEO satellite network architecture containing LEO, HEO, and GEO satellite layers, which yields the near-global coverage with 24 hour uninterrupted over the areas varying from 75° S to 90° N. On the basis of the novel hierarchical architecture, we propose a QoS-guaranteed secure multicast routing protocol (QGSMRP) for satellite IP networks using the concept of logical locations to isolate the mobility of LEO and HEO satellites. In QGSMRP, through the pairwise key pre-distribution, we employ the asymmetric cryptography to secure the data integrity of the control messages, and present a least cost tree (LCT) strategy to construct the multicast tree under the condition that the QoS constraints are guaranteed, aiming to minimize the tree cost. Simulation results demonstrate that the proposed QGSMRP owns the better performance including the end-to-end tree delay, the tree cost, and the failure ratio of multicasting connections by comparison with the conventional shortest path tree (SPT) strategy.

## 7. Acknowledgements

## 8. References

[1] R. Chowdhury, J. S. Baras, M. Hadjitheodosiou and S. Papademetriou, "Security Issues in Hybrid Networks with a Satellite Component," *IEEE Wireless Communications*, Vol. 12, No. 6, December 2005, pp. 50-61.

[2] N. Boudriga, "Security of Mobile Communications," Boca Raton, Auerbach Publications, FL, 2009, pp. 474-479.

[3] L. H. Sahasrabuddhe and B. Mukherjee, "Multicast Routing Algorithms and Protocols: A Tutorial," *IEEE Network*, Vol. 14, No. 1, January/February 2000, pp. 90-102.

[4] U. Varshney, "Multicast over Wireless Networks," *Communications of the ACM*, Vol. 45, No. 12, December 2002, pp. 31-37.

[5] E. Ekici, I. F. Akyildiz and M. D. Bender, "A Multicast Routing Algorithm for LEO Satellite IP Networks," *IEEE/ACM Transactions on Networking*, Vol. 10, No. 2, April 2002, pp. 183-192.

[6] E. Ekici, I. F. Akyildiz and M. D. Bender, "A Distributed Routing Algorithm for Datagram Traffic in LEO Satellite Networks," *IEEE/ACM Transactions on Networking*, Vol. 9, No. 2, April 2001, pp. 137-147.

[7] D. Yang and W. Liao, "On Multicast Routing Using Rect-Ilinear Steiner Trees for LEO Satellite Networks," *IEEE Transactions on Vehicular Technology*, Vol. 57, No. 4, July 2008, pp. 2560-2569.

[8] L. Chen, J. Zhang and K. Liu, "Core-Based Shared Tree Multicast Routing Algorithms for LEO Satellite IP Networks," *Chinese Journal of Aeronautics*, Vol. 20, No. 4, August 2007, pp. 353-361.

[9] I. F. Akyildiz, E. Ekici and G. Yue, "A Distributed Multicast Routing Scheme for Multi-Layered Satellite IP Networks," *Wireless Networks*, Vol. 9, No. 5, September 2003, pp. 535-544.

[10] J. A. Bondy and U. S. R. Murty, "Graph Theory with Applications," The Macmillan Press, Great Britain, 1976, pp. 15-20.

◆◆ Scientific
◆◆ Research

# Team Spirit Model Using MPEG Standards for Video Delivery

**Neetu Singh, Piyush Chauhan, Nitin Rakesh, Nitin Nitin**
*Department of CSE & IT, Jaypee University of Information Technology, Waknaghat, India*
*Email*: *nitin.rakesh@gmail.com, delnitin@ieee.org*

## Abstract

Today the multimedia content delivery chain faces many challenges. There are increasing terminal diversity, network heterogeneity and pressure to satisfy the user preferences. Therefore, the need of customized contents comes in light to provide users the best possible experience. In this paper, we address the problem of multimedia customization. For the customized content, we suggest the team spirit model (TSM) that uses the web edition content description, MPEG-4 standards and the MPEG-21 multimedia framework. It efficiently implements the video customization

## 1. Introduction

Users are already swamped by the quantity of material available on conventional mass media. In today's world, the larger range of multimedia material is on hand .This can only serve up to offer customers with so much choice that they are overwhelmed by it. In addition, obviously they expect to be provided with services that grant access to information. The service might be due to the increased accessibility of technologies on a broad level. It arises the need to customize those services with regarding to the user's receiving device, which could be from a handheld to a widescreen terminal. If not so then the provided information could be inadequate. If at users widescreen device the least standard is applied, it would suffer under-represented information while a maximum standard of handheld devices might not be able to properly view or use the service at all [1]. That is why the obvious way of optimizing and adapting are needed. The provided service matches to the device to maintain or even enhance user-friendliness. It is the only way to cope with the technical convergence of devices, which would provide access to the similar services. In this   paper, our aim probe into the use of dynamic components involved, which should take into user characteristics and assure an efficient way to "access any information from any terminal", eventually after some kind of adaptation.

   In this paper, we are presenting the team spirit model, which take input as web edition content that stores all

content information, MPEG-4 Advanced video coding (AVC), and File Format for audio and video and lastly the MPEG-21 usage environment for other necessary description to efficiently implement adaption and finally achieve the customization in the well-mannered way. In TSM scenarios, it is essential to customize more easily and efficiently for the desired content. The contents have available descriptions of the parts that fit to be matched/ bridged: the content and the usage environment. The major objective of this paper probes into the use and development of a TSM system that has capable of managing and customizing content so that any information can be delivered to different terminals and networks.

## 2. Related Work

Video customization is demand of today's world. The customization can be achieved efficiently by using various models like hybrid multimedia access (HMA) Model and Universal Multimedia Access (UMA) Model using standards MPEG-7 and MPEG-21 [1-4]. In this paper, we are focusing on Team Spirit as an adaptation engine to implement the video customization using some advance technology based on MPEG standards. The major objective of this paper is to discuss the role of the various MPEG standards in the context of multimedia customization adaptation and to contribute for a better organization and understanding of the multimedia customization for TSM.

## 3. MPEG Standards Used

MPEG-4 enables different software and hardware developers to create multimedia objects possessing better abilities of adaptability and flexibility to improve the quality of such services and technologies as digital television, graphics, the World Wide Web and their extensions. This standard enables developers to better control their content and to fight more effectively against copyright violation. Data network providers can use MPEG-4 for data transparency. With the help of standard procedures, MPEG-4 data can be interpreted and transformed into other signal types compatible with any available network.
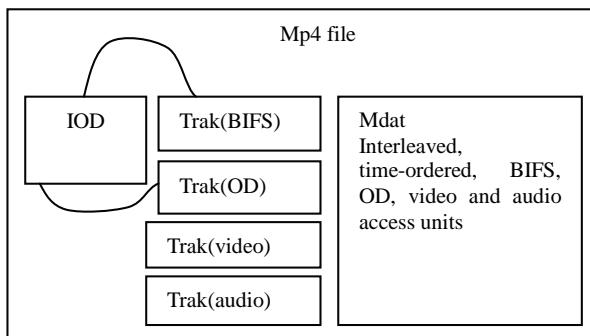
### 3.1. MPEG-4 (Part-14) File Format

The MP4 file format is intended to contain the information regarding the media of an MPEG-4 presentation in a flexible, extensible format that in turn facilitates interchange, management, editing, and presentation of the media. The presentation may be view as 'local' to the system containing the presentation, or may be via a network or other stream delivery mechanism. The file format is designed to be independent of any particular delivery protocol while enabling efficient support for delivery in general. The **Figure 1**, gives an example of a simple interchange file, containing three streams.
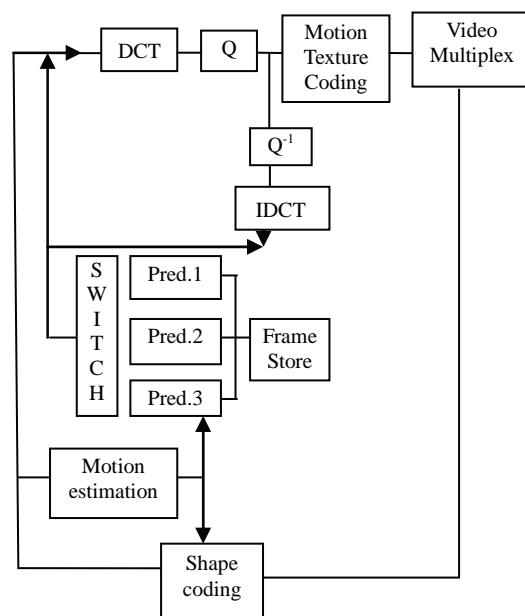
The composition of the Mp4 file format is of object-oriented structures called 'atoms'. A unique tag and a length identify each atom. Most atoms describe a hierarchy of metadata giving information such as index points, durations, and pointers to the media data. Now the 'movie atom' can be defined as the collection of these atoms. The media data itself is located elsewhere; it can be in the MP4 file, contained in one or more 'mdat' or media data atoms, or located outside the MP4 file and referenced via URL's [5].

### 3.2. MPEG-4 Video Image Coding Scheme

**Figure 2** below outlines the description of the MPEG-4



Figure 1. Example of a simple interchange file.



**Figure 2. Basic block diagram of MPEG-4 video coder.**

video algorithms to encode rectangular as well as arbitrarily shaped input image sequences.

The basic coding structure involves shape coding (for arbitrarily shaped VOs) and motion compensation as well as DCT-based texture coding (using standard $8 \times 8$ DCT or shape adaptive DCT). The basic coding structure involves shape coding (for arbitrarily shaped VOs) and motion adaptive DCT). The basic coding structure involves shape coding (for arbitrarily shaped VOs) and motion compensation as well as DCT-based texture coding (using standard $8 \times 8$ DCT or shape adaptive DCT).

An important advantage of the content-based coding approach MPEG-4 is that the compression efficiency can be significantly improved for some video sequences by using appropriate and dedicated object-based motion prediction "tools" for each object in a scene.

A number of motion prediction techniques can be used to allow efficient coding and flexible presentation of the objects [5]:

• Standard $8 \times 8$ or $16 \times 16$ pixel block-based motion estimation and compensation, with up to ¼ pel accuracy.

Global Motion Compensation (GMC) for video objects: Encoding of the global motion for a object using a small number of parameters. GMC is based on global motion estimation, image warping, motion trajectory coding and texture coding for prediction errors.
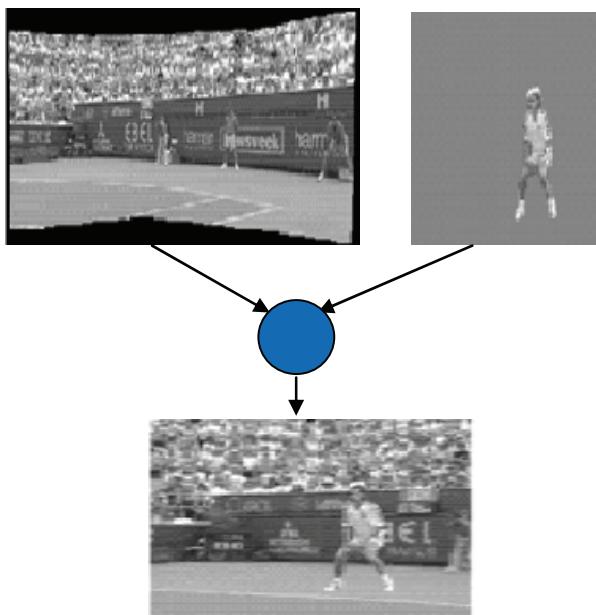
• Global motion compensation based for static "sprites". A static sprite is a possibly large still image, describing panoramic background. For each consecutive image in a sequence, only 8 global motion parameters describing camera motion are coded to reconstruct the object. These parameters represent the appropriate affine transform of the sprite transmitted in the first frame.

- Quarter Pel Motion Compensation enhances the precision of the motion compensation scheme, at the cost of only small syntactical and computational overhead. A accurate motion description leads to a smaller prediction error and, hence, to better visual quality.

- Shape-adaptive DCT: In the area of texture coding, the shape-adaptive DCT (SA-DCT) improves the coding efficiency of arbitrary shaped objects. The SA-DCT algorithm is based on predefined orthonormal sets of one-dimensional DCT basis functions.

**Figure 3** depicts the basic concept for coding an MPEG-4 video sequence using a sprite panorama image. It is assumed that the foreground object (tennis player, image top right) can be segmented from the background and that the sprite panorama image can be extracted from the sequence prior to coding. (A sprite panorama is a still image that describes as a static image the content of the background over all frames in the sequence). The large panorama sprite image is transmitted to the receiver only once as first frame of the sequence to describe the background - the sprite remains is stored in a sprite buffer. In each consecutive frame, only the camera parameters relevant for the background are transmitted to the receiver. This allows the receiver to reconstruct the background image for each frame in the sequence based on the sprite [5]. The moving foreground object is transmitted separately as an arbitrary-shape video object. The receiver composes both the foreground and background images to reconstruct each frame (bottom picture in figure below). For low delay applications, it is possible to transmit the sprite in multiple smaller pieces over consecutive frames or to build up the sprite at the decoder progressively.

Subjective evaluation tests within MPEG have shown



**Figure 3. Sprite coding of video sequence**.

that the combination of these techniques can result in a bit stream saving of up to 50% compared with the version 1, depending on content type and data rate.

### 3.3. MPEG-4 (Part-10) Avc/H.264

The intent of the H.264/AVC project was to create a standard capable of providing good video quality at substantially lower bit rates than previous standards without increasing the complexity of design so much that it would be impractical or excessively expensive to implement. An additional goal was to provide enough flexibility to allow the standard to be applied to a wide variety of applications on a wide variety of networks and systems, including low and high bit rates, low and high-resolution video, packet networks. The H.264 standard is a "family of standards": the members of which are the profiles described below.

A specific decoder decodes at least one, but not necessarily all, profiles. The decoder specification describes which of the profiles can be decoded. Scalable video coding as specified. H.264/AVC allows the construction of bitstreams that contain sub-bitstreams that conform to H.264/AVC. For temporal bitstream scalability, *i.e.*, the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, *i.e.* the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL (Network Abstraction Layer) removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, *i.e.*, the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding.

### 3.4. MPEG-21 Multimedia Framework

MPEG-21 framework is based on two essential concepts: the definition of a fundamental unit of distribution and transaction (the Digital Item) and the concept of Users interacting with Digital Items. The Digital Items can be considered the "what" of the Multimedia Framework (e.g., a video collection, a music album) and the Users can be considered the "who" of the Multimedia Framework. In practice, a Digital Item is a combination of resources, metadata, and structure. The resources are the individual assets or (distributed) resources. The metadata comprises informational data about or pertaining to the Digital Item as a whole or to the individual resources included in the Digital Item.

Finally, the structure relates to the relationships among the parts of the Digital Item, both resources and metadata. Within MPEG-21, a resource is defined as an individu-

ally identifiable asset such as a video or audio clip, an image, or a textual asset. Hence, terms like multimedia content, media data, image, graphics, video, movie, visual content, audio data, speech content, etc. become unnecessary and should be avoided in the context of MPEG-21 when referring to resources with no specific context. Within any system (such as MPEG-21) that proposes to facilitate a wide range of actions involving "Digital Items", there is a need for a very precise description for defining exactly what constitutes such an "item".

Clearly, there are many kinds of content, and probably just as many possible ways of describing it to reflect its context of use. This presents a strong challenge to lay out a powerful and flexible model for Digital Items, which can accommodate the myriad forms that content can take (and the new forms it will assume in the future). Such a model is only truly useful if it yields a format that can be used to represent any Digital Items defined within the model unambiguously and communicate them, and information about them, successfully. The Digital Item Declaration specification (part 2 of ISO/IEC 21000) provides such flexibility for representing Digital Items. Part 7 of the MPEG-21 standard Digital Item Adaptation (D-IA) have been created mainly to address the usage environment description problem [6].

### 3.5. Web Edition Content Description

Web edition is a database-grounded web content management system where in a central database; it stores all "content information" including text, images, graphics, flash movies, style sheets, etc. The plus point of this system is that the page itself does not have to be edited if the content has to be changed. As an alternative, only the database entry has to be changed and the corresponding page re-saved. In addition, the content stored in the database can be quickly searched or used in other ways. Therefore, it can be added that a central database contains all the content on the web site user wish to manage with web edition. However, user hardly notices the database, because it runs initially in the background. Users only see their web sites and the input boxes on the actual web pages. The web edition program only has to be installed on the provider's (ISP) server once [7].

### 4. Team Spirit Engine

Team Spirit® **Figure 4** Mobile is a comprehensive voice and video engine SDK, including a library of highly optimized low-MIPS consuming voice and video codecs to enable mobile video over IP communication on a broad range of handsets, speech enhancement and network adaptation algorithms [8]. To speed up deployment and secure flawless voice quality on every handset, the engine can use tuning Wizard for automatic adjustment of AEC

parameters. MHz-efficiency and Wi-Fi adaptation are extremely important for resource constrained mobile devices, intended for usage in public Wi-Fi networks, more susceptible to delays, jitter, packet loss, aggregated echo, noise and congestion. The SPIRIT's Engine handles Wi-Fi-specific network impairments and makes video transmission over Wi-Fi/WiMAX more stable and reliable. This Engine can be run even on 200 MHz processors. Simultaneously, the Engine supports both software and hardware accelerators inside popular application processors to deliver CIF quality video on mobile devices.

## 5. Team Spirit Model for Video Customization

Starting with the more traditional coding tools such as MPEG-1, and MPEG-2, the recent scalable video coding tools is MPEG-4 Fine-Grain-Scalability (FGS), and passing through web edition content description.MPEG standardization culminates with the MPEG-21 multimedia framework which offers a wrapper to allow all the pieces in a multimedia customization chain to integrate and interact with each other.
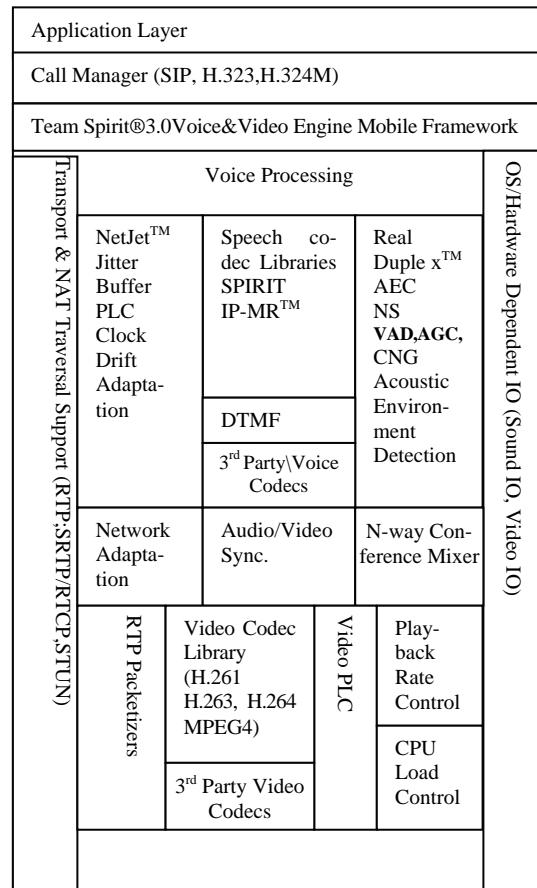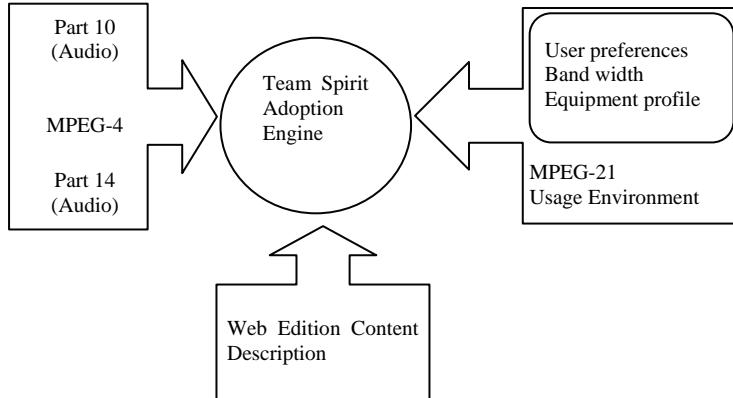


**Figure 4. Layered architecture of team spirit.**

**Figure 5. Team spirit model.**

**Figure 5** depicts a possible configuration of a multimedia customization chain using all MPEG standards Basically adaptation engine require three inputs, *i.e.,* firstly, MPEG-4 AVC and file format for audio and video processing so that whatever the audio and video requirement like formatting, compression, coding and other essentials regarding both. Secondly, content description web Edition is a database-grounded web content management system. Web edition stores all "content information". Additionally, the content stored in the database can be rapidly searched or used in other ways. Lastly, at the end user side, the MPEG-21 UED describes the environment (terminal, network, natural environment, and user preferences) where the content is to be consumed.

When the user performs a query or defines a choice, the UED, thus enabling a customizing application to explore this information to create the right content variation to provide the best possible user experience, accompanies request. Finally, the TSM adaptation engine at the center of in Figure E is responsible for matching the user query either by selecting the most adequate available variation, or by performing some adaptation. As processing a user query, the customizing application creates an adapted variation of the Item to be sent to the user-the new variation and its corresponding description may be added to the resources available at the server. The user query response may be delivered through a network, eventually using a real-time connection. In this case, the streaming module will stream the scalable or non-scalable content to the user; in the case real-time transcoding is been performed, it may happen that real time adjustments to the transcoding process are implemented using measures which characterize, for example, the network fluctuations. This is how the customization is achieved.

## 6. Conclusions and Future Work

In this paper, we present a team spirit model for efficient video customization using MPEG standards and Web

edition content description. The MPEG-4 (part-10) is used for advance video coding and MPEG-4 (part-14) for file formatting. For the usage environment description, we have MPEG-21 standards. Hence, these standards act as a base for development of a TSM system.

The TSM systems are capable of managing and customizing content efficiently so that any information can be delivered to different terminals and networks like for television, phones and over other communication media where video customization is needed. In future, after adding some more functionality to it, the model is capable to deploy all essential features related to video delivery in smart phones, iPhones etc.

## 7. References

[1]  S. Kim and Y. Yoon, "Video Customization System Using MPEG Standards," *International Conference on Multimedia and Ubiquitous Engineering*, 2008.

[2]  S. Kim and Y. Yoon, "Universal Multimedia Access Model for Video Delivery," *The 3rd International Conference on Grid and Pervasive Computing - Workshops*, 2008.

[3]  B. S. Manjunath, P. Salembier and T. Sikora, "Introduction to MPEG 7: Multimedia Content Description Language," Wiley, New York, 2002.

[4]  "MPEG Requirement Group, MPEG-21 Multimedia Framework," *Part 1: Vision, Technologies and Strategy, Proposed Draft Technical Report*, 2nd Edition, MPEG Waikaloa Meeting, USA, December 2003.

[5]  R. Koenen, "MPEG Requirement Group, MPEG-4 Overview," V.21–Jeju Version, March 2002.

[6]  "MPEG MDS Group, Multimedia Framework," *Part 7: Digital Item Adaptation, Final Draft International Standard*, MPEG Waikaloa Meeting, USA, December 2003.

[7]  http://www.macupdate.com/info.php/id/7902/webedition

[8]  "Spirit DSP, Voice and Engine Expert Group," *Datasheet*, Juniper Research, February 2002.

Scientific
Research

# A Novel Approach towards Cost Effective Region-Based Group Key Agreement Protocol for Ad Hoc Networks Using Elliptic Curve Cryptography

**Krishnan Kumar[1], J. Nafeesa Begum[1], V. Sumathy[2]**
[1]*Government College of Engineering, Bargur, India*
[2]*Government College of Technology, Coimbatore, India*
*Email*: *pkk_kumar@yahoo.com, nafeesa_jeddy@yahoo.com, sumi_gct2001@yahoo.co.in*

## Abstract

This paper addresses an interesting security problem in wireless ad hoc networks: the dynamic group key agreement key establishment. For secure group communication in an ad hoc network, a group key shared by all group members is required. This group key should be updated when there are membership changes (when the new member joins or current member leaves) in the group. In this paper, we propose a novel, secure, scalable and efficient region-based group key agreement protocol for ad hoc networks. This is implemented by a two-level structure and a new scheme of group key update. The idea is to divide the group into subgroups, each maintaining its subgroup keys using group elliptic curve diffie-hellman (GECDH) Protocol and links with other subgroups in a tree structure using tree-based group elliptic curve diffie-hellman (TGECDH) protocol. By introducing region-based approach, messages and key updates will be limited within subgroup and outer group; hence computation load is distributed to many hosts. Both theoretical analysis and experimental results show that this Region-based key agreement protocol performs well for the key establishment problem in ad hoc network in terms of memory cost, computation cost and communication cost.

## 1. Introduction

Wireless networks are growing rapidly in recent years. Wireless technology is gaining more and more attention from both academia and industry. Mostly wireless networks are used today e.g. the cell phone networks and the 802.11 wireless LAN, are based on the wireless network model with pre-existing wired network infrastructures. Packets from source wireless hosts are received by near-by base stations, then injected into the underlying network infrastructure and then finally transferred to destination hosts.

Another wireless network model, which is in active research, is the ad-hoc network. This network is formed only by mobile hosts and requires no pre-existing network infrastructure. Hosts with wireless capability form an ad-hoc network, some mobile hosts work as routers to relay packets from source to destination. It is very easy and economic to form an ad-hoc network in real time.

Ad-hoc network is ideal in situations like battlefield or rescuer area where fixed network infrastructure is very hard to deploy.

A mobile ad hoc network is a collection of autonomous nodes that communicate with each other. Mobile nodes come together to form an ad hoc group for secure communication purpose. A key distribution system requires a trusted third party that acts as a mediator between nodes of the network. Ad-hoc networks characteristically do not have a trusted authority. Group key agreement means that multiple parties want to create a common secret key to be used to exchange information securely. Furthermore, group key agreement also needs to address the security issue related to membership changes due to node mobility. The membership change requires frequent changes of group key. This can be done either periodically or updating every membership changes. The changed group key ensures backward and forward secrecy. With frequent changes in group memberships, the

recent researches began to pay more attention on the efficiency of group key update. Recently, collaborative and group–oriented applicative situations like battlefield, conference room or rescuer area in mobile ad hoc networks have been a current research area. Group key agreement is a building block in secure group communication in ad hoc networks. However, group key agreement for large and dynamic groups in ad hoc networks is a difficult problem because of the requirements of scalability and security under constraints of node available resources and node mobility.

We propose a communication and computation efficient group key agreement protocol in ad-hoc network. In large and high mobility ad hoc networks, it is not possible to use a single group key for the entire network because of the enormous cost of computation and communication in rekeying. So, we divide the group into several subgroups; let each subgroup has its subgroup key shared by all members of the subgroup. Each group has sub group controller node and gateway node, in which the sub group controller node is controller of subgroup and gateway node is controller among subgroups. Let each gateway member contributes a partial key to agree with a common outer group key among the subgroups.

The contribution of this work includes

1) In this paper, we propose a new efficient method for solving the group key management problem in ad-hoc network. This protocol provides efficient, scalable and reliable key agreement service and is well adaptive to the mobile environment of ad-hoc network.

2) We introduce the idea of subgroup and subgroup key and we uniquely link all the subgroups into a tree structure to form an outer group and outer group key. This design eliminates the centralized key server. Instead, all hosts work in a peer-to-peer fashion to agree on a group key. We use region-based group key agreement (RBGKA) as the name of our protocol. Here we propose a region based group key agreement protocol for ad hoc networks using elliptic curve cryptography called region-based GECDH and TGECDH protocol.

3) We design and implement region-based group key agreement protocol using Java and conduct extensive experiments and theoretical analysis to evaluate the performance like memory cost, communication cost and computation cost of our protocol for Ad-Hoc network.

The rest of the paper is as follows, Section 2 presents the elliptic curve cryptography schemes. Section 3 presents the proposed schemes. Section 4 describes the performance analysis and finally Section 5 concludes the paper.

## 2. Elliptic Curve Cryptography

ECC [1-3] has become the cryptographic choice for ad hoc networks and communication devices due to its size and efficiency benefits. Elliptic curve cipher uses very small keys and is computationally very efficient, which makes it ideal for the smaller, less powerful devices being used today by majority of individuals to access network services. The elliptic curve cryptosystem (ECCS) is a crypto-algorithm method of utilizing a discrete logarithm problem (DLP) over the points on an elliptic curve.

An elliptic curve is usually defined over two finite fields: the prime finite field $F_p$ containing $p$ elements and the characteristic finite field containing $2^m$ elements. This paper focuses on the prime finite field. Let $F_p$ *be* a prime finite field, $p$ is an odd prime number, and $a,b \ \varepsilon \ F_p$ satisfy $4a^3 + 27b^2 \neq 0 (\text{mod } p)$, then an elliptic curve $E(F_p)$ over $F_p$ defined by the parameters $a,b \ \varepsilon \ F_p$ consists of the set of solutions or points $P = (x, y)$ for $x, y \ \varepsilon \ F_p$ to the equation:

$$y^2 = x^3 + ax + b (\text{mod } p) \qquad (1)$$

The equation $y^2 = x^3 + ax + b$ (mod $p$) is called the defining equation of $E(F_p)$. For a given point $P = (x_P, y_P)$, $x_P$ is called the *x*-coordinate of $P$, and $y_P$ is called the *y*-coordinate of $P$.

Cryptographic schemes based on ECC rely on scalar multiplication of elliptic curve points. Given an integer $k$ and a point $P \ \varepsilon \ E(F_p)$, scalar multiplication is the process of adding $P$ to itself $k$ times. The result of this scalar multiplication is denoted $k * P$ or $kP$. Scalar multiplication of elliptic curve points can be computed efficiently using the addition rule together with the double-and-add algorithm or one of its variants.

### 2.1. Two Parties Elliptic Curve Diffie-Hellman Protocol

Similar to DLP-based Diffie-Hellman key exchange agreement, a key exchange between users $A$ and $B$ using elliptic curve Diffie-Hellman (ECDH) [2,3] can be accomplished as follows:

1) $A$ selects an integer $n_A$ less than $p$, this is $A$'s private key. $A$ then generates a public key $P_A = n_A * G$; the public key is a point in $Ep(a, b)$.

2) $B$ similarly selects a private key $n_B$ and computes a public key $P_B = n_B * G$.

3) $A$ and $B$ generates the secret key $K = n_A * P_B$, $K = n_B * P_A$ respectively.

The two calculations in step 3 produce the same result because

$$n_A * P_B = n_A * (n_B * G) = n_B * (n_A * G) = n_B * P_A.$$

The secret key $K$ is a point in the elliptic curve.

# 3. Proposed Scheme

## 3.1. Motivation

There has been a growing demand in the past few years for security in collaborative environments deployed for emergency services where our approach can be carried out very efficiently are shown in **Figure 1**. Confidentiality becomes one of the top concerns to protect group communication data against passive and active adversaries. To satisfy this requirement, a common and efficient solution is to deploy a group key shared by all group application participants. Whenever a member leaves or joins the group, or whenever a node failure or restoration occurs, the group key should be updated to provide forward and backward secrecy. Therefore, a key management protocol that computes the group key and forwards the rekeying messages to all legitimate group members is central to the security of the group application.

In many secure group applications [4,5], a Region based contributory GKA schemes may be required. In such cases, the group key management should be both efficient and fault-tolerant. In this paper, we describe a military scenario (**Figure 2**). A collection of wireless mobile devices are carried by soldiers or battlefield tanks. These mobile devices cooperate in relaying packets to dynamically establish routes among themselves to and processing

form their own network "on the fly". However, all nodes except the one with the tank, have limited battery power capacities. For the sake of power-consumption and computational efficiency, the tank can work as the Gateway member while a contributed group key management scheme is deployed.

## 3.2. System Model

### 3.2.1. Overview of Region-Based Group Key Agreement Protocol

The goal of this paper is to propose a communication and computation efficient group key establishment protocol in ad-hoc network. The idea is to divide the multicast group into several subgroups, let each subgroup has its subgroup key shared by all members of the subgroup. Each Subgroup has subgroup controller node and a gateway node, in which Subgroup controller node is the controller of subgroup and a Gateway node is controller of subgroups controller.

