蔡育仁 個人代表著作

(Yuh-Ren Tsai)

國立清華大學

電機工程學系暨通訊工程研究所

Yuh-Ren Tsai, “Spatial Correlation Models for Total Co-channel Interference and Carrier-to-Interference Ratio in Mobile Cellular Systems,”

IEEE Transactions on Wireless Communications **(Accepted)**

(SCI, EI) (Impact Factor 1.649)

Spatial Correlation Models for Total Co-channel Interference and Carrier-to-Interference Ratio in Mobile Cellular Systems

Yuh-Ren Tsai, *Member, IEEE,*

***Abstract*— Shadow fading, an inevitable phenomenon caused by obstructions in the propagation path, will cause large variations in the received signal strength for mobile radio environments. The spatial correlation of the shadowing effects is exponentially decayed by the increase in distance between any two separate positions. Variations in the shadowing effects are generally modeled as a Gaussian-Markov stochastic process. In this work, two Gaussian-Markov stochastic models have been proposed to characterize the spatial correlation properties and the variations in the total co-channel interference (CCI) as well as in the carrier-to-interference ratio (CIR) for downlink channels in mobile cellular systems. The numerical and simulation results show that these two models precisely characterize the spatial correlation properties of total CCI and CIR; good accuracy is guaranteed for different propagation environments, and the estimation errors for the standard deviations of the variations are limited to about 0.2 dB. The proposed models can be applied to the applications, such as the design of handoff schemes, link performance prediction, link adaptation control, and performance simulations.**

***Index Terms*— Shadowing, Cochannel Interference (CCI), Carrier-to-Interference Ratio (CIR), Spatial Correlation, Gaussian-Markov Model.**

I. Introduction

In mobile radio environments, received signal strength depends on the propagation loss, the shadowing effects, and the fast multipath fading effects. Shadowing effects, also known as slow fading, are caused by obstructions in the propagation path, and will induce a large variation in the received signal strength for mobile radio environments. The variation, due to shadowing effects, is generally characterized as a log-normally distributed random variable (RV) [1]. Furthermore, the spatial correlation properties of shadowing effects are interesting for some applications, such as the design of diversity and handover schemes [2]. A simple and realistic autocorrelation model, based on experimental data, was proposed by Gudmundson [3], in which the spatial correlation of the shadowing effects is exponentially decayed with the increase in distance between any two separate positions. The variation in shadowing effects is generally modeled as a Gaussian-Markov stochastic process [4].

For mobile cellular systems, the link quality, depending on desired signal strength, is generally limited by the amount of total received co-channel interference (CCI). Total CCI is the sum of multiple log-normally distributed interferers, and there is no closed expression, for the probability density function (pdf) of the total CCI. By modeling the total CCI as a log-normally distributed RV, several approximation methods have been developed [5]–[10]. These methods have been widely applied to the evaluation of system performance, mostly based on the outage probability, for wireless networks [11]–[22]. Furthermore, the second order statistics of CIR were studied based on the assumptions of the second order statistics of shadowing effects. Hence, the level crossing problem was well investigated, and the average duration of an outage, the frequency of an outage and the probability of an outage were evaluated [23]–[25]. Due to the shadowing effects, large variations in desired signal strength and total CCI are expected. Therefore, as the mobile station (MS) travels, the propagation environment changes accordingly, and so does the link quality, which is generally defined as the carrier-to-interference power ratio (CIR). Seeing the uncertainty of the radio link quality, link adaptation techniques have been proposed to prevail over channel conditions, and to maximize the transmission data rate and spectral efficiency in mobile cellular systems. For example, in third-generation CDMA systems, such as cdma2000 and UMTS, the link adaptation mechanism adapts the spreading factor, the transmission power, and the code rate to overcome channel conditions and co-channel interferences [26], [27]. Good predictions of the link quality can benefit the service quality prediction and the link adaptation control. Hence, it is interesting to investigate the spatial correlation properties of link quality, for mobile cellular systems.

So far, studies on the spatial correlation properties of radio propagation environments have mainly been focused on shadowing effects and fast fading effects [28]–[33]. In the literature, there is currently no spatial correlation model for total CCI or CIR, and this issue has not been previously analytically addressed. In this work, the spatial correlation properties of total CCI and CIR for mobile cellular systems are investigated, based on the model proposed by Gudmundson. Two simple Gaussian-Markov stochastic models are proposed to characterize the variations of total CCI and CIR received by an MS. These proposed models can be applied to such applications as the design of handoff schemes, link performance prediction, link adaptation control, and performance simulations. However, these models are not proposed for

This paper was presented in part at the IEEE Globecom'04 Conference, Dallas, November, 2004. This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC 93-2752-E-007-003-PAE.

Y.-R. Tsai is with the Department of Electrical Engineering and the Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: yrtsai@ee.nthu.edu.tw)

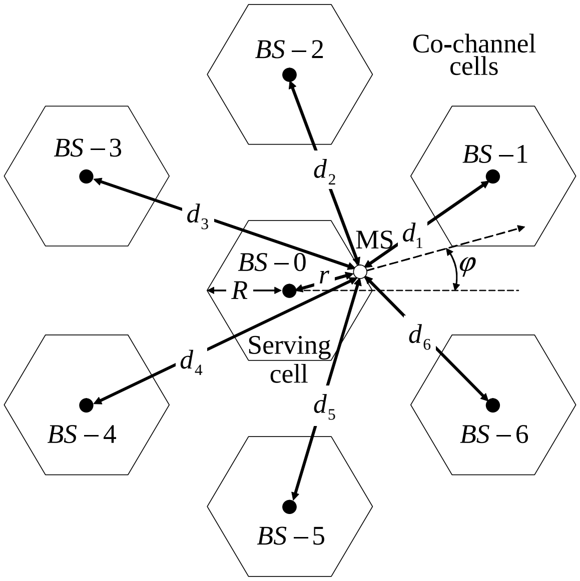


Fig. 1. System model for mobile cellular systems.

evaluating the second order statistics of shadowing processes as in [23]–[25]. The remainder of this paper is organized as follows: Section II illustrates the system and channel models; Section III and Section IV concentrate on the spatial correlation models for total CCI and CIR, respectively; in Section V, the numerical and simulation results are presented; and finally, the conclusions are offered in Section VI.

II. System and Channel Models

A. System Model

In mobile cellular systems, total CCI is accumulated from all co-channel cells. To simplify the calculations, the system is assumed to be with uniform grid of hexagonal cells, and only the N nearest co-channel cells with co-channel interferences $\{I_1, I_2, \ldots, I_N\}$ are considered. Shown in Fig. 1 is the system model for $N = 6$, where R denotes the cell radius and d_i denotes the distance between the base station-i (BS-i) and the desired MS. By geometric calculations, d_i, $i = 1, \ldots, N$ can be represented as functions of $d_0 \equiv r$ and the azimuth φ. Where r, ranging from 0 to R, is the distance between the MS and the serving BS, and φ, uniformly distributed over $(-\pi/6, \pi/6]$, denotes the traveling direction of the MS. For the case with $\varphi > \pi/6$ or $\varphi \le -\pi/6$, it can be recast to the same results as $-\pi/6 < \varphi \le \pi/6$. Consequently, if the distance r and the azimuth φ are known, the relative distances between the MS and all co-channel BSs can be determined for a specific frequency reuse factor K.

B. Channel Model

Owing to the adoption of interleaving with combining error correction coding techniques in radio transmission, only the path loss and the shadowing effects should be considered from the viewpoint of average link performance, which depends not on instantaneous CIR but on average CIR. Therefore, the channel model is assumed to be with the large-scale fading effects. If the small-scale fading effects are averaged out, the CCI power I_i from BS-i is a log-normally distributed RV. By using decibel (dB) units, we can model the mean CCI power as

$$X_i = 10 \log_{10}(I_i) = m_i + \chi_i, \qquad (1)$$

where m_i is the mean CCI power depending on the location of the MS, and χ_i is a Gaussian RV, with the mean zero and variance σ_i^2, representing the shadowing effects. Accordingly, X_i is a Gaussian RV with the mean m_i and variance σ_i^2. The mean power m_i is a function of d_i, the path loss exponent n, the transmission power, and the transmitter and receiver antenna gains. The standard deviation σ_i is dependent on the propagation environment, and is in the range of 5 to 12 dB. For macrocellular applications, a typical value of 8 dB is applied.

Furthermore, because different co-channel interferences, coming from different directions, may still be attenuated by the same obstructions surrounding the MS, we assumed that all co-channel interferences are correlated to each other; the correlation coefficient between X_i and X_j is defined as

$$\rho_{i,j} = \frac{E[(X_i - m_i)(X_j - m_j)]}{\sigma_i \sigma_j}. \qquad (2)$$

C. Gaussian-Markov Model for Shadowing Effect

As in [3], the spatial correlation properties of the shadowing effects are characterized by a Gaussian-Markov random process. For any two separate locations k and $k+1$ with a spatial distance small enough, it is reasonable to assume that the path losses at locations k and $k+1$ are almost the same, except for the shadowing effects. Therefore, the spatial correlation of the received CCI power (in dB units) from BS-i can be modeled as a Gaussian-Markov random process, i.e.

$$X_i[k+1] = \zeta X_i[k] + (1 - \zeta) V_i[k], \quad i = 1, \ldots, N, \qquad (3)$$

where $X_i[k]$ is the received CCI power at location k with the mean assumed to be $m_i[k] = m_i$ and variance σ_i^2; ζ is the spatial correlation coefficient of the shadowing effects; and $V_i[k]$ is a Gaussian RV with the mean m_i and variance

$$\sigma_{V_i}^2 = (1 + \zeta)\sigma_i^2 / (1 - \zeta). \qquad (4)$$

In [3], this spatial correlation model is verified by fitting it to the experimental data, which are obtained via measuring in the real propagation environments. This model shows good accuracy for spatial distances up to approximately 500 m. The spatial correlation coefficient ζ depends on the propagation environment and the spatial distance between locations k and $k+1$. For $\sigma_i \approx 8$ dB , ζ is estimated to be 0.82 at a distance of 100 m [3]. If the velocity of an MS is ν and the signal strength is sampled per T seconds, the value of ζ can be expressed as

$$\zeta = (0.82)^{D/100} = (0.82)^{\nu T/100}, \qquad (5)$$

where $D = \nu T$ is the spatial distance between two adjacent samples. Furthermore, it is noted that $X_i[k+1]$ is a Gaussian distributed RV with the mean $m_i[k+1] \approx m_i[k] = m_i$ and variance σ_i^2. The correlation coefficient between $X_i[k+1]$ and $X_j[k+1]$ is still $\rho_{i,j}$. This implies that $V_i[k]$, $i = 1, \ldots, N$ are correlated RVs with $\rho_{V_i, V_j} = \rho_{i,j}$.

III. Model for Total Co-Channel Interference

To simplify the complexity of our models, we ignored the effects of φ and permanently set φ to be 0, as shown in Fig. 1, so our models are independent of φ. Nevertheless, for the simulation results, the azimuth is treated as a RV, uniformly distributed over $(-\pi/6, \pi/6]$. Moreover, it is assumed that the distance r between the serving BS and the MS can be obtained via positioning technologies. This implies that the mean received CCI powers m_i, $i = 1, \ldots, N$ are available.

A. Gaussian-Markov Model for Total CCI

The total CCI, denoted as I_t, is the linear sum of N log-normally distributed interferers from neighboring co-channel cells, i.e. $I_t = \sum_{i=1}^{N} I_i$, and is generally approximated as another log-normally distributed RV $\hat{I}_t$. According to (1), we have

$$I_t = \sum_{i=1}^{N} I_i = \sum_{i=1}^{N} 10^{X_i/10} \cong \hat{I}_t = 10^{\Lambda/10}, \tag{6}$$

where $\Lambda = 10\log_{10}(\hat{I}_t)$ is the total received CCI in dB units. Assuming that the total received CCI at location k is $\Lambda[k]$, it can be approximated as a Gaussian distributed RV with the mean m_Λ and the variance σ_Λ^2, and we can model the spatial correlation of Λ as a Gaussian-Markov random process similar to (3), i.e.

$$\Lambda[k+1] = \alpha\Lambda[k] + (1-\alpha)U[k], \tag{7}$$

where α denotes the spatial correlation coefficient of the total received CCI, and $U[k]$ is a Gaussian RV with the mean $m_U \approx m_\Lambda$ and the variance

$$\sigma_U^2 = (1+\alpha)\sigma_\Lambda^2/(1-\alpha). \tag{8}$$

It is noted that the distribution of Λ highly depends on the location of the MS [34]. Hence, it is a location dependent model, and parameters α, m_U and σ_U^2 are functions of r. We need to determine m_U, σ_U^2 and α to complete the spatial correlation model for the total received CCI.

B. Determining the Mean and the Variance

To find the mean and the variance of Λ, two methods are proposed in literature to deal with the problem of the sum of multiple correlated log-normal RVs — the Fenton-Wilkinson method, which provides a fast and easy way by matching the first and second moments, and the Schwartz-and-Yeh method, which provides nesting and recursion techniques to find the accurate solutions. The details of these two methods can be found in [5], [6] and [8]. Therefore, m_Λ and σ_Λ^2 can be easily obtained by either one of these two methods.

C. The Adaptation of Spatial Correlation Coefficient α for Total CCI

The spatial correlation coefficient α depends on the location of the MS, and should be adapted to the value of r. We utilize the conditional first or second moment of the total CCI to figure out the value of α. According to (3) and (4), the first moment of the total CCI at location $k+1$, i.e. $I_t[k+1] = \sum_{i=1}^{N} 10^{X_i[k+1]/10}$, conditioning on a set of $X_i[k] = m_i$, $i = 1, \ldots, N$, is

$$\begin{aligned} \boldsymbol{M}_1 &= E\left[\sum_{i=1}^{N} \exp(\xi X_i[k+1]) \middle| X_i[k] = m_i\right] \\ &= \sum_{i=1}^{N} \exp(\xi\zeta m_i) \times \Phi_{V_i}\big(\xi(1-\zeta)\big) \\ &= \sum_{i=1}^{N} \exp\Big[\xi m_i + \xi^2(1-\zeta^2)\sigma_i^2/2\Big] \equiv \beta, \end{aligned} \tag{9}$$

where $\xi = (\ln 10)/10$, and $\Phi_{V_i}(\cdot)$ is the moment generating function of V_i. For a Gaussian distributed RV G with the mean m_G and variance σ_G^2, it is well known that the moment generating function is $\Phi_G(s) = E[e^{sG}] = \exp(s\,m_G + s^2\sigma_G^2/2)$. The value of β depends only on m_i and σ_i^2, and can be easily obtained if the location of the MS is known. Similarly, according to (7) and (8), the first moment of $\hat{I}_t[k+1] = 10^{\Lambda[k+1]/10}$ conditioning on a set of $X_i[k] = m_i$, i.e. conditioning on $\Lambda[k] = \tilde{\Lambda} = 10\log_{10}\left(\sum_{i=1}^{N} 10^{m_i/10}\right)$, is

$$\begin{aligned} \hat{\boldsymbol{M}}_1 &= E\left[\exp\left(\xi\Lambda[k+1]\right) \middle| \Lambda[k] = \tilde{\Lambda}\right] \\ &= \exp\left(\xi\alpha\tilde{\Lambda}\right) \times \Phi_U\big(\xi(1-\alpha)\big) \\ &= \exp\Big[\xi\alpha\tilde{\Lambda} + \xi(1-\alpha)m_\Lambda + \xi^2(1-\alpha^2)\sigma_\Lambda^2/2\Big], \end{aligned} \tag{10}$$

where $\Phi_U(\cdot)$ is the moment generating function of U. It is noted that the spatial correlation coefficient ζ for shadowing effects is non-negative. Since the total CCI is the linear sum of all co-channel interferences which are suffered by different shadowing effects, it is reasonable to conclude that the spatial correlation coefficient α is also non-negative. Therefore, we have the value of α being within $[0, 1]$. Subsequently, by equating (9) and (10), and constraining $0 \leq \alpha \leq 1$, we have

$$\alpha = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \tag{11}$$

where $a = \xi^2\sigma_\Lambda^2/2$, $b = \xi(m_\Lambda - \tilde{\Lambda})$ and $c = \ln\beta - \xi^2\sigma_\Lambda^2/2 - \xi m_\Lambda$.