For example, in **Figure 3**, all member nodes are divided into number of subgroups and all subgroups are linked in a tree structure as shown in **Figure 4**.
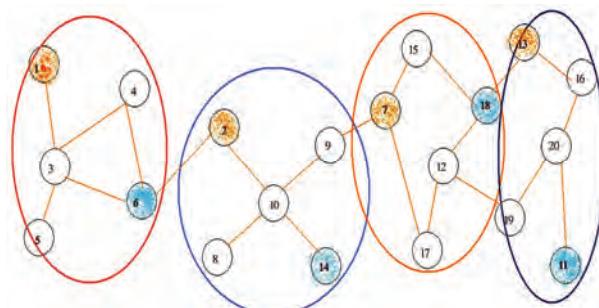
One of the members in the subgroup is subgroup controller. The last member joining the group acts as a subgroup controller. Each outer group is headed by the outer group controller. In each group, the member with high processing power, memory, and Battery power acts as a
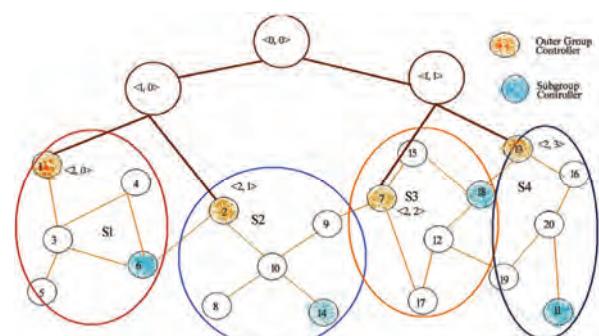


**Figure 1. Secure group applications.**



**Figure 2. Battlefield scenario.**



**Figure 3. Members of group are divided into subgroups.**



**Figure 4. Subgroups link in a tree structure.**

gateway member. Outer group messages are broadcast through the outer group and secured by the outer group key while subgroup messages are broadcast within the subgroup and secured by subgroup key.

Let $N$ be the total number of group members, and $M$ be the number of the subgroups in each subgroup, then there will be $N/M$ subgroups, assuming that each subgroup has the same number of members.

There are two shared keys in the Region-Based Group Key Agreement Scheme:

1) Outer Group Key (KG) is used to encrypt and decrypt the messages broadcast among the subgroup controllers.

2) The Subgroup Key (KR) is used to encrypt and decrypt the Sub Group level messages broadcast to all sub group members.

In our Region-Based Key Agreement protocol shown in **Figure 5**. a Subgroup Controller communicates with the member in the same region using a Regional key (*i.e.* Subgroup key ) KR. The Outer Group key KG is derived from the Outer Group Controller. The Outer Group Key KG is used for secure data communication among subgroup members. These two keys are rekeyed for secure group communications depending on events that occur in the system. The layout of the network is as shown in below **Figure 5**.

Assume that there are total $N$ members in Secure Group Communication. After sub grouping process (algorithm 1), there are $S$ subgroups $M_1, M_2 \ldots M_s$ with $n_1, n_2 \ldots n_s$ members.

---

**Algorithm 1. Region-Based Key Agreement Protocol**

1) The Subgroup Formation
The number of members in each subgroup is

**N / S < 100.**

where,
  N – is the group size. and
    S – is the number of subgroups.
 Assuming that each subgroup has the same number of members.
 2) The Contributory Key Agreement protocol is implemented among the group members. It consists of three stages.
  a. To find the Subgroup Controller for each subgroups.
  b. GECDH protocol is used to generate one common key for each subgroup headed by the subgroup controller.
  c. Each subgroup gateway member contributes partial keys to generate a one common backbone key (*i.e.* Outer group Key (KG)) headed by the Outer Group Controller using TGECDH protocol.
 3) Each Group Controller (sub /Outer) distributes the computed public key to all its members.

---
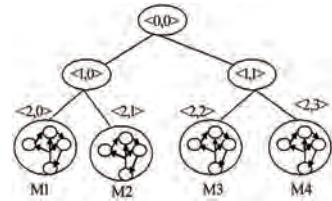


**Figure 5. Region based group key agreement.**

A Regional key KR is used for communication between a subgroup controller and the members in the same region. The Regional key KR is rekeyed whenever there is a membership change event and subgroup joins/leaves and member failure. The Outer Group key KG is rekeyed whenever there is a join/leave subgroup controllers and member failure to preserve secrecy.

The members within a subgroup use Group Elliptic Curve Diffie-Hellman (GECDH) Contributory Key Agreement. Each member within a subgroup contributes his share in arriving at the subgroup key. Whenever membership changes occur, the subgroup controller or previous member initiates the rekeying operation.

The gateway member initiates communication with the neighboring member belonging to another subgroup and mutually agree on a key using Tree-Based Group Elliptic Curve Diffie-Hellman (TGECDH) Contributory Key Agreement protocol to be used for inter subgroup communication between the two subgroups. Any member belonging to one subgroup can communicate with any other member in another subgroup through this member as the intermediary. In this way adjacent subgroups agree on outer group key. Whenever membership changes occur, the outer group controller or previous group controller initiates the rekeying operation.
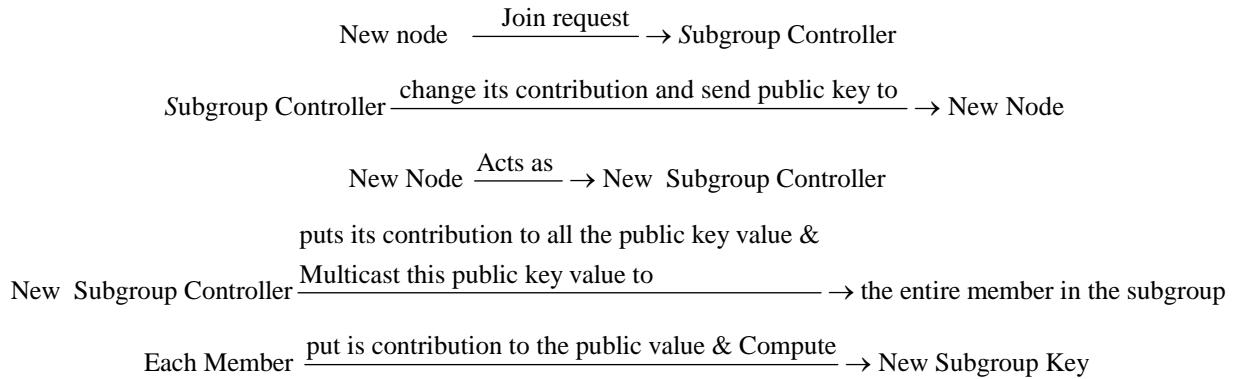
Here, we prefer the subgroup key to be different from the key for backbone. This difference adds more freedom of managing the dynamic group membership. In addition by using this approach can potentially save the communication and computational cost.

## 3.3. Network Dynamics

The network is dynamic in nature. Many members may join or leave the group. In such case, a group key management system should ensure that backward and forward secrecy is preserved.

### 3.3.1. Member Join
When a new member joins, it initiates communication with the subgroup controller. After initialization, the subgroup controller changes its contribution and sends public key to this new member. The new member receives the public key and acts as a group controller by initiating the rekeying operations for generating a new key for the subgroup. The rekeying operation is as follows.

$$\text{New node} \xrightarrow{\text{Join request}} \textit{S}\text{ubgroup Controller}$$

$$\textit{S}\text{ubgroup Controller} \xrightarrow{\text{change its contribution and send public key to}} \text{New Node}$$

$$\text{New Node} \xrightarrow{\text{Acts as}} \text{New Subgroup Controller}$$

$$\text{New Subgroup Controller} \xrightarrow[\text{Multicast this public key value to}]{\text{puts its contribution to all the public key value \&}} \text{the entire member in the subgroup}$$

$$\text{Each Member} \xrightarrow{\text{put is contribution to the public value \& Compute}} \text{New Subgroup Key}$$

### 3.3.2. Member Leave

3.3.2.1. When a Subgroup member Leaves
When a member leaves the Subgroup Key of the subgroup to which it belongs must be changed to preserve the forward secrecy. The leaving member informs the subgroup controller. The subgroup controller changes its private key value, computes the public value and broadcasts the public value to all the remaining members. Each member performs rekeying by putting its contribution to public value and computes the new Subgroup Key. The rekeying operation is as follows.
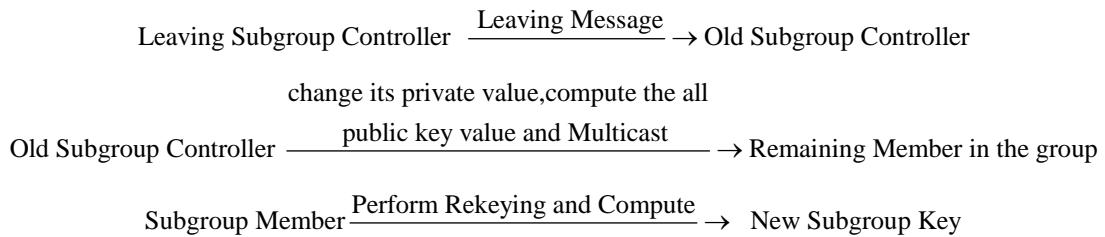
$$\text{Leaving Node} \xrightarrow{\text{Leaving Message}} \text{Subgroup Controller}$$

$$\text{Subgroup Controller} \xrightarrow[\text{Multicast the public key value to}]{\text{changes its private key value, compute the public key value and}} \text{All the remaining Member}$$

$$\text{Each Member} \xrightarrow{\text{Performs Rekeying and Compute}} \text{New Subgroup Key}$$

3.3.2.2. When Subgroup Controller Leaves
When the Subgroup Controller leaves, the Subgroup key used for communication among the subgroup controller needs to be changed. This Subgroup Controller informs the previous Subgroup Controller about its desire to leave the subgroup which initiates the rekeying procedure. The previous subgroup controller now acts as a Subgroup controller. This Subgroup controller changes its private contribution value and computes all the public key values and broadcasts to all the remaining members of the group. All subgroup members perform the rekeying operation and compute the new subgroup key. The rekeying operation is as follows.
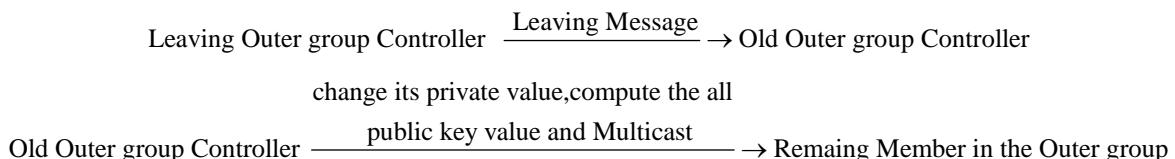
$$\text{Leaving Subgroup Controller} \xrightarrow{\text{Leaving Message}} \text{Old Subgroup Controller}$$

$$\text{Old Subgroup Controller} \xrightarrow[\text{public key value and Multicast}]{\text{change its private value,compute the all}} \text{Remaining Member in the group}$$

$$\text{Subgroup Member} \xrightarrow{\text{Perform Rekeying and Compute}} \text{New Subgroup Key}$$
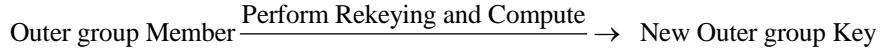
3.3.2.3. When Outer Group Controller Leaves
When the Outer group Controller leaves, the Outer group key used for communication among the Outer group need to be changed. This Outer group Controller informs the previous Outer group Controller about its desire to leave the Outer group which initiates the rekeying procedure. The previous Outer Group controller now becomes the New Outer group controller. This Outer group controller changes its private contribution value and computes the public key value and broadcast to the entire remaining member in the group. All Outer group members perform the rekeying operation and compute the new Outer group key. The rekeying operation is as follows.

$$\text{Leaving Outer group Controller} \xrightarrow{\text{Leaving Message}} \text{Old Outer group Controller}$$

$$\text{Old Outer group Controller} \xrightarrow[\text{public key value and Multicast}]{\text{change its private value,compute the all}} \text{Remaing Member in the Outer group}$$

$$\text{Outer group Member} \xrightarrow{\text{Perform Rekeying and Compute}} \text{New Outer group Key}$$

### 3.3.2.4. When Gateway Member Leaves

When a gateway member leaves the subgroup, it delegates the role of the gateway to the adjacent member having high processing power, memory, and Battery power and acts as a new gateway member. Whenever the gateway member leaves, all the two keys should be changed. These are

   i. Outer group key among the subgroup.

   ii. Subgroup key within the subgroup.

   In this case, the subgroup controller and outer group controller perform the rekeying operation. Both the Controller leaves the member and a new gateway member is selected in the subgroup, performs rekeying in the subgroup. After that, it joins in the outer group. The procedure is same as joining the member in the outer group.

$$\text{Source Member} \xrightarrow{\text{E}_{\text{KR}}[\text{Message}] \ \& \ \text{Multicast}} \text{Destination Member}$$

$$\text{Destination Member} \xrightarrow{\text{D}_{\text{KR}}\left[\text{E}_{\text{KR}}[\text{Message}]\right]} \text{Original Message}$$

### 3.4.2. Communication among the Subgroup

The sender member encrypts the message with the subgroup key (KR) and multicasts it to all members in the subgroup. One of the members in the subgroup acts as a gate way member. This gateway member decrypts the message with subgroup key and encrypts with the outer group key (KG) and multicast to the entire gateway member among the subgroup. The destination gateway

## 3.4. Communication Protocol

The members within the subgroup have communication using subgroup key. The communication among the subgroup members takes place through the gateway member.

### 3.4.1. Communication within the Subgroup

The sender member encrypts the message with the subgroup key (KR) and multicasts it to all member in the subgroup. The subgroup members receive the encrypted message, perform the decryption using the subgroup key (KR) and get the original message. The communication operation is as follows.

member first decrypts the message with outer group key. Then encrypts with subgroup key and multicast it to all members in the subgroup. Each member in the subgroup receives the encrypted message and performs the decryption using subgroup key and gets the original message. In this way the region-based group key agreement protocol performs the communication. The communication operation is as follows.

$$\text{Source Member} \xrightarrow{\text{E}_{\text{KR}}[\text{Message}] \ \& \ \text{Multicast}} \text{Gateway Member}$$

$$\text{Gateway Member} \xrightarrow{\text{D}_{\text{KR}}\left[\text{E}_{\text{KR}}[\text{Message}]\right]} \text{Original Message}$$

$$\text{Gateway Member} \xrightarrow{\text{E}_{\text{KG}}[\text{Message}] \ \& \ \text{Multicast}} \text{Gateway Member [ Among Subgroup]}$$

$$\text{Gateway Member} \xrightarrow{\text{D}_{\text{KG}}\left[\text{E}_{\text{KG}}[\text{Message}]\right]} \text{Original Message}$$

$$\text{Gateway Member} \xrightarrow{\text{E}_{\text{KR}}[\text{Message}] \ \& \ \text{Multicast}} \text{Destination Member}$$

$$\text{Destination Member} \xrightarrow{\text{D}_{\text{KR}}\left[\text{E}_{\text{KR}}[\text{Message}]\right]} \text{Original Message}$$

## 3.5. Applying Elliptic Curve Based Diffie-Hellman Key Exchange

### 3.5.1. Member Join

User A and user B are going to exchange their keys (**Figure 6**): Take p = 211, Ep = (0,-4), which is equivalent to the curve $y^2 = x^3 - 4$ and G = (2,2). A's private key is nA = 47568, so A's public key PA = 47568(2,2) = (206, 121),
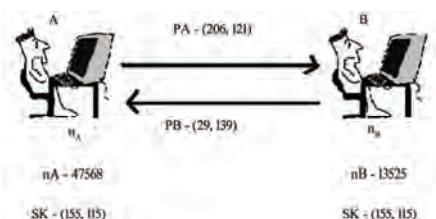


**Figure 6. User-A and User –B join the group.**
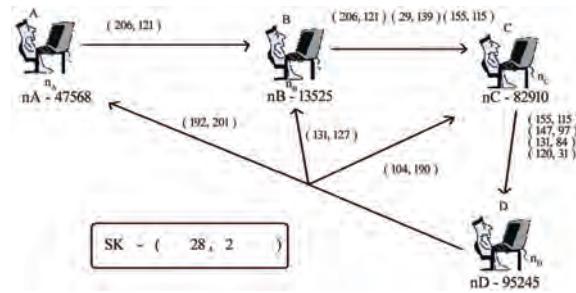
B's private key is nB = 13525,so B's public key PB = 13525(2,2) = (29,139). The group key is computed (**Figure 6**) as User A sends its public key (206,121) to user B, then user B computes their Subgroup key as nB (A's Public key) = 13525(206,121) = (**155,115**). User B sends its public key (29,139) to User A, then User A compute their Subgroup key as nA (B's Public key) = 47568(29,139) = (**120,180**)

When User C is going to join in the group, C's private key becomes nC = 82910. Now, User C becomes a Subgroup Controller. Then, the key updating process will begin as follows: The previous Subgroup Controller User B sends the intermediate key as (B's Public key $ A's Public Key $ Group key of A&B) = ((29,139) $ (206,121) $ (155,115)) User C separates the intermediate key as B's Public key, A's Public Key and Group key of A&B = (29,139), (206,121) and (155,115). Then, User C generates the new Subgroup key as nC (Subgroup key of A&B) = 82910(155,115) = (**120,31**). Then, User C broadcasts the intermediate key to User A and User B. That intermediate key is ((Public key of B & C) $ (Public key of A & C)) = ((131,84) $ (147, 97)). Now, User B extracts the value of public key of A & C from the value sent by User C. Then User B compute the new Subgroup key as follows: nB (Public key of A&C) = 13525 (147,97) = (**120, 31**). Similarly, User A extracts the value of public key of B & C from intermediate key, sent by User C. Then User A compute the new Subgroup key as follows: nA (public key of B&C) = 47568(131,84) = (**120,31**). Therefore, New Subgroup Key of A, B and C = (**120, 31**) as shown in the **Figure 7**.
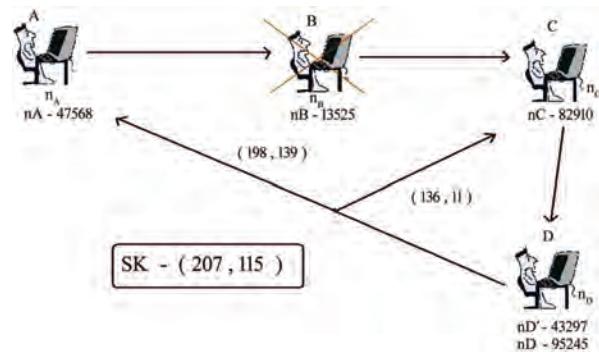
The same procedure is followed when User D joins as shown in the **Figure 8**.

### 3.5.2. Member Leave

When a user leaves (**Figure 9**) from the Subgroup, then the Subgroup controller changes its private key. After that, it broadcasts its new public key value to all users in the Subgroup. Then, new Subgroup key will be generated. Let us consider, User B is going to leave, then the Subgroup Controller D changes its private key nD' = 43297, so public key of User A & User C = (198,139) $(136,11). Then the new Subgroup Key generated is = 43297(198,139) = (**207,115**). Then, User A & User C



**Figure 8. User-D joins in the group.**
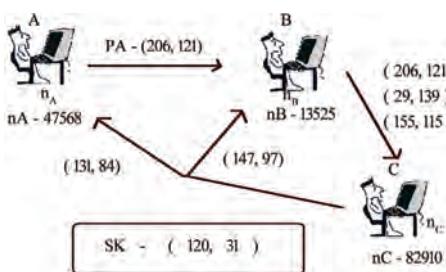


**Figure 9. User –B leaves from the group.**
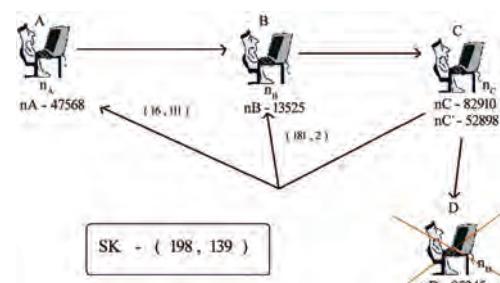
computes the new Subgroup Key by using new public key. Therefore, the new Subgroup Key is (**207,115**).

### 3.5.3. Group Controller Leave

When a Subgroup controller leaves (**Figure 10**) from the group, then the previous Subgroup controller changes its private key. After that, it broadcasts its new public key value to all users in the group. Then, new Subgroup key will be generated. Let us consider, Subgroup Controller User D going to leave, then the previous Subgroup controller User C act as Subgroup Controller and changes its private key nC' = 52898, and computes the public key of B&C $ A&C = (16,111)$(181,2). Then the new Subgroup Key generated is = 52898(21,103) = (**198,139**). Then, User A & User B compute the new Subgroup Key by using new public key. Therefore, the new Subgroup Key is (**198,139**).



**Figure 7. User - C joins in the group.**



**Figure 10. Group controller leaves from the group.**

## 3.6. Tree-Based Group Elliptic Curve Diffie-Hellman Protocol

The proposed protocol (**Figure 11**), Tree-based group Elliptic Curve Diffie-Hellman (TGECDH), is a variant of TGDH based on ECDLP. In TGECDH, a binary tree is used to organize group members. The nodes are denoted as <l, v>, where $0 <= v <= 2l. - 1$ since each level l hosts at most 2l. nodes. Each node <l, v> is associated with the key K<l,v> and the blinded key BK<l,v> = F(K<l,v>) where the function F() is scalar multiplication of elliptic curve points in prime field. Assuming a leaf node <l, v> hosts the member Mi, the node <l, v> has Mi's session random key K<l,v>. Furthermore, the member Mi at node <l. v> knows every key in the key-path from <l, v> to <0, 0>. Every key K<l,v> is computed recursively as follows:

$$K_{<l,v>} = K_{<l+1,2v>} BK_{<l+1,2v+1>} \bmod p$$
$$= K_{<l+1,2v+1>} BK_{<l+1,2v>} \bmod p$$
$$= K_{<l+1,2v>} K_{<l+1,2v+1>} G \bmod p$$
$$= F(K_{<l+1,2v>} K_{<l+1,2v+1>})$$

It is not necessary for the blind key BK<l,v> of each node to be reversible. Thus, simply use the x-coordinate of K<l,v> as the blind key. The group session key can be derived from K<0,0>. Each time when there is member join/leave, the outer group controller node calculates the group session key first and then broadcasts the new blind keys to the entire group and finally the remaining group members can generate the group session key.
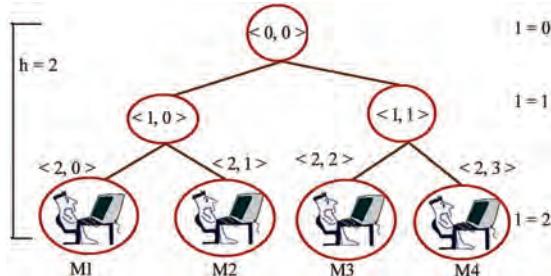


**Figure 11. Key tree.**
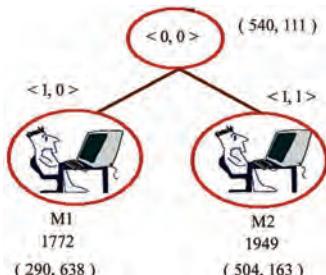


**Figure 12. User $M_1$ and $M_2$ join the group.**

### 3.6.1. When Node $M_1$&$M_2$ Join the Group

User $M_1$ and User $M_2$ are going to exchange their keys: Take **p = 751**, **Ep = (1,188)**, which is equivalent to the curve $y^2 = x^3 + x + 188$ and **G = (0,376)**. User $M_1$'s private key is **1772**, so $M_1$'s public key is (**290,638**). User $M_2$'s private key is **1949**, so $M_2$'s public key is (504,163). The Outer Group key is computed (**Figure 12**) as User $M_1$ sends its public key (**290,638**) to user $M_2$, the User $M_2$ computes their group key as PV(0,0) = $X_{co}$(PV(1,0) *PB(1,1)) and PB(0,0) = PV(0,0)*G = (**540, 111**). Similarly, User $M_2$ sends its public key (**504,163**) to user $M_1$, and then the user $M_1$ computes their group key as (**540,111**). Here, Outer Group controller is User $M_2$.

### 3.6.2. When $3^{rd}$ Node Join

When User $M_3$ joins the group, the old Outer group controller $M_2$ changes its private key value from **1949** to **2835** and passes the public key value and tree to User $M_3$. Now, $M_3$ becomes new Outer group controller. Then, $M_3$ generates the public key (**623, 52**) from its private key as **14755** and computes the Outer group key as (**664,736**) shown in **Figure 13**. $M_3$ sends Tree and public key to all users. Now, user $M_1$ and $M_2$ compute their group key. The same procedure is followed by joining the User $M_4$ as shown in **Figure 14**.
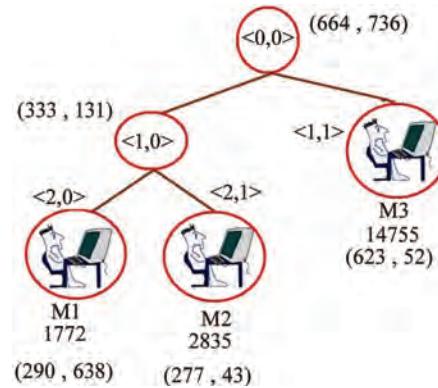


**Figure 13. User $M_3$ joins the group.**
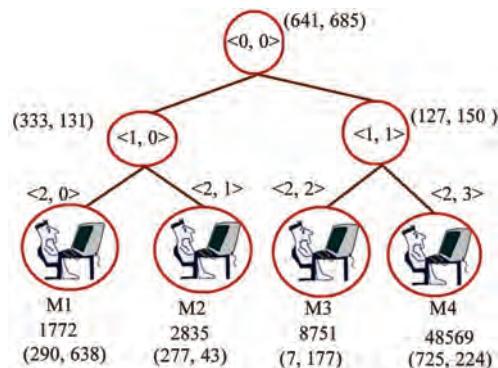


**Figure 14. User $M_4$ joins the group.**

      

### 3.6.3. Leave Protocol

There are two types of leave, 1) Gateway Member Leave and 2) Outer Group Controller Leave.

#### 3.6.3.1. Gateway Member Leave

When user $M_3$ leaves (**Figure 15**) the Outer group, then the Group controller changes its private key **48569** to **98418** and outer group key is recalculated as **(428,686)**. After that, it broadcasts its Tree and public key value to all users in the Outer group. Then, the new Outer group key will be generated by the remaining users.

#### 3.6.3.2. When an Outer Group Controller Leaves

When an Outer Group Controller Leaves (**Figure 16**) from the group, then its sibling act as a New Outer Group Controller and changes its private key value 8751 to 19478 and recalculates the group key as (681,475). After that, it broadcast its Tree and public key value to all users in the Outer group. Then, the new Outer group key will be generated by the remaining users.

## 4. Performance Analysis

The performance of the proposed scheme is analyzed in terms of the storage overhead, communication overhead and the computation overhead. The storage overhead
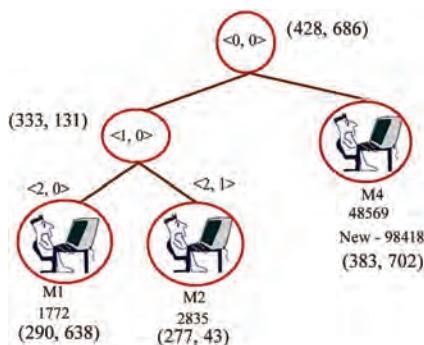


**Figure 15. User M₃ leaves from the group.**
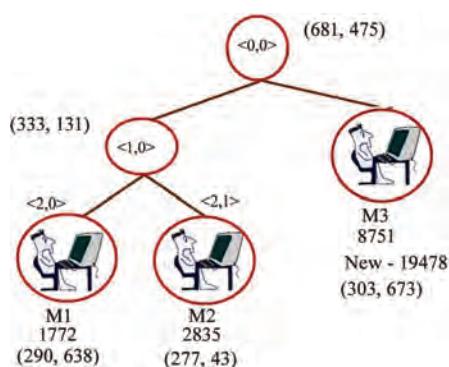


**Figure 16. Outer group controller leaves from the group.**

represents the storage capacity needed for storing the key information. The communication overhead corresponds to the rekeying messages that are transmitted for group communication. The computation overhead is the computation involved in maintaining membership changes.

### 4.1. Storage Overhead

Storage overhead can be considered as the memory capacity required for maintaining the keys, which is directly proportional to the number of members if the key size are same. In this section, the storage cost is formulated, both at gateway member and at each member. Thus our approach consumes very less memory when compared to TGDH and GDH. TGDH and GDH occupy large memory when members go on increasing. But our Region-based Approach takes very less memory even when the members get increased. Consider (**Table 1** and **Figure 17**) there are 1024 members in a group our Region-based approach consumes 10% of memory when compared to GDH and 5% when compared to TGDH. The ratio of memory occupied is very less in our approach.

**Table 1. Memory cost.**

| Protocol | | Keys | Public |
|---|---|---|---|
| GDH | Concretely | 2 | $N+1$ |
| TGDH | Per(L,V) | $L+1$ | $2N-2$ |
| | Averagely | $[\log_2 N]+1$ | $2N-2$ |
| Region Based Protocol (GECDH &TGECDH) | Member/Subgroup Controller | 2 | $X+1$ |
| | Gateway Member/Outer group Controller | $2+M$ | $X+2Y-1$ |
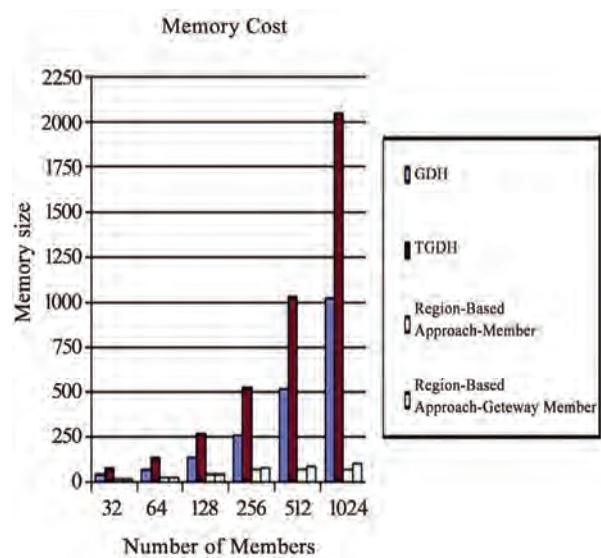


**Figure 17. Memory cost.**
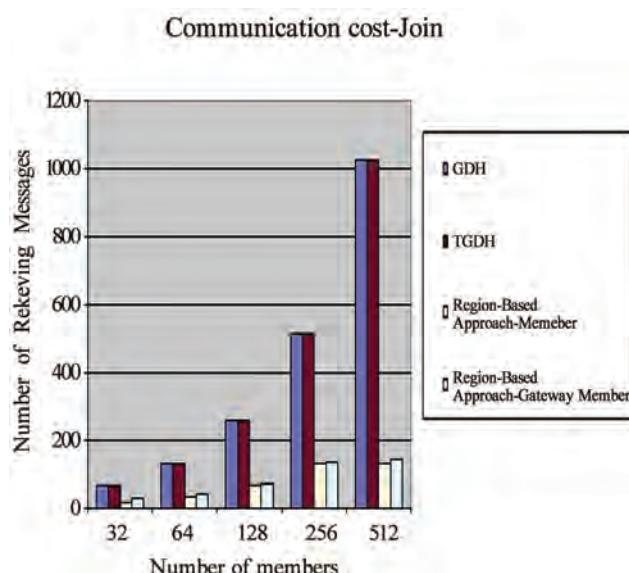
## 4.2. Communication Overhead

GDH is the most expensive protocol. TGDH consumes more bandwidth. The communication and computation of TGDH depends on trees height, balance of key tree, location of joining tree, and leaving nodes. Hence GDH has more communication efficiency than TGDH protocol. But our approach depends on the number of members in the subgroup, number of Group Controllers, and height of tree. So the amount spent on communication is very much less when compared to GDH and TGDH.
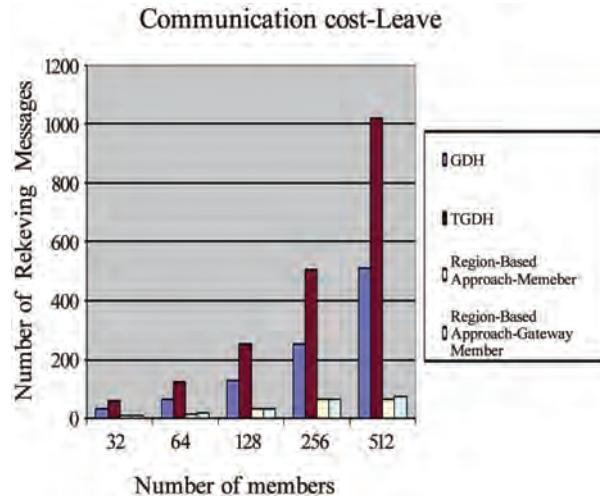
Consider (**Table 2** & **Figures 18** & **19**) there are 512 members in a group our approach consumes only 10% of Bandwidth when compared to GDH and TGDH. So our approach consumes low Bandwidth.

**Table 2. Communication cost.**

| Protocol Suite | | Protocol | Rounds | Communication Cost | |
|---|---|---|---|---|---|
| | | | | Unicast Size | Broadcast Size |
| GDH | | Join | 2 | $N+1$ | $N+1$ |
| | | Leave | 1 | 0 | $N-1$ |
| TGDH | | Join | 2 | 0 | $2N+2$ |
| | | Leave | 1 | 0 | $2N-4$ |
| Region Based Protocol (GECDH & TGECDH) | Member/ Subgroup controller | Join | 2 | $X+1$ | $X+1$ |
| | | Leave | 1 | 0 | $X-1$ |
| | Gateway Member/ Outer Group Controller | Join | 2 | $X+1$ | $X+2Y+3$ |
| | | Leave | 1 | 0 | $X+2Y-5$ |



**Figure 18. Communication cost – join.**
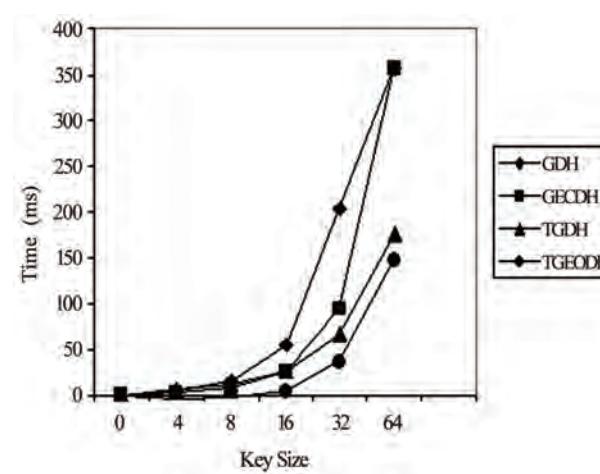


**Figure 19. Communication cost – leave.**

## 4.3. Computation Overhead
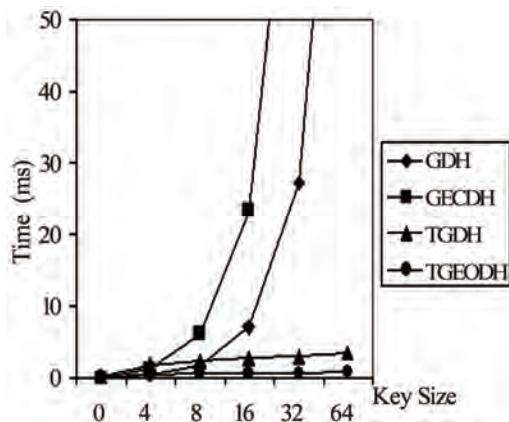
The **Figure 20** shows that GECDH and TGECDH schemes have lower computation time than Group Diffie-Hellman (GDH) schemes for member join operations. The computation time of GDH is **2.2** times that of GECDH and TGDH is **1.7** times that of TGE-CDH on average for member join operations. The computation time for member leave operations of TGE-CDH schemes are far better than group Diffie-Hellman schemes for member leave operations as shown in the **Figure 21**. Performance wise our approach leads the GDH & TGDH methods, even for the very large groups.

The performance of GECDH & TGECDH over wireless ad hoc Networks can be summarized as follows:
1) It uses smaller keys.
2) It uses less computation time than the DLP-based scheme for the same security level.



**Figure 20. Computation time for member join.**

**Figure 21. Computation time for member leave.**

3) Smaller packets are used to handle high bit error rate Wireless links.

## 5. Conclusions

In this paper, a region-based key agreement scheme has been proposed and implemented, which can enhance the secure group communication performance by using multiple group keys. In contrast to other existing schemes using only single key, the new proposed scheme exploits asymmetric key, *i.e.* an outer group key and multiple Subgroup keys, which are generated from the proposed region-based key agreement algorithm. By using a set comprising an outer group key and subgroup keys a re- gion-based scheme can be efficiently distributed for multiple secure groups. Therefore, the number of rekeying messages, computation and memory can be dramatically reduced. Compared with other schemes, the new proposed region-based scheme can significantly reduce the storage and communication overheads in the rekeying process, with acceptable computational overhead. It is expected that the proposed scheme can be the practical solution for secure group applications, especially for battlefield scenario.

## 6. References

[1] Kefa Rabah, "Theory and Implementation of Elliptic Curve Cryptography," *Journal of Applied Sciences*, Vol. 5, No. 4, 2005, pp. 604-633.

[2] W. Stallings, "Cryptography and Network Security Principles and Practices," 3rd Edition, Pearson Education.

[3] Y. Wang, B. Ramamurthy and X. K. Zou, "The Performance of Elliptic Curve Based Group Diffie-Hellman Protocols for Secure Group Communication over Ad Hoc Networks," *IEEE International Conference on Communication*, Vol. 5, 2006, pp. 2243-2248.

[4] Y. Amir, Y. Kim, C. Nita-Rotaru and G. Tsudik, "On the Performance of Group Key-Agreement Protocols," *Proceedings of the 22nd IEEE International Conference on Distributed Computing Systems*, Viena, Austria, 2002.

[5] M. Steiner, G. Tsudik and M. Waidner, "Key Agreement in Dynamic Peer Groups," *IEEE Transaction on Parallel and Distributed Systems*, Vol. 11, No. 8, 2000, pp. 769-780.

Scientific
Research

# On Channel Estimation of OFDM-BPSK and -QPSK over Generalized Alpha-Mu Fading Distribution

**Neetu Sood, Ajay K. Sharma, Moin Uddin**
*National Institute of Technology, Jalandhar, India*
*Email*: *soodn@nitj.ac.in, sharmaajayk@nitj.ac.in, director@nit.ac.in*

## Abstract

This paper evaluates the performance of OFDM-BPSK and -QPSK system in $\alpha$-$\mu$ distribution. A fading model which is based on the non-linearity present in the propagation medium is utilized here for generation of $\alpha$-$\mu$ variants. Different combinations of $\alpha$ and $\mu$ provides various fading distributions, one of which is Weibull fading. Investigations of channel estimation schemes gave an idea of further reducing the BER as to improve the performance of OFDM based systems. In flat fading environment, channel estimation is done using phase estimation of the transmitted signal with the help of trained symbols. Final results show the improvement in BER. However, the amount of results improved depends upon the amount of trained symbols. The more trained symbols will result into more improved BER.

## 1. Introduction

In recent years, OFDM have been studied very widely and deeply in wireless communication systems because of bandwidth efficiency and its robustness to channel fading and Inter Symbol Interference (ISI). OFDM system is capable of mitigating a frequency selective channel to a set of parallel fading channels, which need relatively simple processes for channel equalization.

There exists a large number of distribution schemes to describe the statistics of mobile radio signal. A key assumption in the theoretical explanation of the Rayleigh, Rician, Nakagami and Weibull distribution was that the statistics of the channel do not change over the small (local) area under consideration. However, to describe the long-term signal variations lognormal distribution is being used [1]. These distributions are helpful in precise designing of wireless systems to make the systems more robust to noise.

Rayleigh and Rician fading channels have already been studied and employed in OFDM systems in frequency selective and flat fading environment. Nakagami-m distribution is another useful and important model to characterize the fading model. A threshold value of m is calculated for both frequency and flat fading environment in [2]. There exist many other fading models in literature, which have been proposed for better fitting of

data while aiming at non-linearity of channel. So our motivation behind this paper to explore the non-linear fading environment. One of the interesting models that we could find in literature is $\alpha - \mu$ distribution [3], which provides the generalized model for fading distribution. Depending upon the value of $\alpha$ and $\mu$, this model can be utilized for the generation of Nakagami-m and Weibull variants. However, it also treats One-Sided Gaussian, Rayleigh and Negative Exponential distributions as its special cases. The generalized fading model using three parameter generalized gamma distribution describing all forms of multipath fading and shadowing in wireless systems is analyzed in [4].

This paper is organized as follows: In Section 2, OFDM system model is discussed. Section 3, describes the generalized model of $\alpha - \mu$ distribution. Flat fading channel model to use in OFDM systems is described in Section 4. In Section 5, channel estimation technique is discussed. The analysis of OFDM system without estimation has been done in Section 6, while results with estimation have been presented in Section 7. Finally Section 8 concludes the paper.

## 2. OFDM Model

A Complex base band OFDM signal with N sub-carriers,

is expressed as

$$s(t) = \sum_{k=0}^{N-1} D_i e^{j2\pi k f_0 t} \qquad 0 \le t \le T \qquad (1)$$

For each OFDM symbol, the modulated data sequences are denoted by $D(0), D(1), \dots D(N-1)$. Here, $f_0$ denote the sub-carriers spacing and is set to $f_0 = \frac{1}{T}$, the condition of orthogonality. After IFFT, the time-domain OFDM signal can be expressed as:

$$S(n) = \frac{1}{N} \sum_{k=0}^{N-1} D_i e^{\frac{j2\pi k f_0 n}{N}} \qquad (2)$$

After IFFT, the modulated signal is up-converted to carrier frequency $f_C$ and then the following signal is produced and transmitted through channel:

$$x(t) = \mathrm{Re}\left\{ \sum_{k=0}^{N-1} D_i e^{j2\pi k(f_0 + f_C)t} \right\} \quad 0 \le t \le T \qquad (3)$$

$x(t)$ represents the final OFDM signal in which sub-carriers shall undergo a flat fading channel.

## 3. The $\alpha - \mu$ Distribution

The $\alpha - \mu$ distribution is a general fading distribution that can be used to represent various fading model. This distribution deals with non-linearity of propagation medium [5]. Fading signal with envelope $r$, an arbitrary constant parameter $\alpha > 0$ and a root mean value $\hat{r} = \sqrt[\alpha]{E(r^\alpha)}$ shall have its probability density function $p(r)$, which is written as:

$$p(r) = \frac{\alpha \mu^\mu r^{\alpha\mu - 1}}{\hat{r}^{\alpha\mu} \Gamma(\mu)} \exp\left(-\mu \frac{r^\alpha}{\hat{r}^\alpha}\right) \qquad (4)$$

Weibull and Nakagami-m distribution can be easily derived from $\alpha - \mu$ distribution as its special cases.

By setting $\mu = 1$, Equation (4) shall reduce to Weibull probability distribution function as:

$$p(r) = \alpha \beta r^{\alpha - 1} \exp(-\beta r^\alpha) \qquad (5)$$

where $\beta = \hat{r}^{-\alpha}$.

Here, by varying the value of $\alpha$ different curves of pdf can be plotted.

From Weibull distribution by setting $\alpha = 2$, the Rayleigh distribution can be obtained as:

$$p(r) = \frac{r}{\gamma^2} \exp\left(-\frac{r^2}{2\gamma^2}\right) \qquad (6)$$

where $\gamma^2 = \hat{r}^2 / 2$.

Now, if we put $\alpha = 1$ in Weibull distribution, it shall reduce to Negative exponential distribution represented as:

$$p(r) = \delta \exp(-\delta r) \qquad (7)$$

where $\delta = \hat{r}^{-1}$

So by keeping the value of $\mu = 1$ and varying the value of $\alpha$ it has generated Rayleigh and Negative exponential distribution. Whereas if we keep $\alpha = 2$ and vary the value of $\mu$, we shall be able to represent this $\alpha - \mu$ distribution as Nakagami-m distribution

In such a case

$$p(r) = \frac{2\mu^\mu r^{2\mu - 1}}{\Omega^\mu \Gamma(\mu)} \exp\left(-\mu \frac{r^2}{\Omega}\right) \qquad (8)$$

By setting $\mu = 1/2$, one-sided Gaussian distribution can be obtained as:

$$p(r) = \sqrt{\frac{2}{\pi \hat{r}^2}} \exp\left(-\frac{r^2}{2\hat{r}^2}\right) \qquad (9)$$

However, for $\alpha - \mu$ distribution the envelope $r$ can be written as:

$$r = \left[ \sum_{i=1}^{N} \left( x_i^2 + y_i^2 \right) \right]^{1/\alpha} \qquad (10)$$

where, $x_i$ and $y_i$ are in-phase and quadrature elements of multipath components represented by symbol $i$.

It was interesting to find that $\alpha = 2$ in Equation (10) shall reduce to the envelope equation of Rayleigh fading distribution [6] described as:

$$r = \left[ \sqrt{\sum_{i=1}^{N} x_i^2 + y_i^2} \right] \qquad (11)$$

Same concept has been shown in Equations (5) and (6) that by putting $\alpha = 2$, Weibull distribution converts to Rayleigh distribution, hence introducing the non-linearity into propagation medium. However, at different values of $\alpha$ different fades can be generated.

## 4. Channel Model

In this paper, the sub-channel spacing is equal to inverse of time period, so that the produced parallel fading sub-channels have flat fading characteristics.

Here $\alpha - \mu$ distribution has been utilized for generation of Weibull distribution by setting $\mu = 1$ and varying the value of $\alpha$.

In flat fading environment, the base-band signal at the input of receiver $y(t)$ is as described as follows:

$$y(t) = x(t)*r(t) + n(t) \tag{12}$$

where, $x(t)$ denotes the base-band transmitted signal, $r(t)$ is the Weibull distributed channel envelope and $n(t)$ is the additive white Gaussian noise with zero mean.

## 5. Channel Estimation

Channel estimation in frequency selective has different approach then compared with flat fading environment. A comparative study using Minimum Mean Squared Error (MMSE) and Least square (LS) estimator in frequency selective fading environment has been presented in [7]. The channel estimation based on comb type pilot arrangement is studied using different algorithms by bahai *et al*. [8]. A novel channel estimation scheme for OFDMA uplink packet transmissions over doubly selective channels was suggested in [9]. The proposed method uses irregular sampling techniques in order to allow flexible resource allocation and pilot arrangement. In flat fading environment, estimation of the channel using trained sequence of data has been studied and implemented in [10]. He presented the channel estimation in flat fading environment using some trained data. Channel phase was estimated during each coherence time. Then pilot data of some required percentage of data length (referred as training percentage in simulation) is inserted into the source data. It is used to estimate the random phase shift of the fading channel and train the decision to adjust the received signal with phase recover. The results obtained showed the great variation in BER for with and without estimation curves. It is clear from literature reviewed that phase estimation using training symbol can be implemented in flat fading environment to improve the performance of system.

In this paper, we have implemented the above described phase estimation technique in flat fading for Weibull fading distribution on OFDM system.

## 6. Results without Estimation
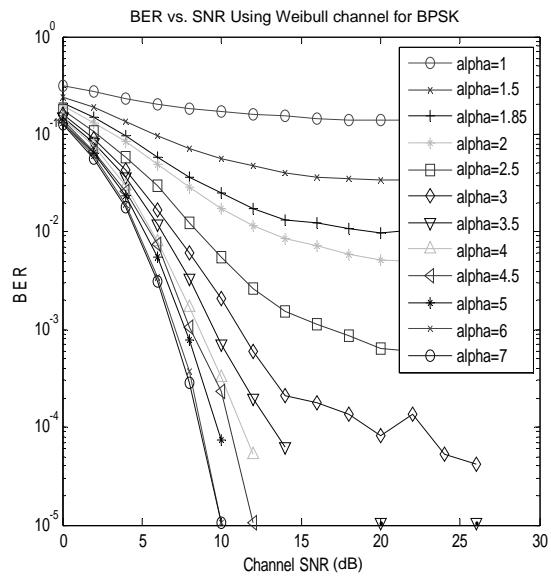
OFDM-BPSK and -QPSK signal is simulated in MATLAB environment by choosing total number of sub-carriers 400, IFFT length 1024 by using guard interval of length 256.

The results presented in **Figures 1** and **2** are simulated by varying the value of $\alpha$ and keeping $\mu = 1$. Here, the BER values have been obtained for varying $\alpha$ over a range of 1 to 7, however, improvement in BER was not significant for higher values of $\alpha$, So range has been
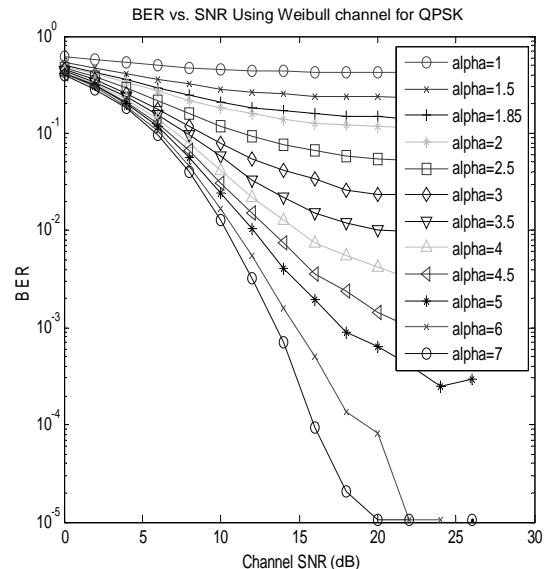
kept from 1 to 7, both for OFDM-BPSK and -QPSK system.

From the simulations, it has been verified that the results for $\alpha = 2$ are same that are obtained by using Rayleigh fading distribution.

It has been observed that if we plot BER curves by changing the value of $\mu$ there is no change in these curves. This is because of fact that in the envelope Equation 10, the variable parameter is $\alpha$ which varies the fading variants and $\mu$ has no role to change this fading envelope and hence no change in BER values.



**Figure 1. BER vs. SNR for OFDM-BPSK system.**



**Figure 2. BER vs. SNR for OFDM-QPSK system.**
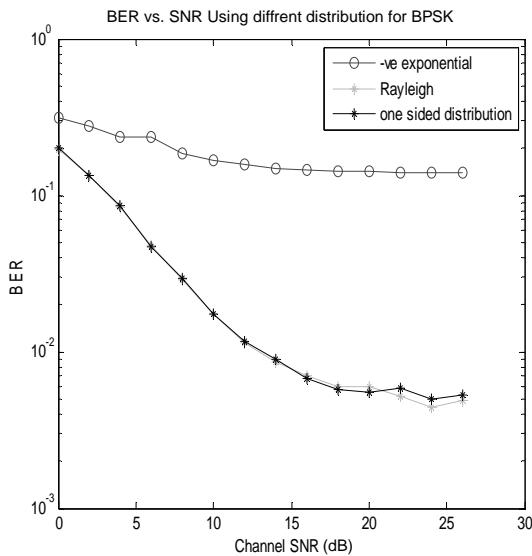
To explore the other side of $\alpha - \mu$, by keeping the value of $\alpha$ fixed and varying the value of $\mu$, we are able to have BER curves for other distributions. In **Figure 3**, Negative exponential distribution has been plotted as special case where $\alpha = 1$ and $\mu$ can vary. Here, it has been plotted for fixed value of 1. Rayleigh distribution is having $\alpha = 2$ and $\mu$ can vary. Here, it has been plotted with $\mu = 1$. One sided distribution has been plotted with $\alpha = 2$ and $\mu = 1/2$.

BER varies in the range of $10^{-1}$ to $10^{-5}$ for OFDM-BPSK and -QPSK for SNR of 0 to 25 dB. In case of OFDM-BPSK the BER value of $10^{-5}$ is obtained at SNR of 10dB. However, for -QPSK case the BER of $10^{-5}$ is obtained at SNR of 20 dB.

Comparison between Negative exponential value, Rayleigh and one sided distribution results clearly reveals the fact that in $\alpha - \mu$ distribution the variation in value of $\alpha$ can change the value of BER, however by changing the value of $\mu$, there is no impact upon the BER. Results obtained without estimation technique has been presented in [11].

## 7. Results with Estimation
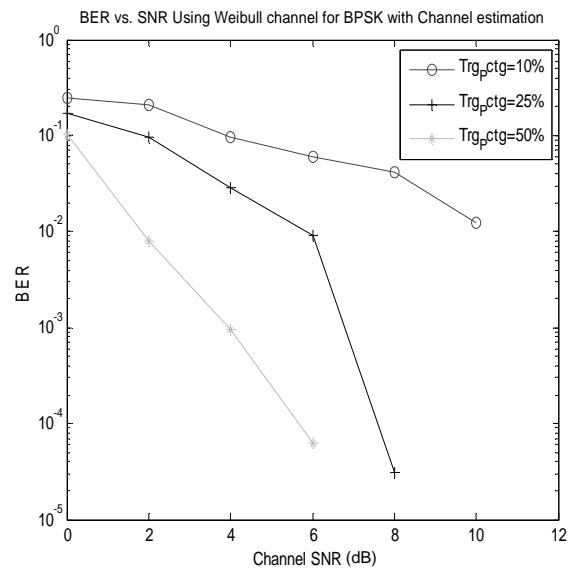
Trained symbols are added to source signal as discussed in Section 5. The percentage of such symbol may be varied depending upon the system response to the trained sequence. We have analyzed the results for various percentage values of trained sequence. We have plotted new graphs with value of $\alpha = 7$ and varying value of training sequence over the range from 10% to 50%. Results

for OFDM-BPSK and -QPSK have plotted in **Figures 4** and **5** respectively.

In depth analysis of these graphs shows that BER decreases, if the training percentage is increased. In **Figure 5**, if we evaluate the reading obtained at SNR = 10 db, BER has decreased from 0.0208 to 0.001 for training percentage of 10 to 50. This means, for the same value of SNR and $\alpha$, different training percentage has resulted into different values of BER. More trained sequence will results into lesser errors. The same has been depicted from **Figure 4**.



**Figure 4. BER vs. SNR for OFDM-BPSK with channel estimation.**



**Figure 3. BER vs. SNR for negative, rayleigh and exponential distributions.**



**Figure 5. BER vs. SNR for OFDM-QPSK with channel estimation.**

## 8. Conclusions

This paper, presents performance analysis of OFDM system with generalized fading model of $\alpha - \mu$ distribution with and without estimation. The non-linearity added in propagation medium has been clearly shown in simulated results, since the BER has significantly reduced by varying $\alpha$ from 1 to 7. However, higher values of $\alpha$ can be used for further reductions in BER. It is clear from the simulations that the result shows significant improvement by applying the phase estimation using trained symbols.

## 9. References

[1]  H. Hashemi, "The Indoor Radio Propagation," *proceedings of IEEE*, Vol. 81, No. 7, July 1993, pp. 943-968.

[2]  D. Zheng, J. L. Cheng and N. C. Beaulieu, "Accurate Error Rate Performance Analysis of OFDM on Frequency Selective Nakagami-m Fadiing Channels," *IEEE Transaction on Communications*, Vol. 54, No. 2, February 2006, pp. 319-328.

[3]  Miechel Daoud Yacoub, "The $\alpha$-$\mu$ Distribution: A General Fading Distribution," 2002, pp. 629-633.

[4]  J. Malhotra, A. K. Sharma and R. S. Kaler, "On the Performance Analysis of Wireless Receiver Using Gene-Ralized-Gamma Fading Model," *Annals of Telecommunication Annals Des Telecommunications*, *International Journal*, Springer, Vol. 64, No. 1-2, November 2008, pp. 147-153.

[5]  M. D. Yacoub, "The $\alpha$-$\mu$ Distribution: A Physical Fading Model for the Stacy Distribution," *IEEE Transactions on Vehicular Technology*, Vol. 56, No. 1, Janurary 2007, pp 27-34.

[6]  G. S. Prabhu and P. M. Shankar, "Simulation of Flat Fading Using MATLAB for Classroom Instruction," *IEEE Transaction on Education*, Vol. 45, No. 1, February 2002, pp. 19-25.

[7]  J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson and P. O. B. Rjesson, "On Channel Estimation in OFDM Systems," *Proceedings of Vehicular Technology Conference* (*VTC Ô95*), Chicago, USA, Vol. 2, September 1995, pp. 815-819.

[8]  S. Coleri, M. Ergen, A. Puri and A. Bahai, "A Study of Channel Estimation in OFDM Systems," *IEEE VTC*, Vancouver, Canada, September 2002.

[9]  P. Fertl and G. Matz, "Multi-User Channel Estimation in OFDMA Uplink Systems Based on Irregular Sampling and Reduced Pilot Overhead," *IEEE ICASSP*, Vol. 3, 2007.

[10] Z. F. Chen, "Performance Analysis of Channel Estimation and Adaptive Equalization in Slow Fading Channel," University of Florida. http://users.ecel.ufl.edu/~zhifeng

[11] N. Sood, A. K. Sharma and M. uddin, "BER Performance of OFDM-BPSK and -QPSK over Generalized Alpha-Mu Fading Distribution," *IEEE International Advance Computing Conference (IACC-2009)*, Patiala, India, 6-7 March 2009, pp. 1197-1199.

Scientific
Research

# A Study on Scalable Video Coding for AMC with Mobile Media-Based Multicast over WiMAX 802.16e

**Tzu-Kai Cheng[1], Feng-Ming Yang[2], Jean-Lien C. Wu[3]**

[1]*Applied Research Center Taiwan Company, Telcordia Technologies, Inc., Taipei, Taiwan, China*
[2]*National Taiwan University of Science and Technology, Department of Electronic Engineering, Taipei, Taiwan, China*
[3]*Department of Computer and Communication Engineering, St. John's University, Taipei, Taiwan, China*
*Email*: *williamcheng@tarc-tw.research.telcordia.com, {d9602206, jcw}@mail.ntust.edu.tw*

## Abstract

Adaptive Modulation and Coding (AMC), which is a technology used when channel condition changes, is adapted in Mobile Worldwide Interoperability for Microwave Access (WiMAX). Scalable Video Coding (SVC) is a video coding scheme used for different users with bandwidth level. SVC encodes a video into a number of layers. Users receive different number of encoded layers based on their channel condition. In this paper, Intermediate Control Server (ICS) is proposed to deal with the signaling between multimedia server and BS. Both AMC and SVC are employed to enhance the user perceived video quality in the system.

**Keywords:** AMC, SVC, Server Signaling, 802.16e, Mobile WiMAX

## 1. Introduction

More television programs can be transmitted because signals are compressed by compression technology in digital television as compared to analog television [1,2]. For example, IPTV (Internet Protocol Television) is a system where a digital television service is delivered by Internet protocols over broadband network and contents could be on-demand video or live streaming. On-demand contents are already stored in the multimedia server after being per-coded and compressed. The term "triple play" means the service provider provides voice, data, and video. The service containing triple play and mobile is called "quadruple play". In [3], a utility-based resource allocation scheme, called U-LEM, for layer-encoded multicast streaming service in fixed WiMAX networks is proposed to discuss the time complexity. This work proves this problem is NP-hard, and can run in polynomial time.

In [4], the authors discussed the key success factors, benefits, and associated challenges of introducing IPTV into fixed WiMAX. A congestion control mechanism is proposed in [5] which gave consideration to fairness and system overhead applied in IEEE 802.16d broadband wireless networks. In [6], the transmission performance of the scalable video streaming services over Mobile WiMAX system is investigated and the failure rate is calculated in two scenarios. The less important layers are transmitted using the second connection with lower pri-

ority as compared to the one connection scenario. A mechanism called QoS-based Active Dropping to the MAC layer is proposed in [7] to deal with the bandwidth utilization. An equation between network loading and dropping probability is investigated to show how to adjust between these two factors.

This paper is intended to enhance the user perceived video quality by multicasting the videos encoded by scalable video coding (SVC) over the Mobile WiMAX [8]. The rest of the paper is organized as follows. Brief introductions of works related to SVC over WiMAX are in Section 2. In Section 3, the system model is described, and there is a detailed description of the proposed SVC+AMC quality enhancement scheme and the intermediate control server (ICS) signaling. Simulation results are shown in Section 4 to stand for the proposed scheme. Finally, this paper is concluded in Section 5.

## 2. Background

### 2.1. AMC (Adaptive Modulation and Coding)

WiMAX is expected to provide higher transmission rate and wider transmission range and mobile WiMAX has support for mobility. The value of signal-to-noise ratio (SNR) varies according to the distance between mobile station (MS) and BS. Modulation is changed according to channel condition in AMC. When the channel condi-

tion becomes worse, the more robust modulation is used. Therefore, AMC can lead to improve system coverage and capacity while maintaining high power and spectrum efficiency. **Table 1** shows the modulation and the receiver SNR [9].

## 2.2. SVC (Scalable Video Coding)

A video sequence is composed of a stream of individual frames. The joint video team (JVT) develops the extension of H.264/AVC (Advanced Video Coding) SVC which encodes a picture into one base layer and several enhancement layers in general [10-12]. SVC has three dimensions in spatial scalable, temporal scalable, and quality scalable. Spatial scalable means the bit-stream can provide different spatial resolutions. Temporal scalable means various frame rates are available. Quality scalable, also called SNR scalable which means the visual quality is scalable. To achieve the scalability, the base layer provides basic video quality with low bit rate for user. Enhancement layer is used to refine the base layer video quality. SVC can provide different video quality for each level user by encoding an image into several layers. The number of layers is not fixed, it is decided in different applications when encoding. The more the layer received, the better the perceived quality.
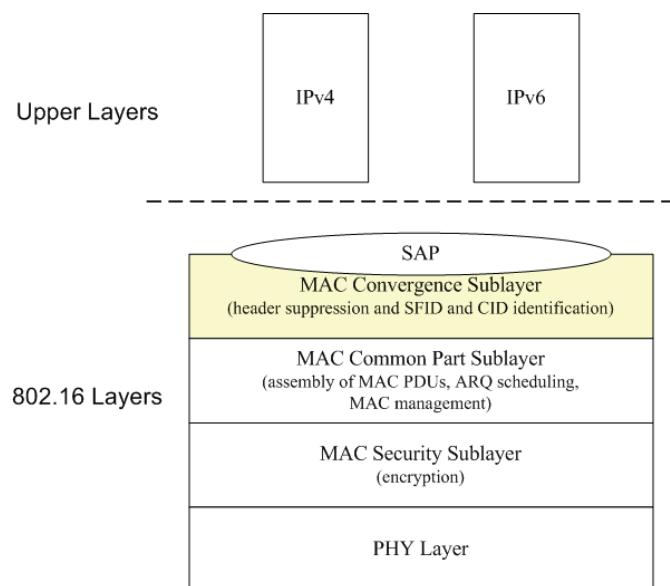
## 2.3. MAC CS Classification

The Medium Control Access (MAC) layer is to provide an interface between the upper transport layers and the PHY layer. The MAC layer receives packets, MAC Service Data Units (MSDUs), from upper layers, and or-ganizes them into MPDUs for transmission over the air. The MAC layer of WiMAX, as shown in **Figure 1**, is divided into three distinct sub-layers: Convergence Sublayer (CS), Common Part Sublayer (CPS), and Security Sublayer (SS). The MAC CS, which is the interface between the MAC layer and layer 3 of the network, supports MSDU header suppression to reduce the higher layer overheads on each packet and provides any transformation or mapping of external network data through CS Service Access Point (CS SAP) into MAC layer. The MAC CPS performs all the packets operations including fragmentation and concatenation of SDUs into MAC PDUs, transmission of MAC PDUs, QoS control, and ARQ. The MAC SS is responsible for encryption, authorization, and keys exchange between BS and MS. Besides the header suppression, MAC CS is also responsible for classification. Classification is the process which a MAC SDU is mapped onto a particular connection for transmission between two peers. The WiMAX MAC layer is connection oriented. Before any data transmission happens, BS and MS first establish a unidirectional logical link, called a connection. Each connection
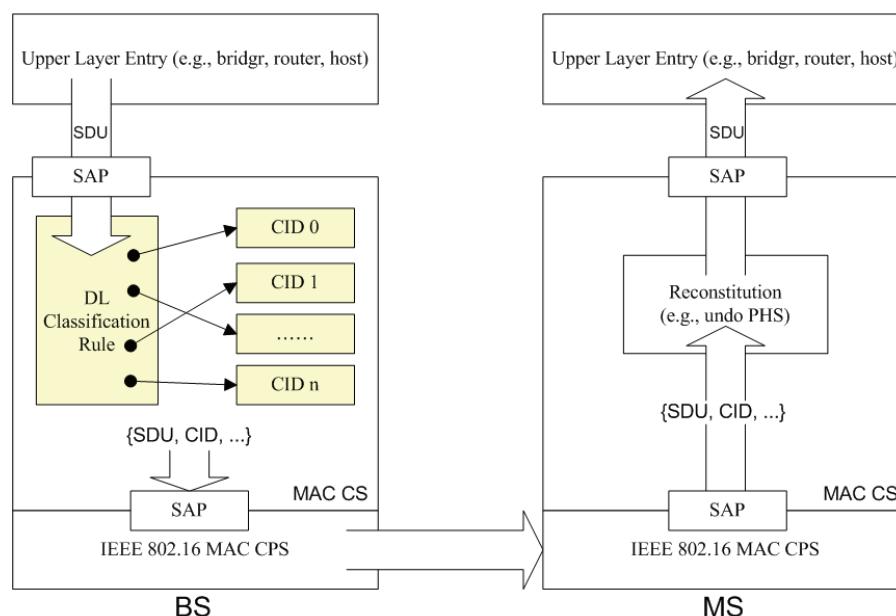
**Table 1. The modulation and the receiver SNR.**

| Modulation | Coding Rate | Receiver SNR (dB) |
|------------|-------------|-------------------|
| BPSK       | 1/2         | 3.0               |
| QPSK       | 1/2         | 6.0               |
|            | 3/4         | 8.5               |
| 16QAM      | 1/2         | 11.5              |
|            | 3/4         | 15.0              |
| 64QAM      | 2/3         | 19.0              |
|            | 3/4         | 21.0              |



**Figure 1. The WiMAX MAC layer.**

is identified by a Connection InDentifier (CID), which is 16-bit long and serves as temporary and dynamic layer 2 addresses for data transmission and control plane traffic over the particular link. MAC CS determines the appropriate CID based on not only the destination address but also various other factors, e.g. source address and Service Flow InDentifier (SFID). A service flow is a unidirectional flow of packets with a particular set of QoS parameters and is identified by a SFID. The BS is responsible for issuing the SFID and mapping it to unique CIDs. The classification and the CID mapping from BS to MS is revealed in **Figure 2**. CS specification is defined in classification. The data from upper layer is classified and mapped to different to different SFID and CID. CS specification is shown in **Table 2**.