The conditional first moment shown in (9) doesn't exhibit the cross-correlations between the CCIs from different cells. If the shadowing effects from different cells are correlated, i.e. $\rho_{i,j} \neq 0$, we may use the conditional second moment to determine the value of α. According to (3), (4) and the fact that $\rho_{V_i,V_j} = \rho_{i,j}$, the second moment of $I_t[k+1] = \sum_{i=1}^{N} 10^{X_i[k+1]/10}$ conditioning on a set of $X_i[k] = m_i$, $i = 1, \ldots, N$, is

$$\begin{aligned} &\boldsymbol{M}_2 \\ &= E\left[\left(\sum_{i=1}^{N} \exp\left(\xi X_i[k+1]\right)\right)^2 \middle| X_i[k] = m_i\right] \\ &= \sum_{i=1}^{N} \exp\left(2\xi\zeta m_i\right) \times \Phi_{V_i}\big(2\xi(1-\zeta)\big) \\ &\quad + \sum_{i=1}^{N} \sum_{j=1; j\neq i}^{N} \exp\big[\xi\zeta(m_i+m_j)\big] \times \Phi_{V_i,V_j}\big(\xi(1-\zeta), \xi(1-\zeta)\big) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \exp\left[2\xi m_i + 2\xi^2(1-\zeta^2)\sigma_i^2\right] \\
&+ \sum_{i=1}^{N} \sum_{j=1;j\neq i}^{N} \exp\left[\xi(m_i+m_j) + \frac{\xi^2(1-\zeta^2)(\sigma_i^2+\sigma_j^2+2\rho_{i,j}\sigma_i\sigma_j)}{2}\right] \\
&\equiv \beta', \qquad (12)
\end{aligned}
$$

where

$$
\begin{aligned}
&\Phi_{V_i,V_j}(s_1,s_2) \\
&= E\Big[\exp(s_1V_i + s_2V_j)\Big] \\
&= \exp\left(s_1 m_{V_i} + s_2 m_{V_j} + \frac{s_1^2\sigma_{V_i}^2 + s_2^2\sigma_{V_j}^2 + 2\rho_{V_i,V_j}s_1s_2\sigma_{V_i}\sigma_{V_j}}{2}\right)
\end{aligned}
$$

is the joint moment generating function of V_i and V_j. The value of β' depends only on m_i and σ_i^2, and can be easily obtained if the location of the MS is known. Similarly, according to (7) and (8), the second moment of $\hat{I}_t[k+1] = 10^{\Lambda[k+1]/10}$ conditioning on $\Lambda[k] = \tilde{\Lambda}$ is

$$
\begin{aligned}
\hat{\boldsymbol{M}}_2 &= E\left[\exp\left(2\xi\Lambda[k+1]\right)\Big|\Lambda[k]=\tilde{\Lambda}\right] \\
&= \exp\left(2\xi\alpha\tilde{\Lambda}\right) \times \Phi_U\left(2\xi(1-\alpha)\right) \\
&= \exp\left[2\xi\alpha\tilde{\Lambda} + 2\xi(1-\alpha)m_\Lambda + 2\xi^2(1-\alpha^2)\sigma_\Lambda^2\right]. \qquad (13)
\end{aligned}
$$

Subsequently, by equating (12) and (13), and constraining $0 \leq \alpha \leq 1$, we have

$$
\alpha = \frac{-b' + \sqrt{b'^2 - 4a'c'}}{2a'}, \qquad (14)
$$

where $a' = 2\xi^2\sigma_\Lambda^2$, $b' = 2\xi(m_\Lambda - \tilde{\Lambda})$ and $c' = \ln\beta' - 2\xi^2\sigma_\Lambda^2 - 2\xi m_\Lambda$.

It is noted that the value of α depends on m_i and σ_i^2, which can be determined if the location of the MS is known.

D. Algorithm

According to m_U and σ_U^2 obtained in Subsection III.*B* and α obtained in Subsection III.*C*, the spatial correlation model, in (7), for the total CCI can be properly determined. The algorithm for finding this model is summarized as follows.

***Algorithm* 1**

1) Determine the MS location parameter r.
2) According to (5), determine the desired spatial correlation coefficient ζ.
3) According to r and the path loss model, find m_i of $X_i[k]$, $i = 1, \ldots, N$.
4) Using the Fenton-Wilkinson or the Schwartz-and-Yeh method, find m_Λ and σ_Λ^2.
5) According to (11) or (14), determine the value of α.
6) According to the results obtained in Step 4, determine $m_U \approx m_\Lambda$ and $\sigma_U^2 = (1+\alpha)\sigma_\Lambda^2/(1-\alpha)$.
7) Based on α, m_U, σ_U^2 and (7), the spatial correlation model for the total CCI is constructed.

IV. Model for Carrier-to-Interference Power Ratio

The CIR is the ratio of the desired signal power to the total CCI power, and is generally used to represent the link quality in mobile cellular systems. If there is no multiple access interference (MAI) from the serving cell, the spatial correlation properties of the CIR can also be modeled as a Gaussian-Markov model.

A. Gaussian-Markov Model for CIR

Since the desired signal power Ω (in dB units), received by the MS, is affected by the path loss and the shadowing effects, we can model Ω as a Gaussian RV with the mean m_0 and the variance σ_0^2. It is also assumed that Ω and the co-channel interference X_i are correlated, with the correlation coefficient $\rho_{i,0}$. Following (3), the spatial correlation properties of Ω are modeled as a Gaussian-Markov random process, i.e.

$$
\Omega[k+1] = \zeta\Omega[k] + (1-\zeta)V_0[k], \qquad (15)
$$

where $\Omega[k]$ is the desired signal power at location k with the mean assumed to be $m_0[k] = m_0$ and the variance σ_0^2; and $V_0[k]$ is a Gaussian distributed RV with the mean $m_{V_0} \approx m_0$ and the variance $\sigma_{V_0}^2 = (1+\zeta)\sigma_0^2/(1-\zeta)$.

According to (6), the CIR (in linear scale) can be represented as

$$
\begin{aligned}
CIR &= \frac{10^{\Omega/10}}{\sum_{i=1}^{N} 10^{X_i/10}} = \frac{1}{\sum_{i=1}^{N} 10^{(X_i-\Omega)/10}} \\
&= \frac{1}{\sum_{i=1}^{N} 10^{\Psi_i/10}}, \qquad (16)
\end{aligned}
$$

where $\Psi_i = X_i - \Omega$ is a Gaussian distributed RV with the mean $m_{\Psi_i} = m_i - m_0$ and the variance $\sigma_{\Psi_i}^2 = \sigma_i^2 + \sigma_0^2 - 2\rho_{i,0}\sigma_i\sigma_0$. The random variables Ψ_i, $i = 1, \ldots, N$ are correlated. The correlation coefficient between Ψ_i and Ψ_j is

$$
\begin{aligned}
\rho_{\Psi_i,\Psi_j} &= \frac{E[(X_i-\Omega)(X_j-\Omega)] - (m_i-m_0)(m_j-m_0)}{\sqrt{(\sigma_i^2+\sigma_0^2-2\rho_{i,0}\sigma_i\sigma_0)(\sigma_j^2+\sigma_0^2-2\rho_{j,0}\sigma_j\sigma_0)}} \\
&= \frac{\rho_{i,j}\sigma_i\sigma_j - \rho_{i,0}\sigma_i\sigma_0 - \rho_{j,0}\sigma_j\sigma_0 + \sigma_0^2}{\sqrt{(\sigma_i^2+\sigma_0^2-2\rho_{i,0}\sigma_i\sigma_0)(\sigma_j^2+\sigma_0^2-2\rho_{j,0}\sigma_j\sigma_0)}}. \qquad (17)
\end{aligned}
$$

A similar result and concept has also been addressed in [19]. One special case is $\rho_{\Psi_i,\Psi_j} = 1/2$ for $\sigma_i = \sigma_j = \sigma_0$ and $\rho_{i,j} = \rho_{i,0} = \rho_{j,0} = \rho$.

By defining $\Theta = 10\log_{10}(\sum_{i=1}^{N} 10^{\Psi_i/10})$, we have Θ (in dB units) as the linear sum of N log-normally distributed RVs; Θ can be approximated as a Gaussian distributed RV with the mean m_Θ and the variance σ_Θ^2. Consequently, the spatial correlation properties of Θ can be modeled as a Gaussian-Markov random process similar to (3), giving us

$$
\Theta[k+1] = \lambda\Theta[k] + (1-\lambda)W'[k], \qquad (18)
$$

where λ represents the spatial correlation coefficient, and $W'[k]$ is a Gaussian RV with the mean $m_{W'} \approx m_\Theta$ and the variance

$$
\sigma_{W'}^2 = (1+\lambda)\sigma_\Theta^2/(1-\lambda). \qquad (19)
$$

Following the procedures shown in Section III, and by substituting Ψ_i for X_i, we can determine the values of m_Θ, σ_Θ^2 and λ.

According to (16), the CIR (in dB units), denoted as Γ, is equal to $-\Theta$. Therefore, Γ can be approximated as a Gaussian distributed RV with the mean $m_\Gamma = -m_\Theta$ and the variance $\sigma_\Gamma^2 = \sigma_\Theta^2$. By substituting $-\Gamma$ for Θ in (18), we have the spatial correlation properties of Γ modeled as a Gaussian-Markov random process, i.e.

$$\Gamma[k+1] = \lambda\Gamma[k] + (1-\lambda)W[k], \tag{20}$$

where λ represents the spatial correlation coefficient of the CIR, and $W[k] = -W'[k]$ is a Gaussian distributed RV with the mean $m_W \approx m_\Gamma$ and the variance

$$\sigma_W^2 = (1+\lambda)\sigma_\Gamma^2/(1-\lambda). \tag{21}$$

It is noted that Γ is also a location dependent process.

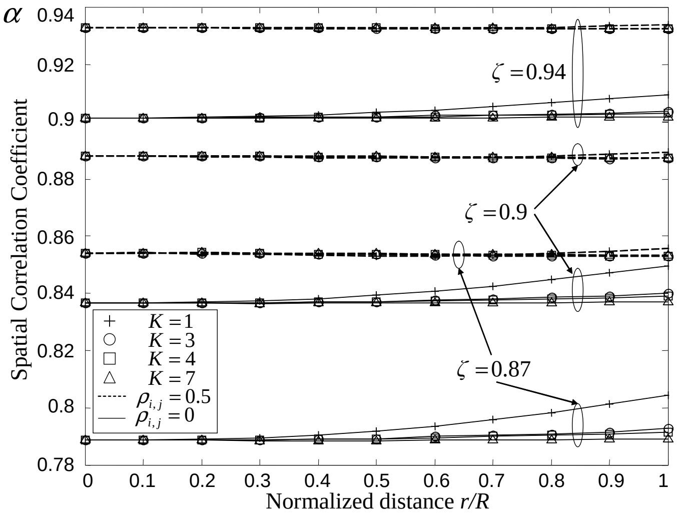


Fig. 2. The spatial correlation coefficient α versus r/R.

B. Algorithm

The algorithm for finding the spatial correlation model for the CIR is summarized as follows.

***Algorithm* 2**

1) Determine the MS location parameter r.
2) According to (5), determine the desired spatial correlation coefficient ζ.
3) According to r and the path loss model, find m_i for $i = 0, \ldots, N$.
4) Define $\Psi_i = X_i - \Omega$ and determine $m_{\Psi_i} = m_i - m_0$ and $\sigma_{\Psi_i}^2 = \sigma_i^2 + \sigma_0^2 - 2\rho_{i,0}\sigma_i\sigma_0$ for $i = 1, \ldots, N$.
5) According to (17), determine ρ_{Ψ_i,Ψ_j}.
6) By substituting Ψ_i for X_i and according to Step 4 to Step 6 of ***Algorithm* 1**, determine m_Θ, σ_Θ^2 and λ.
7) Determine $m_\Gamma = -m_\Theta$, $\sigma_\Gamma^2 = \sigma_\Theta^2$, $m_W \approx m_\Gamma$ and $\sigma_W^2 = (1+\lambda)\sigma_\Gamma^2/(1-\lambda)$.
8) Based on λ, m_W, σ_W^2 and (20), the spatial correlation model for the CIR is constructed.

V. Numerical and Simulation Results

In the numerical results, it was assumed that the standard deviation of the shadowing effects is $\sigma_i = 8$ dB for all i, the path loss exponent is $n = 4$, and only the first-tier of co-channel interfering cells with the same transmission power was considered, i.e. $N = 6$. In addition, the cross-correlation of the shadowing effects was assumed to be homogenous, i.e. $\rho_{i,j}$ is a constant. Since the spatial correlation model for the shadowing effects was shown to be able to accurately characterize the real propagation environments [3], the simulation results based on (3) can represent the real propagation environment very well, and are therefore used to verify the accuracy of the proposed models for total CCI and CIR. The simulation results were obtained via maintaining six CCI processes, based on (3) and with $X_i[k]$ and $V_i[k]$ treated as RVs, to simulate the total CCI in real propagation environments. For the simulation results of the CIR, a desired signal process based on (15) was also maintained. Furthermore, the azimuth φ was assumed to be uniformly distributed over $(-\pi/6, \pi/6]$ for the simulation results. In the literature, it has been shown that the Schwartz-and-Yeh method provides better accuracy than the Fenton-Wilkinson method. Therefore, the following numerical results were basically obtained by employing the Schwartz-and-Yeh method, since this method provided better approximation results.

Fig. 2 shows the spatial correlation coefficient α versus r/R for $\zeta = 0.94$, 0.9, or 0.87, $K = 1$, 3, 4, or 7, and $\rho_{i,j} = 0$ or 0.5. Recalling Subsection III.*C*, there are two possible ways to figure out the spatial correlation coefficient α—by matching the first moments, i.e. (11), or the second moments, i.e. (14). It is noted that (11) is only feasible for $\rho_{i,j} = 0$, and (14) is feasible for all $\rho_{i,j}$. In this figure, (11) is applied to figuring out the values of α for the case $\rho_{i,j} = 0$, whereas (14) is applied for the case $\rho_{i,j} = 0.5$. The spatial correlation coefficient α of the total CCI was found to be much smaller than that of the shadowing effects ζ for $\rho_{i,j} = 0$. The diminution in spatial correlation coefficient is due to the total CCI being the sum of six CCIs from different cells, which results in a smaller spatial correlation for the output of the Gaussian-Markov random process. Since the total CCI will be dominated by one or two interferers when the MS approaches the cell border, we also found that α slightly increases as r/R approaches 1, especially for $K = 1$. Furthermore, α is close to ζ for a large value of the cross-correlation $\rho_{i,j}$, since a high degree of cross-correlation implies good stability on the channel variation. However, the value of ζ will always be the upper bound of α.

From the viewpoint of link performance, it is interesting to predict the variation of the total CCI in a prospective time interval. Shown in Fig. 3 is the standard deviation of $\Lambda[k]\big|\Lambda[k+1]$, denoted as $\sigma_{\Delta\Lambda}$, versus r/R for $\zeta = 0.94$, 0.9, or 0.87, $K = 1$, 3, 4, or 7, and $\rho_{i,j} = 0$. Recalling (7) and (8), we have $\sigma_{\Delta\Lambda} = (1-\alpha)\sigma_U = \sqrt{1-\alpha^2}\sigma_\Lambda$. Moreover, (11) was applied in the numerical results since the cross-correlation $\rho_{i,j} = 0$. Comparing the numerical and simulation results, we found that the proposed model provided very good accuracy, and the errors were only about 0.1 dB for all cases.

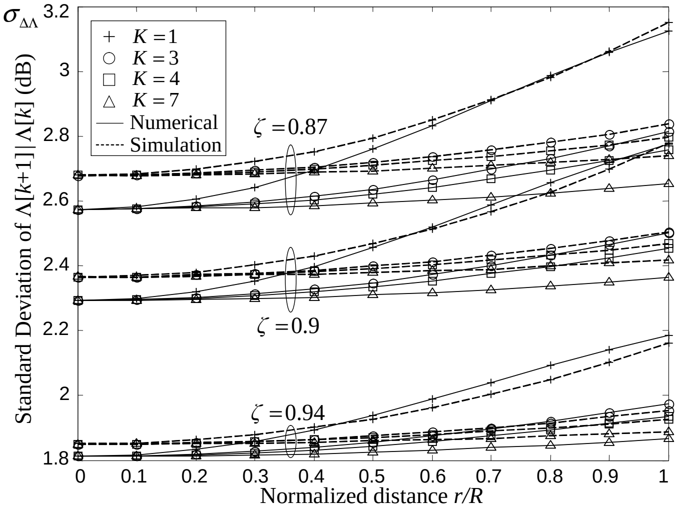


Fig. 3. The standard deviation $\sigma_{\Delta\Lambda}$ versus r/R with $\rho_{i,j}=0$.