## 2.4. Network Reference Model (NRM)

The WiMAX Forum defines itself as an industry-led non-profit organization to promoting and certifying interoperable WiMAX products and also determined a network architecture which showed how a WiMAX network connects with other networks [13,14]. The NRM is a logical representation of the network architecture. The NRM identifies key functional entities and reference points between functional entities over which a network interoperability framework is defined. The objective of the NRM is to allow multiple implementation options for a given functional entity, and achieve interoperability among different functional entities. NRM structure is represented in **Figure 3**.



**Figure 2. Classification and the CID mapping from the BS to the MS.**

**Table 2. IEEE 802.16e CS specification.**

| Type | Length | Value | Scope |
|------|--------|-------|-------|
| [145/146].28 | 1 | 0: Reserved<br>1: Packet, IPv4<br>2: Packet, IPv6<br>3: Packet, IEEE 802.3/Ethernet<br>4: Packet, IEEE 802.1Q VLAN<br>5: Packet, IPv4 over IEEE 802.3/Ethernet<br>6: Packet, IPv6 over IEEE 802.3/Ethernet<br>7: Packet, IPv4 over IEEE 802.1Q VLAN<br>8: Packet, IPv6 over IEEE 802.1Q VLAN<br>9: ATM<br>10: Packet, IEEE 802.3/Ethernet[a] with ROHC header compression<br>11: Packet, IEEE 802.3/Ethernet[b] with ECRTP header compression<br>12: Packet, IP2 with ROHC header compression<br>13: Packet, IP2 with ECRTP header compression<br>14-255: Reserved | DSx-REQ |

## 3. SVC+AMC Quality Enhancement and ICS Signaling

### 3.1. System Model

When an MS requests a video, it sends the messages to the BS, and the BS then forwards the request to the multimedia server. When adapting SVC, which encodes a frame into several layers, the multimedia server which stores the videos performs the encoding procedure and transmits the layers to the BS. For all MSs who are watching the same video, the BS grouped them based on their feedback channel condition. The group with better channel condition gets more layers. On the contrary, the group with worse channel condition gets fewer layers. However, the BS can not distinguish the layer ID received from the multimedia server. In order to solve this problem, an Intermediate Control Server (ICS) is newly added in the system, see **Figure 4**. A WiMAX network is served by an ICS. After executing the videos in the multimedia server, the ICS forwards the SVC-encoded layers to BS.
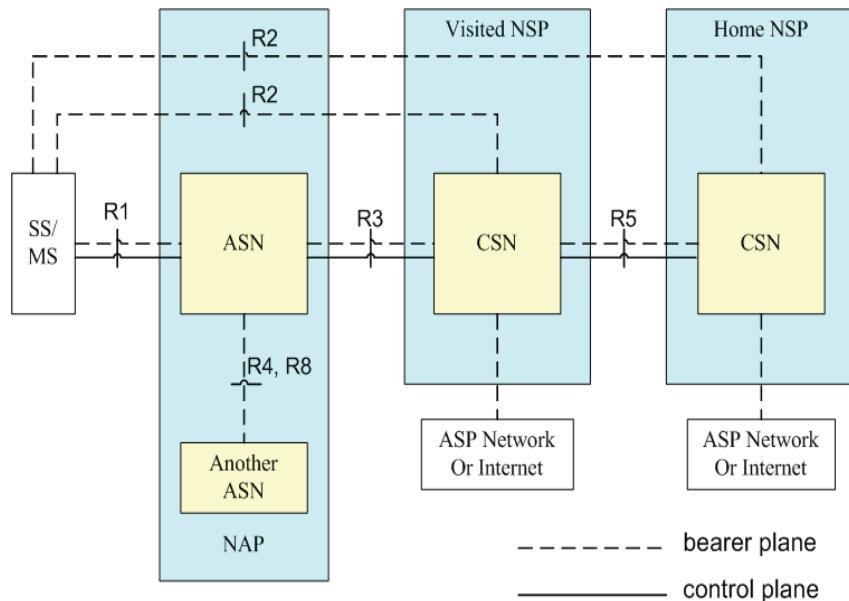


**Figure 3. Network reference model.**
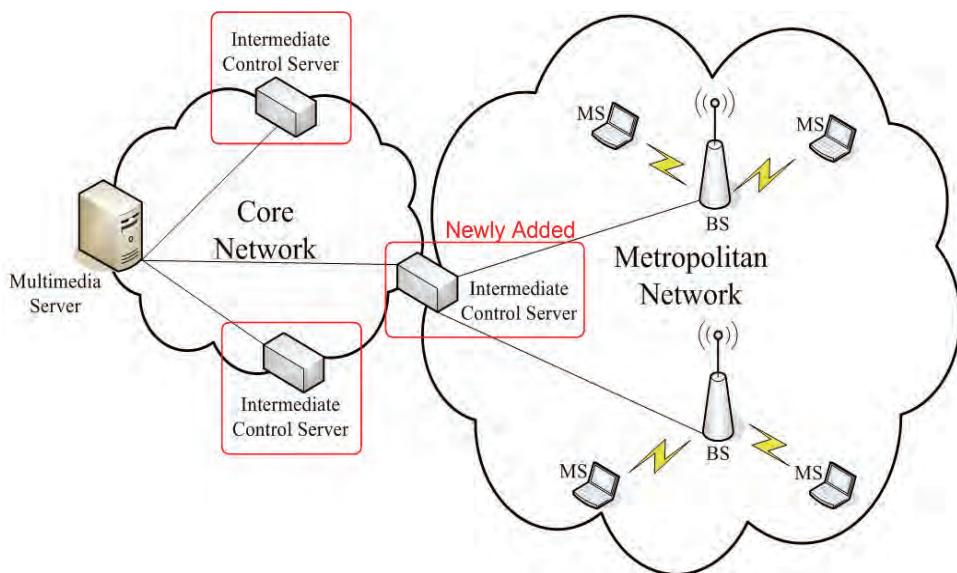


**Figure 4. The proposed system model with a newly added ICS.**

In our proposed scheme, SVC is implemented to encode a frame into three layers. Layer 0 is meant to be the base layer, which contains the most important information. Layer 1 and layer 2 belong to the enhancement layer, which carries less information as compared to layer 0. In order to assure the video quality perceived by user within worse channel condition, layer 0 is encoded to carry 50% information of the original frame. After receiving and decoding layer 1 with layer 0, the containing information is refined to 75%. Users with best channel condition get all three layers, the frame which is decoded by combining layer 0, layer 1, and layer 2 is the same as the original transmitted frame. The retrieval frames are shown in **Figure 5**.
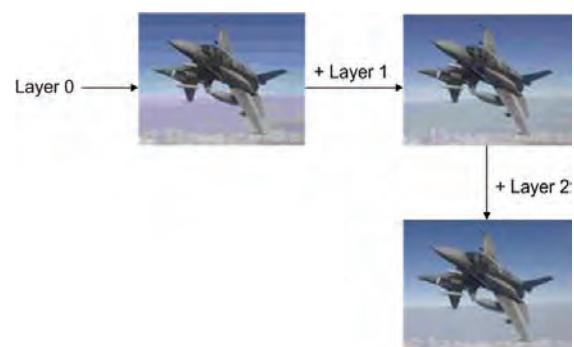
## 3.2. The ICS Signaling

The multimedia server encodes the video into three layers in the application layer. All layers are packetized and added different protocol headers from the upper layer to lower layer, the BS could not distinguish the layer ID. An ICS is used to solve this problem. An ICS is individually set between multimedia server and the BS for flexibility and extensibility of many applications provided by different operators in the future. The major task of ICS is to re-packetize the packets received from multimedia server so that the BS recognizes the layer IDs. The proposed system model with network layers is shown in **Figure 6**.
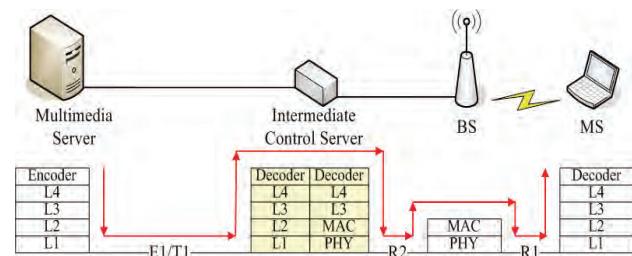
When the ICS receives the SVC-encoded layer packet, it decodes the packet in the application layer and gets the layer ID information. While transmitting to the BS, the ICS re-packages the SVC-encoded frame by adding the layer ID information in the reserved filed of each network protocol header. The Network Abstraction Layer (NAL) unit [15], shown in **Figure 7**, defines the interface between the encoded video data itself and the possible transport layers or storage formats. Each NAL unit regardless of its type is encapsulated in the RTP/UDP/IP packet. The encapsulation of NAL unit is shown in **Table 3**. Three bits in NAL unit are used to indicate the layer ID.

When the packet transmits by RTP (Real-time Transport Protocol), there are seven bits in the payload type filed [16]. The RTP header format is presented in **Figure 8**. The reserved values in payload type, which is dis-

played in **Table 4**, are 72 to 76 [17]. We use three values, 72, 73, and 74, to indicate the layer ID. In UDP (User Datagram Protocol) header format which is shown in **Figure 9**, there are 8-bit protocol numbers [18]. Part of the assigned protocol numbers is presented in **Table 5**, the values 72-76 and 80-254 are unassigned [19], and three values, 72, 73, and 74, are used to denote the layer ID. Continuing transmitting to lower layer, the IPv4 (Internet Protocol version 4) header is displayed in **Figure 10** [20]. 2-bit reserved field could be used to show the layer ID. '00' denotes layer 0, '01' denotes layer1, and '10' denotes layer2 separately. In this way, MAC layer can clearly distinguish what the received layer ID is.
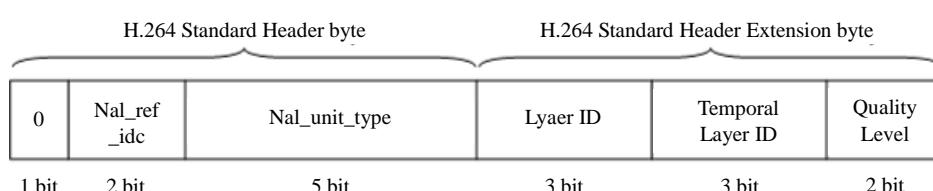


**Figure 5. The retrieval frame of SVC.**



**Figure 6. The proposed system model with network layers.**

**Table 3. The encapsulation of NAL unit in RTP/UDP/IP header.**

| IP header | UDP header | RTP header | NAL unit header | NAL unit payload |
|-----------|------------|------------|-----------------|------------------|



|  | H.264 Standard Header byte | | | H.264 Standard Header Extension byte | | |
|---|---|---|---|---|---|
| 0 | Nal_ref _idc | Nal_unit_type | Lyaer ID | Temporal Layer ID | Quality Level |
| 1 bit | 2 bit | 5 bit | 3 bit | 3 bit | 2 bit |

**Figure 7. NAL unit header format.**
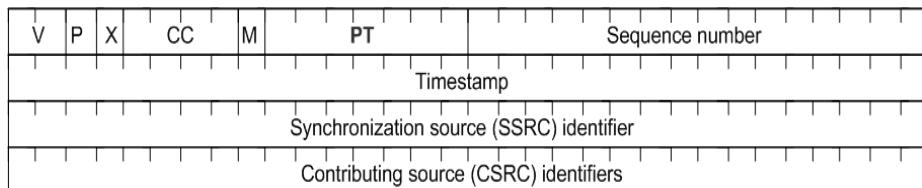
**Figure 8. RTP header format.**
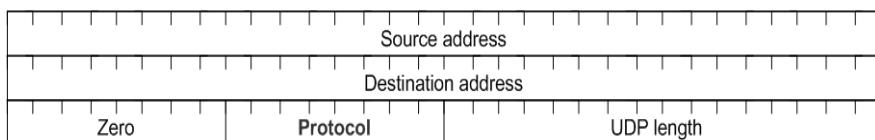


**Figure 9. UDP header format.**
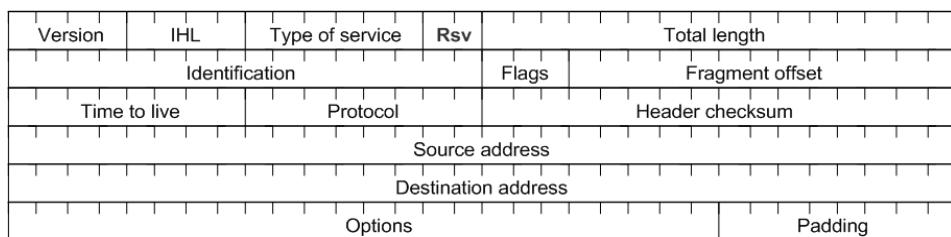


**Figure 10. IPv4 header format.**

**Table 4. Partial RTP payload type.**

| PT | Encoding Name | Media Type | Clock Rate (Hz) |
|----|--------------|-----------|-----------------|
| 34 | H.263 | V | 90,000 |
| 35-71 | Unassigned | ? | |
| **72-76** | **Reserved** | **N/A** | **N/A** |
| 77-95 | Unassigned | ? | |
| 96-127 | Dynamic | ? | |

**Table 5. Partial UDP protocol number.**

| Decimal | Octal | Protocol Numbers |
|---------|-------|------------------|
| 69 | 105 | SIMP monitoring |
| 70 | 106 | SIMP polling |
| 71 | 107 | SIMP packet core/U |
| **72-76** | **110-114** | **Unassigned** |
| 77 | 115 | Backroom SIMP polling |
| 78 | 116 | Backroom SIMP monitoring |
| 79 | 117 | SIMP message generators |
| **80-254** | **120-376** | **Unassigned** |
| 255 | 377 | Reserved |

### 3.3. The SVC+AMC Quality Enhancement Scheme

In the SVC+AMC quality enhancement scheme, the BS multicasts different layers using different modulations based on different connections. The mapping of the layer to Connection IDentifiers (CIDs) is decided by the classification rule in the MAC CS specification defined in the IEEE 802.16e standard. Three reserved value 14-255 are used to indicate the layer IDs which are "Packet, IPv4 over 802.3/Ethernet ECRTP compression SVC-encoded layer 0", "Packet, IPv4 over 802.3/Ethernet ECRTP compression SVC-encoded layer 1", and "Packet, IPv4 over 802.3/Ethernet ECRTP compression SVC- encoded layer 2". Based on the CS classification rule, the layers are easily mapped to different CIDs. In the 802.16e OFDMA DL-MAP_IE format, there are 8-bit CIDs assigned for this Information Element (IE). Three values in the unassigned CIDs are chosen to indicate these three connections.

The SVC+AMC quality enhancement scheme groups MSs who are receiving the same channel into three groups according to their fast-feedback channel condition to the BS. The fast-feedback CQICH information is allocated in the uplink subframe in OFDMA frame. These MSs who subscribe the same video are all in group 0 to receive layer 0. If there are some with better channel condition, they are grouped into group 1 to re-

ceive layer 1, so they are in group 0 and group 1 at the same time. For those MSs with the best channel condition, they are grouped into group 2 to receive layer 2, therefore, they are in group 0, 1, and 2. While transmitting, the BS multicasts layer 0 to group 0 using QPSK through CID 0, multicasts layer 1 to group 1 using 16QAM through CID 1, and multicasts layer 2 to group 2 using 64QAM through CID 2. For this reason, users with better channel condition could enjoy better perceived video quality.

## 4. Simulation

### 4.1. Simulation Parameters

The fundamental simulation parameters [21-23] are shown in **Table 6**. Network congestion is considered in the system. Assume packet loss rates in the channel are 0.30%, 0.24%, and 0.21% while using QPSK, 16QAM, and 64QAM respectively.

PSNR (Peak Signal-to-Noise Ratio) is a conventional assessment method used to assess the quality of the reconstructed images. PSNR is used to compare the similarity between the original frame and the retrieval frame. Each pixel in the retrieval frame is compared with each related pixel in the original frame. The equation of PSNR is shown in following. In general, when the value of PSNR reaches 30dB, it means the difference between original frame and retrieval frame is hardly recognized by human eyes.

$$PSNR = 10 \times \log_{10} \frac{(255)^2}{MSE} dB \qquad (1)$$

$$MSE = \frac{1}{(w \times h)} \sum_{i=1}^{w} \sum_{j=1}^{h} \left( \alpha(i,j) - \beta(i,j) \right)^2 \quad (2)$$

### 4.2. Simulation Results

Simulation results are used to show the SVC+AMC quality enhancement scheme improves the user perceived video quality effectively. There are three methods considered in two scenarios. The first method, referred as Normal scheme, transmits the original video to MSs using 16QAM without applying SVC and AMC. The second method, referred as the AMC scheme, transmits the original video which is not encoded by SVC to MSs according to their channel condition. MSs in worse channel condition receive the video using QPSK. The BS multicasts the video to MSs in normal channel condition using 16QAM and multicasts to MSs with better channel condition using 64QAM. The last one is the SVC+AMC quality enhancement scheme proposed in this paper. In SVC+AMC, the BS transmits the SVC-encoded layers to all MSs using QPSK, to MSs with normal channel condi-

tion using 16QAM, and to MSs in better channel condition using 64QAM.

We setup two different scenarios and use two videos to simulate. Scenario1 has 20 MSs which are investigated to the PSNR rate on Mobile WiMAX system. There is only one connection between the multimedia server and the BS, all layer-encoded videos are transmitted via this connection. The ICS forwards the SVC-encoded layers to the BS. Scenario2 is basically the same as Scenario1 except the channel conditions of MSs.

In Scenario1 of the first video, 20% of MSs are in worse channel condition, 40% of MSs are in normal channel condition, and 40% of MSs are in better channel condition. Therefore, there are 20 MSs in group 0, 16 MSs in group 1, and 8 MSs in group 2. The above-mentioned three methods are employed to simulate the scenario. **Figure 11** shows the simulation results. The average values of PSNR in normal, AMC, and SVC+AMC scheme are 26.22, 27.58, and 32.18, separately. The AMC scheme is compared to the normal scheme first. The average value of PSNR in the AMC scheme is just a little higher than the normal scheme. While in the normal scheme, all MSs receive the non-encoded video using 16QAM. In normal scheme, the average PSNR in the system becomes lower because 20% of MSs receive the video in an unsuitable modulation and coding scheme and 40% of MSs with better channel condition should deserve better video quality. In the AMC scheme, all MSs receive the full video in a suitable packet loss rate according to their channel condition. Packet loss rate becomes smaller while transmitting to MSs with better channel condition using 64QAM. Packet loss occurs more often in MSs with worse channel condition because of the more robust modulation and coding scheme. For that reason, PSNR is only a little better in the AMC scheme than in normal scheme. The SVC+AMC scheme, which receives the SVC-encoded video based on their channel condition, gets the highest average PSNR. MSs in worse channel condition receive layer 0 using QPSK resulting in less packet loss, but the perceived quality is still embraceable. Two layers are transmitted to MSs

**Table 6. Fundamental simulation parameters.**
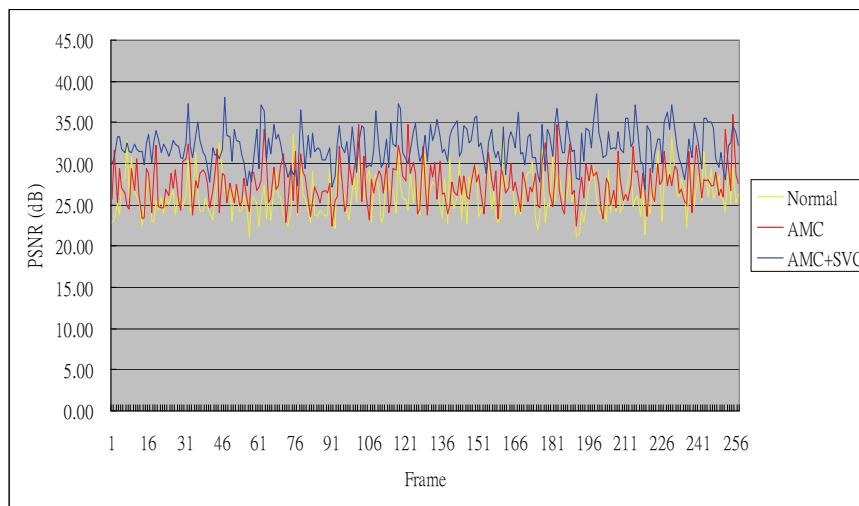
| Parameter | Value |
| --- | --- |
| Channel Bandwidth | 10 MHz |
| Size of FFT | 1024 |
| Number of Sub-channels | 30 |
| OFDMA Symbol Time | 102.9 μs |
| Frame Duration | 5 ms |
| Modulation Scheme | QPSK, 16QAM, 64QAM |
| Mobility Model | Random Waypoint |
| Number of MSs | 20 |

who are in normal channel condition using 16QAM also cause less packet loss. The result of better channel condition MSs in SVC+AMC is the same to the AMC scheme, because the retrieval frame is the same as the non-encoded one.
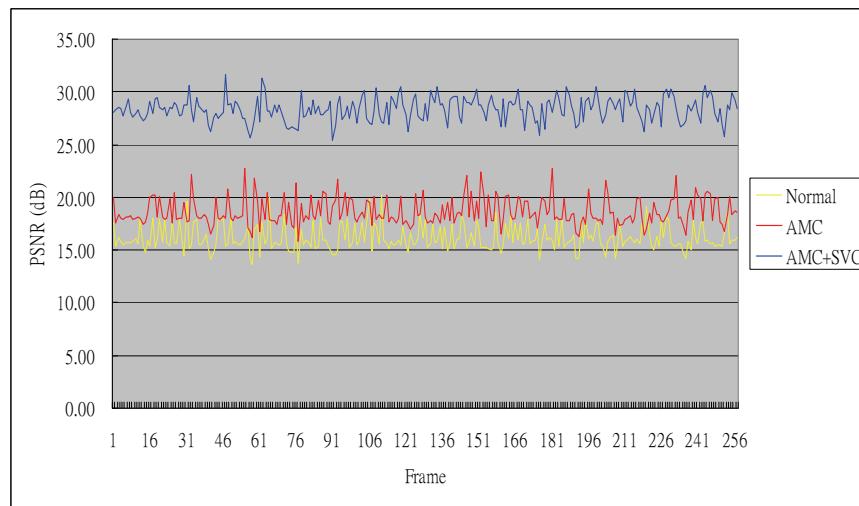
In Scenario2 of the first video, 80% of MSs are in worse channel condition, 10% of MSs are in normal channel condition, and 10% of MSs are in better channel condition. Therefore, there are 20 MSs in group 0, 4 MSs in group 1, and 2 MSs in group 2. The average values of PSNR in three schemes are 16.15, 18.60, and 28.37. The simulation results of Scenario2 employed with these three methods is shown in **Figure 12**. The reason why the PSNR in the AMC scheme is higher than that in the

normal scheme, and the SVC+AMC scheme gets the best quality is the same as Scenario1.

However, the variation between the two scenarios is obvious. In Scenario1, PSNR ranges from 26.22 to 32.18, but PSNR ranges from 16.15 to 28.37 in Scenario2. Each MS in better channel condition whether the percentage is 10% or 40% receives almost the same level of the video quality. But PSNR is influenced more because 80% of MSs with worse channel condition. Once the packet loss rate gets higher, the value of PSNR becomes lower. Therefore, the proposed SVC + AMC scheme enhances the user perceived video quality better in the situation that the percentage of worse channel condition is larger in the system.



**Figure 11. 20% worse channel condition MSs, 40% normal channel condition MSs, and 40% better channel condition MSs in normal, AMC, and SVC + AMC methods in the first video.**
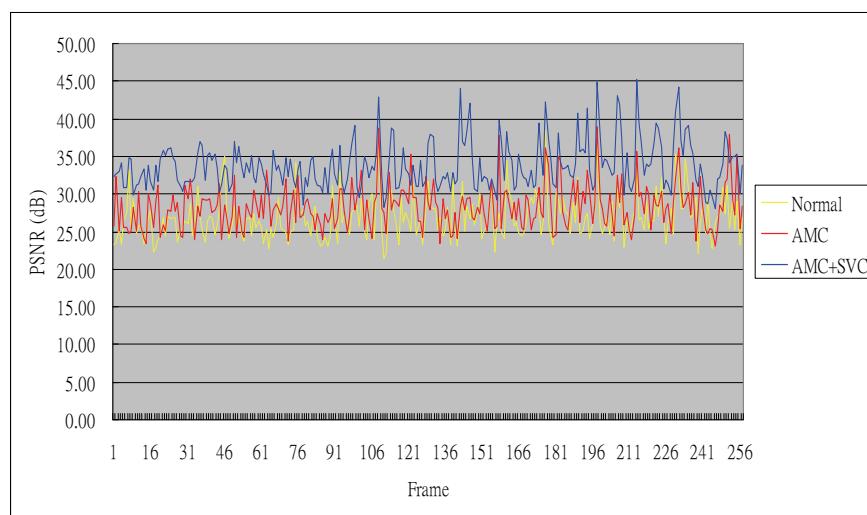


**Figure 12. 80% worse channel condition MSs, 10% normal channel condition MSs, and 10% better channel condition MSs in normal, AMC, and SVC + AMC methods in the first video.**
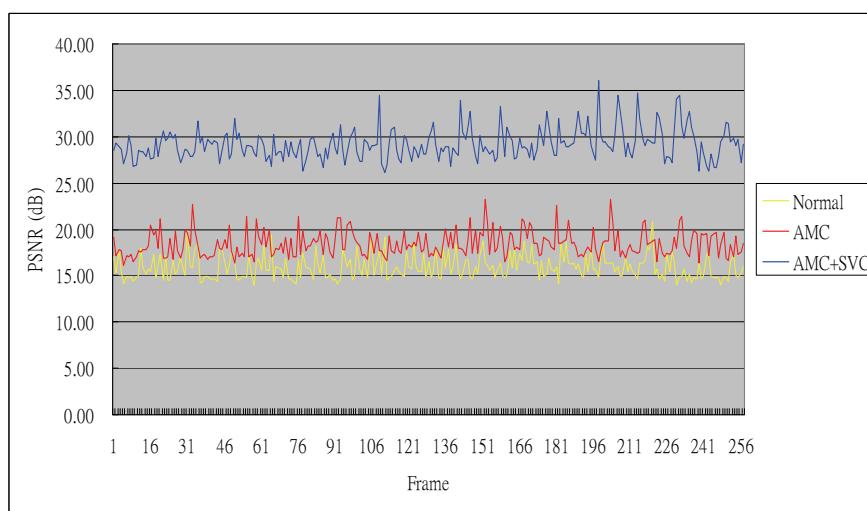
      

Three methods are also considered with two scenarios in the second video. The simulation results are shown separately in **Figure 13** and **Figure 14**. In Scenario1 of the second video, 20% worse channel condition MSs, 40% normal channel condition MSs, and 40% better channel condition MSs in normal, AMC, and SVC+ AMC methods. Therefore the average values of PSNR in three schemes are 26.90, 28.30, and 33.76. In Scenario2 of the second video, 80% worse channel condition MSs, 10% normal channel condition MSs, and 10% better channel condition MSs in normal, AMC, and SVC + AMC methods. Therefore the average values of PSNR ranges are 15.91, 18.42, and 29.19. The reason why the SVC + AMC scheme performs better than the other two

schemes is the same when simulating in the first video in despite Scenario1 or Scenario2, and the reason why Scenario1 performs better than Scenario2 is also the same for the first video.

In summary, no matter how the percentage of the MSs in worse, normal, and better channel condition changes, the proposed SVC + AMC scheme can further enhance the user perceived video quality. When a full video adapts AMC without SVC, the video quality is almost the same as the normal scheme. According to the simulation results, SVC + AMC improves the user perceived video quality well, it improves more significantly when there are more users with worse channel condition.



**Figure 13. 20% worse channel condition MSs, 40% normal channel condition MSs, and 40% better channel condition MSs in normal, AMC, and SVC + AMC methods in the second video.**



**Figure 14. 80% worse channel condition MSs, 10% normal channel condition MSs, and 10% better channel condition MSs in normal, AMC, and SVC + AMC methods in the second video.**

# 5. Conclusions

We proposed an effective video quality enhancement approach applied on mobile media-based multicast over WiMAX. A signaling mechanism is designed between the multimedia server and the BS. The channel condition varies when MSs move. The BS multicasts different number of layers to MSs within different channel condition using different modulation and coding scheme. MSs with worse channel condition receive fewer layers by more robust modulation and coding scheme. The signal mechanism between the multimedia server and the BS is performed by an ICS between the multimedia server and the BS. The main task of ICS is to decode the SVC-encoded layers from multimedia server and re-package the packet with layer ID information in different protocol headers. The mapping of the layer and CID is discussed and performed in the MAC CS specification.

Simulation results show that the proposed SVC together with AMC can indeed effectively improve the user perceived video quality, even when there are many users with bad channel condition. The ICS signaling resolves the problem that the BS can not distinguish the layer ID received from the multimedia server.

# 6. References

[1] B. Fox, "Digital TV Comes Down to Earth," *IEEE Spectrum Magazine*, Vol. 35, October 1998, pp. 23-29.

[2] R. M. Rast, "The Dawn of Digital TV," *IEEE Spectrum Magazine*, Vol. 42, October 2005, pp. 26-31.

[3] W. H. Kuo, T. H. Liu and W. J. Laio, "Utility-Based Resource Allocation for Layer-Encoded IPTV Multicast in IEEE 802.16 (WiMAX) Wireless Networks," *In Proceedings of IEEE ICC*, June 2007, pp. 1754-1759.

[4] J. She, F. Hou, P. H. Ho and L. L. Xie, "IPTV over WiMAX: Key Success Factors, Challengers, and Solutions," *IEEE Communication Magazine*, Vol. 45, August 2007, pp. 87-93.

[5] O. I. Hillestad, A. Perkis, V. Genc, S. Murphy and J. Murthy, "Adaptive H.264/MPEG-4 SVC Video over IEEE 802.16 Broadband Wireless Networks," *In Proceedings of IEEE Packet Video*, November 2007, pp. 26-35.

[6] H. H. Juan, H. C. Huang, C. Y. Huang and T. H. Chiang,

"Scalable Video Streaming over Mobile WiMAX," *In Proceedings of IEEE ISCAC*, May 2007, pp. 3463-3466.

[7] J. C. Chiang, H. F. Lo and W.T. Lee, "Scalable Video Coding of H.264/AVC Video Streaming with QoS-Based Active Dropping in 802.16e Networks," *In Proceedings of IEEE WAINA*, March 2008, pp. 1450-1455.

[8] T. K. Cheng, F. M. Yang, J. L. C. Wu and I. C. Lin, "Adaptive Modulation and SVC-Encoded Video IPTV Multicast over Mobile WiMAX," Computing and Communication Technologies, IEEE-RIVF 2009 International Conference, July 2009, pp. 1-4.

[9] IEEE 802.16e-2005, "IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems," February 2006.

[10] "Advanced Video Coding for Generic Audiovisual Services," ITU-T and ISO/IEC JTC1, ITU-R Recommendation H.264-ISO/IEC 14496-10 AVC, March 2003.

[11] "Joint Draft 5: Scalable Video Coding," ITU-T and ISO/IEC JTC1, JVT-R201, January 2006.

[12] "Joint Scalable Video Model JSVM-5," ITU-T and ISO/IEC JTC1, JVT-R202, January 2006.

[13] "Network Working Group Stage 2 Specification, Release 1.1," Published by WiMAX Forum, September 2007.

[14] "Network Working Group Stage 3 Specification, Release 1.1," Published by WiMAX Forum, September 2007.

[15] "RFC 3984 on RTP Payload Format for H.264 Video," February 2005.

[16] "RFC 3550 on RTP: A Transport Protocol for Real-Time Applications," July 2003.

[17] "RFC 3551 on RTP Profile for Audio and Video Conferences with Minimal Control," July 2003.

[18] "RFC 768 on User Datagram Protocol," August 1980.

[19] "RFC 762 on Assigned Numbers," January 1980.

[20] "RFC on Internet Protocol," September 1981.

[21] "Multi-Hop Relay System Evaluation Methodology (Channel Model and Performance Metric)," IEEE 802.16j-06/013r3, February 2007.

[22] "Mobile WiMAX−Part 1: A Technical Overview and Per- formance Evaluation," Published by WiMAX Forum, August 2006.

[23] L. H. Wan, W. C. Ma and Z. H. Guo, "A Crosslayer Packet Scheduling and Subchannel Allocation Scheme in 802.16e OFDMA System," *In Proceedings of IEEE WCNC*, March 2007, pp. 1865-1870.