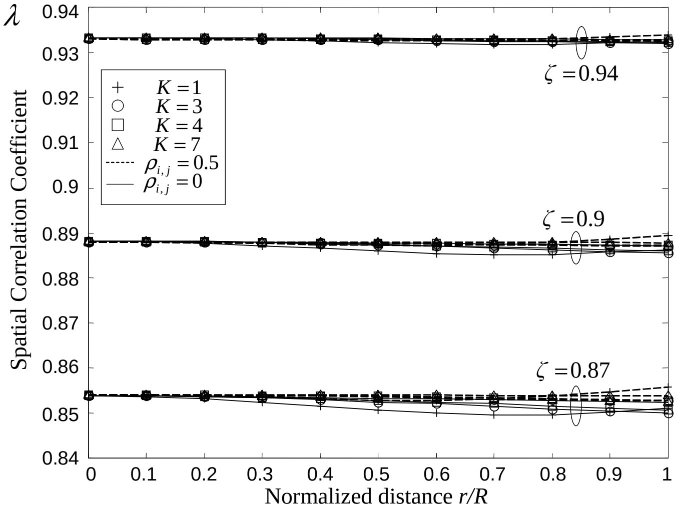


Fig. 4. The spatial correlation coefficient λ versus r/R.

It is reasonable that a large value of $\sigma_{\Delta\Lambda}$ was observed for a small value of ζ. Moreover, the standard deviation increases as r/R increases. This implies that a large variation of the total CCI is experienced when the MS approaches the cell border. For example, if an MS with the velocity ν = 60 Km/hr is at $r=0.8R$ and the prediction interval is 4 sec, the value of ζ will be 0.87. According to Fig. 3, the variation of the total CCI for $K=1$ after 4 sec will be a Gaussian RV with a standard deviation of 3 dB, and the probability of the total CCI increasing by more than 3 dB is 0.159.

Fig. 4 shows the spatial correlation coefficient λ versus r/R for $\zeta=0.94$, 0.9, or 0.87, $K=1$, 3, 4, or 7, and $\rho_{i,j}=0$ or 0.5. Recalling (17), the cross-correlation between Ψ_i and Ψ_j is non-zero for all $\rho_{i,j}$, and therefore (14) should be applied to find the spatial correlation coefficient λ. We found that the spatial correlation coefficient λ is smaller than that of the shadowing effects ζ, but much larger than α, especially for $\rho_{i,j}=0$. Since the value of CIR (in dB units) is obtained by subtracting the total CCI power from the received desired signal power, λ depends both on the spatial correlation coefficients ζ and α. Therefore, a spatial correlation coefficient between ζ and α is expected for the CIR.

Fig. 5 shows the standard deviation of $\Gamma[k{+}1]\big|\Gamma[k]$, denoted as $\sigma_{\Delta\Gamma}$, versus r/R for $\zeta=0.94$, 0.9, or 0.87, $K=1$, 3, 4, or 7, and $\rho_{i,j}=0$. Recalling (20) and (21), we have $\sigma_{\Delta\Gamma}=(1-\lambda)\sigma_W=\sqrt{1-\lambda^2}\sigma_\Gamma$. Moreover, (14) was applied in the numerical results. The proposed model provides very good accuracy, and the approximation errors are smaller than 0.2 dB for all cases. It is reasonable that a large value of the standard deviation was observed for a small value of ζ. In addition, the standard deviation increases as r/R increases. For an MS under the above-mentioned conditions, the variation of the CIR will be a Gaussian RV with a standard deviation of 5 dB, and the probability of the CIR degrading more than 3 dB is 0.274.

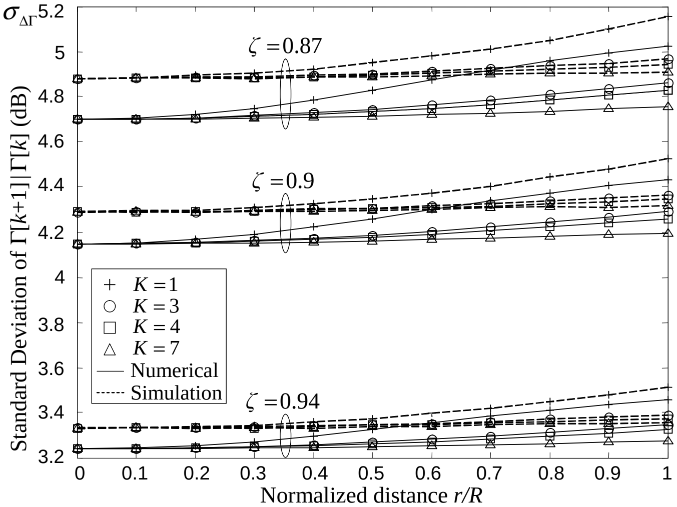


Fig. 5. The standard deviation $\sigma_{\Delta\Gamma}$ versus r/R with $\rho_{i,j}=0$.

To examine the fitness of the proposed models for different environments, we show the results of the normalized estimation error of $\sigma_{\Delta\Lambda}$ and $\sigma_{\Delta\Gamma}$, defined as the absolute value of the difference between the numerical and simulation results normalized to the simulation result, for different values of cross-correlation $\rho_{i,j}$. Fig. 6 shows the normalized estimation error of $\sigma_{\Delta\Lambda}$ versus $\rho_{i,j}$ for $\zeta=0.94$, 0.9, or 0.87, $K=1$ and $r/R=0.5$. We found that the normalized estimation error is inversely proportional to the values of $\rho_{i,j}$. In other words, better estimation can be obtained for the environment with a high degree of cross-correlation of shadowing effects. This figure also shows that the results based on the Schwartz-and-Yeh method provide good accuracy, since this method provides good accuracy on the approximation of the mean and the variance. The results based on the Fenton-Wilkinson method also show good estimation accuracy, although the approximation of the mean and the variance has a large deviation.

Fig. 7 shows the normalized estimation error of $\sigma_{\Delta\Gamma}$ versus $\rho_{i,j}$ for $\zeta=0.94$, 0.9, or 0.87, $K=1$ and $r/R=0.5$. We found that the normalized estimation error is also inversely proportional to the values of $\rho_{i,j}$ for the results based on the Schwartz-and-Yeh method. Furthermore, the results based on both the Schwartz-and-Yeh and the Fenton-Wilkinson methods

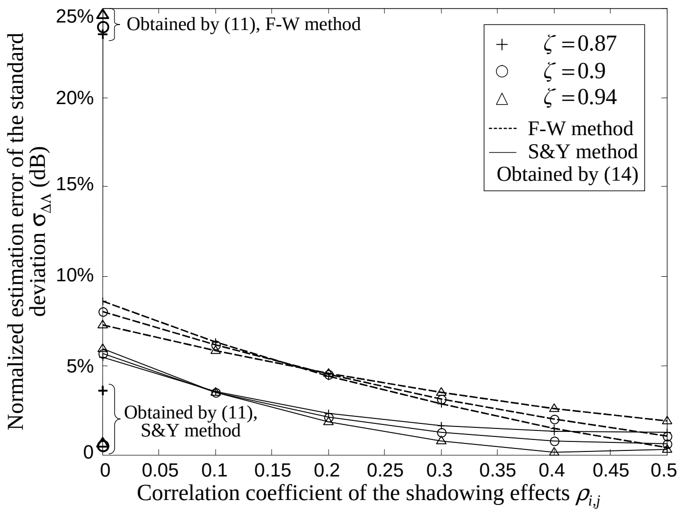


Fig. 6. Normalized estimation error of the standard deviation $\sigma_{\Delta\Lambda}$ versus $\rho_{i,j}$ with $r/R = 0.5$ and $K = 1$.

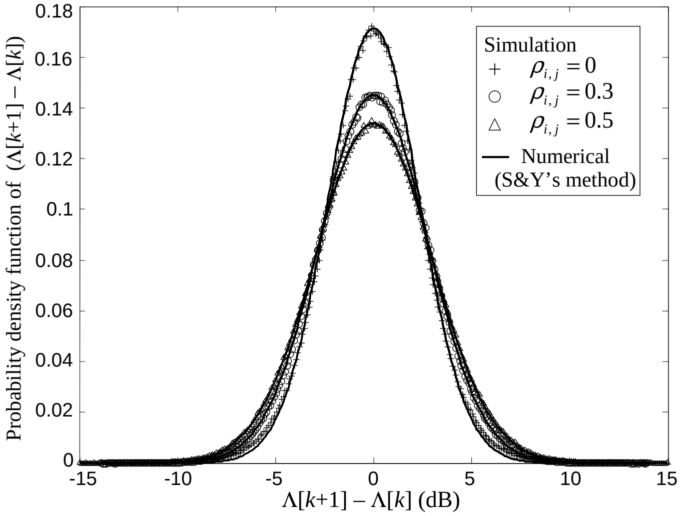


Fig. 8. The probability density function of $\Lambda[k+1] - \Lambda[k]$ for $r/R = 0.5$, $K = 1$ and $\zeta = 0.9$.

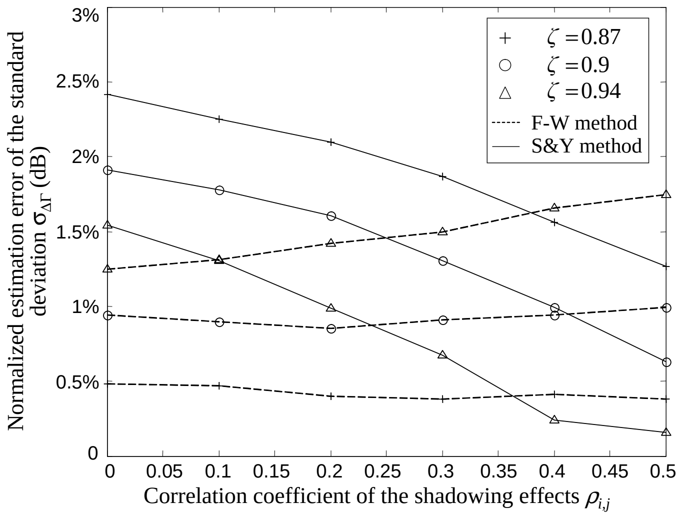


Fig. 7. Normalized estimation error of the standard deviation $\sigma_{\Delta\Gamma}$ versus $\rho_{i,j}$ with $r/R = 0.5$ and $K = 1$.

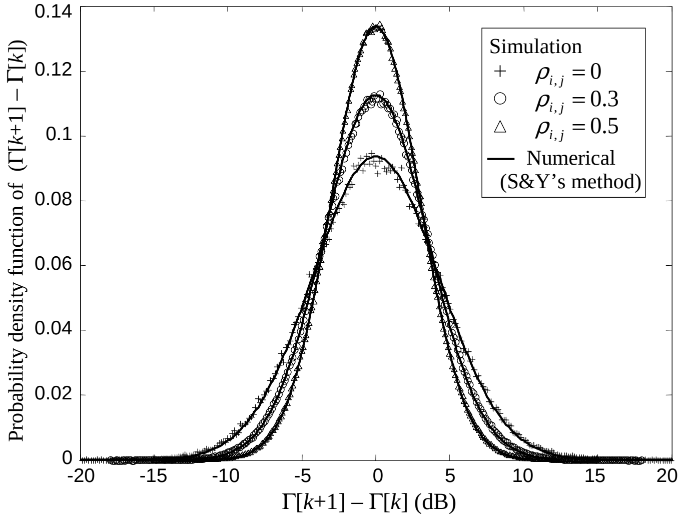


Fig. 9. The probability density function of $\Gamma[k+1] - \Gamma[k]$ for $r/R = 0.5$, $K = 1$ and $\zeta = 0.9$.

provided good accuracy. It is of note, however, that the computational complexity of the Fenton-Wilkinson method is much lower than that of the Schwartz-and-Yeh method. If one desires only a prediction of the variation of total CCI or CIR, the Fenton-Wilkinson method should be applied in order to reduce computational complexity. However, the Schwartz-and-Yeh method should be applied when constructing the complete spatial correlation models, since it provides good accuracy on the approximation of the mean and the variance.

To further verify the accuracy of the proposed models, we examined the conditional distributions of total CCI and CIR. Shown in Fig. 8 is the probability density function (pdf) of $\Lambda[k+1] - \Lambda[k]$ for $r/R = 0.5$. $K = 1$ and $\zeta = 0.9$. All numerical results were obtained by applying the method based on matching the second moments, i.e. (14). We found that the numerical results, obtained from the proposed model for total CCI, fitted in very well with the simulation results for different values of $\rho_{i,j}$. Furthermore, a large variance was observed for a large value of $\rho_{i,j}$, since a large variance of the total CCI had been expected for a large value of $\rho_{i,j}$. Fig. 9 shows the pdf of $\Gamma[k+1] - \Gamma[k]$ for $r/R = 0.5$, $K = 1$ and $\zeta = 0.9$. The numerical results, obtained from the proposed model for CIR, still fitted in very well with the simulation results for different values of $\rho_{i,j}$. The high degree of coincidence, between the numerical and simulation results, in distribution functions implies that the proposed models can describe the spatial correlation properties of total CCI and CIR very well.

VI. Conclusions

In this work, two Gaussian-Markov stochastic models were proposed to characterize the spatial correlation properties and variations of total CCI and CIR for mobile cellular systems. The simulation results, based on the existing spatial correlation model for shadowing effects, were also presented to verify the accuracy of the proposed models. The numerical and simulation results showed that these two models precisely characterize the spatial correlation properties of the total CCI and CIR. More important, for the prediction of variations of in total CCI and CIR, good accuracy can be guaranteed for different propagation environments, with the estimation errors in the standard deviations of the variations being limited to about 0.2 dB for total CCI and CIR. Considering the computational complexity of the Fenton-Wilkinson method, it can be applied if only the prediction of the variation of total CCI or CIR is desired. However, if complete spatial correlation models are required, the Schwartz-and-Yeh method is preferred, since it provides good accuracy on the approximation of the mean and the variance.

By using these proposed models, the variations in total CCI and CIR can be predicted, according to user mobility and location. The results can be applied to such applications as the design of handoff schemes, link performance prediction, link adaptation control, and performance simulations.

REFERENCES

[1] G. L. Stuber, *Principles of Mobile Communication*, 2nd ed. Boston, MA: Kluwer Academic Publishers, 2001.

[2] M. Gudmundson, "Analysis of handover algorithms," in *Proc. IEEE Vehicular Technology Conf.*, May 1991, pp. 537–542.

[3] ——, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[4] H. Stark and J. W. Woods, *Probability and Random Processes with Application to Signal Processing*, 3rd ed. Prentice Hill, 2002.

[5] L. F. Fenton, "The sum of log-normal probability distributions in scatter transmission systems," *IEEE Trans. Commun.*, vol. 8, no. 1, pp. 57–67, Mar. 1960.

[6] P. Cardieri and T. S. Rappaport, "Statistics of the sum of lognormal variables in wireless communications," in *Proc. IEEE Vehicular Technology Conf., Spring*, vol. 1, Tokyo, Japan, May 15–18, 2000, pp. 1823–1827.

[7] S. C. Schwartz and Y. S. Yeh, "On the distribution function and moments of power sums with lognormal components," *Bell Syst. Tech. J.*, vol. 61, no. 7, pp. 1441–1462, Sept. 1982.

[8] A. Safak, "Statistical analysis of the power sum of multiple correlated log-normal components," *IEEE Trans. Veh. Technol.*, vol. 42, no. 1, pp. 58–61, Feb. 1993.

[9] A. Abu-Dayya and N. C. Beaulieu, "Outage probabilities in the presence of correlated lognormal interferers," *IEEE Trans. Veh. Technol.*, vol. 43, no. 1, pp. 164–173, Feb. 1994.

[10] N. C. Beaulieu, A. Abu-Dayya, and P. J. MacLane, "Estimating the distribution of a sum of independent lognormal random variables," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2869–2873, Dec. 1995.

[11] Y. S. Yeh and S. C. Schwartz, "Outage probability in mobile telephony due to multiple log-normal interferers," *IEEE Trans. Commun.*, vol. 32, no. 4, pp. 380–388, Apr. 1984.

[12] A. Safak and R. Prasad, "Effects of correlated shadowing signals on channel reuse in mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 708–713, Nov. 1991.

[13] M. J. Ho and G. L. Stuber, "Co-channel interference of microcellular systems on shadowed Nakagami fading channel," in *Proc. IEEE Vehicular Technology Conf.*, May 18–20, 1993, pp. 568–571.