Scientific
Research

# Radio Frequency Modelling for Future Wireless Sensor Network on Surface of the Moon

**Jayesh P. Pabari[1], Yashwant B. Acharya[1], Uday B. Desai[2], Shabbir N. Merchant[3], Barla Gopala Krishna[4]**

[1]*PLANEX, Physical Research Laboratory, Ahmedabad, India*
[2]*Indian Institute of Technology, Hyderabad, India*
[3]*Indian Institute of Technology Bombay, Mumbai, India*
[4]*Space Application Centre, Ahmedabad, India*
*Email*: *jayesh@prl.res.in*

## Abstract

In order to study lunar regolith properties, wireless sensor network is planned to be deployed on surface of the Moon. This network can be deployed having few wireless sensor nodes capable of measuring soil properties and communicating results, as and when ready. Communication scenario on lunar surface is quite different as compared to that on the Earth, as there is no atmosphere and also there are lots of craters as well as various terrain topologies. Since the deployment of sensors on the Moon is a challenging and difficult task, it is advisable to predict the behaviour of communication channel on lunar surface. However, communication models like Irregular Terrain Model used for terrestrial communication networks are not directly applicable for Unattended Ground Sensor type sensor networks and need modifications according to lunar surface conditions and lunar environment. Efforts have been put to devise a model of radio frequency environment on the Moon using basic equations governing various physical phenomena occurring during radio propagation. The model uses Digital Elevation Model of four sites of the Moon, measured by Terrain Mapping Camera on board Chandrayan-1, a recent Indian mission to the Moon. Results presented in this paper can provide understanding of percentage area coverage for given minimum received signal strength, potential sites for sensor deployment assuring wireless communication, decision whether a given sensor node can work and can provide suggestion for possible path of rover with cluster head to remain in contact with the nodes. Digital Elevation Model based results presented here can provide more insight in to the communication scenario on the Moon and can be very useful to mission planners.

**Keywords:** Wireless Model, Moon, DEM, Coverage

## 1. Introduction

In order to detect few regolith properties on lunar surface, a Wireless Sensor Network (WSN) is planned to be deployed near lunar South Pole, which utilizes wireless sensors capable of working in harsh environmental conditions. Main property of interest is permittivity obtainable from impedance sensor. Since the deployment of network on the Moon is a challenging and difficult task, it is advisable to predict the behaviour of communication channel on the lunar surface by some channel model.

There have been various models for radio communication on the Earth, for example Irregular Terrain Model (ITM) also known as Longley-Rice model given

by G. Hufford, A. Longley and W. Kissick [1] and model given by Durkin [2,3]. ITM is a good model for terrestrial communication network and it uses minimum antenna height as 0.5 m for simulation while Durkin's model does not consider multipath effect in the simulation. Hata [4] had proposed an empirical formulation of graphical path loss valid from about 150 MHz to 1500 MHz and it was extended up to 2 GHz. Walfisch and Bertoni [5] gave a model to consider impact of rooftops and building height by using diffraction to predict average signal strength at street level. Alberto Cerpa *et al*. [6] describes statistical model of lossy links in wireless sensor networks. S. Willis and C. J. Kikkert [7] have given radio propagation model for long range wireless sensor networks. Chirag Patel [8]

has described wireless channel modeling in his thesis. Vishwanath Chukkala *et al.* [9] and Anirudh Daga *et al.* [10] gave modeling and simulation of Radio Frequency (RF) environment of Mars.

In case of present application scenario, operating frequency is to be 2.4 GHz and sensors are to be deployed on lunar surface, where operating conditions are different than that on the Earth. It is known that there is no atmosphere on the Moon and there exists very high vacuum of the order of $10^{-12}$ Torr [11,12]. The communication models used for terrestrial communication networks are not directly applicable to Unattended Ground Sensor (UGS) type sensor network, like that planned on the Moon and need modification according to lunar surface condition. This has motivated us to carry out the work presented in this paper. Efforts have been put to derive modeling of radio frequency environment on the Moon, using basic equations governing various physical phenomena occurring during radio propagation. Our work uses Digital Elevation Model (DEM) of four sites of the Moon, measured by Terrain Mapping Camera (TMC) on board Chandrayan-1, a recent Indian mission to the Moon. Results presented in this paper can provide understanding of percentage area coverage for given minimum received signal strength, potential sites for sensor deployment assuring wireless communication, decision whether a given sensor node could be used and suggestion for possible path of rover (carrying cluster head) to remain in contact with the nodes.

Section 2 presents suitability of existing propagation models for lunar application, Section 3 describes physical phenomena which can occur on the Moon and associated path losses, Section 4 gives lunar wireless model, Section 5 gives details of DEM of selected sites on lunar surface, Section 6 shows results and paper ends with conclusion.

## 2. Suitability of Existing Propagation Models for Lunar Application

It is expected that the communication would be better on the Moon as compared to that on the Earth, due to absence of atmosphere on the Moon. Few existing propagation models were developed for communication on the Earth and need to be reviewed for the Moon in view of their applicability.

### 2.1. Irregular Terrain Model

Irregular Terrain Model (ITM) was suggested by Rice *et al.* [13] and is also known as Longley-Rice model. Hufford [14] informed that it can be used in area prediction mode and point-to-point mode. ITM takes terrain and other parameters as input and produces output as signal distribution in a given area. There are certain limitations of ITM that it can be used for minimum antenna height of 0.5 metre and minimum distance for communication as 1 km. The Point-To-Point (PTP) model given by Wong [15] is based on Longley-Rice model. However, PTP model describes method to obtain diffraction loss. As the wireless sensor network involves deployment of sensors on the ground with very small antenna heights especially at 2.4 GHz operating frequency, the ITM can't be used for such applications.

### 2.2. Two-Ray Model

For a line-of-sight communication, a two-ray model was given by Neskovic *et al.* [16]. This theoretical model incorporates reflection by the reflection coefficient, which is calculated from incidence angle, dielectric constant, surface conductivity and polarization of antenna. The signal strength at the receiver as given in (1) is shown in Willis [7] using free space loss and reflection.

$$P_r = \frac{P_t G_t G_r \cdot \lambda^2}{16 \cdot \pi^2} \cdot | \frac{1}{d_1} \exp(-jkd_1) + \Gamma(\psi) \frac{1}{d_2} \exp(-jkd_2) |^2 \tag{1}$$

where $d_1$ and $d_2$ are lengths of first and second path respectively. The two ray model is mostly used for direct and ground reflected rays. In case of lunar wireless network, sensor nodes are to be the surface with very small antenna heights and therefore ground reflected signal is not expected at the receiver.

### 2.3. VSS Multipath Model

Signals reflected from reflectors would reach the receiver with different strengths and phases with reference to direct signal. At receiver, equivalent signal strength is equal to superposition of varying amplitude and phase signals and it is possible to get overall signal strength as reduced or improved. The multipath channel model in AWR Visual System Simulator [17] uses Equations (2) and (3) to obtain signal strength for a sample *k*. The multipath model was implemented by Willis [18] in MATLAB.

$$x(k) = \sum_{i=1}^{N} path(k,i) \tag{2}$$

$$path(k,i) = A_i \cdot x(k-d_i) \cdot \exp(\frac{j2\pi \cdot V \cdot f_c \cdot Cos\theta_i}{C}) \tag{3}$$

where $x(k)$ is *k*th sample, $N$ is number of multipath signals, $path(k,i)$ is contribution of *i*th multipath signal determined by (3), $A_i$ is gain of *i*th path, $x(k-d_i)$

is delayed sample associated with path $i$ and exponential term represents doppler shift due to receiver movement. This model may be used for moving nodes, which is not the case for lunar wireless sensor network.

## 2.4. Multipath Signal Distribution

Instantaneous receiver signal strength (envelope) is represented by either Rayleigh or Rician distribution as shown by Hernando *et al.* [19]. Rayleigh distribution considers only multipath components available and does not take direct path. However, Rician distribution considers direct as well as multipath components. Communication on lunar surface is expected to be line-of-sight and also multipath due to reflections from terrain and therefore, Rician distribution is more suited for lunar application. Rician distribution is shown by Hernando *et al.* [19] and given below:

$$p(r) = \frac{r}{\sigma^2} \cdot \exp[-\frac{r^2+a^2}{2\cdot\sigma^2}] \cdot I_0(\frac{r\cdot a}{\sigma^2}) \quad \text{for} \ \ r \geq 0 \tag{4}$$

where $I_0$ is modified 0th order Bessel function and value of $a$ depends on direct component. The Rician function is usually expressed in terms of carrier-to-multipath ratio or $k$ factor as

$$k = \frac{c}{m} = \frac{a^2}{2\cdot\sigma^2} \tag{5}$$

where $c$ and $m$ are strengths of carrier and multipath components respectively. As $k$ increases, received signal strength increases and for larger values of $k$, $p(r)$ becomes Gaussian.

## 2.5. Willis Multipath Model

Willis [18] has given extended version of two ray model as multipath model, shown in Equation (6) for terrestrial application on the Earth.

$$P_r = \frac{P_t G_t G_r \cdot \lambda^2}{16 \cdot \pi^2} \cdot | L_{d1}\Gamma_1(\psi_1)\frac{1}{d_1}exp(-jkd_1)+... \\ + L_{dn}\Gamma_n(\psi_n)\frac{1}{d_2}exp(-jkd_2)|^2 \tag{6}$$

## 3. Physical Phenomena and Path Loss on the Lunar Surface

Since the Earth based model cannot be directly used for lunar applications, we have examined various physical phenomena occurring during wave propagation in general and those applied to wave propagation on the Moon. Such phenomena and losses associated with them are

given below and we have combined them to arrive at the path loss given in Section 4.

## 3.1. Free Space Loss

The plan is to use compact wireless impedance sensors (along with other types of sensors) for detection of electrical properties of lunar Regolith having small size quarter wave antenna (3.125 cm). The idea is to use line of sight propagation technique for small distances of up to few hundred metres. Under ideal communication conditions, the power radiated from antenna is omni directional in the plane of interest, which is horizontal here and power is inverse square function of the distance for free space wave propagation given below [19] by Friis formula.

$$P_r = \frac{P_t G_t G_r \lambda^2}{16 \cdot \pi^2 d^2} \tag{7}$$

where $P_t$ and $P_r$ are transmitter and receiver powers, $G_t$ and $G_r$ are gains of transmitter and receiver antennas, $\lambda$ is wavelength of operation and $d$ is distance from the transmitter. Free space loss is a basic loss for communication on the lunar surface, since there is a vacuum on the Moon.

## 3.2. Reflection

For communication on the Earth, there can be factors like steady and moving reflectors and scatterers, atmospheric absorption etc and can lead to multipath components at the receiver. Due to multipath components, signal strength can vary at the receiver due to moving scatterers. However, in case of lunar application, there are no moving objects on the surface and therefore there can be multipath signals at the receiver due to surface topography, but received signal strength is not expected to vary randomly due to steady terrain, but it can show periodic variations due to signals travelling in different time durations. Also, there is no atmosphere on the Moon and therefore there are no atmospheric losses, which are present on the Earth. The main possibility of signal getting affected is that of the surface reflections due to uneven terrain structure. This can cause multipath propagation and signal reaching at the receiver by direct path will be modified due to multipath components. The expected number of multipath components is few as the mission lander should land on comparatively plane or smooth surface, where rover can move easily. The wavelength of signal is 12.5 cm at 2.4 GHz wireless operation and objects should be of larger size to cause the reflection. Parsons [20] derived the received signal strength using reflection coefficient (Γ) from the reflector as given below.

$$P_r = \frac{P_t G_t G_r \lambda^2}{16 \cdot \pi^2 d^2} \cdot | 1 + \Gamma \exp(j\Delta) |^2 \qquad (8)$$

where $\Gamma = \dfrac{a \cdot Sin\psi - \sqrt{(\varepsilon_r - jx) - \cos^2\psi}}{a \cdot Sin\psi + \sqrt{(\varepsilon_r - jx) - \cos^2\psi}}$ is reflection

coefficient

$\Delta$ = phase shift created due to reflected signal

$a$ = 1 or $(\varepsilon_r - jx)$ for horizontal or vertical polarization respectively

$\psi$ = angle of incidence with vertical from transmitter antenna to the reflector

$\varepsilon_r$ = relative dielectric constant of the ground

$x = \dfrac{\sigma}{2 \cdot \pi \cdot f \cdot \varepsilon_0}$

$\sigma$ = conductivity of the ground

$f$ = frequency of operation

$\varepsilon_0$ = free space permittivity

It is also possible to calculate received signal strength by the Equation (9) given by Hernando *et al.* [19] derived from two ray model for longer link distances and lower antenna heights at both ends having direct visibility of each other.

$$P_r = \frac{P_t G_t G_r h_t^2 h_r^2}{d^4} \qquad (9)$$

where $h_t$ and $h_r$ are heights of transmitter and receiver antennas respectively. However, for sensors being on surface, Equation (9) may not be used. Reflections are expected from nearby ground terrain on the Moon. Infact the major component for channel fading on the Moon is due to multipath created by reflections from craters and surface irregularities. Also, transceiver used for ground sensors should not use horizontal polarization; otherwise signal would be attenuated [21] very near the transmitter. Instead, vertical polarization should be used, which can provide sufficient signal at farther distances.

### 3.3. Reflection Scattering

When signal is reflected from lunar surface, it is likely that the ray may be scattered due to dispersed signal. Gibson [22] has suggested this specular reflection and stated that roughness of a surface can be classified by Rayleigh criterion given below:

$$h_c = \frac{\lambda}{8 \cdot \cos\theta_i} \qquad (10)$$

where $\theta_i$ is angle of incidence at the reflector. Gibson [22] gave a parameter $h$, which represents minimum to maximum deviation about mean terrain height. If $h > h_c$,

then terrain is considered as rough terrain and the loss for it is considered by multiplying reflection coefficient by a scattering loss factor $\rho_s$ calculated by Bothias' equation as below:

$$\rho_s = \exp[-8 \cdot (\frac{\pi\sigma_h \cos\theta_i}{\lambda})^2] \cdot I_0[8(\frac{\pi\sigma_h \cos\theta_i}{\lambda})^2] \qquad (11)$$

where $\sigma_h$ is standard deviation of surface height about the mean surface height and $I_0$ is the $0^{th}$ order Bessel function of the first kind. On the lunar surface, the regions are mostly rough and may have varying sized objects, which can lead to scattering losses.

### 3.4. Diffraction

For line of sight communication, if an obstacle of size comparable to wavelength is present in between transmitter and receiver, then diffraction loss can occur at the edge of an obstacle. The signal may be scattered and attenuated before reaching to the receiver. Diffraction from knife-edge obstacle can cause signal to bend and Wong [15] stated that bending of signal due to knife-edge obstacle causes higher signal strength as compared to that due to rounded. On the lunar surface it is likely to have such obstacles in between transmitter and receiver, which may be considered only for direct path but the deployment is supposed to be in almost plane terrain for smooth movement of the rover and hence possibility of occurrence of such loss is rare and may be neglected.

## 4. Lunar Wireless Model

Equation (6) given by Willis [18] does not include possibility of direct path, but it is expected to be present in case of wireless sensor networks. Also, diffraction loss given by Wong [15] may be considered for direct path from transmitter to receiver and not for multipath for operation at 2.4 GHz frequency. Considering diffraction loss only for direct path and including reflection scattering loss factor, following equation is given for obtaining area coverage by wireless sensor network on the lunar surface:

$$\begin{aligned}
P_r = \frac{P_t G_t G_r \cdot \lambda^2}{16 \cdot \pi^2} \cdot | L_{dd} \frac{1}{d_d} \exp(-jkd_d) \\
+ \rho_{s1}\Gamma_1(\psi_1)\frac{1}{d_1}\exp(-jkd_1) \\
+ \rho_{s2}\Gamma_2(\psi_2)\frac{1}{d_2}\exp(-jkd_2)+... \\
+ \rho_{sn}\Gamma_n(\psi_n)\frac{1}{d_n}\exp(-jkd_n) |^2
\end{aligned} \qquad (12)$$

where $L_{dd}$ is diffraction loss for direct path, $d_d$ is direct path distance, $k$ is phase constant, $\rho_{sn}$ is reflection scattering loss factor of $n$th multipath, $\Gamma_n$ is reflection coefficient of $n$th multipath, $\psi_n$ is angle of incidence at $n$th reflector and $d_n$ is distance of $n$th multipath component.

## 5. Digital Elevation Model of Lunar Surface
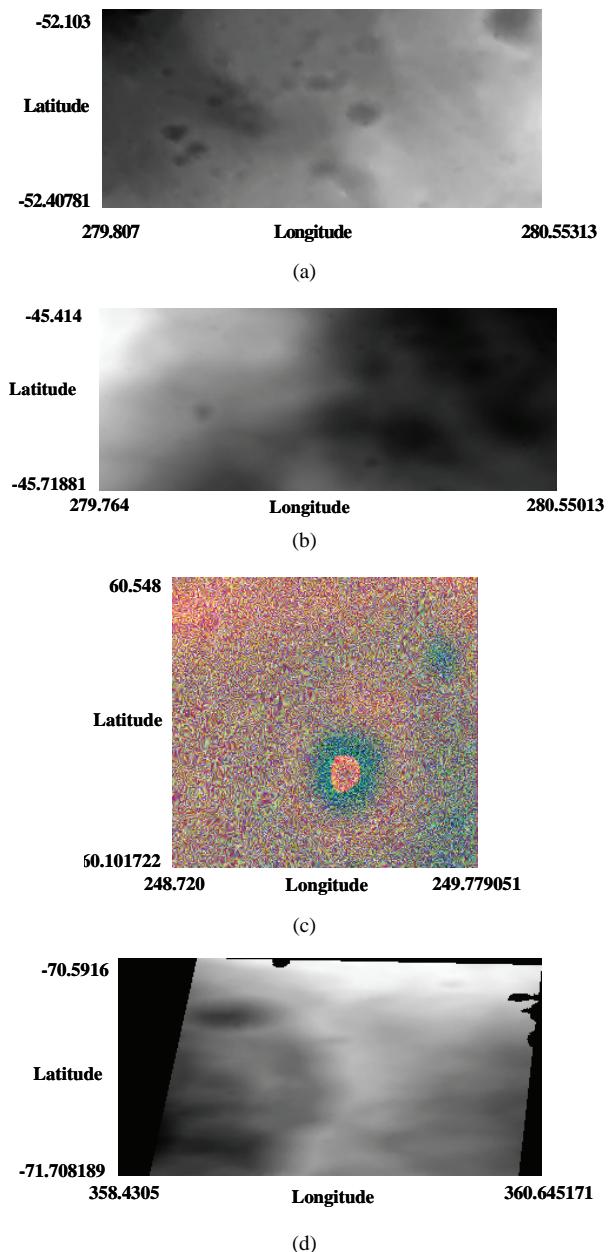
Recent mission to the Moon from India, Chandrayan-1, had Terrain Mapping Camera (TMC) on board for deriving Digital Elevation Model (DEM) of the Moon and has provided good quality data during the mission time. Four sample sites at various locations on the Moon have been selected considering almost plane surface, region with some peaks and region with few craters in order to study lunar radio propagation model. **Table 1** shows details of site used in deriving communication area coverage. DEM data for various sites were obtained during different orbits and resolutions were different. Sites 1 and 2 have resolution of 56.1 m in both directions; site 3 has resolution of 122.67 m for horizontal (longitude) direction and 53.34 m for vertical (latitude) direction, while site 4 has 185.07 m for both directions. **Figures 1(a)** to **1(d)** show images of all these four sites respectively, taken by TMC on board Chandrayan-1.

## 6. Results

Lunar wireless model is given in Equation (12), which considers all possible phenomena and losses for wireless sensor network on the lunar surface. Diffraction loss is expected to be very low in value as compared to direct path and multipath reflections for targeted application. Since, major interest is in establishing direct line-of-sight communication between transmitter and receiver, such loss is neglected. As readily available software cannot be used for wireless sensor network for lunar applications, MATLAB code was developed and DEM data were taken as input in the programme along with values shown in **Table 2** and results of radio coverage are



(a)



(b)



(c)



(d)

**Figure 1. (a) Lunar site 1; (b) Lunar site 2; (c) Lunar site 3; (d) Lunar site 4.**

**Table 1. Lunar sites for RF model.**
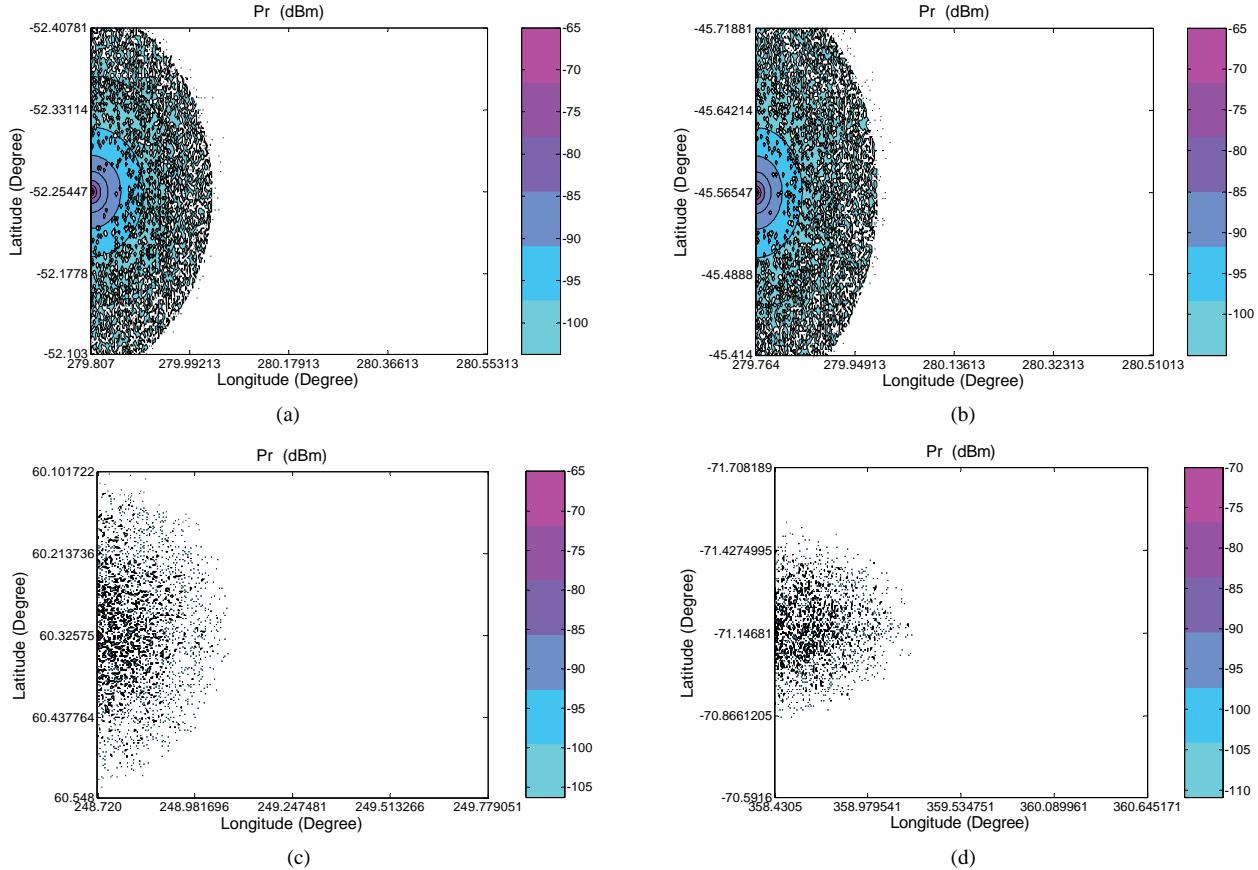
| Site No. | Crater Near the Site Location | Lunar Latitude | Lunar Longitude |
|---|---|---|---|
| 1 | Catalan | − 52.103 to − 52.40781 | 279.807 to 280.55313 |
| 2 | Baade | − 45.414 to − 45.71881 | 279.764 to 280.51013 |
| 3 | Zsigmondy | 60.548 to 60.101722 | 248.720 to 249.779051 |
| 4 | Moretus | − 70.5916 to − 71.708189 | 358.4305 to 360.645171 |

**Table 2. Input parameters for lunar radio frequency model.**

| Parameter for Model | Value |
|---|---|
| Moon Radius | 1737.4 km |
| Transmitter Power | 0 dBm |
| Transmitter Antenna | Quarter Wave Monopole |
| Receiver Antenna | Quarter Wave Monopole |
| Frequency of Operation | 2.4 GHz |
| Lunar Regolith Dielectric Constant [24] | 4 |
| Lunar Regolith Conductivity [25] | $10^{-8}$ s/m |
| Wave Polarization | Vertical |
| Transmitter Location | Middle Left |

Figure 2. (a) RF model for site 1; (b) RF model for site 2; (c) RF model for site 3; (d) RF model for site 4.

depicted in **Figures 2(a)** to **2(d)** respectively; for sites 1, 2, 3 and 4. For a given area on the Moon, there can be many multipath components depending up on terrain and if all are allowed to contribute in the code, then computational complexity is increased highly due to involvement of higher order matrix. Moreover, it needs highly sophisticated computing facility and it takes very long time for computation. Number of multipath components may, therefore, be restricted to smaller numbers, which is reasonably justified, as distance travelled by signal along multipath is quite large in most of the regions as compared to that for the direct path and hence the contribution of those multipath components would be very less as compared to contribution from the direct path. Normally, sensor node operating in 2.4 GHz ISM band has 0 dBm as output power and its receiver has sensitivity of about -100 dBm [23]. We have calculated radio coverage of the sites as a ratio of occupied region having more than -100 dBm received signal strength to total region under consideration. **Table 3** shows percentage radio coverage of all sites for 250 kbps link, effectively showing useful area for sensor network deployment.

Sites 1 and 2 have been selected with comparatively plane surface having lesser undulations on the surface,

while sites 3 and 4 have different profiles, for example, site 4 is at the edge of Moretus crater near lunar South Pole. Due to this, site coverage is very less for sites 3 and 4 as compared to that for sites 1 and 2, as expected. From the derived radio coverage at selected lunar sites, one can know the areas having more than some specific value of received power, for example, region receiving more than -100 dBm power. Results indicate the percentage coverage of a given site with known topography and also the possibility of checking if a given sensor node would be useful on the Moon, as far as power is concerned. Sensor nodes should preferably be deployed in the region where sufficient amount of power is available (pink in the coverage patterns). In case of lunar wireless sensor network, rover is supposed to carry the cluster

**Table 3. Site coverage.**

| Site | Site Coverage (250 kbps link) |
|------|------------------------------|
| 1 | 11.27 % |
| 2 | 10.02 % |
| 3 | 2.73 % |
| 4 | 1.01 % |

head and radio coverage patterns can suggest possible path for the rover to move, assuring intact communication link with the nodes.

## 7. Conclusions

In this paper, we have investigated that existing radio frequency models cannot be directly applied to lunar wireless sensor network and arrived at lunar wireless model from fundamental physical phenomena occurring during wave propagation on the Moon. We have presented radio signal coverage patterns of four lunar sites by using actual DEM data of the Moon. We have used 0 dBm transmitter power at 2.4 GHz frequency with quarter wave antennas for transmitter and receiver. The results show percentage coverage for 250 kbps links on the lunar surface, suggesting possible use of commercially available transceiver in sensor node as well as possible deployment sites and rover paths to assure wireless connectivity.

## 8. References

[1] G. A. Hufford, A. G. Longley and W. A. Kissick, "A Guide to the Use of the ITS Irregular Terrain Model in the Area Prediction Mode," *National Telecommunication and Information Administration*, Report 82-100, NTIS Document PB82-217977, April 1982.

[2] R. Edwards and J. Durkin, "Computer Prediction of Service Areas for VHF Mobile Radio Networks," *Proceedings of IEE*, Vol. 116, September 1969, pp. 1493-1500.

[3] J. Durkin, "Computer Prediction of Service Areas for VHF and UHF Land Mobile Radio Services," *IEEE Transactions on Vehicular Technology*, November 1977, pp. 323-327.

[4] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services," *IEEE Transactions on Vehicular Technology*, August 1980, pp. 317-325.

[5] J. Walfisch and H. L. Bertoni, " A Theoretical Model of UHF Propagation in Urban Environments," *IEEE Transactions on Antennas Propagation*, Vol. 36, 1988, pp. 1788-1796.

[6] C. A. E., Wong, J. L., Kuang, L. Potkonjak and M. D. Estrin, "Statistical Model of Lossy Links in Wireless Sensor Networks," *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks*, Center for Embedded Network Sensing, University of California Los Angeles, 2005. http://escholarship.org/uc/ item/4s2698fs

[7] S. L. Willis and C. J. Kikkert, "Radio Propagation Model for Long-Range Ad Hoc Wireless Sensor Network," *International Conference on Wireless Networks, Communications and Mobile Computing*, Vol. 1, No. 13-16, 2005, pp. 826-832. http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2F10391%2F33022%2F01549514.pdf%3Farnumber%3D1549514&authDecision=-203

[8] C. S. Patel, "Wireless Channel Modeling, Simulation, and Estimation," Ph. D. Thesis, *School of Electrical and Computer Engineering, Georgia Institute of Technology*, May 2006.

[9] V. Chukkala, P. DeLeon, S. Horan and V. Velusamy, "Radio Frequency Channel Modeling for Proximity Networks on The Martian Surface," *Computer Networks*, Vol. 47, 2005, pp. 751-763.

[10] A. Daga, G. R. Lovelace, D. K. Borah and P. L. De Leon, "Terrain-Based Simulation of IEEE 802.11a and b Physical Layers on the Martian Surface," *IEEE Transactions on Aerospace and Electronic System*, Vol. 43, No. 4, October 2007, pp. 1617-1624.

[11] P. D. Spudis, "Introduction to the Moon, Moon 101, NASA Johnson Space Centre," 17 December 2009. http://www.spudislunarresources.com/moon101.htm

[12] "Anonymous," 17 December 2009. http://en.wikipedia.org/wiki/Moon

[13] P. L. Rice, A. G. Longley, K. A. Norton and A. P. Barsis, "Transmission Loss Predictions for Tropospheric Communication Circuits," *Technical Note* 101, National Bureau of Standards, Vol. 1-2, 1967.

[14] G. Hufford, "The ITS Irregular Terrain Model Version 1.2.2 - The Algorithm," 12 December 2009. http://flattop.its.bldrdoc.gov/itm/itm_alg.pdf

[15] H. Wong, "Field Strength Prediction in Irregular Terrain - The PTP Model," 12 December 2009. http://www.fcc.gov/oet/fm/ptp/report.pdf

[16] A. Neskovic, N. Neskovic and G. Paunovic, "Modern Approaches in Modeling of Mobile Radio Systems Propagation Environment," *IEEE Communications Surveys*, Vol. 3, No. 3, 2000, pp. 1-12.

[17] "Microwave Office," 12 December 2009. http://web.awrcorp.com/products/mwoffice/

[18] S. L. Willis, "Investigation in to Long Range Wireless Sensor Networks," Ph. D. Thesis, James Cook University, Townsville, Australia, December 2007.

[19] J. M. Hernando and F. Perez-Fontan, "Introduction to Mobile Communications Engineering," Artech House, Boston, 1999.

[20] J. D. Parsons, "The Mobile Radio Propagation Channel," 2nd Edition, Wiley, New York, 2000.

[21] G. Kennedy, "Electronic Communication Systems," 3rd Edition, McGraw-Hill Book Company, 1985.

[22] J. D. Gibson, "The Mobile Communications Handbook," CRC Press Inc., Florida, 1996.

[23] Atmel, 17 December 2009. http://www.atmel.com/dyn/products/product_card.asp?part_id=3941

[24] M. G. Buehler, H. Bostic, K. B. Chin, T. McCann, D. Keymeulen, R. C. Anderson, S. Seshadri and M. G. Schaap, "Electrical Properties Cup (EPC) for Characterizing Water Content of Martian and Lunar Soils," *IEEE Aerospce Conference*, 2006. http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2F11012%2F34697%2F01655752.pdf&authDecision=-203

[25] D. H. Chung, W. B. Westphal and G. R. Olhofet, "Di-Electric Properties of Apollo 14 Lunar Samples," *Proceedings of the Third Lunar Science Conference*, The MIT Press, Cambridge, Vol. 3, 1972, pp. 3161-3172.

Scientific
Research

# Oilfield GIS Service Based on Mobile Platform

**Lingling Guo[1], Man Yuan[1], Shaobin Hu[2]**

[1]*School of Computer & Information Technology, Daqing Petroleum Institute, Daqing, China*
[2]*Key Laboratory of Educational Ministry for Improving Oil and Gas Recovery, Petroleum Engineering Department,*
*Daqing Petroleum Institute, Daqing, China*
*Email*: *rainmanqing@yahoo.com.cn, hsbdqpi@yahoo.com.cn*

## Abstract

With the development of mobile technology, especially the development of 3G and mobile IP, the computational capacity of handsets is becoming more and more powerful, which provides a new method to solve the difficulties encountered in real time GIS accessing caused by the characteristics of mobility and remoteness of fieldwork in oilfield. On the basis of studying in-depth on the technologies of J2ME platform, Mobile SVG and mobile data transfer, etc., and in accordance with the actual situation of oilfield, the design framework of oilfield mobile GIS service is put forward and the schemes of key technologies are given in this paper, which establishes the technical foundation for the construction of "Digital Oilfield".