[14] R. Prasad and A. Kegel, "Effects of rician faded and log-normal shadowed signals on spectrum efficiency in microcellular radio," *IEEE Trans. Veh. Technol.*, vol. 42, no. 3, pp. 274–281, Aug. 1993.

[15] A. Safak, "Optimal channel reuse in cellular radio systems with multiple correlated log-normal interferers," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 304–312, May 1994.

[16] F. Graziosi and F. Santucci, "Analysis of second order statistics of the SIR in cellular mobile networks," in *Proc. IEEE Vehicular Technology Conf., Fall*, vol. 3, Amsterdam, Netherlands, Sept. 19–22, 1999, pp. 1316–1320.

[17] A. Ligeti, "Outage probability in the presence of correlated lognormal useful and interfering components," *IEEE Commun. Lett.*, vol. 4, no. 1, pp. 15–17, Jan. 2000.

[18] G. Karmani and K. N. Sivarajan, "Capacity evaluation for CDMA cellular systems," in *Proc. IEEE INFOCOM'01*, vol. 1, Anchorage, Alaska, Apr. 22–26, 2001, pp. 601–610.

[19] F. Graziosi, L. Fuciarelli, and F. Santucci, "Second order statistics of the SIR for cellular mobile networks in the presence of correlated co-channel interferers," in *Proc. IEEE Vehicular Technology Conf., Spring*, vol. 4, Rhodes Island, Greece, May 6–9, 2001, pp. 2499–2503.

[20] V. Emamian, M. Kaveh, and M.-S. Alouini, "Outage probability with transmit and receive diversity in a shadowing environment," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC2002)*, vol. 1, Orlando, FL, Mar. 17–21, 2002, pp. 54–57.

[21] F. Berggren and S. B. Slimane, "A simple bound on the outage probability with lognormally distributed interferers," *IEEE Commun. Lett.*, vol. 8, no. 5, pp. 271–273, May 2004.

[22] T. Piboongungon and V. A. Aalo, "Outage probability of l-branch selection combining in correlated lognormal fading channels," *Electron. Lett.*, vol. 40, no. 14, pp. 886–887, July 2004.

[23] N. B. Mandayam, P.-C. Chen, and J. M. Holtzman, "Minimum duration outage for cellular systems: a level crossing analysis," in *Proc. IEEE Vehicular Technology Conf., Spring*, vol. 2, Atlanta, Georgia, Apr. 1996, pp. 879–883.

[24] F. Santucci, M. Pratesi, M. Ruggieri, and F. Graziosi, "A general analysis of signal strength handover algorithms with cochannel interference," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 231–241, Feb. 2000.

[25] F. Graziosi and F. Santucci, "On SIR fade statistics in Rayleigh-lognormal channels," in *Proc. ICC 2002*, vol. 3, New York, NY, Apr. 2002, pp. 1352–1357.

[26] *Physical layer standard for cdma2000 spread spectrum systems*, 3GPP2 C.P0002-A, 3rd Generation Partnership Project 2 Std., Oct. 1999.

[27] *Technical Specification Group Radio Access Networks*, 3G TS 25.XXX (Release 1999), 3rd Generation Partnership Project Std., June 2000.

[28] S. R. Saunders and B. G. Evans, "The spatial correlation of shadow fading in macrocellular mobile radio systems," in *IEE Colloquium on Propagation Aspects of Future Mobile Systems*, Oct. 25, 1996, pp. 2/1–2/6.

[29] Y. Karasawa and H. Iwai, "Formulation of spatial correlation statistics in Nakagami-Rice fading environments," *IEEE Trans. Antennas Propagat.*, vol. 48, no. 1, pp. 12–18, Jan. 2000.

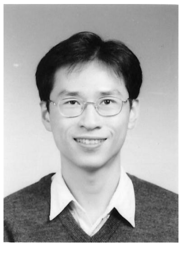
[30] J.-K. Han, J.-G. Yook, and H.-K. Park, "A deterministic channel simulation model for spatially correlated Rayleigh fading," *IEEE Commun. Lett.*, vol. 6, no. 2, pp. 58–60, Feb. 2002.

[31] A. Giorgetti, M. Chiani, M. Shafi, and P. J. Smith, "Level crossing rates and MIMO capacity fades: impacts of spatial/temporal channel correlation," in *Proc. ICC 2003*, vol. 5, Anchorage, Alaska, May 11–15, 2003, pp. 3046–3050.

[32] X. Cai and G. B. Giannakis, "A two-dimensional channel simulation model for shadowing processes," *IEEE Trans. Veh. Technol.*, vol. 52, no. 6, pp. 1558–1567, Nov. 2003.

[33] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350–1359, July 2004.

[34] Y.-R. Tsai and K.-J. Yang, "Available data rate variations for third generation CDMA mobile cellular systems," in *Proc. IEEE Vehicular Technology Conf., Spring*, vol. 5, Milan, Italy, May 17–19, 2004, pp. 2512–2516.



Yuh-Ren Tsai received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1989, and the Ph.D. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1994. From 1994 to 2001, he was a Researcher in Telecommunication Laboratories of Chunghwa Telecom Co., Ltd., Taiwan. Since 2001, he has been with the Department of Electrical Engineering and the Institute of Communications Engineering at National Tsing Hua University, Taiwan, and is currently an Assistant Professor. His research interests include sensor networks, mobile cellular systems, CDMA technology and cryptography.

Yuh-Ren Tsai, and Cheng-Ju Chang, “SIM-based Subscriber Authentication Mechanism for Wireless Local Area Networks,” *Computer Communications* **(Accepted)**

(SCI, EI) (Impact Factor 0.574)

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Computer Communications xx (2006) 1–10

computer communications

www.elsevier.com/locate/comcom

SIM-based subscriber authentication mechanism for wireless local area networks[☆]

Yuh-Ren Tsai [*], Cheng-Ju Chang

Institute of Communications Engineering, National Tsing Hua University 101, Sec. 2, Kuang-Fu Rd., Hsinchu 300, Taiwan ROC

Received 27 July 2004; received in revised form 16 September 2005; accepted 27 September 2005

Abstract

Authentication and roaming are two critical issues for the integration of heterogeneous networks, such as the integration of WLAN access networks and mobile cellular networks. Due to the strong points of mobile cellular networks, it is favorable to integrate the authentication mechanism of WLAN access networks into that of mobile cellular networks. For GSM/GPRS networks, the subscriber identity module (SIM) card is used for user identification, authentication and message encryption. Therefore, it is feasible to authenticate the subscribers in WLAN via exchanging the authentication information between mobile cellular networks and subscribers' SIM cards. In this work, the issue of the subscriber authentication for WLAN access networks is investigated. By integrating the authentication mechanism of WLAN into that of GSM/GPRS networks, a GSM/GPRS SIM-based authentication mechanism for WLAN as well as the related protocols is proposed. Furthermore, the implementation issues are well investigated, and an experimental system is implemented to verify the feasibility of this authentication mechanism.

Keywords: Authentication; Subscriber identity module; Wireless local area network; Mobile cellular network

1. Introduction

For mobile data communications, seamless access and a high transmission rate are highly desirable. Due to the original design objective and limited resources, a single system cannot satisfy all the requirements for mobile data communications. Therefore, the integration of heterogeneous networks becomes an important issue for future mobile data communications. Wireless local area networks (WLANs), such as IEEE 802.11 [1], can offer a maximum available data rate over 11 Mbps, and is suitable for unlicensed indoor wireless data applications. On the other hand, mobile cellular networks can provide comprehensive radio coverage with a limited available data rate. Since these two kinds of networks are complementary to each other in the aspects of available data rates and radio coverage, the integration of WLAN access networks and mobile cellular networks is very attractive for wireless data services [2,3].

Mobile cellular networks, such as the GSM/GPRS (Global System for Mobile communications/General Packet for Radio Service) systems and the 3G systems, have several advantages over WLAN access networks, including numerous existing subscribers, existing IP (Internet Protocol) networks investment, well established subscriber roaming agreements, robust and secure subscriber identification and authentication mechanism, well defined service entries (such as billing systems), and well defined inter-system protocols for subscriber information exchanging and services providing. Therefore, integrating WLAN access networks into mobile cellular networks, such as GSM/GPRS systems, will strongly enhance the capability of WLAN, and is worth studying in detail.

Authentication and roaming are critical issues for the integration of heterogeneous networks. Due to the advantages of mobile cellular networks, it is favorable to integrate the authentication mechanism of WLAN access networks into that of mobile cellular networks. Here we take GSM/GPRS, which is a mature and most popular mobile cellular technology, as an example. The removable user identity module (R-UIM) has been widely used in second-generation mobile cellular networks, and will become a standard function in third-generation

[☆] This work was supported in part by the National Science Council, Taiwan, ROC, under Grant NSC 94-2752-E-007-003-PAE, and the Taiwan Cellular Corp., Taiwan, ROC, under Grant PCSF-91-002.

[*] Corresponding author. Tel.: +886 35715131/62210; fax: +886 35751787.
E-mail address: yrtsai@ee.nthu.edu.tw (Y.-R. Tsai).

0140-3664/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2005.09.016

mobile cellular systems. In GSM/GPRS systems, the R-UIM, referred to as the subscriber identity module (SIM) card, is a smart card, which stores the subscription-related information used for user identification, authentication, and message encryption. Hence, it is feasible to authenticate the GSM/GPRS subscribers in WLAN access networks via exchanging the authentication information between GSM/GPRS networks and subscribers' SIM cards. In [2], the concept of using SIM-based authentication for WLAN access networks was proposed. Some research works also focused on the authentication issue for the integration of GSM and WLAN networks, including the WLAN authentication standard 802.1x and the EAP-SIM protocol [4–6]. In addition, some industry corporations flung into this issue, and developed experimental systems for technical trial [7–9]. Some focused on the issue of the SIM access for an open platform [7], and some are developed under the cooperation with existing cellular operators [8–9]. In this work, a SIM-based subscriber authentication mechanism with detailed protocols for WLAN access networks is proposed. Based on this authentication mechanism, a mobile station (MS) with a GSM/GPRS SIM card can attach to a WLAN access network to obtain services, without the need of pre-subscription to the WLAN access network. Furthermore, the implementation issues are well investigated, and an experimental system is implemented to verify the feasibility of this authentication mechanism.

The remainder of this paper is organized as follows. Section 2 introduces the authentication mechanism of GSM/GPRS systems. Section 3 proposes the SIM-based authentication mechanism and protocols for WLAN access networks. In Section 4 some implementation-related issues are discussed. Finally, the conclusions are drawn in Section 5.

2. GSM/GPRS authentication mechanism

In this section, the authentication mechanism adopted in GSM/GPRS systems is briefly introduced.

2.1. Definitions of authentication-related parameters

Some authentication-related parameters and terminologies in GSM/GPRS systems are introduced as follows: [10–12]

- IMSI (International Mobile Subscriber Identity): Each GSM/GPRS subscriber is uniquely assigned an IMSI, which is stored both in the subscriber's SIM card and in the HLR (Home Location Register) entity of the home GSM/GPRS network. IMSI is used as the unique identity of a subscriber in GSM/GPRS networks, and is composed of three parts: Mobile Country Code (MCC) identifying the country of the subscriber, Mobile Network Code (MNC) identifying the home GSM network of the subscriber, and Mobile Subscriber Identification Number (MSIN) identifying the subscriber within a GSM network. Therefore, a subscriber can be globally identified by the unique IMSI, and the home network can be determined by MCC and MNC.
- Ki (Individual subscriber authentication key): Each GSM/GPRS subscriber is assigned a Ki, which is the base of the authentication mechanism. The Ki is 128 bits in length, and is stored both in the SIM and in the Authentication Center (AuC) of the home GSM/GPRS system.
- Algorithms A3 and A8: A3 and A8 algorithms are implemented both in the SIM and in the GSM AuC. A3 and A8 are all one-way functions, which are used to authenticate the subscriber and to generate the ciphering key Kc, respectively.
- Authentication triplets (RAND, SRES, and Kc): The authentication triplets are essential components for the authentication procedures in GSM/GPRS systems. All the parameters in an authentication triplet are described as follows.
    - □ RAND (Random number): RAND is a random challenge with 128 bits in length. Different values of RAND are used in different authentication procedures.
    - □ SRES (Signed response): SRES, with 32 bits in length, is obtained via A3 algorithm by using RAND and Ki as inputs. SRES is used for subscriber identity verification.
    - □ Kc (Ciphering key): Kc, with 64 bits in length, is obtained via A8 algorithm by using RAND and Ki as inputs. Kc is the session key for message ciphering. In each successful authentication procedure, a new Kc is generated to replace the old one.

2.2. Subscriber authentication procedure for GSM/GPRS systems

For GSM/GPRS networks, the subscriber authentication is based on the individual subscriber authentication key Ki. The GSM/GPRS network will verify whether the subscriber, i.e. the SIM, is with the same Ki as the one stored in the database of GSM AuC or not. The authentication procedure of GSM/GPRS systems is briefly described as follows: [13]

Step 1. When a GSM/GPRS MS is powered on, it will search for an available GSM/GPRS network, and send a registration message, which includes the individual IMSI read out from the SIM, to initiate the authentication procedure.

Step 2. The GSM/GPRS base station system (BSS) pass the IMSI received from the MS to the home HLR/AuC to ask for the authentication triplets (RAND, SRES, Kc) coupled to this IMSI.

Step 3. Upon the received IMSI, the home HLR/AuC can find out the Ki corresponding to the subscriber/SIM. By using the A3 and A8 algorithms, the home HLR/AuC generates several authentication triplets, and sends back to BSS in reply.

Step 4. After the authentication triplets have been received, the BSS selects a triplet and sends the RAND to the MS via the wireless link.

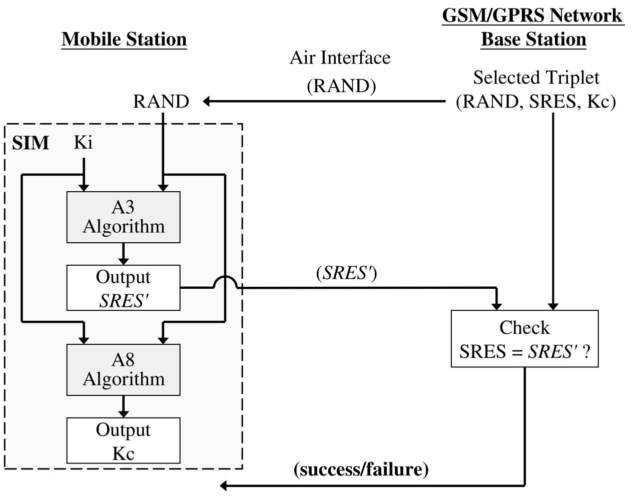


Fig. 1. The authentication procedure of GSM/GPRS systems.

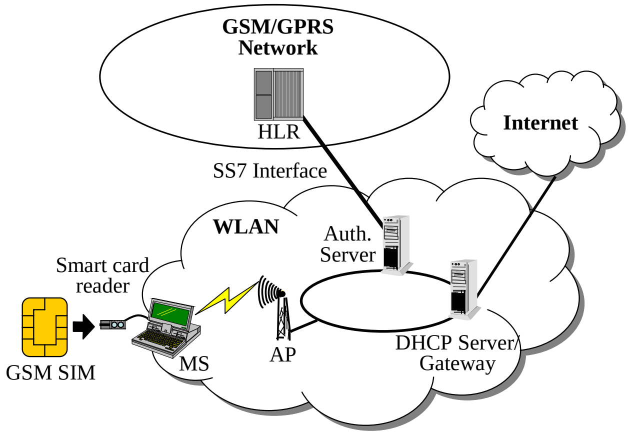


Fig. 2. System architecture for the integration of WLAN and GSM/GPRS networks.

Step 5. Upon the RAND received from BSS and Ki stored in the SIM, the MS uses the A3 and A8 algorithms within SIM to calculate the corresponding SRES and Kc. The MS replies the calculated SRES to manifest that the subscriber couples to the claimed identity IMSI.

Step 6. The BSS verifies the equivalence of the SRES received from the MS and the SRES in the selected triplet. If they are equivalent, the authentication procedure is claimed to be successfully completed; otherwise it is claimed to be failed.

Steps 4–6 are illustrated in Fig. 1. If the authentication procedure is claimed to be successfully completed, the MS can access to the GSM/GPRS network. On the contrary, if it is claimed to be failed, the GSM/GPRS network will deny further request from the MS or re-initiate the authentication procedure.

3. SIM-based authentication for WLAN

In this section, the SIM-based subscriber authentication mechanism for WLAN access networks is presented. It is assumed that the WLAN access networks do not have any subscription information of GSM/GPRS subscribers, including IMSI, Ki and authentication algorithms. However, it is assumed that the WLAN access networks and the GSM/GPRS networks can interoperate and exchange system information via the GSM-MAP (Mobile Application Part) [14] interface based on SS7 (Signaling System No.7) interface [15–19].

3.1. Definition and description of entities

The system architecture is shown in Fig. 2. Some important entities are described as follows.

- *GSM HLR*: HLR is a database in charge of the management of GSM/GPRS subscribers. All subscription-related and location-related information is stored in HLR. HLR can communicate with the AuC to acquire the authentication-related information.
- *SIM*: SIM contains user’s subscription information as well as the A3 and A8 algorithms [20]. As in the WLAN access network, the MS (shown to be a notebook computer) can access the SIM via a smart card reader.
- *WLAN card and access point (AP)*: The WLAN card and AP construct a WLAN access network. The WLAN card, installed in the MS, provides the wireless communication capability connecting to the AP, which offers the physical connection to the IP network.
- *Dynamic Host* Configuration Protocol (DHCP) Server: The DHCP Server provides a framework for passing network configuration information to hosts on a TCP (Transmission Control Protocol)/IP network [21]. The configuration information consists of the IP address, the subnet mask, the DNS (Domain Name System) server address, the gateway address, and the lease time of the IP address, etc.
- *Authentication Server*: The Authentication Server performs the user authentication mechanism in the WLAN access network. In order to acquire the authentication information from the GSM/GPRS network, the Authentication Server possesses an SS7 interface as well as the IP interface. From the viewpoint of GSM/GPRS networks, the Authentication Server is regarded as a VLR (Visitor Location Register) entity in GSM/GPRS networks.
- *Gateway*: The Gateway controls the incoming and outgoing packets of the WLAN access network corresponding to a specific IP address. The packets corresponding to an unauthorized user will be blocked.

In general, the DHCP Server, Authentication Server, and Gateway can be implemented in the same hardware entity or in several separate entities.

3.2. Implementation philosophy

Considering the implementation issue, following subjects are important and should be considered. For the DHCP Server, there are two different policies to allocate the network configuration parameters—static configuration and dynamic configuration. Since the WLAN access networks do not have any subscription information of GSM/GPRS subscribers, the dynamic configuration should be employed. Any user with an arbitrary Ethernet address may request the configuration parameters from the DHCP Server, and the allocated IP address is randomly selected from a set of reserved IP addresses.

For a subscriber newly attaching a WLAN access network, a set of temporary configuration parameters, including a temporary IP address, must be assigned to this client. Accordingly, this client can use this temporary IP address to perform the authentication procedure with the Authentication Server. However, an illegal user may request for a temporary IP address, and then refuse to perform the authentication procedure. Consequently, this illegal user will seize the IP address and deplete the available IP resource. Therefore, the authentication protocols must eliminate this malignant attack, and protect the availability of IP addresses.

Before the identity of a subscriber is successfully verified, the WLAN access network should forbid this subscriber to access Internet or other networks. Hence, the Gateway should block the incoming and outgoing packets corresponding to an unauthenticated user. Furthermore, the interoperation between WLAN access networks and GSM/GPRS networks is essential for this integrated authentication mechanism. However, it is impractical that the implementation of this authentication mechanism needs to modify the existing GSM/GPRS protocols or entities.

3.3. Authentication protocols

According to the above-mentioned subjects, the protocols for the SIM-based authentication mechanism are proposed. The goal of the proposed protocols is to authenticate a GSM/GPRS subscriber in a WLAN access network via the GSM/GPRS SIM card. This goal is achieved by exchanging some information to verify that the client and the GSM/GPRS system have the same secret. Hence, following demands and functionalities must be fulfilled. The MS must be able to acquire the subscriber authentication-related information from the SIM, and a SIM_Access program is implemented in the MS to fulfill this demand. The WLAN access network must be able to send the authentication request to, and to receive the authentication parameters from the GSM/GPRS network. A MAP_Authentication program is implemented in the Authentication Server to fulfill this demand. Furthermore, the MS and the WLAN access network must be able to exchange and verify the authentication-related information. Two programs, Authentication_Client and Authentication_Server, are implemented to fulfill this demand.

The procedure of the proposed SIM-based authentication mechanism is divided into two phases: *Temporary IP Address Acquisition Phase* and *Subscriber Identity Verification Phase*. In *Temporary IP Address Acquisition Phase*, the MS attaches to the WLAN access network, and discovers the DHCP Server to acquire the IP network configuration parameters. Subsequently, in the *Subscriber Identity Verification Phase*, the MS exchanges the authentication information with the WLAN Authentication Server to manifest the subscriber's identity.

Before presenting the authentication protocols, some related messages are defined in Table 1.

The procedure of *Temporary IP Address Acquisition Phase* is shown in Fig. 3, and is presented as follows:

Step 1. When an MS with a GSM/GPRS SIM card attaches to a WLAN access network, it broadcasts a *DHCP_DISCOVER* message, containing its MAC (media access control) address, to find an available DHCP server in this network.

Step 2. Upon the received *DHCP_DISCOVER* message, the DHCP Server responds to the MS with a *DHCP_OFFER* message, which contains an available IP address and other configuration parameters.

Table 1
The definitions of the messages in the SIM-based authentication mechanism

Message	Description
DHCP_DISCOVER	An MS broadcasts to locate the available DHCP servers.
DHCP_OFFER	In response to the *DHCPDISCOVER* message, a DHCP Server sends to an MS with the offer of the configuration parameters.
DHCP_REQUEST	An MS broadcasts to all DHCP servers to request the offered configuration parameters from one server, and to implicitly decline the offers from all others.
DHCP_ACK	The chosen DHCP Server sends to an MS with the acceptance of the request of the configuration parameters.
DHCP_NAK	The chosen DHCP Server sends to an MS with the rejection of the request of the configuration parameters.
AUTH_Notification	The chosen DHCP Server sends to the Authentication Server to inform the enrollment of a new client.
Registration	An MS sends to the Authentication Server to initiate the authentication procedure.
MAP_SEND_AUTH_ INFO (Request)	The Authentication Server sends to the GSM HLR to invoke the authentication triplets.
MAP_SEND_AUTH_INFO (Response)	The GSM HLR sends to the Authentication Server in reply to the request of the authentication triplets.
AUTH_Request	The Authentication Server sends to an MS to challenge the subscriber's identity.
AUTH_Response	An MS sends to the Authentication Server in reply to the challenge from the Authentication Server.
AUTH_Confirmation	The Authentication Server sends to an MS to inform the result of the authentication procedure.
AUTH_Verdict	The Authentication Server sends to the DHCP Server/Gateway to inform the result of the authentication procedure.

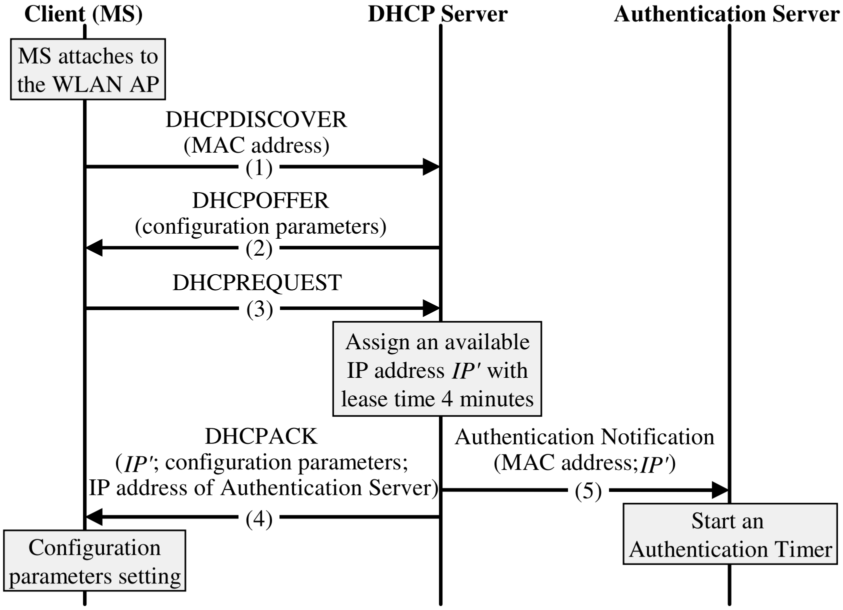

Client (MS)
DHCP Server
Authentication Server
MS attaches to the WLAN AP
DHCPDISCOVER
(MAC address)
(1)
DHCPOFFER
(configuration parameters)
(2)
DHCPREQUEST
(3)
Assign an available IP address IP′ with lease time 4 minutes
DHCPACK
(IP′; configuration parameters; IP address of Authentication Server)
(4)
Authentication Notification
(MAC address; IP′)
(5)
Start an Authentication Timer
Configuration parameters setting

Fig. 3. The procedure of Temporary IP Address Acquisition Phase.

Step 3. If the MS decides to adopt the configuration parameters offered by the DHCP Server, it broadcasts a *DHCP_REQUEST* message to the DHCP Server.

Step 4. If the DHCP Server accepts the *DHCP_REQUEST* message, a temporary IP address, denoted as IP', with a very short lease time, say 4 min, is allocated to this MS. Then the DHCP Server replies to the MS with a *DHCP_ACK* message, containing the configuration parameters and the IP address of the Authentication Server. If the DHCP Server cannot accept the request, a *DHCP_NAK* message will be sent to the MS.

Step 5. In addition, the DHCP Server forwards an *AUTH_-Notification* message, containing the MAC address of the MS and the assigned IP', to inform the Authentication Server. After being informed by the DHCP Server, the Authentication Server starts and maintains an Authentication Timer, corresponding to this MAC address, to prevent a malignant request. If the Authentication Timer expires before the authentication procedure is successfully completed, the Authentication Server claims a failure of authentication, and then informs the DHCP Server to retrieve the assigned IP address.

After the *DHCP_ACK* message is received, the MS configures its network configuration parameters, and originates the procedure of *Subscriber Identity Verification Phase*. On the other hand, if a *DHCP_NAK* message is received, the MS may re-initiate the procedure of *Temporary IP Address Acquisition Phase*. The procedure of *Subscriber Identity Verification Phase* is shown in Fig. 4, and is presented as follows:

Step 1. The MS performs the SIM_Access program to activate the SIM and to read out the IMSI via a smart card reader. By performing the Authentication_Client program, the MS sends a *Registration* message, containing the IMSI and its MAC address, to the Authentication Server to initiate the authentication procedure. It is noted that the incoming and outgoing packets corresponding to IP' is currently blocked by the Gateway. Therefore, the MS cannot use IP' to communicate with other networks in current status.

Step 2. Upon the received *Registration* message, the Authentication Server determines the subscriber's home GSM/GPRS network, which can be identified by the received IMSI. Then the Authentication Server performs the MAP_Authentication program to send the *MAP_SEND_AUTH_INFO* (Request) message to the GSM HLR via the SS7 interface. This message contains the received IMSI, and is used to invoke the authentication triplets (RAND, SRES, Kc) coupled to this IMSI. It is noted that this message can be regarded as the registration of this subscriber in the GSM/GPRS network. Therefore, the HLR will store the location information, which indicates that the subscriber is under the coverage of the WLAN access network.

Step 3. According to the received IMSI, the GSM HLR/AuC searches out the Subscriber Authentication Key Ki, and calculates multiple sets of corresponding authentication triplet. Then the GSM HLR/AuC replies to the Authentication Server with the *MAP_SEND_AUTH_INFO* (Response) message, containing the authentication triplets.

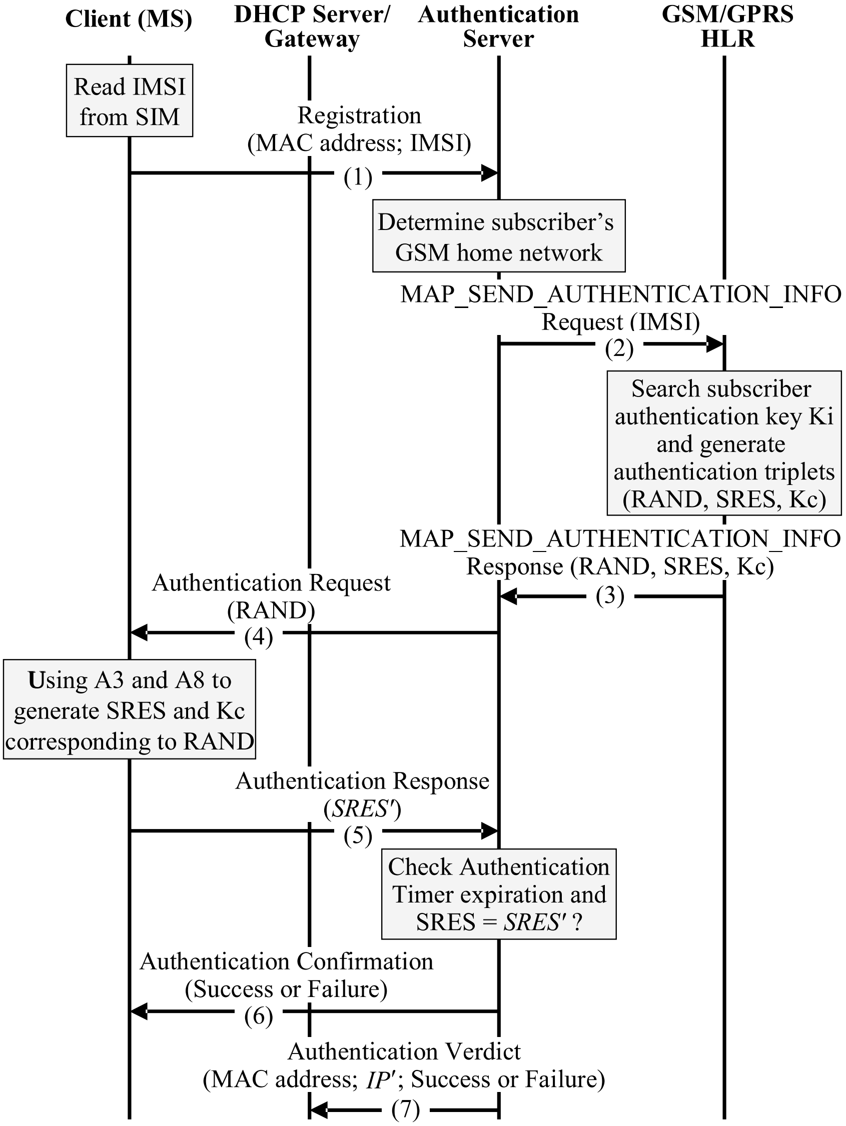

Client (MS)
DHCP Server/ Gateway
Authentication Server
GSM/GPRS HLR
Read IMSI from SIM
Registration (MAC address; IMSI)
(1)
Determine subscriber's GSM home network
MAP_SEND_AUTHENTICATION_INFO Request (IMSI)
(2)
Search subscriber authentication key Ki and generate authentication triplets (RAND, SRES, Kc)
MAP_SEND_AUTHENTICATION_INFO Response (RAND, SRES, Kc)
(3)
Authentication Request (RAND)
(4)
Using A3 and A8 to generate SRES and Kc corresponding to RAND
Authentication Response (SRES')
(5)
Check Authentication Timer expiration and SRES = SRES' ?
Authentication Confirmation (Success or Failure)
(6)
Authentication Verdict (MAC address; IP'; Success or Failure)
(7)

Fig. 4. The procedure of Subscriber Identity Verification Phase.

Step 4. After the authentication triplets have been received, the Authentication Server performs the Authentication_Server program, and randomly selects a set of triplet for the present run of authentication. Subsequently, by using IP' as destination, the Authentication Server sends the *AUTH_Request* message, containing the selected RAND, to challenge the MS.

Step 5. Based on the received RAND, the MS activates the SIM and performs SIM_Access program to calculate the SRES and Kc, by using Ki and the A3 and A8 algorithms within the SIM. Then the MS sends the *AUTH_Response* message, containing the calculated signed response $SRES'$, in reply to the challenge from the Authentication Server.