## 1. Introduction

Oilfields, especially the peripheral oilfields, are generally located in remote areas and the workplaces scatter widely and span large regions. The field operations, such as geophysical prospecting, drilling, logging, well testing have the characteristic of strong mobility. In this case, it is difficult to build wired network. But on the other hand, it is needed to access and control the oilfield GIS information real-timely by the "Digital Oilfield". In such a case, oilfields proposed higher requirements on the WWW-based GIS platform. In recent years, with the development of mobile communication and mobile computing technologies, it is possible for the traditional GIS to develop into mobile GIS [1-3]. The mobile GIS with handsets as the terminal is becoming a research hotspot in GIS domain. Value added service of mobile GIS in oilfield is emerging as the times require, it is meaningful to carry out the research on mobile platform-based oilfield GIS service.

Mobile GIS services are based on the technologies of Mobile SVG, J2ME platform, etc. SVG, the abbreviation of scalable vector graphics, is the formal recommendation standard of internet (W3C), and a language which uses XML to describe the two-dimensional diagrams [4]. In order to meet the requirements in wireless domain, the SVG workgroup of W3C established the specific standard Mobile SVG for the mobile application domain. For mobile developer, vector graphic has the advantages of small

space usage and nice flexibility, so it is widely used in the mobile network. An important application of Mobile SVG is the location-based services, an important part of which is the handset-based map service, the product of the combination of the technologies of spatial information, mobile communication, mobile computing and so on. As a new emerging service, it has the characteristics of mobility, diversity of handsets, thin client/fat server, and limited bandwidth. J2ME is the abbreviation of Java 2 Micro Edition released by SUN, is a development platform aiming at the embedded consumption electronic products, and is a whole collection of technology and specification [5]. The discussion of this contribution is how to realize oilfield GIS service on the basis of Mobile SVG specification and J2ME platform.

Compared with the services provided by sitemap, such as the well-known Google Map, oilfield GIS service is committed to different objectives and the user groups are not the same. Microsoft, Google and Yahoo are consumer-centered to a large extent. They take mapping and geo-spatial visualization as an aspect of the search technology, trying to achieve the visualization and interrelation of global base map. While the oilfield GIS service is oilfield oriented. By using the professional data of oilfield, oilfield GIS services create the contents that can be visualized on the mobile terminal, supporting the professional and technical personnels to carry out various analyzing and decision making.

## 2. Overall Design of Mobile Platform-Based Oilfield GIS Service

The prototype of the mobile platform-based oilfield GIS service applies C/S structure. It is divided into three layers: data layer, middle layer and presentation layer, the hierarchical structure is showed in **Figure 1**.

Presentation layer: This layer implements the graphics user interface, the data representation (map display) and the simple operations on map (map zoom in or out, etc.) for various kinds of handset equipments. Because of the limited process and memory capability, most of the computing work must be implemented at the server end, while the handset can only implement the visualization of spatial data and the communication between the user and server. Due to the limited bandwidth, the SVG file transferred by the server to the handset can only contain graphic data. When map service is required by user, it sends HTTP request to the server, the server responses the user request and sends the SVG-formatted spatial data to the client, and then the client handset parses and generates the graphic map [6]. If the user needs other information query or data processing, communication with the server is required again.
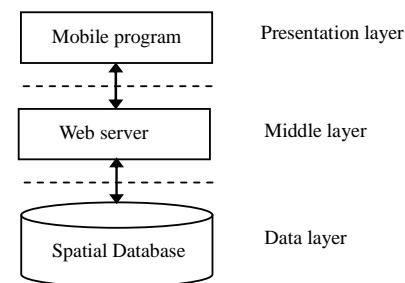
Middle layer: Middle layer is made up by web server. It requests data from the data layer in accordance with the user demand, implements necessary information process and communicates with the client end. Web server acepts the service request sent by the client based on HTTP protocol, implements the data query in the back-end spatial database, then sends the query results back to the client after compression.

Data layer: It is the spatial database on the server end. It stores various geographic data needed by the mobile platform-based oilfield GIS services. After development for scores of years, oilfields have accumulated mass of data, including raw data in exploration and exploitation, results of processing and interpretation and information of surface engineering, etc., which are the basis of the spatial database. These data are classified into three sorts: basic geographic information, metadata and oilfield professional data (including gathering and transportation system for crude and gas, water injection system, water supply system, drainage system, road system, power supply system, corrosion protection and thermal insulation, communica-

tion system, exploration and exploitation related mining right, reservoir boundary, basin and formation boundary, ground, fire fighting, sanitation and so on). The structure diagram of oilfield spatial attribute database is shown in **Figure 2**.
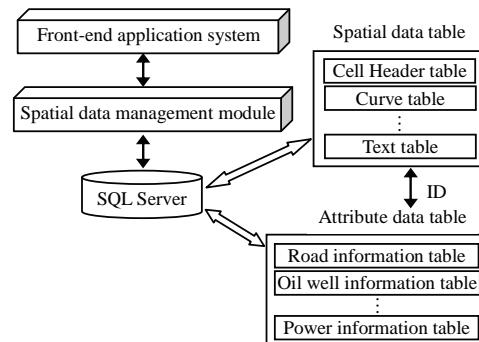
## 3. Key Technologies
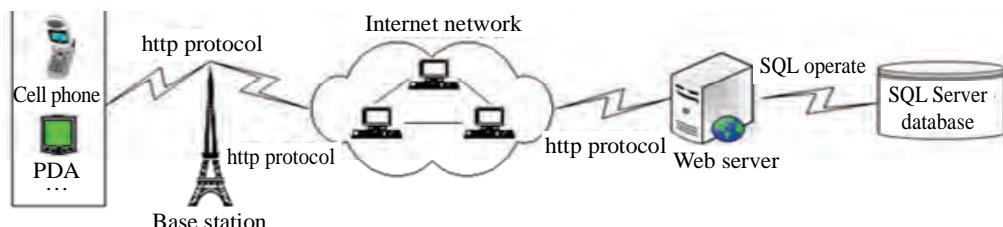
The mobile platform-based oilfield GIS service prototype is made up by two parts: client end handset application program and Web server program. The handset application program mainly implements the interactive graphic user interface, SVG file parsing and the map display and operations (map translation, zoom in or out, etc.). The functions of Web server are mainly to accept the handset's HTTP service request, implement the data query, generate and compress the SVG map file, then send the map data file back to the client handset [7,8]. The structure of the prototype is shown in **Figure 3**.



**Figure 1. Layer structure of mobile platform-based oilfield GIS service.**



**Figure 2. Structure of spatial attribute database.**



**Figure 3. System architecture diagram of the prototype.**

### 3.1. Generation of SVG File at the Server End

At the server end, the file data are written to SVG file in accordance with the Mobile SVG specification, all the related geographic information needed in the SVG file are obtained by querying the spatial database, and these data are assigned to the corresponding attributes of the elements in the SVG file.

Based on the characteristics of Mobile SVG, the ground objects can be classified into point entity, linear entity and area entity according to their geometrical shape [9]. These entities are represented by corresponding graphic coding in Mobile SVG. In Mobile SVG, basic graphic elements (rectangular, circle, ellipse, line, folding line, polygon, etc.) are supported, so as the path. Various complex graphics can be plotted with path elements.

The data fields of various entity objects in the spatial database are basically the same, while the types of the spatial data of the entity objects are different. Thus, the data query should be implemented in three ways according to the entity objects. The queried contents mainly include ID, name and spatial coordinate data. As for the attributes, such as color of the entity objects, it can be acquired by querying the entity description table in the database. After the correlated data of each entity are acquired, the corresponding group and element can be generated in accordance with the Mobile SVG normative standard, and written to the SVG file as the values of the attributes of various elements, generating the demanded SVG map file.

### 3.2. SVG File Compression

Compared with other graphic formats, the SVG technology itself has the advantage of small data amount, but for the bandwidth-limited wireless network, the less the transmitted data amount, the better. As a result, compression treatment is implemented on the SVG file by using GZIP format at the server end [10].

### 3.3. SVG File Parsing at the Client End

At the mobile terminal, the SVGZ file from the server should be decompressed into SVG file. SVG file is based on XML. Presently, there are special-handset-aimed XML parsers, such as KXML [11]. Tinyline viewer is a special browser to browse the SVG file on handsets, which can be installed on handsets directly. On the other hand, one can develop the SVG file browser by himself, too. Due to the limited capacity of Tinyline viewer, a self developed browser is applied in this prototype [12].

During the process of file parsing, a third party developed SVG file parsing package, tinyline.zip, is applied. This parsing package can parse the graphic elements and their attributes in the Mobile SVG standardaccorded
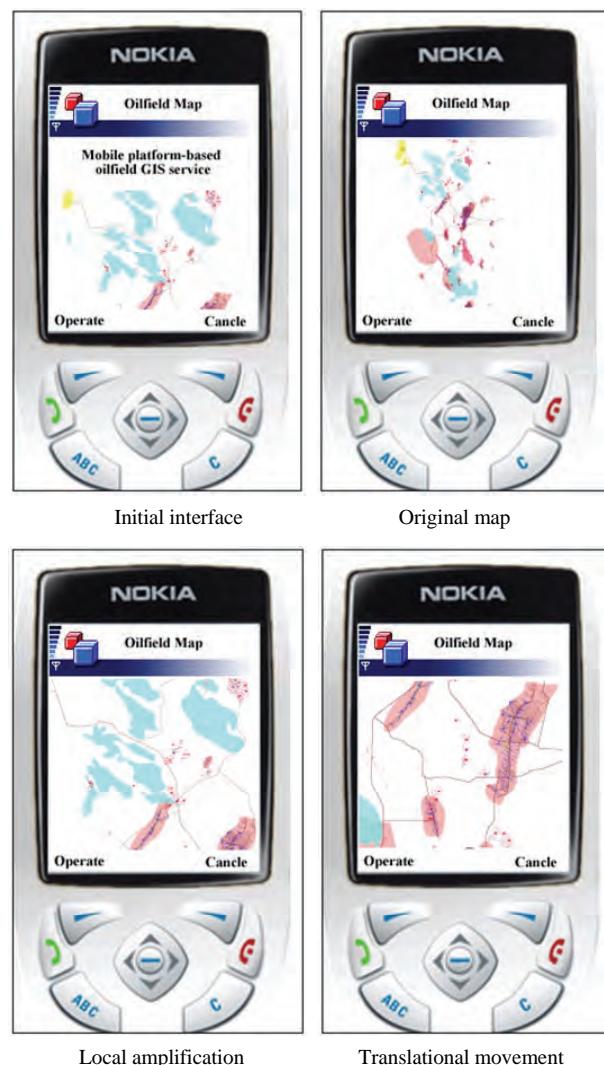
SVG file. It transforms them into the J2ME graphics and displays on the handset screen.
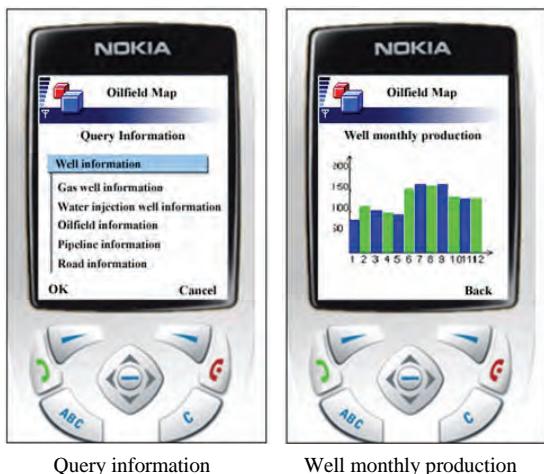
### 3.4. Map Display on Client Handset and Simple Operations

On the client handset, simple operations on the parsed SVG file can be implemented by the user, such as map zoom in or out, translational movement, and so on.

## 4. Application Example

Through the study on the technologies of J2ME and Mobile SVG, a mobile platform-based oilfield GIS service prototype was realized by using the geographic information data of an oil production plant in Daqing oilfield. Using this prototype, one can real-timely query the surface engineering map of the plant and carry out simple operations on it, as shown in **Figure 4**.



| Initial interface | Original map |

| Local amplification | Translational movement |

Query information          Well monthly production

**Figure 4. Operation interface of surface engineering map of oilfield.**

## 5. Conclusions

The design and realization of mobile platform-based oilfield GIS service prototype are introduced in this contribution. Making use of the characteristics of Mobile SVG and its advantage in the wireless graphic application, combining the J2ME development platform, the prototype organizes and codes the oilfield spatial data, then transmits the file through Web server to the client end to display. The graphics generated by using the Mobile SVG technology described in this contribution has the advantages of small size, fast access speed and easily process, therefore it is the first choice for graphic displaying at the client end of the mobile platform-based oilfield GIS service.

The study on the mobile platform-based oilfield GIS service in this contribution plays the navigation role in the "Digital Oilfield" based applications, and establishes the technical foundation for the construction of "Digital Oilfield" by using the front technology of mobile GIS.

## 6. Acknowledgment

## 7. References

[1]  R. J. Zhang, H. Qi and W. J. Han, *et al.* "Design and Implement of Mobile GIS Based J2ME/Mobile SVG," *Microcomputer Information,* Vol. 3, No. 3, 2006, pp. 164-166.

[2]  B. Z. Wu and B. Xia, "Mobile Phone GIS Based on Mobile SVG," *IEEE International Geoscience and Remote Sensing Symposium*, Seoul, Korea, Vol. 2, January 2005, pp. 889-892.

[3]  X. Q. Zuo and Q. Q. Li, "The Deliver and Visualization of Geospatial Information in Mobile GIS," *International Conference on Wireless Communications, Networking and Mobile Computing,* Wuhan, September 2005, pp. 1348-1351.

[4]  "W3C. Scalable Vector Graphics (SVG) 1.1 Specification [EB/OL]." http://www.w3.org/TR/SVG/,2003-01/2009-02

[5]  "Java 2 Platform, Micro Edition (J2ME) Overview, Sun Microsystems," 10 November 2004. http://java.sun.com/j2me/

[6]  "W3C. Mobile SVG Profiles: SVG Tiny and SVG Basic [EB/OL]." http://www.w3.org/TR/2002/PR-SVGMobile-20021115,2002–11/2009–02

[7]  Y. G. Li, X. F. Zhou and M. Xu, "Design Mobile Commerce Security Project by Technology of J2ME and XML," *Application Research of Computers,* Vol. 11, 2006, pp. 105-108.

[8]  "The Jakarta Site - The Apache Jakarta Tomcat, the Apahce Jakarta Project," 10 November 2004. http://jakarta. apache. org/

[9]  W. Y. Yan, "Mobile Map Service with Scalable Vector Graphics, Geoscience and Remote Sensing Symposium, 2004," IGARSS'04, *Proceedings 2004 IEEE International,* Anchorage, Alaska, USA, Vol. 5, January 2004, pp. 2967-2970.

[10]  F. L. Zhang, X. G. He, Z. G. Qin and M. T. Zhou, "Location Management in Mobile Environment," *International Conference on Communications, Circuits and Systems,* January 2004, pp. 1491-1496.

[11]  L. L. Guo, Z. G. Qin and M. Yuan, *et al.* "Mechanism of Mobile Data Transfer Based on XML," *Journal of Daqing Petroleum Institute,* Vol. 30, No. 5, 2006, pp. 122-124.

[12]  J. H. Shin, B. J. Yi and J. J. Song, "A Development of the Mobile Computing System for Repair and Patrol of Electric Power Facilities," *Fourth Annual ACIS International Conference on Computer and Information Science,* January 2005, pp. 622-627.

Scientific
Research

# Unified Performance and Availability Model for Call Admission Control in Heterogeneous Wireless Networks

**Ramesh Babu H. Siddamallaiah[1], Gowrishankar Subramanian[2], Piriyapatna S. Satyanarayana[3]**

[1]*Department of Information Science and Engineering, Acharya Institute of Technology, Bangalore, India*
[2]*Department of Computer Science and Engineering, B.M.S. College of Engineering, Bangalore, India*
[3]*Department of Electronics and Communication Engineering, B.M.S. College of Engineering, Bangalore, India*
*Email*: *rameshbabu@acharya.ac.in, gowrishankar.cse@bmsce.ac.in, pssvittala.ece@bmsce.ac.in*

## Abstract

The system capacity of wireless networks varies temporally. This may be due to the dynamic allocation of the channels and also the mobility of the users. The change in capacity will create greater impact on the system performance parameters. This variation of capacity particularly poses a greater challenge to the research community to ensure the Quality of Service (QoS) as it affects the call blocking probability which is one of the important QoS parameters. This paper proposes a performance model for call admission control and the availability model for a heterogeneous wireless network environment. The proposed model is able to handle three types of traffic considered for the study includes conversation traffic, interactive traffic and background traffic. The unified performance-availability model is developed using the Stochastic Area Networks (SAN). The performance of both analytical models and the SAN based performance-capacity models are verified by taking the call blocking probabilities for all the three types of traffics.

**Keywords:** Call Admission Control, Call Blocking Probability, Heterogeneous Wireless Networks, Stochastic Activity Networks, Quality of Service

## 1. Introduction

There is tremendous increase in the growth of wireless communication technologies which is evidently indicated by exponential increase in the wireless communication technologies usage. This popularity obviously demands the next stage beyond third-generation (3G) networks to include multiple wireless access technologies, all of which will coexist in a heterogeneous wireless access network environment [1,2] and use a common IP core to realize user-focused service delivery. The coexistence of Heterogeneous radio access technologies (RATs) is becoming mandatory to fulfill the needs of the growing users' community of wireless technologies. This in turn will noticeably amplify the intensity in development of different high-speed multimedia services, such as video on demand, mobile gaming, Web browsing, video streaming, voice over IP and e-commerce etc. Seamless intersystem roaming across heterogeneous wireless access networks will be a major feature in the architecture of next generation wireless networks [3]. It is very well

evident that no single RAT can provide ubiquitous coverage and continuously high quality service (QoS), the mobile users may have to roam among various radio access technologies to keep the network connectivity active to meet the applications/users requirements. With the increase in offered services and access networks, efficient user roaming and management of available radio resources becomes decisive in providing the network stability and QoS provisioning. The future users of mobile communication look for always best connected (ABC) anywhere and anytime in the Complementary access technologies like Wireless Local Area Networks (WLAN), Worldwide Inter operability for Microwave Access (WiMAX) and Universal Mobile Telecommunication Systems (UMTS) which may coexist with the satellite networks [4-6].

With this prevalent scenario of user mobility, ensuring the required radio resources to the users is a highly challenging task. In spite of failures in hardware, software, or combination of these two may still make the wireless networks to work and support the users. But there will be

a definite effect on the system capacity, *i.e.*, the number of channels available and number of users that can be accommodated by the wireless networks may decrease and the throughput of the system will come down. This is an important aspect concerned to performance modelling. There is a strong need of performance modelling which should be able to take care of not only the pure performance but also the availability and reliability model of the system. But most of the performance models will consider this contention and generally over estimate the situation and will not consider the failure–recovery of the systems resources in turn the traditional availability models are not considering the performance metrics. There is a need for a composite model which will work on failure recovery model.

Researchers have achieved success in developing techniques for modelling the performance, availability and reliability of communication systems in a unified way. There are good number of approaches that have proved the need of unified performance model and availability models of any stochastic system under study. In this paper we have proposed composite Performance and Availability models and are evaluated using the stochastic petrinet.

The further sections of the paper are organized as follows. The Section 2 presents the performance model for call admission control. Section 3 focuses on the Availability model. The Section 4 represents SAN based performance-availability of the CAC system and Section 5 is devoted to discuss the simulation results and Section 6 concludes the paper.

## 2. Performance Model

In this paper we propose a novel analytic performance model admission control mechanism for reducing the call blocking probability there by increasing the resource utilization. This would achieve the objective of guaranteeing the user QoS requirements. The proposed model is able to handle three types of traffic considered for the study includes conversation traffic, interactive traffic and background traffic. All of this traffic represents different QoS service class of traffic with the following QoS parameters.
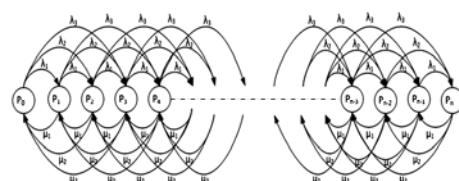
The conversational traffic is sensitive to transfer delay and jitter. It demands guaranteed bit rate and low bit error rate. The examples of the applications belonging to this category are video-conferencing and audio conferencing. The interactive traffic is a QoS class that is not sensitive to transfer delay and Jitter but demands low bit error rate. The applications of this QoS class do not need guaranteed bit rate for example web browsing, interactive chats and interactive games. The background traffic QoS class is not sensitive to transfer delay and jitter but needs low bit error rate from the network and these ap-

plications do not depend on guaranteed bit rate. The examples belonging to this group are e-mail, SMS applications. The assumption made for the design and development of analytical CAC model was type3 traffic would require three channels to be assigned in the system and type2 traffic demands two channels and type1 traffic needs one channel.

The proposed model is developed keeping in mind the WCDMA, WiFi, and WiMax. The CAC mechanism proposed is focused only on the system's ability to accommodate newly arriving users in terms of the total channel capacity which is needed for all terminals after the inclusion of the new user. In the case when the channel load with the admission of a new call was precompiled (or computed online) to be higher than the capacity of the channel the new call is rejected, if not, the new call could be admitted. The decision of admitting or rejecting a new call in the network will be made only based on the capacity needed to accommodate the call.

We consider a heterogeneous network which comprises a set of RATs $R_n$ with co-located cells in which radio resources are jointly managed. Cellular networks such as Wireless LAN and Wi-Max can have the same and fully overlapped coverage, which is technically feasible, and may also save installation cost. $H$ is given as $H$ {RAT 1, RAT 2, RAT $k$} where $k$ is the total number of RATs in the heterogeneous cellular network. The heterogeneous cellular network supports $n$-classes of calls, and each RAT in set $H$ is optimized to support certain classes of calls.

The analytical model for call admission control mechanism in heterogeneous wireless networks is modelled using higher order Markov model and is as shown in **Figure 1**. In the proposed model it is assumed that, whenever a new user enters the network will originate the network request at the rate $\lambda_i$ and is assumed to follow a poisson process. The service time of the different class of traffic and types of calls is $\mu_i$. The mean service time of all types of users were assumed to follow negative exponential distribution with the mean rate $1/\mu$. Since voice traffic is Erlang distributed, the condition that is considered for simulation is negative exponential distribution. The total number of virtual channels in the system are $N$. When the numbers of available channels are below the specified threshold the system will drop the calls. The threshold limit is determined by three positive integers $A_1$, $A_2$ and $A_3$. These are called as utilization rates, where $A$ is represented as



**Figure 1. Analytical model for CAC.**

　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　*IJCNS*

$$A = \lambda / \mu$$

Similarly,

$$A_1 = \frac{\lambda_1}{\mu_1}, \quad A_2 = \frac{\lambda_2}{\mu_2}, \quad A_3 = \frac{\lambda_3}{\mu_3} \tag{1}$$

are the utilisation rate of type1 traffic, type2 traffic and type3 traffic respectively. In general the values of the utilisation rate in a steady state system will be with in 1.

When the available number of channels falls below the threshold $A_3$ the proposed system will accept only the voice calls and web browsing. When the available number of channels falls below the threshold $A_2$ the proposed system will accept only the voice calls. If the available number of channels falls below the threshold $A_1$ the proposed system will not accept any calls as it reaches the stage where there will be no channels available to allocate to the incoming calls and leads to system blocking. The $P(0)$ is the probability that there are no allocated channels in the designated system. The parameters of analytical performance model are also called as Performance model parameters. The parameters are number of virtual channels ($N$), user arrival rate ($\lambda$), arrival rate of type1 call ($\lambda_1$), arrival rate of type2 call ($\lambda_2$) arrival rate of type3 call ($\lambda_3$) and service time of the calls is taken as $\mu_1$, $\mu_2$ and $\mu_3$.

Assuming that the arrival time of all the types of traffic are equal *i.e.*, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$ and the service time for the types of traffic are equal *i.e.*, $\mu_1 = \mu_2 = \mu_3 = \mu$, the call blocking probability for type1 traffic could be expressed as

$$P_n = \frac{a}{3}(P_{n-1} + P_{n-2} + P_{n-3}) \tag{2}$$

where $a = \lambda / \mu$ which should be generally less than one for the system stability. Similarly, the call blocking probability for type2 traffic $P_{n-1}$ is

$$P_{n-1} = \frac{a}{3}(P_{n-2} + P_{n-3} + P_{n-4}) \tag{3}$$

And the call blocking probability for type3 traffic $P_{n-2}$ is represented as

$$P_{n-2} = \frac{a}{3}(P_{n-3} + P_{n-4} + P_{n-5}) \tag{4}$$

The call blocking probability for the overall system traffic $P_{nb}$ can be expressed as

$$P_{nb} = \frac{a}{3}(P_n + P_{n-1} + P_{n-2}) \tag{5}$$

## 3. System Availability Model

The system availability model indicates the availability

of the channels in the proposed system. When all $N$ channels are failed then the system is unavailable for data transmission. Here the system is modelled as independent failure - repair model [7] which is also iterated in our previous work [8]. Each virtual channel is subjected to time and frequency selective fading and multipath fading then the virtual channel will be unavailable for data transmission. The individual channel is available for use by changing the mobile terminal position or by channel equalization technique [7]. The channel recovery model is developed as independent repair facility [9]. The failure rate of the channel is Poisson distribution with the rate $\gamma$. The channel recovery is exponential repair time distribution of parameter distribution with the parameter $\tau$. The $N$ independent channel failure and recovery can be represented as a single dimension Markov Chain. The system availability can be modelled as Markov chain. The steady state probability $P(i)$, where $i$ is the number of available channel in the system and given by

$$p(i) = \frac{\left(\dfrac{\tau}{v}\right)^i i!}{\displaystyle\prod_{j=0}^{i-1}(N-j)} p(0) \tag{6}$$
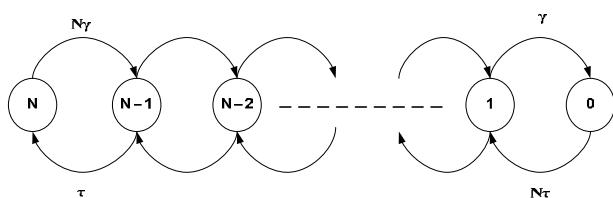
The steady state unavailability $P(0)$ of the system is given by

$$P(0) = \left[ \sum_{i=0}^{N} \frac{\left(\dfrac{\tau}{v}\right)^i i!}{\displaystyle\prod_{j=o}^{i-1}(N-j)} \right]^{-1} \tag{7}$$

The important observations from Performance-availability model is, increasing the number of channels will decrease the call blocking probability $P_{nb}$. The system availability model is as shown in **Figure 2**.

## 4. Composite Performance and Availability Model

The performance-availability model is based on Stochastic Activity Networks (SAN). The SAN is a stochastic extension of Stochastic Petri Networks (SPN) in which the capacity to define temporary characteristics with statistical parameters has been added. The SAN exhibits the
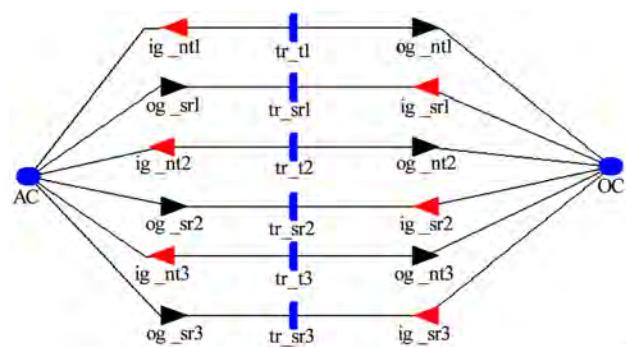
**Figure 2. System availability model.**
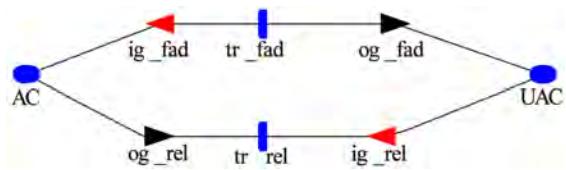
innovative graphics which allows the researchers to represent a model with a high level of formal specification, expression of behaviour and dependency of the system in an uncomplicated and straightforward way. SPNs can be in general considered to constitute a method to model distributed, asynchronous concurrent systems, which have parallel characteristics. It is possible to study the performance and evolution of the system easily using Petri nets as Petri Nets combine graphic design and extensive mathematical theory to represent a system model.

A SAN performance model for CAC is represented in **Figure 3** as channel usage model and the channel availability model which is also called as system capacity model is represented in **Figure 4** for the aggregate traffic. The performance model of the proposed system is shown in **Figure 5** and primitive components used in the proposed model are shown in **Table 1**. The activities tr_$t_1$, tr_$t_2$ and tr_$t_3$ represents the new user arrival/call arrival of traffic type1, traffic type2 and traffic type3 respectively which are timed activities and the firing distribution is a Poisson distribution. The new traffic arrivals have an inhibitory input from the input gate $ig\_nt_1$, $ig\_nt_2$ and $ig\_nt_3$ when the number of virtual channels is less than $A_1$ channels, $A_2$ and $A_3$ respectively. The transition tr_$sr_1$, tr_$sr_2$ and tr_$sr_3$ represent user Service requests from traffic type1, type2 and type3 to system. Service requests are hyper-exponential distribution. The places *AC* and *OC* in the channel usage model indicate *available channels* and *occupied channels /used channels.*
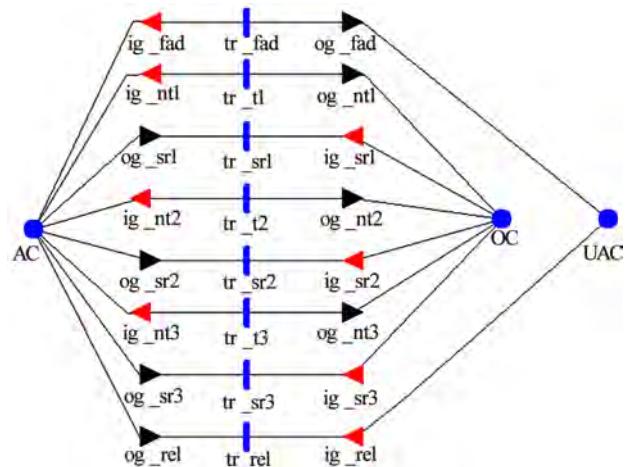
The activity tr_$t_1$ represents the call arrival of traffic type1 on firing of transition tr_$t_1$, the output gate $og\_nt_1$ shown in the Traffic type1 SAN performance model will function and removes single token from place AC and deposit a single token in place *OC* as shown in **Figure 3**. The transition tr_$sr_1$ represents user Service requests from traffic type1 to system. Service requests are hyper-exponential distribution. After the call is serviced the channel is released to the timed activity tr_$sr_1$ through input gate $ig\_sr_1$. On firing tr_$sr_1$ the output gate $og\_sr_1$ will draw a token from OC and deposit the token in place *AC.*



Figure 4. Channel availability/ fading model.



Figure 5. Performance-availability model.

**Table 1. Structural components of traffic models.**

| Symbol | SAN object | Description |
|---|---|---|
| AC | Place | Available Virtual Channels |
| OC | Place | Used / Consumed Virtual Channels |
| tr_$t_1$ tr_$t_2$ tr_$t_3$ | Transitions | Type1 call arrival, type2 call arrival, type3 call arrival respectively |
| tr_$sr_1$ tr_$sr_2$ tr_$sr_3$ | Transitions | Service completion of type1 call/traffic, Service completion of type2 call/traffic, Service completion of type3 call/traffic |
| $ig\_nt_1$ $ig\_nt_2$ $ig\-nt_3$ | Input gate | Input predicate for type1 Traffic arrival, Input predicate for type2 Traffic arrival, Input predicate for type3 Traffic arrival, |
| $og\_nt_1$ $og\_nt_1$ $og\_nt_1$ | Output gate | Output function for type1 traffic arrival, Output function for type2 traffic arrival, Output function for type3 traffic arrival |
| $ig\_sr_1$ $ig\_sr_2$ $ig\-sr_3$ | Input gate | Input predicate for type1 Traffic service, Input predicate for type2 Traffic service, Input predicate for type3 Traffic service, |
| $og\_sr_1$ $og\_sr_1$ $og\_sr_3$ | Output gate | Output function for type1 traffic service, Output function for type2 traffic service, Output function for type3 traffic service |



Figure 3. Channel usage model.

The performance model for type2 call is represented in **Figure 2**. On firing of transition $tr\_t_2$ the output gate $og\_nt_2$ of the traffic type2 SAN performance model will function and removes single token from place *AC* and deposit a single token in place *OC*. The transition $tr\_sr_2$ represents user Service requests from traffic type2 to system. Service requests are hyper-exponential distribution. After the call is serviced the channel is released to the timed activity $tr\_sr_2$ through input gate $ig\_sr_2$. On firing $tr\_sr_2$ the output gate $og\_sr_2$ will draw a token from *OC* and deposit the token in place *AC*.