Step 6. The Authentication Server verifies the equivalence of the $SRES'$ received from the MS and the SRES in the selected triplet. According to the result of comparison, one of the following two verdicts is claimed:

- *Authentication procedure successfully completed*: If these two values are equivalent and the Authentication Timer does not expire, the authentication procedure is claimed to be successfully completed. Then the Authentication Server sends the *AUTH_Confirmation* message with the indication of success to the MS.
- A*uthentication procedure failed*: If these two values are not equivalent or the Authentication Timer expires, the authentication procedure is claimed to be failed. Then the *AUTH_Confirmation* message with the indication of failure is sent to the MS.

Step 7. Furthermore, the Authentication Server sends the *AUTH_Verdict* message, containing the MAC address of the MS and IP', to inform the DHCP

Server/Gateway of the authentication verdict. According to the authentication verdict, the DHCP Server/Gateway takes one of the following two actions:

- *Authentication procedure successfully completed*: This event is recorded in the DHCP Server, and the restriction of IP' is revoked on the Gateway. Consequently, the MS can communicate with other networks by using IP'. In addition, when the lease time of IP' expires, the MS may send a new *DHCP_REQUEST* message to the DHCP Server requesting for a new lease, and the DHCP Server will assign a long lease time for this MS.
- *Authentication procedure failed*: The DHCP Server records this authentication failure event, and starts a Penalty Timer. The packets corresponding to IP' are still blocked by the Gateway. When the lease time of IP' expires, the MS may send a new *DHCP_REQUEST* message to the DHCP Server requesting for a new lease. According to the MAC address, the DHCP Server will check the database and refuse/ignore any new request until the Penalty Timer expires.

4. Implementation

Based on the protocols proposed in Section 3, we have implemented an experimental system. The implementation of this authentication mechanism consists of three parts:

(1) SIM Authentication Information Access Part,
(2) WLAN Client/Server Part, and
(3) GSM-MAP Information Request Interface.

The details are described as follows.

4.1. SIM Authentication information access part

The GSM/GPRS SIM is a smart card with the physical characteristics specified in [22–25]. The SIM supports the use of CHV (Card Holder Verification) to authenticate the user, and therefore provides well protection against the illegal use of stolen cards. There are two CHV values, denoted as CHV1 and CHV2, stored in the SIM, and each CHV is formed by 4–8 decimal digits. The SIM_Access program is implemented to acquire the authentication-related information from the SIM, including the user interface and the SIM access interface. By using a smart card reader, this program can verify the authorization of the user, acquire the GSM/GPRS subscriber identity (IMSI), and input the RAND parameter to obtain the corresponding authentication information (SRES and Kc).

Shown in Fig. 5 is the user interface by which the user can connect or disconnect the SIM card and input the CHV1. Furthermore, the acquired IMSI, RAND, SRES, and Kc are also shown in this user interface. Fig. 6 shows the flowchart of this part. Firstly, the SIM_Access program activates the SIM, and reads out the CHV enabled/disabled indicator. If the CHV is enabled, the user interface requests the client to input the CHV1, and sends the input into the SIM for user verification; otherwise, the SIM_Access program can access the SIM immediately. Afterward, the SIM_Access program reads out the IMSI, and the Authentication_Client program of *WLAN Client/Server Part* will deliver the IMSI to the Authentication Server by using IP'. When the authentication challenge RAND is received, the SIM_Access program requests the execution of A3 and A8 algorithms to obtain the corresponding $SRES'$ and Kc. Finally, the SIM_Access program deactivates the SIM.

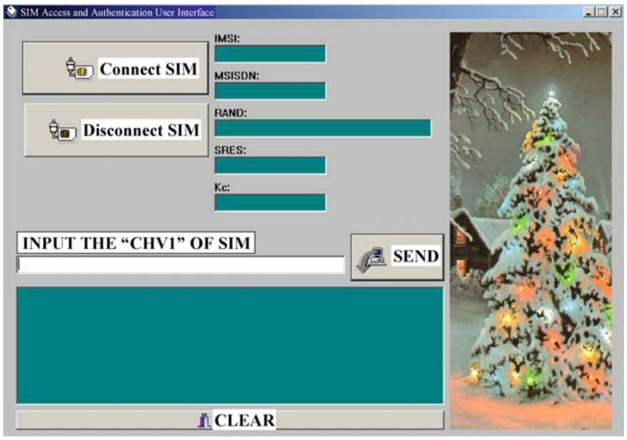


Fig. 5. User interface of SIM Authentication Information Access Part.

4.2. WLAN client/server part

This part includes four programs: *Parameter_Assignment*, *Authentication_Client*, *Authentication_Server*, and *Packet_-Filtering*. The DHCP protocols are well defined for the IP address request services. However, in this work,

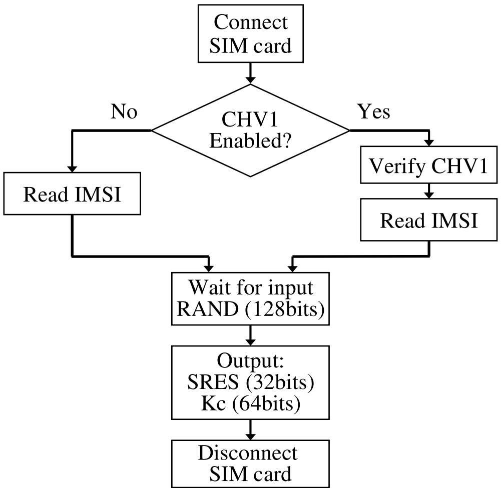
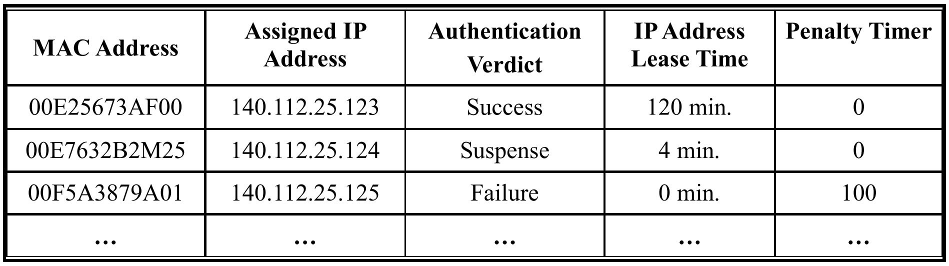


Fig. 6. Flowchart of SIM Authentication Information Access Part.

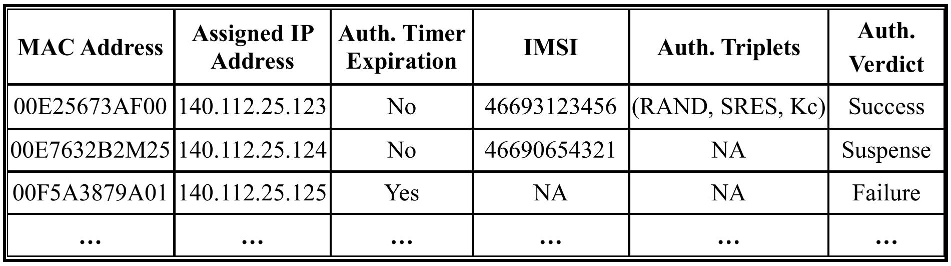


MAC Address	Assigned IP Address	Authentication Verdict	IP Address Lease Time	Penalty Timer
00E25673AF00	140.112.25.123	Success	120 min.	0
00E7632B2M25	140.112.25.124	Suspense	4 min.	0
00F5A3879A01	140.112.25.125	Failure	0 min.	100
…	…	…	…	…

Fig. 7. Database of DHCP Server.

Table 2
The parameters stored in the database of the DHCP Server

Parameter	Description
MAC address	The MAC address of the attaching WLAN card used to represent the identity of the MS.
Assigned IP address	The IP address assigned to the MAC address, i.e. IP'.
Authentication verdict	The verdict of the authentication procedure informed by the Authentication Server. There are three possible states in this field—success, failure, and suspense, i.e. the authentication procedure is not finished yet.
IP address lease Time	The lease time of the assigned IP address.
Penalty timer	The time duration before the expiration of the Penalty Timer. If the Penalty Timer is not zero, the request from the corresponding MAC address will be ignored.



MAC Address	Assigned IP Address	Auth. Timer Expiration	IMSI	Auth. Triplets	Auth. Verdict
00E25673AF00	140.112.25.123	No	46693123456	(RAND, SRES, Kc)	Success
00E7632B2M25	140.112.25.124	No	46690654321	NA	Suspense
00F5A3879A01	140.112.25.125	Yes	NA	NA	Failure
…	…	…	…	…	…

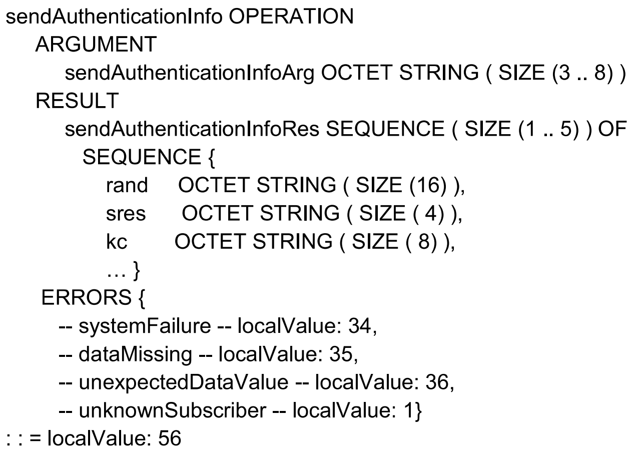
Fig. 8. Database of Authentication Server.

the DHCP protocols require some modifications. The *Parameter_Assignment* program, implemented in the DHCP Server, is used to support the assignment of network configuration parameters for a newly attaching MS. This program follows the procedure of *Temporary IP Address Acquisition Phase* as shown in Fig. 3. Furthermore, a database regarding each attaching subscriber is maintained. As shown in Fig. 7, the database of the DHCP Server consists of the following fields (Table 2).

The *Authentication_Client* and *Authentication_Server* programs are implemented to support the messages exchanging between the Authentication Server and the MS, and to determine the success or failure of the authentication procedure. The procedure of *Subscriber Identity Verification Phase*, as shown in Fig. 4, is followed. Furthermore, a database regarding each attaching subscriber is maintained. As shown in Fig. 8, the database of the Authentication Server consists of following fields (Table 3).

Table 3
The parameters stored in the database of the Authentication Server

Parameter	Description
MAC address	The MAC address of the attaching WLAN card used to represent the identity of the MS.
Assigned IP address	The IP address assigned to the MAC address, i.e. IP'.
Authentication timer expiration	The status of the Authentication Timer. If the timer expires, the Authentication Verdict is claimed to be failed.
IMSI	The GSM/GPRS subscriber identification claimed by the MS.
Authentication triplets	The authentication information obtained from the GSM/GPRS network. The Authentication Server may store multiple sets of authentication triplet.
Authentication verdict	The verdict of the authentication procedure—success, failure or suspense.



```
sendAuthenticationInfo OPERATION
   ARGUMENT
      sendAuthenticationInfoArg OCTET STRING ( SIZE (3 .. 8) )
   RESULT
      sendAuthenticationInfoRes SEQUENCE ( SIZE (1 .. 5) ) OF
        SEQUENCE {
           rand     OCTET STRING ( SIZE (16) ),
           sres      OCTET STRING ( SIZE ( 4) ),
           kc       OCTET STRING ( SIZE ( 8) ),
           ... }
    ERRORS {
      -- systemFailure -- localValue: 34,
      -- dataMissing -- localValue: 35,
      -- unexpectedDataValue -- localValue: 36,
      -- unknownSubscriber -- localValue: 1}
: : = localValue: 56
```

Fig. 9. Representation of *MAP_SEND_AUTH_INFO* service.

For an unauthenticated MS, the corresponding packets should be blocked. Therefore, the *Packet_Filtering* program is implemented in the Gateway to filter the passing packets. If the status of the Authentication Verdict is not ‘success’, all the packets corresponding to this restricted IP address will be blocked by the Gateway.

4.3. GSM-MAP information request interface

The GSM-MAP, specified by ETSI (European Telecommunication Standards Institute), is the standard protocols for GSM/GPRS core networks [14]. The GSM-MAP service *MAP_SEND_AUTH_INFO* is used for a GSM VLR to acquire the subscriber information from the GSM HLR. In the Authentication Server, a GSM-MAP interface is implemented, which is based on the SS7 interface and connects to the GSM HLR. The MAP_Authentication program is implemented to support this service, and thus the Authentication Server can acquire the authentication triplets from the GSM/GPRS network. The GSM-MAP protocols are abstractly represented by the ASN.1 (Abstract Syntax Notation One) syntax [26–29]. The representation of the *MAP_SEND_AUTH_INFO* service is shown in Fig. 9, and the details of this service are described as follows (Table 4).

Table 4
The definitions of the parameters in the *MAP_SEND_AUTH_INFO* service

Parameter	Description
ARGUMENT	The ‘sendAuthenticationInfoArg’, sent from the VLR to the HLR, is an octet string with the size ranges from 3 to 8 octets. The content of this string is the value of the IMSI corresponding to the requesting subscriber.
RESULT	The ‘sendAuthenticationInfoRes’ sequence, sent from the HLR to the VLR, consists of RAND, SRES and Kc. The size of ‘sendAuthenticationInfoRes’ ranges from 1 to 5.
ERRORS	If the service cannot be successfully completed, an error code is sent from the HLR to the VLR. Several error codes representing the reasons of incompletion are defined in the protocols.

4.4. Examples

Assume that there are three clients MS1, MS2 and MS3, attaching to the WLAN access network with the corresponding MAC addresses being ‘00E25673AF00’, ‘00E7632B2M25’ and ‘00F5A3879A01’, respectively. According to the proposed protocols, these clients perform the steps presented in *Temporary IP Address Acquisition Phase* to acquire the network configuration parameters. As shown in Figs. 7 and 8, these three clients are in different statuses. Assume that MS1 has successfully completed the authentication procedure. Therefore, in the database of the Authentication Server, the field of Authentication Timer Expiration is set ‘No’; the IMSI and authentication triplets corresponding to MS1 are stored; and the field of Authentication Verdict is ‘Success’. Similarly, in the database of the DHCP Server, the field of Authentication Verdict is ‘Success’; the IP Address Lease Time is set a very long time interval, e.g. 120 min; and the field of Penalty Timer is zero.

For MS2, it is assumed that the authentication procedure is currently in the Step 3 of *Subscriber Identity Verification Phase*. So, in the database of Authentication Server, the field of Authentication Timer Expiration is ‘No’; the IMSI corresponding to MS2 is stored; the authentication triplets are not available yet; and the field of Authentication Verdict is ‘Suspense’. In the database of the DHCP Server, the field of Authentication Verdict is ‘Suspense’; the IP Address Lease Time is still 4 min; and the field of Penalty Timer is zero.

For a malicious client MS3, it is assumed that the *Temporary IP Address Acquisition Phase* is completed; yet it does not send the *Registration* message to the Authentication Server. After the Authentication Timer expires, the authentication procedure is claimed to be failed. Consequently, in the database of the Authentication Server, the field of Authentication Timer Expiration is ‘Yes’; the IMSI and authentication triplets corresponding to MS3 are not available; and the field of Authentication Verdict is ‘Failure’. In the database of the DHCP Server, the field of Authentication Verdict is ‘Failure’; the field of IP Address Lease Time is zero; and the field of Penalty Timer is set a very large value, e.g. 100 min.

5. Conclusions

In this work, we have proposed a GSM/GPRS SIM-based authentication mechanism for WLAN access networks. In addition, an experimental system used to verify the feasibility of this authentication mechanism is implemented. The proposed authentication mechanism has the following advantages. There is a considerable business potential for WLAN access networks, since there are numerous existing subscribers in GSM/GPRS networks. New subscription is unnecessary for GSM/GPRS subscribers to have services in the WLAN access networks. It is possible to reuse the existing GSM/GPRS billing system by sending the accounting information to GSM/GPRS networks via the SS7 interface. The proposed mechanism is suitable for all subscribers from various GSM/GPRS networks, and can also be applied to other mobile cellular systems with the functionality of R-UIM, such as 3G networks.

Nevertheless, there are still some important issues should be studied further.