The activity $tr\_nt_3$ in **Figure 3** represents the call arrival of traffic type3 and on firing of transition $tr\_t_3$ the output gates $og\_nt_3$ in of the traffic type3 SAN performance will function and removes single token from place *AC* and deposit a single token in place *OC*. The transition $tr\_sr_3$ represents user Service requests from traffic type3 to system. Service requests are hyper-exponential distribution. After the call is serviced the channel is released to the timed activity $tr\_sr_3$ through input gate $ig\_sr_3$. On firing $tr\_sr_3$ the output gate $og\_sr_3$ will draw a token from OC and deposit the token in place AC.

The transition $tr\_t_1$ represents an event of call arrival of type1 traffic to the system. The transition new call/ user arrival of traffic type1 has an inhibitory input from the input gate $ig\_nt_1$, when the total numbers of available channels are less than A1 the transition $tr\_t_1$ is disabled. The transition $tr\_t_2$ represents an event of call arrival of type2 traffic to the system. The transition new user arrival of traffic type2 has an inhibitory input from the input gate $ig\_nt_2$, when the total numbers of available channels are less than *A2* the transition is $tr\_t_2$ is disabled. The $tr\_t_3$ is timed transition that represents the event of arrival of type3 traffic to the system. The transition $tr\_nt_3$ is disabled when the available channel falls below *A3*.

The **Figure 4** represents the fading model of the channel which gives the availability of the channel which represents the system capacity. The *AC* and *UAC* are the places in fading model and will represent channel availability and channel non-availability respectively in the proposed system. The transition $tr\_fad$ represent the fading rate in wireless network and fading rate generally follows Weibull distribution. The transition $tr\_fad$ is fired if and only if the tokens are available in the place *AC* and this condition is implemented through input gate $ig\_fa$d. Transition $tr\_rel$ is the channel recovery process and is assumed to be exponential distribution. When the Timed activity $tr\_rel$ is fired the output gate $og\-rel$ will draw a token from the place UAC and send it to AC. This is nothing but when a channel fades then the channel will be in UAC state and when channel comes out of fading state it will trigger the transition $tr\_rel$ and place the token in AC. In other words the channel after coming out of fading state UAC will enter the available channel state AC.

The SAN based performance-capacity model is represented in **Figure 5**. which is a composite architecture of the composite performance-availability model developed by combining the channel usage model and channel availability model/channel fading model.

## 5. Simulation Results and Discussion

In this section, we present the numerical results and compare the call blocking probabilities of the different types of traffic. A set of experiments were conducted varying the number of channels and the call blocking probability was compared for SAN performance-Availability model and the analytical models for all three types of traffic.

The parameters of analytic performance-capacity model are divided into performance model parameters and availability model parameters. Performance model parameters are number of channels ($N$), total user arrival rate of type1 traffic ($\lambda1$), total user arrival rate of type2 traffic ($\lambda2$), total user arrival rate of type3 traffic ($\lambda3$) and service time of the users ($\mu$).

The system capacity model parameters are channel failure rate ($\gamma$) and channel recovery rate ($\tau$). Parameters Values are 50($N$), 7($\lambda1$), 3($\lambda2$), 1($\lambda3$), 2($\mu$), 3($\gamma$) and 2($\tau$).

The Parameters of SAN Performance-capacity model is divided into channel usage model parameters and channel fading model parameters. The channel usage parameters are number of virtual channels represented by number of tokens in place AC, firing rate of type1 traffic arrival transition ($tr\_t1$), firing rate of type2 traffic arrival transition ($tr\_t2$), firing rate of type3 traffic arrival transition ($tr\_t3$) and channel usage service rates for type1, type2, type3 traffic are $tr\_sr 1$, $tr\_sr 2$, $tr\_sr 3$ respectively. The channel usage rate is hyper-exponential distribution and classified into three types of usage such as low, medium and high with the service rate μL, μM and
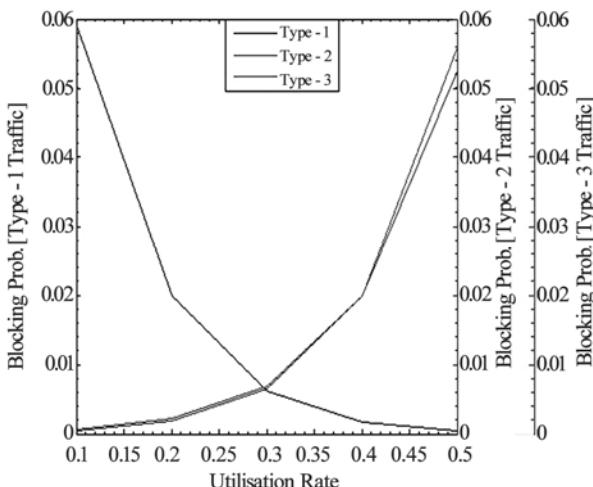
**Table 2. Fading model structural components.**

| Symbol | SAN object | Description |
|--------|-----------|-------------|
| *AC* | Place | Channel Availability |
| *UAC* | Place | Unavailability of Channel |
| $tr\_fad$ | Transition | Channel fading Rate |
| $tr\_rel$ | Transition | *Channel recover y/release rate* |
| $ig\_fad$ | Input gate | Input predicate for channel fading |
| $ig\_rel$ | Input Gate | Service completion |
| $og\_fad$ | Output gate | Output function for fading transition |
| $og\_rel$ | Output gate | Output function for recovery transition |
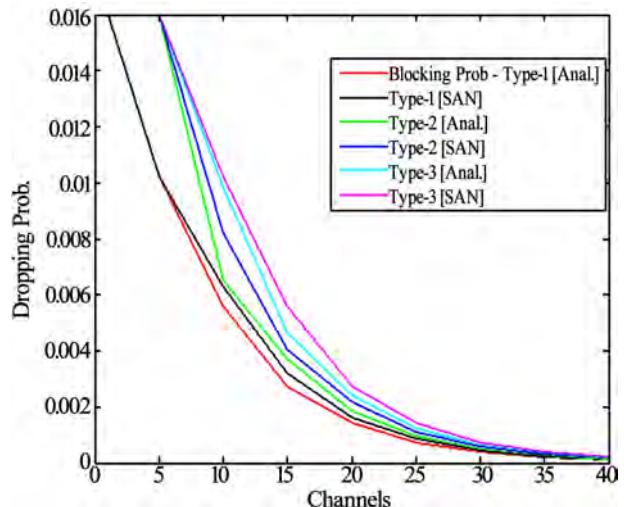
μH respectively. The probability of low, medium and high usage are PL, PM and PH respectively. The channel fading model parameters are number of channels represented by number of tokens in place AC. The transition (tr_fad) represents the fading rate of the channel and recovered channel rate (tr_rel). Parameters values are 50(AC), 0.25(tr_t1), 0.5(tr_t2), 0.7(tr_t3), 1(μL), 2(μM), 3(μH), 0.5(PL), 0.3(PM), 0.2(PH), 0.34(tr_fad) and 0.5(tr_rel).

The first set of experiments is indicated by the simulation result shown in **Figure 6**. The call blocking probability for a system with N channels which supports three types of traffic is conducted. The experiment considers that, whenever a new user enters the network will originate the network request at the rate $\lambda_1$ for type1 traffic, $\lambda_2$ for type2 traffic and $\lambda_3$ for type3 traffic and is assumed to follow a Poisson process. The service time of the different types of traffic based calls is considered as $\mu_1$ for type1 traffic, $\mu_2$ for type2 traffic and $\mu_3$ for type3 traffic and is assumed to follow a Lognormal random Process. For the first set of experiments we have considered the arrival rate of all the three types of traffic as $\lambda$ and service rate of all the three type of calls is same and is equal to $\mu$.

The second set of experiments conducted will present the numerical results and compare the call blocking probabilities of the different types of traffic obtained for performance-capacity model and the analytical model. The proposed a performance-availability model for call admission control mechanism in the heterogeneous RATs environment is analysed for the call blocking probability, by having variation in the number of channels. The graph obtained for the experiment setup conducted considering both the analytical model and SAN performance–availability model for the blocking probability of type1, type2, and type3 traffic is plotted. The horizontal axis shows the number of channels while the vertical axis shows the call blocking probability of all types of traffic.



Figure 7. Comparison of call blocking probability of SAN model v/s analytical model.

The simulation results show that the call blocking probability of the different types of traffic will decrease with the increase in the number of channels in the system. The simulation results are shown in **Figure 7**. The simulation results show that the call blocking probability of the different types of traffic will decrease with the increase in the number of channels in the system.

The behaviour of analytical model and the performance model developed using stochastic activity network and simulated using the Mobius simulator behaves identically.

## 6. Conclusions

In this paper, the performance of analytical model for CAC system for next generation networks is compared and validated with the system performance-capacity model developed using SAN. The Performance of both call admission control models in the heterogeneous RATs are studied pitching upon the call blocking probability by varying the number of channels. The increase in number of channels in the system decreases the call blocking probability of all traffic types. The results obtained for analytical model is in line with the performance model results where both the models behave in the similar fashion in the experiments conducted.

## 7. References

[1]  Sanmateu, *et al.* "Seamless Mobility across IP Networks Using Mobile IP," *International Journal of Computers and Telecommunications Networking*, Vol. 40, September

**Figure 6. Call blocking probablity of varying traffic.**

2002, pp. 181-190.

[2] G. Wu, *et al.* "MIRAI Architecture for Heterogeneous Networks," *IEEE Communications,* Vol. 40, No. 2, 2002, pp. 126-134.

[3] S. Y. Hui and K. H. Yeung, "Challenges in the Migration to 4G Mobile Systems," *IEEE Communications*, Vol. 41, No. 12, 2003, pp. 54-59.

[4] X. G. Wang, J. E. Mellor and G. Min, "A QoS-Based Bandwidth Management Scheme in Heterogeneous Wireless Networks," *International Journal of Simulation Systems, Science and Technology,* Vol. 5, No. 1-2, 2004, pp. 9-17.

[5] E. Vanem, S. Svaet and F. Paint, "Effects of Multiple Access Alternatives in Heterogeneous Wireless Networks," *IEEE Wireless and Networking*, Vol. 3, 2003, pp. 1696-1700.

[6] K. Murray, R. Mathur and D. Pesch, "Network Access and Handover Control in Heterogeneous Wireless Networks for Smart Space Environments," *First International Workshop on Management of Ubiquitous Communications and Services*, MUCS ,Waterford, Ireland, December 11, 2003.

[7] Gowrishankar, H. S. R. Babu, G. T. Raju and P. S. Satyanarayana, "Performability Model of Vertical Handoff in Wireless Data Networks," *4th IEEE International Conference on Wireless and Mobile Communication (ICWMC2008)*, Athens, Greece, 27 July-1 August 2008, pp. 309-314.

[8] K. S. Trevedi, "Probability and Statistics with Reliability, Queuing and Computer Science Application," 2nd Edition, John Wiley and Sons, New York, 2001.

[9] Mulgrw, "Applying Radial Basis Functions," *IEEE Signal Processing Magazine*, Vol. 13, No. 2, 1996, pp. 50-65.

Scientific
Research

# MIMO Multi Carrier Interleave Division Multiple Access System with Multiuser Detection: Performance Results

**Prabagarane Nagaradjane[1], Sai N. Chandrasekaran[2], Kuttathatti Srinivasan Vishvaksenan[3]**

[1]*Department of Electronics and Communication Engineering, SSN College of Engineering,
SSN Institutions, Chennai, India*
[2]*The University of Kansas, Kansas, USA*
[3]*Department of ECE, SSN College of Engineering, SSN Institutions, Chennai, India*
Email: *prabagaranen@ssn.edu.in, csnivi@gmail.com, vishvaksenanks@ssn.edu.in*

## Abstract

This paper provides the performance analysis of multi input multi output (MIMO) Multi carrier (MC) interleaver based multiple-access system with multiuser detection. Interleave-division multiple access (IDMA) is a scheme in which users are separated by employing different interleavers instead of different signatures as in a conventional code-division multiple-access (CDMA) scheme. The basic principle of IDMA is that the interleaver is different for different users. Interleavers are generated independently and randomly. At the receiver, ZF, LLSE, VBLAST/ZF/MAP and VBLAST/LLSE/MAP detectors are employed. The performance of the system is analyzed for different channel conditions using extensive simulation runs based on Monte Carlo simulation trials. Simulation results divulge that VBLAST/LLSE/MAP detector results in significant performance amelioration compared to other detectors as applied to MC/IDMA system.

## 1. Introduction

This paper presents an approach to asynchronous multiple access scheme called interleave division multiple-access (IDMA) [1] system with MIMO support. In an IDMA scheme, different interleavers are used to distinguish users as against different codes in a conventional code division multiple access (CDMA) system. In a conventional CDMA scheme, interleavers are placed before the spreaders and they are effective only when used in conjunction with channel coding [2,3]. Recently, a very interesting technique using chip-level interleavers was addressed in [1], which aims at mitigating inter-symbol interference (ISI) in multipath fading environments. Many papers have discussed the role of interleavers in a multiple access systems [1,3]. IDMA inherits many benefits of CDMA; in particular, path diversity and mitigation of intra cell interference. Also all the users employ a common spreading sequence. Recent research on wireless communication systems has shown that using multiple antennas at both transmitter and receiver offers the possibility of communications at higher data rates compared to single antenna systems [4-6]. Multi input

and multi output (MIMO) system has proved in the recent past to provide very high capacity without any increase in the transmission bandwidth and power. The information-theoretic capacity of these multiple-input multiple-output channels was shown to grow linearly with smaller numbers of transmit and receive antennas in rich scattering environments, and at sufficiently high signal-to-noise (SNR) ratios [7-9]. MIMO profile in any wireless communication system can be realized by two schemes namely, 1) using the classical BELL labs architecture (VBLAST–Vertical Bell Laboratories Space Time architecture) and 2) Space Time Block Codes. And OFDM techniques can compensate for multi-path fading effects, which cause time dispersion, and out of band emission of the received signal. So combining these techniques with the IDMA system can result in MIMO MC/IDMA that can offer bandwidth efficiency, space diversity and lower speed parallel type of signal processing and interference rejection capability (ISI reduction) in high data-rate transmission. Of late, significant progress has been made in multi-user detection for code-division multiple-access (CDMA) systems [6]. In this paper, we investigate the performance of the MIMO as-

sisted Multi Carrier interleave division multiple access (MIMO MC/IDMA) scheme with multiuser detection. We will show that the MIMO MC/IDMA scheme can provide better performance with the aid of VBLAST/ LLSE/MAP [5,6] detection technique. Also this approach is independent of the number of users and gives better performance in multipath environment. The rest of the paper is organized as follows - the system model is presented in Section 2. Section 3 gives the gist of signal detection techniques. Performance results will be presented in Section 4. Eventually conclusions are drawn in Section 5.

## 2. System Model

In this section, we introduce the IDMA scheme with Multi carrier modulation. The block diagram of MIMO MC/IDMA scheme is elucidated in **Figure 1**. In IDMA the interleavers are different for different users. Interleavers are generated independently and randomly. Different interleavers separate the users and all the users employ a common spreading sequence. Interleavers, which are usually placed between forward error correction (FEC) coding and spreading, is used to combat fading effect in CDMA, whereas the arrangement of interleaving and spreading is reversed in IDMA. Now, different interleavers distinguish distinct data streams. In IDMA, encoder block may do FEC encoding and spreading jointly. The spreader has no special function. In this paper we consider a single user MIMO MC IDMA system. Each of the synchronous user is equipped with $M$ transmit antennas and they communicate to a single base station, equipped with $N$ receive antennas. MIMO profile is realized by using VBLAST (Vertical Bell Laboratories

Space Time architecture).

## 3. Symbol Detection Techniques
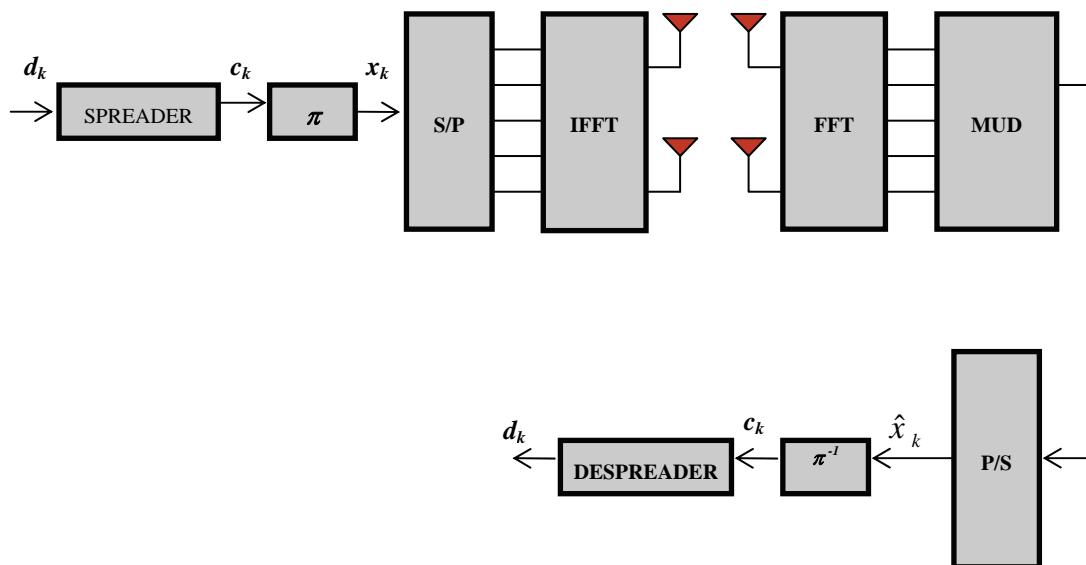
Considering a synchronous signal over a time variant channel, the received signal at the time instant $j$ is given as

$$r(j) = \sum_{k=1}^{\min(m,n)} \sum_{k=1}^{k} H_k x_k(j) + n(j), j = 1, 2, ..., J \quad (1)$$

where $H_k$ is the channel co-efficient for user $k$ and $\{n(j)\}$ are samples of an AWGN process with variance $\sigma^2 = No/2$. Here perfect knowledge of the channel state information is assumed to be available at the receiver. In each use of the MIMO channel, a vector $x = (x_1, x_2...x_M)^T$ of complex numbers is sent and a vector $r = (r_1, r_2,..., r_N)^T$ of complex numbers is received. We assume an input-output relationship of the form

$$r = Hx + n \quad (2)$$

where $H$ is a $N \times M$ matrix ($N$ = number of transmit antennas and $M$ = number of receive antennas) representing the scattering effects of the channel and $n = (n_1, n_2,..., n_N)^T$ is the noise vector. Throughout, we assume that $H$ is a random matrix with independent complex Gaussian elements $\{h_{ij}\}$ with mean 0 and unit variance, denoted by $h_{ij} \sim N(0, 1)$. We also assume throughout that $n$ is i.i.d complex Gaussian random vector. The symbol detection problem considered in this paper is the problem of estimating the MIMO channel input vector $x$ given the received vector $r$ assuming that the receiver has perfect knowledge of $H$. The decision is made on a symbol-by-symbol basis. The next section discusses the various VBLAST detectors employed in our work.



**Figure 1. MIMO MC/IDMA system.**

## 3.1. Zero-Forcing (ZF) Receiver

Zero-forcing (ZF) receiver is a low-complexity linear detection algorithm that outputs,

$$\hat{x} = Q(\hat{x}_{ZF})  \qquad (3)$$

where

$$\hat{x}_{ZF} = H^+ \; r  \qquad (4)$$

and $H^+$ denotes the Moore-Penrose pseudo inverse of $H$, which is a generalized inverse that exists even when $H$ is rank-deficient.

## 3.2. Linear Least Square Error Receiver (LLSE)

LLSE receiver is also a linear detection algorithm that outputs

$$\hat{a} = Q(\hat{a}_{LLSE})  \qquad (5)$$

where

$$\hat{a}_{LLSE} = Wr  \qquad (6)$$

and when $H$ and $n$ are gaussian, $W$ is the weight vector given by

$$W = \frac{\rho}{M} H^+ (\frac{\rho}{M} HH^+ + N_o I_N)^{-1}  \qquad (7)$$

## 3.3. V-BLAST Detection

The V-BLAST detection algorithm [6] is a recursive procedure that extracts the components of the transmitted vector $x$ according to a certain ordering $(k_1, k_2,...,k_M)$ of the elements of $x$, where, $(k_1, k_2,...,k_M)$ is a permutation of $(1,..., M)$. In VBLAST, this permutation depends on $H$ (which is known at the receiver by assumption) but not on the received vector $r$. The VBLAST/ZF algorithm is a variant of VBLAST derived from ZF rule. The algorithm determines the order of channels to be detected, performs *nulling* and computes the decision statistics. It then slices the computed decision statistics and yields the decision by performing *cancellation* with the help of decision feedback, and finally computes the new pseudo-inverse for the next iteration. V-BLAST/ZF may be seen as a successive-cancellation scheme derived from the ZF scheme discussed in the previous section. The ZF rule creates a set of subchannels by forming

$$\hat{x} = (H^+ H)x + H^+ n  \qquad (8)$$

The $j$'th such sub-channel has noise variance $\|(\mathrm{H}^+)_j\|^2 N_0$ and the order selection rule prioritizes the sub-channel with the smallest noise variance.

## 3.4. VBLAST/MAP Detection Algorithm

In this section, a detection algorithm for MIMO channels, called VBLAST/MAP that combines the features of VBLAST and MAP rules is used for symbol detection in a MC/IDMA system. This algorithm does not use the layered structure of VBLAST, but uses a different strategy for channel processing order, inspired by the MAP rule. This algorithm is an extension of the well-known VBLAST algorithm. This VBLAST/MAP combines the elements of VBLAST algorithm and the maximum a-posteriori (MAP) rule. The original VBLAST algorithm is a multi-layer symbol detection scheme which detects symbols transmitted at different transmit antennas successively in a certain data-independent order. The VBLAST/ MAP algorithm differs from VBLAST only in the ordering strategy of the symbols detected. The complexity of the VBLAST/ MAP is higher than that of VBLAST; however, the performance improvement is significant.

## 3.5. V-BLAST/ZF/MAP Detection Algorithm

Using the same notation of VBLAST algorithm, VBLAST/ZF/MAP algorithm may be described, as follows-VBLAST/ZF/MAP algorithm is identical to VBLAST/ ZF except for the ordering in which symbols are detected. Instead of selecting the next symbol to be detected according to the rule, here the set of all probable symbol decisions are ranked with respect to their a-posteriori probabilities of being correct, as estimated by $p_{ij}$. Thus, it is important to emphasize that $p_{ij}$'s are not true MAP probabilities but approximations to how probable it is that $s_{ij} = x_j$. Note that the index permutation $(k_1, k_2,... ,k_M)$ produced by VBLAST/ZF/MAP depends on both $H$ and $r$, unlike VBLAST/ZF where the permutation depends only on $H$. Complexity of VBLAST/ZF/MAP is increased with respect to that of VBLAST/ZF by the computation steps involved. One major point about complexities of VBLAST/ZF and VBLAST/ZF/MAP is that, the former allows pre-computation of all weighting vectors whereas in the latter the weighting vector must be computed in real-time since it also depends on $r$. This increased complexity of VBLAST/ZF/MAP is justified by performance improvements. The algorithm is summarized as follows:

*Initialization*

$$W_1 = H^+$$
$$i = 1  \qquad (9)$$

*Recursion*

$$y_i = W_i r_i$$
$$s_i = Q(y_i)$$
$$p_{ij} = \frac{f_{ij}(y_{ij}|s_{ij})}{\sum_{s' \in A} f_{ij}(y_{ij}|s')}, j \notin \{k_1...k_{i-1}\}$$
$$k_i = \begin{array}{c} \arg\max \\ j \notin \{k_1...k_{i-1}\} \end{array} \{p_{ij}\}$$
$$\hat{x}_{k_i} = S_i k_i$$
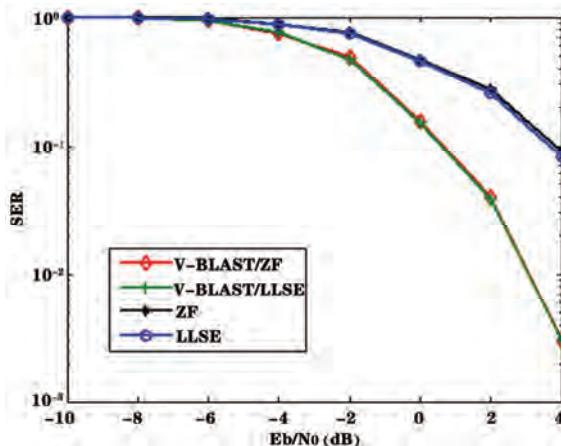
$$r_{i+1} = r_i - \hat{x}_{k_i}(H)_{k_i}$$

$$W_{i+1} = H_{\overline{k_i}}^{+}$$

$$i = i + 1 \qquad\qquad (10)$$

### 3.6. V-BLAST/LLSE/MAP Detection Algorithm
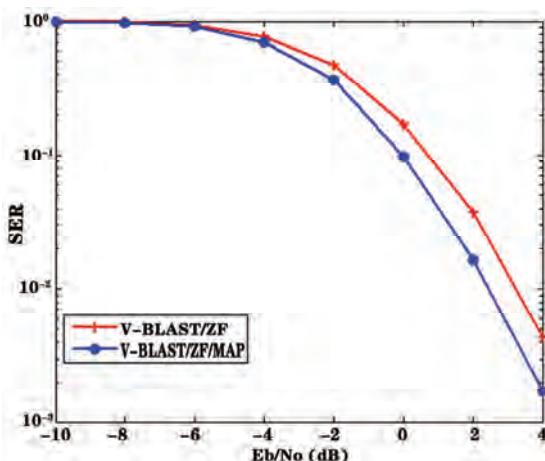
This detection algorithm is akin to V-BLAST/ZF/MAP detection except the weight vector defined by the Equation (7).

## 4. Performance Results
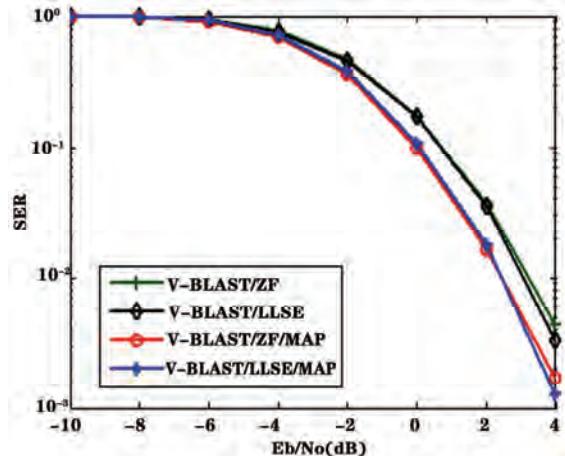
**Figure 2** expounds the simulation results for an $8 \times 12$ MIMO MC/IDMA system employing ZF, LLSE, VBLAST/ZF and VBLAST/LLSE detectors. The modulation technique considered is 16-QAM and the $E_b/N_0$ ranges between -10 dB and 4 dB. The symbol error rate SER is

**Figure 2. Symbol error rates (SER) of VBLAST/ZF, VBLAST/LLSE, ZF and LLSE receiver.**

**Figure 3. Symbol error rates (SER) of VBLAST/ZF/MAP and VBLAST/ZF receiver.**

**Figure 4. Symbol error rates (SER) of VBLAST/ZF, VBLAST/LLSE, VBLAST/ZF/MAP and VBLAST/LLSE/MAP receiver.**

calculated by performing 10,000 trials at each $E_b/N_0$ point. A new realization of **H** was chosen in each trial and for each $E_b/N_0$ value. The performance curves in **Figure 2** show that VBLAST/LLSE provides significant improvement in SER compared to ordinary VBLAST/ZF, ZF and LLSE detector. **Figure 3** elucidates the simulation results for an $8 \times 12$ MIMO MC/IDMA system employing VBLAST/ZF and VBLAST/MAP detectors. It is discerned that VBLAST/MAP outperforms VBLAST/ZF detector. **Figure 4** evinces the simulation results for an $8\times 12$ MIMO MC/IDMA system employing VBLAST/ZF, VBLAST/LLSE, VBLAST/ZF/MAP and VBLAST/LLSE/MAP receiver with 16QAM modulation. It is discerned that VBLAST/LLSE/MAP outperforms VBLAST/ZF/MAP detector in terms of SER performance.

## 5. Conclusions

In this paper a multiple access scheme for 4G wireless communications has been considered based on interleavers. Problems in two major areas have been identified in this paper: efficient transmission technique and symbol detection in the presence of multi stream interference. A solution has been conceived to achieve the above. Simulation results have proved that the MIMO MC/IDMA in combination with VBLAST/LLSE/MAP multiuser detector can achieve significant performance improvement compared to ZF, LLSE, and VBLAST/ZF/MAP.

## 6. References

[1]  H. Schoeneich and P. A. Hoeher, "Adaptive Interleave-Division Multiple Access - A Potential Air Interface for 4G Bearer Services and Wireless LANs," *International Conference on Communications*, 2004.

[2]  C. S. Nivedita, S. Padmini and C. Vijay, "Interleaver

Based Multiple Access System: A Proposal for 4G Mobile Communication with Iterative Chip-by-Chip Multi-User Detection," *Project Report*, *SSN Institutions*, 2006.

[3]  N. Prabagarane, C. S. Nivedita, S. Padmini and C. Vijay, "Interleave Division Multiple Access System with Iterative Multiuser Detection," *IEEE International Conference and Exposition on Advanced Ceramics and Composites*, February 2007.

[4]  N. Prabagarane, M. Ramakrishnan, T. Divya, M. Thomas and B. Jalaja, "Performance Evaluation of Multi Stage Receivers for Coded Signals in MIMO Channels," *4th IEEE WiCOM*, October 2008.

[5]  T. Divya, M. Thomas and B. Jalaja, "Performance Evaluation of Multi Stage Receivers for Coded Signals in MIMO Channels," *Project Report, SSN Institutions*, 2008.

[6]  V. Arun, A. S. S. Vasan and L. Krishnan, "Investigations on the Performance of MIMO Assisted Multi Carrier DS/CDMA System with Multiuser Detection for 4G Mobile Communications," *Project Report*, *SSN Institutions*, 2009.

[7]  I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *Europe Transportation and Telecommunication*, Vol. 10, No. 6, November-December 1999, pp. 585-595.

[8]  P. Nagaradjane, A. S. S. Vasan and L. Krishnan, "A Robust Space Time Co-Channel Interference Mitigation and Detection Technique for Multiuser MIMO Multicarrier DS/CDMA Systems," *Proceedings of IEEE International Conference on Wireless Vitae*, Montreal, 2009.

[9]  P. Nagaradjane, A. S. S. Vasan, L. Krishnan and A. Venkataswamy "Joint VBL-AST/STBC Assisted MC DS/CDMA System with Multiuser Detection," *Proceedings of IEEE International Conference*, *WiCOM*, October 2009.

Scientific
Research

# Comparative Study of Different Space-Time Coding Schemes for MC-CDMA Systems

**Lokesh Kumar Bansal[1], Aditya Trivedi[2]**
[1]*Department of ECE, Nikhil Institute of Engineering and Management, Mathura, India*
[2]*Department of Information and Communication Technology, ABV-Indian Institute of Information Technology and Management, Gwalior, India*
*Email*: *lokesh_bansal@rediffmail.com, atrivedi@iiitm.ac.in*
*Received January* 5, 2010; *revised February* 9, 2010; *accepted March* 10, 2010

## Abstract

In this paper, performance of space-time trellis-code (STTC), space-time block code (STBC), and space-time trellis-code concatenated with space-time block code (STTC-STBC) for multi-carrier code-division multiple-access (MC-CDMA) system are studied. These schemes are considered by employing different detection techniques with various multi input multi output (MIMO) antenna diversity for different number of states in multi-path fading channel. The corresponding bit error rate (BER) is obtained using simulation for minimum mean-square error (MMSE), maximum-ratio combining (MRC), and equal-gain combining (EGC) receivers employing Viterbi decoder. The simulation results show that the STTC-STBC MC-CDMA system perform better compared to other schemes considered in this paper using MMSE detection and it is also observed that the performance can also be enhanced by increasing diversity using more transmitter and receiver antennas. However, this improvement in performance comes at the cost of increased computational complexity, which is calculated for different transmitting and receiving antennas.

## 1. Introduction

Cellular services are now being used every day by millions of people worldwide and the demand is increasing exponentially. Also, there is a demand for integration of a variety of multimedia services namely short messaging, voice, data, and video. Consequently, the bit rate is required to vary from 1.2 kbps for paging to several Mbps for video transmission. The data rate of the second generation is limited up to 9.6 kbps [1].