- To access the information from SIM cards, a smart card reader is required. In the future, it is favorable to integrate a smart card reader into a WLAN card or a notebook computer in order to reduce the cost of hardware. Another possible yet complex approach is to permit the access of the SIM via a mobile handset. Accordingly, the modifications on the mobile handset standards are essential.
- Currently, the SIM should be removed from the GSM/GPRS mobile handset in order to access the SIM via a smart card reader. Consequently, the GSM/GPRS services are interrupted. However, after the authentication procedure is successfully completed, the GSM HLR records that the subscriber is under the coverage of the WLAN access network. Then the GSM/GPRS-related services may be provided via the WLAN access network. If the voice service is the major concern, the VoIP (Voice over IP) technology will be a critical issue.
- To acquire the authentication information from GSM/GPRS networks, an SS7 interface, including an SS7 interface card and a dedicated lease-line, must be provided in the Authentication Server. This implies that considerable cost and complexity is taken. However, SS7 over IP, an emerging technology in telecommunications, can greatly reduce the cost and complexity, and should be adopted in the future.

Furthermore, the session key Kc, obtained in the authentication procedure, can be applied to message ciphering or other applications in WLAN access networks.

References

[1] IEEE Standard 801.11b, IEEE standard for wireless LAN medium access control (MAC) and physical layer (PHY) specifications, Jan. 2000.

[2] J. Ala-Laurila, J. Mikkonen, J. Rinnemaa, Wireless LAN access network architecture for mobile operators, IEEE Communication Magazine November (2001) 82–89.

[3] K. Ahmavaara, H. Haverinen, R. Pichna, Interworking architecture between 3GPP and WLAN systems, IEEE Communication Magazine November (2003) 74–81.

[4] H. Wang, A.R. Prasad, P. Schoo, K.M. Bayarou, S. Rohr, Security mechanisms and security analysis: hotspot WLANs and inter-operator roaming, in Proceedings of IEEE Vehicular Technology Conference, VTC 2004-Spring, 17–19, May 2004, pp. 2492–2496.

[5] J.-S. Leu, R.-H. Lai, H.-I Lin, W.-K. Shih, Practical considerations on end-to-end cellular/PWLAN architecture in support of bilateral roaming, in Proceedings of IEEE Wireless Communications and Networking Conference, 13–17, March 2005, pp. 1702–1707.

[6] R. Huber, N. Jordan, An Experimental Study of a Business Domain Independent Application Level and Internet Access Authentication and Authorization Concept, in Proceedings of International Conference on Mobile Business, ICMB, 11–13, July 2005, pp. 35–41.

[7] A. Ahmad, R. Chandler, A.A. Dharmadhikari, U. Sengupta, SIM-based WLAN authentication for open platforms, Technology@Intel Magazine, August 2003.

[8] http://www.ericsson.com/press/20040223-132604.html

[9] http://www.tatarasystems.com/contentmgr/showdetails.php/id/305

[10] European Telecommunication Standard, GSM 03.03: Digital cellular telecommunications system (Phase 2+); Numbering, addressing and identification.

[11] European Telecommunication Standard, GSM 03.08: Digital cellular telecommunications system (Phase 2+); Organization of subscriber data.

[12] European Telecommunication Standard, GSM 02.09: Digital cellular telecommunications system (Phase 2+); Security aspects.

[13] European Telecommunication Standard, GSM 03.20: Digital cellular telecommunications system (Phase 2+); Security related network functions.

[14] European Telecommunication Standard, GSM 09.02: Digital cellular telecommunications system (Phase 2+); Mobile Application Part (MAP) specification.

[15] ITU-T Q.700, Introduction to CCITT Signaling System No. 7, March 1993.

[16] ITU-T Q.771, Specifications of Signaling System No. 7—Transaction capabilities application part; Functional description of transaction capabilities, June 1997.

[17] ITU-T Q.772, Specifications of Signaling System No. 7—Transaction capabilities application part; Transaction capabilities information element definitions, June 1997.

[18] ITU-T Q.773, Specifications of Signaling System No. 7—Transaction capabilities application part; Transaction capabilities formats and encoding, June 1997.

[19] Travis Russell, Signaling System no. 7, 4th ed. 2002.

[20] European Telecommunication Standard, GSM 11.11: Digital cellular telecommunications system (Phase 2+); Specification of the Subscriber Identity Module-Mobile Equipment (SIM-ME) interface.

[21] RFC 1541 of IETF, Dynamic Host Configuration Protocol, Oct. 1993.

[22] ISO/IEC 7816-1, Identification cards—Integrated circuit(s) cards with contacts—Part 1: Physical characteristics, 1998.

[23] ISO/IEC 7816-2, Information technology—Identification cards—Integrated circuit(s) cards with contacts—Part 2: Dimensions and location of the contacts, 1999.

[24] ISO/IEC 7816-3, Information technology—Identification cards—Integrated circuit(s) cards with contacts—Part 3: Electronic signals and transmission protocols, 1997.

[25] European Telecommunication Standard, GSM 02.17: Digital cellular telecommunications system (Phase 2+); Subscriber Identity Modules (SIM); Functional characteristics.

[26] ITU-T X.208, Open systems interconnection model and notation; Specification of AbstractSyntax Notation One (ASN.1), Dec. 1997.

[27] ITU-T X.680, OSI networking and system aspects—Abstract Syntax Notation One (ASN.1) Information technology—Abstract Syntax Notation One (ASN.1): Specification of basic notation, Dec. 1997.

[28] John Larmouth, ASN.1 Complete, 1999.

[29] Olivier Dubuisson, ASN.1-Communication with Heterogeneous Systems, June 2000.

Yuh-Ren Tsai, and Kai-Jie Yang, “Coverage Shrinking and Available Data Rate Variations for 3G CDMA Mobile Cellular Systems,”

IEICE Transactions on Communications, vol. E89-B, no. 3, pp. 739-747, March 2006.

(SCI, EI) (Impact Factor 0.330)

PAPER

Coverage Shrinking and Available Data Rate Variations for 3G CDMA Mobile Cellular Systems

Yuh-Ren TSAI[†*a)] *and* Kai-Jie YANG[†], *Nonmembers*

SUMMARY In 3G CDMA mobile communication systems, high data rate services are essential for many key applications. When an MS approaches the cell border, link performance is degraded and more power should be allocated to maintain the link performance. Since the maximum available signal power is limited, the link adaptation mechanism may diminish the data rate to maintain link performance. This implies that the valid coverage shrinks when the data rate increases. The shrinking of valid coverage under a predetermined data rate will strongly impact on the reliability of high data rate services. In this work, the encoded bit error probabilities of 3G CDMA mobile communication systems, over large-scale and large-small-scale fading channels, were analyzed based on SGA and SIGA methods. Analytic methods were also proposed to investigate the issues of coverage shrinking and service data rate variations. Furthermore, the outage probability, cell coverage percentage and the staying probabilities of available data rates were well examined. The proposed analytic methods can be applied, as a preliminary research, to the design of cellular-system-related techniques, such as QoS control, available data rate prediction, power reservation, and service adaptation.

key words: *Mobile Cellular, Cell Coverage, Code Division Multiple Access, Outage Probability, Link Adaptation*

1. Introduction

Code division multiple access (CDMA) technology, based on direct-sequence spread spectrum (DSSS) technique [1]–[2], has become the main stream of third-generation (3G) mobile communication systems, including cdma2000 [3], UMTS [4], and LAS-CDMA [5]. Since a DSSS-CDMA system is an interference-limited system, the system capacity and performance are no longer explicit, and strongly depend on interferences, including the multiple access interference (MAI) from the serving cell and co-channel interferences (CCI) from the adjacent cells. To evaluate the performance of a DSSS-CDMA system, the carrier to interference ratio (CIR) and the processing gain are the major concerns. CIR is defined as the ratio of the received desired signal power to the total interference power; while the processing gain is defined as the ratio of the spread code chip rate to the information rate.

In mobile radio environments, the received signal power depends on the path loss, shadowing fading, and fast multipath fading. Seeing the uncertainty of the radio propagation channel, link adaptation techniques have been proposed to prevail over channel conditions, and to maximize the transmission data rate and spectral efficiency. In 3G CDMA systems, such as cdma2000 and UMTS, the link adaptation mechanism adapts the spreading factor, the transmission power, and the code rate to overcome channel conditions and co-channel interferences [3]–[4]. Furthermore, due to the complexity of performance analysis, some approximation methods were proposed to approximate the exact results, including standard Gaussian approximation (SGA) [6]–[7], improved Gaussian approximation (IGA) [8] and simplified improved Gaussian approximation (SIGA) [9]–[10] methods. In recent years, some researches have worked on extensions and applications of these methods. Generalization of SGA for band-limited DS-CDMA systems was proposed in [11]–[13], and the results of IGA and SIGA cases were proposed as well in [14]–[19].

Since the occupied transmission bandwidth is held to be invariable for different data rates, the variations in transmission data rates are accomplished by applying different spreading factors. When a high data rate is desired, the increase in the information rate will depress the processing gain and encoded bit energy. The term 'encoded bits' denotes the bit stream after transmitter's channel coding and before receiver's channel decoding. If a predetermined link performance, i.e., a encoded bit error rate (BER) or a frame error rate (FER), is desired, a corresponding encoded bit energy-to-total interference power density ratio should be maintained to sustain the radio link. When a mobile station (MS) is near the serving base station (BS), the desired signal experiences a small amount of total CCI and a less propagation loss; thus a good CIR can be obtained. On the contrary, when an MS approaches the cell border, link performance is degraded and more power should be allocated to maintain the link performance. Since the maximum available signal power is limited, the link adaptation mechanism may readjust the spreading factor, i.e., diminish the information rate, to maintain

Manuscript received January ??, 2005.
Manuscript revised June ??, 2005.
Final manuscript received November ??, 2005.

[†]The authors are with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan.

[*]This paper was presented in part at the IEEE 59th Semiannual Vehicular Technology Conference, Milan, Italy, May 2004. This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grants NSC 94-2752-E-007-003-PAE and NSC 94-2213-E-007-008.

a) E-mail: yrtsai@ee.nthu.edu.tw

link performance. This implies that the valid coverage shrinks when the information rate increases.

The shrinking of valid coverage strongly impacts on service reliability, and thus service adaptation mechanisms should be employed to assure the service continuity. In this viewpoint, it is interesting to investigate the issues of coverage shrinking and data rate variations. Some research works related to these issues can be found in literature [20]–[23]. Especially in [20] and [21], the impact on cell coverage both in forward and reverse links was investigated; however, the results were obtained via Monte Carlo simulations, and no analytic method was proposed. Furthermore, the adopted system was a second generation CDMA system, i.e., IS-95B [24], having multiple code channels assigned to increase the transmission data rate.

In the literature SGA and SIGA methods were proposed as analytic tools for the performance evaluation of CDMA systems. These research works were generally focused on the link performance analysis. However, some characteristics of cellular systems, such as the combination of multiple shadowed interferences, are not considered; and some interesting issues, such as valid cell coverage and available data rate variations, are not studied. In addition, even though several research works focused on valid cell coverage and available data rate variations issues for 2G or 3G CDMA systems, the results were generally obtained via Monte Carlo simulations, and no analytic methods were proposed. Therefore the main contribution of this work is to provide a simple and systematic way to investigate the above-mentioned issues for 3G CDMA systems. As analytic tools, SGA and SIGA approaches are applied in the performance analysis in order to simplify and obtain the analytic results. Considering the attributes of high data rate services (e.g. web browsing, file downloading and video streaming services), we focus only on the downlink. The remainder of this paper is organized as follows: Section 2 illustrates the system and channel models; Section 3 concentrates on the analytic methods; in Section 4, the numerical results are presented; and finally, the conclusions are offered in Section 5.

2. System and Channel Models

In this section, we briefly introduce the system and channel models adopted in this work.

2.1 Channel Model

The channel is assumed to be with large- and small-scale fading effects. Large-scale fading includes the path loss and log-normal shadowing, whereas the small-scale fading focuses on fast multipath fading effects. The channel impulse response can be modeled as

$$h_k\left(t\right) = L_k A_k \mathrm{e}^{j\beta_k} \delta\left(t-\tau_k\right), \quad k=0,\cdots,m, \qquad (1)$$

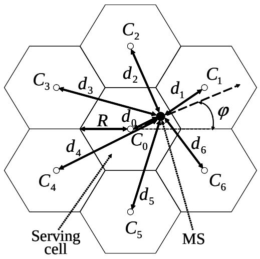


Fig. 1 The system model for mobile cellular CDMA system.

where the subscript k indicates different propagation channels; $\{L_k\}_{k=0}^m$ and $\{A_k\}_{k=0}^m$ are random variables (RVs) accounting for the large- and small-scale fading, respectively; $\{\beta_k\}_{k=0}^m$, assumed to be uniformly distributed within $[0,2\pi)$, represent the random phases introduced by the fading channel; and $\{\tau_k\}_{k=0}^m$ denote the transmission delays corresponding to different propagation paths. The fast fading is assumed to be Rayleigh fading.

The large-scale fading effect can be represented as

$$L_k = \sqrt{PL_{ref}\ d_k^{-n}\chi_k}\ , \qquad (2)$$

where PL_{ref} is a reference value depending on the propagation environment; $\{d_k\}_{k=0}^m$ denote the propagation distances; n is the path loss exponent ranging from 3 to 4; and $\{\chi_k\}_{k=0}^m$ are i.i.d. (independently and identically distributed) log-normally distributed RVs, representing shadowing effects, with a zero mean and the variance σ_χ^2 in decibel (dB) units. The distribution of χ_k is expressed as

$$f_{\chi_k}(x_k)=\left(\frac{1}{x_k\xi\sqrt{2\pi}\sigma_\chi}\right)\exp\left[\frac{-(10\log_{10}\chi_k)^2}{2\sigma_\chi^2}\right], \quad (3)$$

where $\xi = \ln(10)/10$. For macrocellular applications, σ_χ is in the range of 5 to 12 dB, and a typical value of 8 dB is applied. In addition, we assume that the propagation channel can be resolved into j multipaths, having relative delays over one chip and equal average power in each path.

2.2 System Model

The system is assumed to have uniform grid of hexagonal cells. To simplify the calculations, only the serving BS, denoted as C_0, and the six nearest co-channel BSs, denoted as $\{C_1,\cdots,C_6\}$, are considered. As shown in Fig. 1, R denotes the cell radius and d_k denotes the distance between C_k and the MS. By geometric calculations, $\{d_k\}_{k=1}^6$ can be represented as functions of $d_0 \equiv r$. The azimuth φ, representing the traveling direction of the MS, is assumed to be uniformly distributed within $[-\pi/6,\pi/6)$. For a case with $\varphi \geq \pi/6$ or $\varphi < -\pi/6$, it can be recast to the case with $-\pi/6 \leq \varphi < \pi/6$. The total transmission power of each BS is assumed to be P_{tot}. The power allocated to a desired user corresponds to a power ratio U, which is defined as the ratio of P_{tot} to the desired signal power. Furthermore, to determine the relationship between P_{tot} and the additive white Gaussian noise (AWGN) spectral density N_0, we introduce a noise-raised factor α, which is defined as the ratio of average total received power from the serving BS to the received AWGN power at the cell boundary, i.e.,

$$\alpha \equiv \frac{P_{tot} \times PL_{ref} \times R^{-n}}{N_0/T_c} \Rightarrow P_{tot} = \frac{\alpha\,(N_0/T_c)}{PL_{ref} \times R^{-n}}\ , \quad (4)$$

where T_c is the chip duration of the spreading sequence.

3. Performance Analysis

According to (1) and (2), for a specific MS with a distance r from the serving BS, the received desired signal power from a specific path is

$$P_0 = L_0^2 A_D^2 P_{tot}/jU = PL_{ref}\, r^{-n} \chi_0 A_D^2 \times P_{tot}/jU, \quad (5)$$

where A_D is the Rayleigh fading gain of this path, and the signal power is assumed to be equally divided into j propagation paths. In a CDMA cellular system, all downlink channels from a BS are well synchronized and are spread by orthogonal channelization codes; hence, the orthogonality between different channels can be maintained. However, due to the multipath propagation phenomenon, the orthogonality is destroyed and the reception of a specific path is interfered by other $j-1$ paths. If the path number is $j=1$, there will be no MAI from the serving BS. Assuming that each multipath possesses equal average power, the MAI power from the serving BS becomes $(1-1/j)\,P_{tot}$. Furthermore, the interference from each co-channel BS is simplified to be a signal source with transmission power P_{tot}. Therefore, the total received interference power, at the receiver front end, can be expressed as

$$\begin{aligned} P_I \;&=\; PL_{ref}\, r^{-n} \chi_0 \sum_{l=2}^{j} A_{0,l}^2 \frac{P_{tot}}{j} \\ &\quad + \sum_{k=1}^{6} PL_{ref} d_k^{-n} \chi_k \sum_{l=1}^{j} A_{k,l}^2 \frac{P_{tot}}{j}\ , \end{aligned} \quad (6)$$

where $A_{k,l}$ represents the Rayleigh fading gain of the lth path from the kth cell.