The third generation cellular systems are being designed to support wideband services like high-speed internet access, video and high quality image transmission with the same quality as the fixed networks. The third generation can provide the maximum data rate ranging from 64 kbps for mobile users to 2 Mbps for stationary users. The drawbacks arising in the third generation system are the limitations of data rate, capacity, inter symbol interference (ISI), and inter chip interference (ICI).

For solving the problems of third generation wireless communication systems, and fulfilling the demand of high performance and broadband internet access with high

spectrum efficiency, the fourth generation systems which may employ multi-carrier code-division multiple-access (MC-CDMA) technology in combination with multi-input multi-output (MIMO) antennas are proposed [2]. In the fourth generation wireless communication systems the data rate may be as high as 1Gbps. Space-time coding techniques may be employed in conjunction with the MC-CDMA system to achieve very high data rate [3,4].

In this paper, the space-time trellis-code (STTC), space-time block code (STBC), and space-time trellis-code concatenated with space-time block code (STTC-STBC) techniques are studied and applied in multi-carrier code-division multiple-access (MC-CDMA) systems. These techniques are considered with multi input multi output (MIMO) antenna diversity in multi-path fading channel. The bit error rate (BER) is obtained for 4, 8, and 16 states using simulation. At the receiver side minimum mean-square error (MMSE), maximum-ratio combining (MRC), and equal-gain combining (EGC) techniques are used by employing Viterbi algorithm.

The organization of the rest of this paper is as follows: in Section 2, MC-CDMA systems and the space-time

code techniques are presented. In Section 3, the mathematical representation for the space-time trellis code (STTC) MC-CDMA system is described. In Section 4, the STTC concatenated with STBC (STTC-STBC) MC-CDMA system is discussed. The MMSE detector is discussed briefly in Section 5. Simulation of the error rate performance of STTC-MC-CDMA systems, STBC- MC-CDMA systems, and STTC-STBC MC-CDMA systems for different states with computational complexity is carried out in Section 6. The conclusions are presented in Section 7.

## 2. MC-CDMA and Space-Time Code Systems

Multi-carrier modulation is being proposed for fourth generation wireless communication systems for high data rate application to reduce the effect of ISI. These systems solve the ISI problem by transmitting the same data symbol over a large number of narrow band orthogonal carriers [5].

An MC-CDMA signal is composed of $N$ narrowband sub-carrier signals each with symbol duration, $T_b$, much larger than the delay spread of the channel, $T_d$, hence MC-CDMA signal does not experience significant ISI. Multiple accesses is achieved with different users transmitting at the same set of sub-carriers but with spreading codes that are different to the codes of other users.

Initially, the data stream is serial to parallel converted to a number of lower rate streams. Each stream feeds a number of parallel streams with the same rate. On each of the parallel streams, bits are interleaved and spread by a PN code with a suitable chip rate. Then, these streams modulate different orthogonal carriers with a successively overlapping bandwidth [6,7].

In recent years, antenna systems which employ multiple antennas at both the base station (BS) and mobile station (MS), have been proposed and demonstrated to significantly increase system performance as well as capacity [8]. The merit of using multiple antennas or space diversity is that no bandwidth expansion or increase in transmitted power is required for capacity and performance improvements.

Space-time (S-T) coding is a technique designed for use with multiple transmits antennas. There are various approaches in coding structures, which include Alamouti space-time code (STC), space-time block code (STBC), space-time trellis code (STTC), space-time turbo trellis codes (STTTC), and layered space-time (LST) codes. STC with multiple transmit and receive antennas minimizes the effect of multi-path fading and improves the performance and capacity of digital transmission over wireless radio channels [9,10].

STBC can achieve a maximum possible diversity advantage with a simple decoding algorithm. It is very at-

tractive because of its simplicity. However, no coding gain can be provided by STBC. Tarokh, Seshadri, and Calderbank first introduced STTC. STTC are able to combat the effects of fading. However, STTC have a potential drawback due to the fact that its decoder complexity (maximum likelihood) grows exponentially with the number of bits per symbol [1].

## 3. STTC-MC-CDMA System Model

Trellis Coded Modulation (TCM) is a bandwidth efficient technique that combines coding and modulation, without reducing the data rate. STTC-MC-CDMA system provides coding gain and diversity. **Figure 1** shows the general block diagram of space-time coded MC-CDMA system. In this, the space-time trellis encoder encodes the source data; next the encoded data is interleaved, and then mapped according to the desired signal constellation. Finally, the space-time trellis encoded data symbols are modulated at each time interval, and transmitted simultaneously over different transmit antennas.

At the receiver, the received data is combined according to the different combining techniques described for MC-CDMA systems. The soft output of the combiner is sent to the deinterleaver and then finally it is applied to a space-time Trellis decoder, employing Viterbi algorithm to decode the data [11,12].

At the transmitter, $K$ users transmit simultaneously the space-time trellis coded information symbols from the two transmit antennas. The frequency selective channel between transmit and receive antennas is divided into $Q$ subchannels such that each subchannel is approximately flat. For the $k$th user, let the input message sequence $A_k$, is given by,

$$A_k(n) = (a_k(1), a_k(2), a_k(3),...,a_k(n),...) \quad (1)$$

where $a_k(n)$ is a group of $m = \log_2 M$ information bits at time $n$ and given by,

$$a_k(n) = (a_k^1, a_k^2, a_k^3,...,a_k^m) \quad (2)$$

The encoder maps the input stream into an *M-ary* PSK modulated signal sequence, which is given by, $\{X(n)\}$ be the ST-Trellis encoded output. For the *kth* user, the output of ST-Trellis encoder, which is modulated by MC-CDMA system, is represented by the following code matrix:

$$X_k(n) = (x_k(0), x_k(1), x_k(3),...,x_k(n),...) \quad (3)$$

where $x_k(n)$ is a $n$th STT coded symbol and is given by,

$$x_k^{n_T}(n) = (x_k^1(n), x_k^2(n),...,x_k^{n_T}(n))^T \quad (4)$$

The modulated signals, $x_k^1(n), x_k^2(n),...,x_k^{n_T}(n)$, are

transmitted simultaneously through $n_T$ transmit antennas. Here each user is assigned two distinct spreading codes to spread symbols transmitted from the two antennas. Let

$$c_k^1(p) = \left[ c_k^1(0), c_k^1(1), ..., c_k^1(p-1) \right]^T$$

and

$$c_k^2(p) = \left[ c_k^2(0), c_k^2(1), ......, c_k^2(p-1) \right]^T$$

be two spreading codes for user $k$ with processing gain $p$ which spread user $k$'s symbols transmitted from Tx1 and Tx2, respectively, where $(.)^T$ denotes vector/matrix transpose. $u_k^1(n)$ is defined as the signal associated with Tx1 after the spreading. Then, we have

$$u_k^1(n) = x_k^1(n)c_k^1 \tag{5}$$

For the ST-MC-CDMA system under study, we assume that the numbers of subcarriers equal to the processing gain $p$. Each element in $u_k^1(n)$ is then modulated onto a sub-carrier with center frequency at $\{\omega\}_{p=0}^{p-1}$, which is common for all $K$ users. The MC-CDMA modulation can be implemented by the inverse fast Fourier transform (IFFT) (consequently, the received signal can be demodulated by FFT). Performing $p$-point IFFT on $u_k^1(n)$ yields

$$z_k^1(n) = F^{-1}u_k^1(n) \tag{6}$$

where $F^{-1}$ denotes the $p \times p$ IFFT matrix. The contributions from all $K$ users associated with Tx1 are given by the superposition

$$Z^1(n) = \sum_{k=1}^{K} z_k^1(n).$$

Following the same procedure, we obtain the signal associated with Tx2 as $Z^2(n)$. Both $Z^1(n)$ and $Z^2(n)$ are converted into a serial sequences before being transmitted through the frequency selective channels [13-15].

For the forward link, let $J$ *be* the number of resolvable paths. If we model the common channel between Tx1 and Rx for all $K$ users as a FIR filter with coefficients:

$$h^1 = \left[ h(0), ....., h(J-1), 0_{p-J} \right]^T \tag{7}$$

where we pad $h^1$ with $p$ - $J$ zeros to make its total length $p$, its frequency response is then given by

$$g^1 = Fh^1 = \left[ g(0), ....., g(p-1) \right]^T \tag{8}$$

where $F$ denotes the $p \times p$ FFT matrix. Similarly, we define the FIR channel, also with length $J$, and its

corresponding frequency response between Tx2 and Rx as $h^2$ and $g^2$, respectively.

$$\begin{aligned} Y(n) &= \sum_{k=1}^{K} \left[ \varphi_k^1 g^1 x_k^1(n) + \varphi_k^2 g^2 x_k^2(n) \right] + v(n) \\ &= \sum_{k=1}^{K} C_k H x_k(n) + v(n) \end{aligned} \tag{9}$$

where $\quad \varphi_k^1 = diag\{c_k^1\}, \varphi_k^2 = diag\{c_k^2\}$ (10)

$$C_k = \left[ \varphi_k^1, \varphi_k^2 \right], \tag{11}$$

$$H = \begin{bmatrix} g^1 & 0 \\ 0 & g^2 \end{bmatrix}, \tag{12}$$

$$x_k(n) = \left[ x_k^1(n), x_k^2(n) \right]^T$$

and where

$$v(n) = \left[ v_0(n), ....., v_{p-1}(n) \right]^T$$

contains samples of the channel noise with zero-mean and variance $\sigma_n^2$. At the receiver, the received signals are first converted into parallel format and demodulated via FFT transform and then applied to deinterleaver. The interleaved data is applied to MMSE detector and then for ST-Trellis decoder for demodulation.

## 4. STTC-STBC MC-CDMA System

Space time block coding is a simple technique to achieve diversity, however there is not much significant coding gain. An outer channel code is required to yield coding gain.

Space time trellis code concatenated with space time block code (STTC-STBC) also provides coding gain. As shown in **Figure 1**, first the STTC encoder encodes the source data. Next, the encoded data is applied to S-T block encoder & interleaver, and then mapped according to the desired signal constellation. Finally at each time interval, the symbols are modulated and transmitted simultaneously over different transmit antennas.

At the receiver, the received data is combined according to the combining techniques described for STBC. The soft output of the combiner is sent directly to the deinterleaver, and then finally, it is applied to a STTC decoder, such as the Viterbi algorithm, to decode the data.

## 5. MMSE Detection

In conventional MC-CDMA systems (without transmit diversity), the demodulated signals are often combined in the frequency domain in order to collect the overall received signal energy scattered on different subcarriers. Typical signal combining schemes include the maximum

        

(a) Transmitter



(b) Receiver
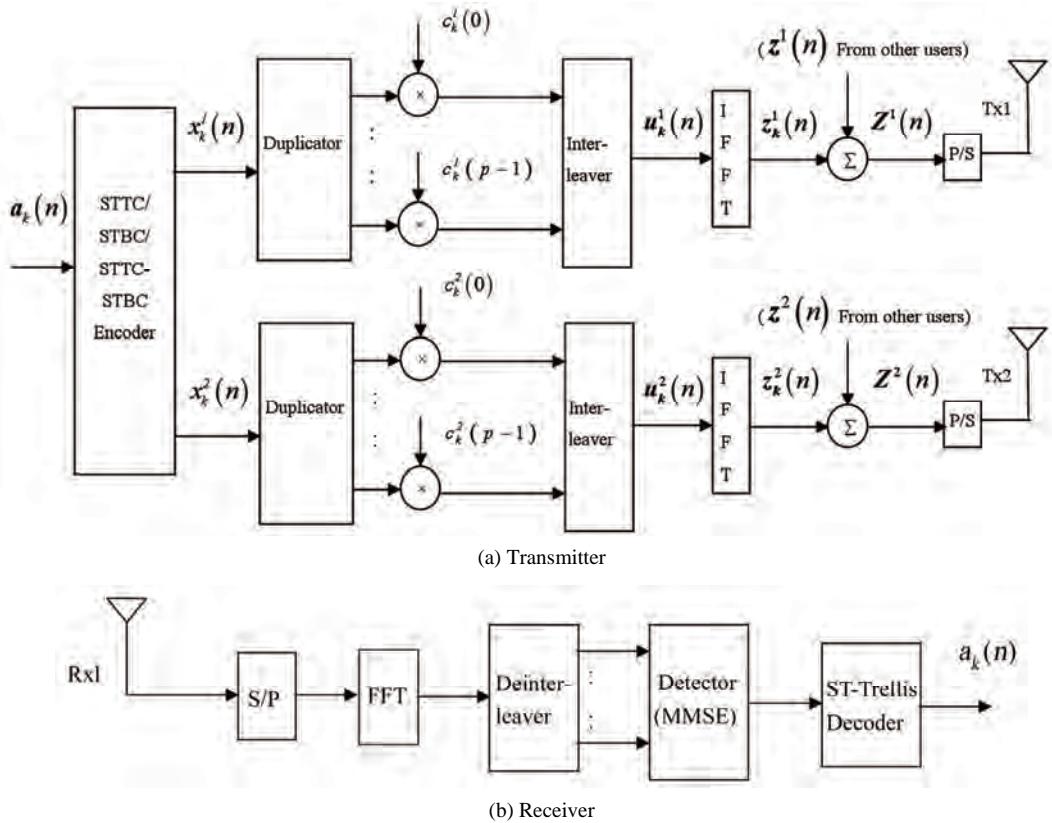
**Figure 1. General block diagram of S-T coded MC-CDMA system.**

ratio combining (MRC) and the equal gain combining (EGC). Both MRC and EGC are *single-user* detection schemes based on per-subcarrier combining, *i.e.* the signals at individual subcarrier are independently weighted and summed to generate decision variables. In [16], a linear multiuser minimum mean-squared error (MMSE) detector for the STC-MC-CDMA systems is presented which performs joint weighting and combining on all subcarriers by utilizing the mean-squared error (MSE) criterion. We consider the two demodulated symbols at a time. Define

$$G^1 = \text{diag}\{g^1\}, \tag{13}$$

$$\Theta^1 = [c_1,...,c_k] \tag{14}$$

And let $G^2$ and $\Theta^2$ be the counterpart of $G^1$ and $\Theta^1$ respectively, which are associated with Tx2.

$$S = \begin{bmatrix} G^1\Theta^1 & G^2\Theta^2 \\ G^{2*}\Theta^2 & -G^{1*}\Theta^1 \end{bmatrix} \tag{15}$$

$$R_z = SS' + \sigma_n^2 I \tag{16}$$

$$R_{zb1} = \begin{bmatrix} G^1 c_1^1 & G^2 c_1^2 \\ G^{2*} c_1^2 & -G^{1*} c_1^1 \end{bmatrix} \tag{17}$$

$$W_1 = R_z^{-1} R_{zb1} \tag{18}$$

By applying $W_1$ on $Y(n)$;

$$D_{MMSE(1)} = W_1' \times Y(n) \tag{19}$$

The output of the MMSE detector is employed to Viterbi decoder for decoding STTC data. This decoded data is then compared with original transmitted sequence $a_k(n)$ for finding bit error rate (BER) [17,18].

## 6. Simulation Results

The simulation is done for STBC MC-CDMA systems, STTC MC-CDMA systems, and STTC-STBC MC-CDMA systems for 4, 8, and 16 states with different combinations of diversity for $K = 5$ users. The user symbols are drawn from a unit-energy BPSK (binary phase shift keying) constellation. Walsh-Hadamard codes with processing gain $P=Q=32$ are used for spreading.

We assume a rich scattering environment and generate the FIR channel coefficients as i.i.d. complex Gaussian random variables with zero mean and variance 1/M [15]. The SNR (signal to noise ratio) is defined as

$$SNR = 10\log_{10} \frac{1}{\sigma_n^2} \text{ in dB.}$$

In **Table 1** we present the bit error rate (BER) of

MMSE detector for different states of STTC-MC-CDMA systems versus the SNR in a Rayleigh fading environment for $K = 5$ users and 10000 bit sequences. 100 independent trials are taken for each SNR and their average is plotted. This table is also shows the BER of two standard single-user signals combining schemes (MRC and EGC) for comparison. From **Table 1**, it is concluded that the MMSE detector outperforms the single-user signals combining schemes in known channel. This is due to the reason that EGC/MRC is basically single user detection techniques and is not able to effectively suppress the MAI as compared to MMSE detector. Here, we have also studied the BER performance of the MMSE detector for

the STTC-MC-CDMA system with the different states of STTC-MC-CDMA systems.
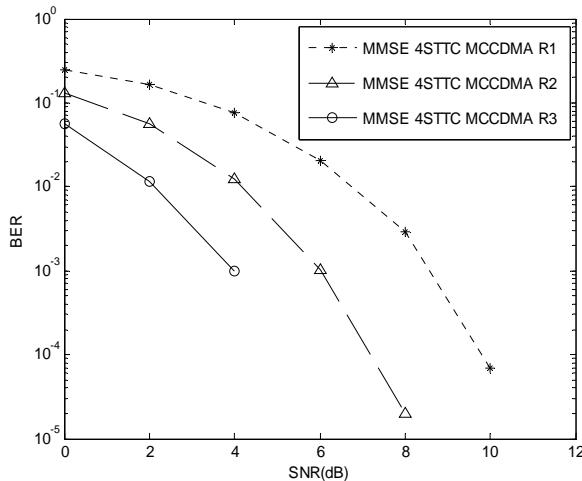
**Table 2** show the performance of STBC MC-CDMA systems, STTC MC-CDMA systems, and STTC-STBC MC-CDMA systems for diversity $2 \times 1$ employing MMSE detector. Simulations results are also shown in **Figures 2**, **3** and **4** for MIMO antenna systems. With all these simulations it is noted that the performance gain can be enhanced up to around 5 dB at $10^{-3}$ BER by increasing the diversity from $2 \times 1$ to $2 \times 3$ in 16 states STTC-STBC MC-CDMA systems with MMSE detector as shown in **Figures 5, 6** and **7**.

**Table 1. Comparison with different detection techniques.**
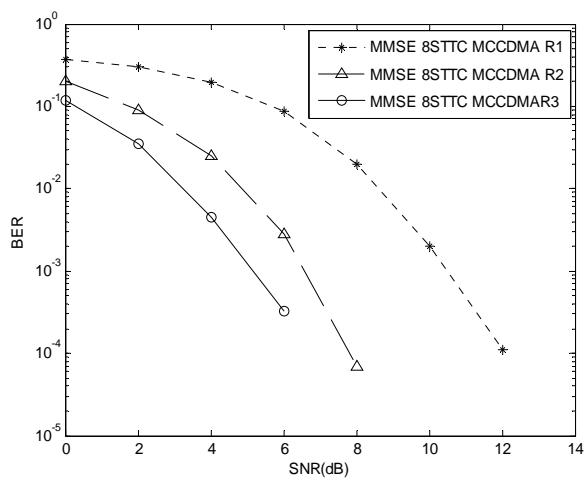
| SNR | BER | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | STTC-MC-CDMA Using MRC Detection | | | STTC-MC-CDMA Using EGC Detection | | | STTC-MC-CDMA Using MMSE Detection | | |
| | 4-State | 8-State | 16-State | 4-State | 8-State | 16-State | 4-State | 8-State | 16-State |
| 0 | 0.3568 | 0.3743 | 0.3882 | 0.2552 | 0.3097 | 0.4053 | 0.3101 | 0.3648 | 0.3978 |
| 2 | 0.3255 | 0.3262 | 0.3389 | 0.1802 | 0.2143 | 0.3454 | 0.2448 | 0.2849 | 0.3236 |
| 4 | 0.2967 | 0.2748 | 0.2703 | 0.1025 | 0.1260 | 0.2649 | 0.1618 | 0.1764 | 0.2072 |
| 6 | 0.2692 | 0.2246 | 0.1993 | 0.0497 | 0.0541 | 0.1712 | 0.0756 | 0.0706 | 0.0823 |
| 8 | 0.2533 | 0.1777 | 0.1412 | 0.0171 | 0.0180 | 0.0901 | 0.0208 | 0.0159 | 0.0170 |
| 10 | 0.2390 | 0.1432 | 0.0996 | 0.0043 | 0.0036 | 0.0402 | 0.0025 | 0.0014 | 0.0015 |

**Table 2. Comparison with different coding schemes.**
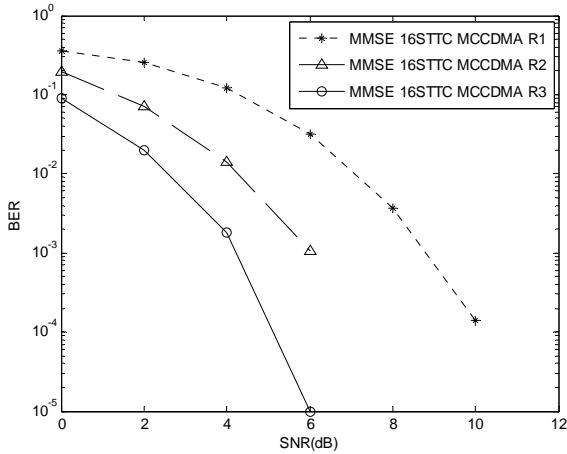
| SNR | BER Using MMSE Detection | | |
|---|---|---|---|
| | STBC-MC-CDMA | STTC-MC-CDMA | STTC-STBC-MC-CDMA |
| | | 4-State | 4-State |
| 0 | 0.0606 | 0.2680 | 0.0684 |
| 2 | 0.0191 | 0.1853 | 0.0174 |
| 4 | 0.0025 | 0.0996 | 0.0020 |
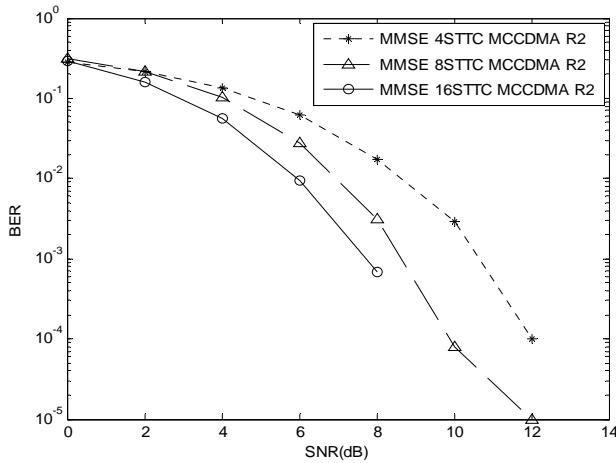


**Figure 2. BER performance for 4 states STTC-MC-CDMA system using 2 Tx and 1, 2, & 3 Rx antennas.**
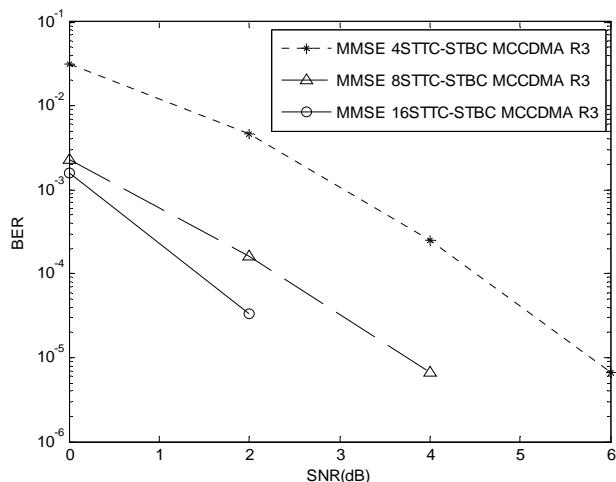


**Figure 3. BER performance for 8 states STTC-MC-CDMA system using 2 Tx and 1, 2, & 3 Rx antennas.**
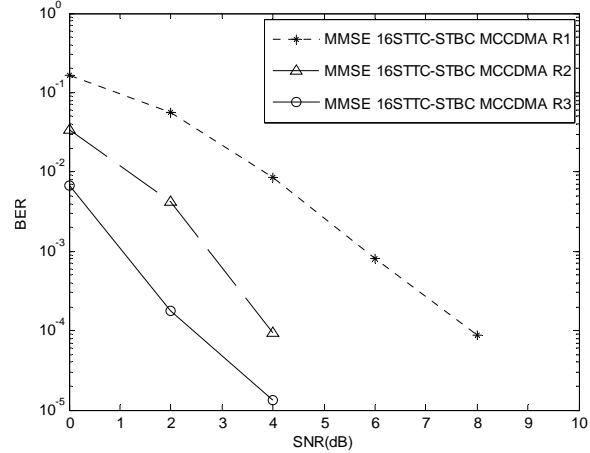
**Figure 4. BER performance for 16 states STTC-MC-CD-MA system using 2 Tx and 1, 2, & 3 Rx antennas.**



**Figure 5. BER performance for 4, 8, & 16 states STTC-MC-CDMA system using 2 Tx and 2 Rx antennas.**



**Figure 6. BER performance for 4, 8, & 16 states STTC-STBC MC-CDMA system using 2 Tx and 3 Rx antennas.**



**Figure 7. BER performance of 16 states STTC-STBC MC-CDMA system using 2 Tx and 1, 2, & 3 Rx antennas.**

## 6.1. Computational Complexity Calculations

Computational complexity in decoding is an important issue. Viterbi algorithm is used to decode the actual trellis codeword. The numeric computational complexity involved is calculated as follows.

Consider $L = 2^v$ is the number of states where $v$ is the number of delay elements in the STTC encoder. The number of operations for multiplication/division are $\{ n_R (n_T + 1) 2^L \}$, comparisons are $\{ 2^L - 1 \}$, and additions/ subtractions are $\{ [n_R (n_T + 1) - 1] 2^L \}$. It may be noted that $n_T$ is the number of transmitting and $n_R$ is the number of receiving antennas.

For example the numeric computational complexity for 4 states, 2 transmitter and 2 receiver antennas are:
Multiplication/Division are

$$= n_R (n_T + 1) 2^L = 96; \qquad (20)$$

Comparisons are $= 2^L - 1 = 15;$ $\qquad (21)$

& Addition/Subtraction are

$$= [n_R (n_T + 1) - 1] 2^L = 80 \qquad (22)$$

Compare to one receive antenna the no. of multiplication/division operations involved are two times (as evident from Equation (20)). However, the improvement in the performance is approximately 2 dB at $10^{-3}$ probability of error employing MMSE detection as shown in the simulation results. This demonstrates the tradeoff involved in computational complexity and performance improvement.

## 7. Conclusions
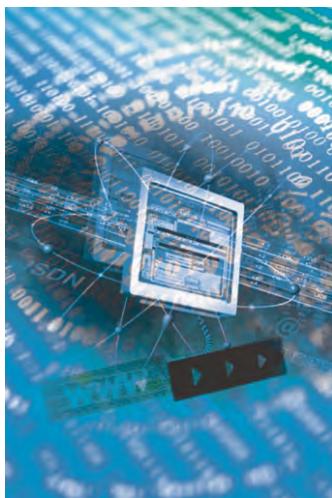
In this paper, BER performance for STBC-MC-CDMA

systems, STTC-MC-CDMA systems, and STTC-STBC MC-CDMA systems is evaluated using simulation for different states in known channel environment. For signal detection MMSE, MRC, and EGC receivers are considered. The simulation results are presented for the evaluation of the multi transmit multi receive antenna systems employing Viterbi decoding.

It is noted that the performance of STTC–STBC MC-CDMA system using MMSE detector is better than other schemes considered in this paper. Performance of all the schemes improves with the increase in number of states and antenna diversity. In particular, it is observed that STTC-STBC MC-CDMA systems using MMSE detector gives advantage of around 5 dB in terms of SNR while increasing diversity from $2 \times 1$ to $2 \times 3$, 16 states and at $10^{-3}$ BER. This improvement in performance is noted at the cost of increased computational complexity which is calculated and mentioned in this paper.

## 8. References

[1] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-Time Codes for High Data rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Transactions on Information Theory*, Vol. 44, No. 2, March 1998, pp. 744-765.

[2] H. Sampath, S. Talwar and J. Tellado, "A Fourth Generation MIMO-OFDM Broadband Wireless System: Design, Performance, and Field Trial Results," *IEEE Communication Magazine*, September 2002, pp. 143-149.

[3] M. Juntti, M. Vehkapera, J. Leinonen, "MIMO MC-CDMA Communications for Future Cellular Systems," *IEEE Communication Magazine*, Feburary 2005, pp. 118-124.

[4] S. Hijazi, B. Natarajan and Z. Wu., "Flexible Spectrum use and Better Coexistence at the Physical Layer of Future Wireless Systems via a Multicarrier Platform," *IEEE Wireless Communication*, April 2004, pp. 64-71.

[5] E. A. Sourour and Masao Nakagawa, "Performance of Orthogonal Multicarrier CDMA in a Multipath Fading Channel," *IEEE Transactions on Communications*, Vol. 44, No. 3, March 1996, pp. 356-366.

[6] S. Hara and R. Prasad, "An Overview of Multi-Carrier CDMA," *IEEE Communication Magazine*, Vol. 35, No. 12, December 1997, pp. 126-133.

[7] S. L. Miller and B. J. Rainbolt, "MMSE Detection of Multicarrier CDMA," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 11, November 2000, pp. 2356-2362.

[8] K. K. Wong, R. D. Murch and K. B. Letaief, "Perform

ance Enhancement of Multiuser MIMO Wireless Communication Systems," *IEEE Transactions on Communications*, Vol. 50, No. 12, December 2002, pp. 1960-1968.

[9] A. F. Naguib, V. Tarokh, N. Seshadri and A. R. Calderbank, "A Space-Time Coding Modem for High-Data-Rate Wireless Communications," *IEEE Journal on Selected Areas in Communication*, Vol. 16, No. 8, October 1998, pp. 1459-1478.

[10] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 8, October 1998, pp. 1451-1458.

[11] A. Trivedi and M. Bansal, "Adaptive Equalization of Space-Time Coded MC-CDMA Systems," *Communication, Computers and Signal Processing, IEEE Pacific Rim*, 22-24 August 2007, pp. 186-189.

[12] W. Sun, H. B. Li and M. Amin, "A Subspace-Based Channel Identification Algorithm for Forward Link in Space-Time Coded MC-CDMA Systems," *IEEE Proceedings,* 2002, pp. 445-448.

[13] F. S. Ostuni, M. R. Nakhai and H. Aghvami, "Iterative Multi-User MMSE Receiver for Space-Time Trellis Coded CDMA over Frequency Selective Channels," *IEEE 14th International Symposium on Personal, Indore and Mobile Radio Communication Proceedings*, Chennai 2003, pp. 1968-1972.

[14] D. Mottier, D. Castelain, J. F. Helard and J. Y. Baudais, "Optimum and Sub-Optimum Linear MMSE Multi-User Detection for Multi-Carrier CDMA Transmission Systems," *IEEE Proceedings*, 2001, pp. 868-872.

[15] Wei Sun, Hongbin Li and Moeness Amin, "MMSE Detection for Space-Time Coded MC-CDMA," *IEEE International Conference on Communications*, Cape Town, Vol. 5, No. 11-15, May 2003, pp. 3452-3456.

[16] L. A. P. Hernandez and M. G. Otero, "A New STB-TCM Coded MC-CDMA Systems with MMSE-SVA Based Decoding and Soft-Interference Cancellation," *IEEE Proceedings*, 2005, pp. 113-116.

[17] A. Trivedi and L. K. Bansal, "Performance Study of Space-Time Trellis Coded MC-CDMA System Employing Different Detection Techniques," *IEEE Proceedings, World Championship Sports Network*, Conducted by Indian Institute of Information Technology, Allahabad, December 27-29, 2008, pp. 137-140.

[18] A. Trivedi and L. K. Bansal, "Comparative Study of Space-Time Trellis Code Concatenated with Space-Time Block Code MC-CDMA System," *IEEE Proceedings, International Anti Corruption Conference*, Conducted by Thapar University, Patiala, March 6-7, 2009, pp. 1099-1102.

# International Journal of

# Communications, Network and System Sciences (IJCNS)

ISSN 1913-3715 (Print)     ISSN 1913-3723 (Online)

http://www.scirp.org/journal/ijcns/

IJCNS is an international refereed journal dedicated to the latest advancement of communications and network technologies. The goal of this journal is to keep a record of the state-of-the-art research and promote the research work in these fast moving areas.

## Editors-in-Chief

| | |
|---|---|
| Prof. Huaibei Zhou | Advanced Research Center for Sci. & Tech., Wuhan University, China |
| Prof. Tom Hou | Department of Electrical and Computer Engineering, Virginia Tech., USA |

## Subject Coverage

This journal invites original research and review papers that address the following issues in wireless communications and networks. Topics of interest include, but are not limited to:

| | |
|---|---|
| MIMO and OFDM technologies | Sensor networks |
| UWB technologies | Ad Hoc and mesh networks |
| Wave propagation and antenna design | Network protocol, QoS and congestion control |
| Signal processing and channel modeling | Efficient MAC and resource management protocols |
| Coding, detection and modulation | Simulation and optimization tools |
| 3G and 4G technologies | Network security |

We are also interested in:

- Short reports—Discussion corner of the journal :
    2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data.
- Book reviews—Comments and critiques.

## Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

## Website and E-Mail

http://www.scirp.org/journal/ijcns          ijcns@scirp.org

# TABLE OF CONTENTS

**Volume 3 Number 4**                                                    **April 2010**