3.1 Error Probability Based on SGA Method

By applying SGA method, the interference is regarded as a Gaussian RV. If only the large-scale fading effects are considered, the average encoded bit error probability is given by

$$\begin{aligned} P_e^{SGA}(r,\varphi) &\approx E[Q(\sqrt{2\gamma_b})] \\ &= \int_0^\infty \cdots \int_0^\infty Q(\sqrt{2\gamma_b}) f_{\chi_0,\ldots,\chi_6}(x_0,\ldots,x_6)dx_0 \cdots dx_6, \end{aligned} \quad (7)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-x^2/2)dx$ is the well-known Q-function; $f_{\chi_0,\ldots,\chi_6}(x_0,\ldots,x_6) = \prod_{k=0}^{6} f_{\chi_k}(x_k)$ denotes the joint probability density function (pdf) of $\{\chi_k\}_{k=0}^{6}$ with $f_{\chi_k}(x_k)$ defined in (3); and γ_b is the encoded bit energy-to-total interference plus noise power density ratio. According to (4), (5) and (6), we have

$$\gamma_b = \frac{\alpha M x_0/jU}{(1-\frac{1}{j})(\frac{2\alpha x_0}{3}) + \frac{2\alpha}{3r^{-n}} \sum_{k=1}^{6} d_k^{-n} x_k + \left(\frac{R}{r}\right)^{-n}}, \quad (8)$$

where M is the spreading factor, i.e., the chip number in a encoded bit.

The total CCI is the sum of multiple log-normally distributed interferers. Currently, there is no closed expression for the pdf of the sum of over two independent log-normal RVs. However, by approximating the total CCI as another log-normal RV, (7) and (8) can be further simplified. Several approximation methods were proposed to deal with this problem, including Fenton-Wilkinson's method [25]–[26], Schwartz-and-Yeh's method [27]–[28], and Farley's method [29]. In this work, we apply Fenton-Wilkinson's method to simplify the distribution of total CCI, since it provides good accuracy with limited computational complexity. In (8), the summation term $\sum_{k=1}^{6} d_k^{-n}\chi_k$ is approximated as a log-normal RV $\widetilde{L}_S$ with the pdf $f_{\widetilde{L}_S}(y)$ [25]. Therefore, the average encoded bit error probability can be further simplified to

$$P_e^{SGA}(r,\varphi) \approx \int_0^\infty \int_0^\infty Q\left(\sqrt{2\widetilde{\gamma}_b}\right) f_{\chi_0}(x) f_{\widetilde{L}_S}(y) dx dy, \quad (9)$$

where x and y are dummy variables corresponding to χ_0 and $\widetilde{L}_S$, respectively; and

$$\widetilde{\gamma}_b = \frac{\alpha M x/jU}{(1-\frac{1}{j})(\frac{2\alpha x}{3}) + \frac{2\alpha y}{3r^{-n}} + \left(\frac{R}{r}\right)^{-n}}. \quad (10)$$

If a large-small-scale fading channel is considered, the average encoded bit error probability for QPSK modulation becomes

$$\begin{aligned} &P_e^{SGA}(r,\varphi) \\ &\quad \approx \int_0^\infty \int_0^\infty \frac{1}{2}\left[1-(\widetilde{\gamma}_b^{-1}+1)^{-\frac{1}{2}}\right] f_{\chi_0}(x) f_{\widetilde{L}_S}(y) dx dy, \end{aligned} \quad (11)$$

where $\widetilde{\gamma}_b$, equivalent to (10), is the average received encoded bit energy-to-total interference plus noise power density ratio. In CDMA systems, RAKE receivers with maximal ratio combining (MRC) are introduced to overcome the multipath fading environments. In cdma2000 or WCDMA systems, 3-path RAKE receivers are employed; thus, the encoded bit error probability becomes [30]

$$P_e^{SGA}(r,\varphi) \approx E\left[\left(\frac{1-\eta}{2}\right)^3 \sum_{k=0}^{2} \binom{2+k}{k} \left(\frac{1+\eta}{2}\right)^k\right]$$

$$= \int_0^\infty \int_0^\infty \left(\frac{1-\eta}{2}\right)^3 \sum_{k=0}^{2} \binom{2+k}{k} \left(\frac{1+\eta}{2}\right)^k \\ \times f_{\chi_0}(x) f_{\widetilde{L}_S}(y) dx dy, \qquad (12)$$

where $\eta = 1/\sqrt{\widetilde{\gamma}_b^{-1}+1}$.

3.2 Error Probability Based on SIGA Method

If SIGA method is applied, the variance of the interference is now treated as a RV. If only the large-scale fading effects are considered, the average encoded bit error probability can be expressed as

$$P_e^{SIGA}(r,\varphi) \\ = \int_0^\infty \cdots \int_0^\infty \left[\frac{2}{3}Q(\frac{1}{\sqrt{a}}) + \frac{1}{6}Q(\frac{1}{\sqrt{b}}) + \frac{1}{6}Q(\frac{1}{\sqrt{c}})\right] \\ \times f_{\chi_0,\ldots,\chi_6}(x_0,\ldots,x_6) dx_0 \cdots dx_6, \qquad (13)$$

where $a = \Gamma$, $b = \Gamma + \Delta$, $c = \Gamma - \Delta$,

$$\Gamma = \frac{jUR^{-n}}{2\alpha M r^{-n}\chi_0} + \frac{jU}{3Mr^{-n}\chi_0}\left[\left(1-\frac{1}{j}\right)r^{-n}\chi_0 + \sum_{k=1}^{6} d_k^{-n}\chi_k\right],$$

$$\Delta = \frac{2jU}{M^2 r^{-n}\chi_0}\left[\frac{M-1}{12}\left(\left(1-\frac{1}{j}\right)r^{-n}\chi_0 + \sum_{k=1}^{6} d_k^{-n}\chi_k\right)^2 \right. \\ \left. + \frac{23M^2+8M-8}{120}\left(\left(1-\frac{1}{j}\right)^2 r^{-2n}\chi_0^2 + \sum_{k=1}^{6} d_k^{-2n}\chi_k^2\right)\right]^{\frac{1}{2}}$$

and $Q(c^{-1/2}) \equiv 0$ for $c < 0$.

For a large-small-scale fading channel, the average encoded bit error probability becomes

$$P_e^{SIGA}(r,\varphi) \\ = \int_0^\infty \cdots \int_0^\infty \left[\frac{1}{3}\left(1-\frac{1}{\sqrt{a'+1}}\right) + \frac{1}{12}\left(1-\frac{1}{\sqrt{b'+1}}\right) \right. \\ \left. + \frac{1}{12}\left(1-\frac{1}{\sqrt{c'+1}}\right)\right] f_{\chi_0,\ldots,\chi_6}(x_0,\ldots,x_6) dx_0 \cdots dx_6, \qquad (14)$$

where $a' = \Gamma'$, $b' = \Gamma' + \Delta'$, $c' = \Gamma' - \Delta'$,

$$\Gamma' = \frac{jUR^{-n}}{\alpha M r^{-n}\chi_0} + \frac{2jU}{3Mr^{-n}\chi_0}\left[\left(1-\frac{1}{j}\right)r^{-n}\chi_0 + \sum_{k=1}^{6} d_k^{-n}\chi_k\right],$$

$$\Delta' = \frac{2jU}{M^2 r^{-n}\chi_0}\left[\frac{M-1}{12}\left(\left(1-\frac{1}{j}\right)r^{-n}\chi_0 + \sum_{k=1}^{6} d_k^{-n}\chi_k\right)^2 \right. \\ \left. + \frac{23M^2+13M-13}{60}\left(\left(1-\frac{1}{j}\right)^2 r^{-2n}\chi_0^2 + \sum_{k=1}^{6} d_k^{-2n}\chi_k^2\right)\right]^{\frac{1}{2}}$$

and $1 - (1/\sqrt{c'+1}) \equiv 0$ for $c'+1 < 0$.

If a 3-path RAKE receiver is employed, the encoded bit error probability becomes

$$P_e^{SIGA}(r,\varphi) \\ \approx \int_0^\infty \cdots \int_0^\infty \left[\frac{2}{3}\left(\frac{1-\eta_1}{2}\right)^3 \sum_{k=0}^{2} \binom{2+k}{k}\left(\frac{1+\eta_1}{2}\right)^k \right. \\ + \frac{1}{6}\left(\frac{1-\eta_2}{2}\right)^3 \sum_{k=0}^{2} \binom{2+k}{k}\left(\frac{1+\eta_2}{2}\right)^k \\ \left. + \frac{1}{6}\left(\frac{1-\eta_3}{2}\right)^3 \sum_{k=0}^{2} \binom{2+k}{k}\left(\frac{1+\eta_3}{2}\right)^k\right] \\ \times f_{\chi_0,\ldots,\chi_6}(x_0,\ldots,x_6) dx_0 \cdots dx_6, \qquad (15)$$

where $\eta_1 = \frac{1}{\sqrt{a'+1}}$, $\eta_2 = \frac{1}{\sqrt{b'+1}}$, $\eta_3 = \frac{1}{\sqrt{c'+1}}$, and $\eta_3 \equiv 0$ for $c'+1 < 0$.

3.3 Valid Cell Coverage and Outage Probability

Assuming that a set of system parameters, including α, U and j, has been determined, the average performance BER P_e^{SGA} can be obtained by applying (9), (11) or (12). For a user at a location with a specific value of r/R, if an average performance constraint P_e is desired, the minimum allowable spreading factor M can be found to satisfy $P_e^{SGA} \leq P_e$. Therefore, the maximum available data rate R_b^{max} can be obtained according to the minimum allowable M. It is of note that R_b^{max} obtained here is only the mean value averaged over the shadowing RVs. The actual value of R_b^{max} is still a RV depending on χ_0 and $\widetilde{L}_S$.

The outage probability, being a widely used expression of QoS (Quality of Service) criterions, is commonly used to express the validity of cell coverage. According to (10), we assume that a predetermined performance Pe is satisfied only when

$$\widetilde{\gamma}_b = \frac{\alpha M x/jU}{(1-\frac{1}{j})(\frac{2\alpha x}{3}) + \frac{2\alpha y}{3r^{-n}} + \left(\frac{R}{r}\right)^{-n}} \geq \Lambda_{th} \\ \Rightarrow Ax \geq By + C, \qquad (16)$$

where Λ_{th} is a threshold corresponding to P_e. The thresholds can be obtained by

$$\Lambda_{th} = \begin{cases} P_1^{-1}(P_e), & \text{for large-scale fading channel} \\ P_2^{-1}(P_e), & \text{for large-small-scale fading channel} \end{cases} \qquad (17)$$

where $P_1^{-1}(\cdot)$ and $P_2^{-1}(\cdot)$ are the inverse functions of $P_1(x) = Q(\sqrt{2x})$ and

$$P_2(x) = \left(\frac{1}{2} - \frac{1}{2\sqrt{x^{-1}+1}}\right)^3 \sum_{k=0}^{2} \binom{2+k}{k}\left(\frac{1}{2} + \frac{1}{2\sqrt{x^{-1}+1}}\right)^k,$$

respectively. Therefore, the outage probability, under specific values of α, U, j and P_e, becomes

$$P_{out}(r,M) = \begin{cases} \int_0^\infty \int_0^\lambda f_{\chi_0}(x) f_{\widetilde{L}_S}(y) dx dy, & A > 0 \\ 1, & A \leq 0 \end{cases} \qquad (18)$$

where $\lambda = \frac{By+C}{A}$, $A = \frac{\alpha M}{jU} - \frac{2\alpha\Lambda_{th}(1-1/j)}{3}$, $B = \frac{2\alpha\Lambda_{th}}{3r^{-n}}$

and $C = \Lambda_{th}\left(\frac{R}{r}\right)^{-n}$. It is noted that (16) will not be satisfied for $A \le 0$, and thus $P_{out}(r, M) = 1$ for $A \le 0$.

The cell coverage can also be represented as the percentage of whole cell area, over which a certain encoded bit rate can be supported. According to (18), we have the percentage of cell coverage being

$$P_{Coverage}(M) = \frac{2}{R^2}\int_0^R r \times [1 - P_{out}(r, M)]dr. \qquad (19)$$

3.4 Variance of Maximum Available Data Rate

In the viewpoint of service adaptation, the variance of the maximum available data rate is a major concern. In CDMA systems, the spreading factor M must be a power of 2, and is only available for some specific values. Assuming that there are m available spreading factors, $M_1,\ M_2, ...,\ M_m$ in an ascending order, and the corresponding data rates are $R_{b1},\ R_{b2}, ...,\ R_{bm}$, the probability of the minimum available spreading factor being M_k is

$$\begin{cases} P(r, M_k) = [1 - P_{out}(r, M_1)], & \text{for } k = 1 \\ P(r, M_k) = P_{out}(r, M_{k-1}) - P_{out}(r, M_k), & \text{for } 2 \le k \le m \end{cases} \qquad (20)$$

Therefore, the mean and the variance of R_b^{max} can be expressed as

$$\mu_{R_b^{max}}(r) = \sum_{k=1}^{m} P(r, M_k) R_{bk} \qquad (21)$$

$$\sigma^2_{R_b^{max}}(r) = \sum_{k=1}^{m} P(r, M_k)(R_{bk} - \mu_{R_b^{max}})^2 \qquad (22)$$

4. Numerical Results and Discussions

Considering the system-related parameters, we took W-CDMA systems as an example in this work. The spreading code chip rate is 3.84 Mcps, QPSK modulation is applied, and the downlink spreading factor M ranges between 4 and 512, i.e., 4, 8, 16, 32, 64, 128, 256 and 512. Thus, the encoded bit rate corresponding to M is $2 \times 3840/M$ kbps [4]. Channel coding combining with interleaving is commonly used to overcome the fast fading characteristics of radio channels. In W-CDMA systems, the convolutional code, with a code rate equal to 1/2 or 1/3, is used [4]. The predetermined link performance applied in the following numerical results is set to be an encoded BER $P_e = 0.1$. If a different predetermined link performance is desired, other values of encoded BER can be applied to the numerical results. Furthermore, it is assumed that the path loss exponent is $n = 4$; the shadowing standard deviation is $\sigma_\chi = 8$dB; and the multipath number is $j = 1$ and 3 for large-scale and large-small-scale fading channels, respectively.

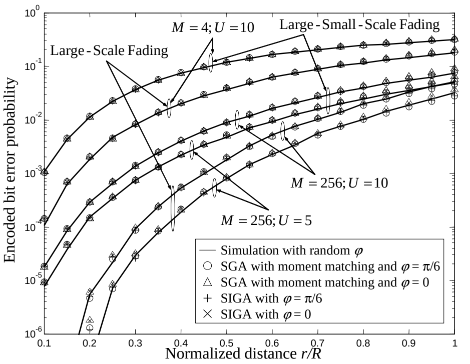
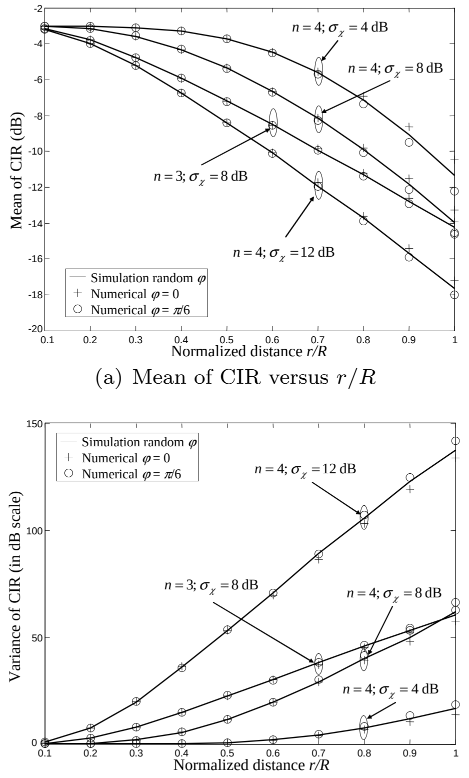


Fig. 2 Encoded bit error rates versus r/R

We compared the numerical results based on SGA and SIGA methods with the simulation results to verify the approximation accuracy. Fig. 2 shows the encoded BER versus r/R with $U = 5$ or 10, $\alpha = 1$ and $j = 1$. For the numerical results the azimuth is set to $\varphi = 0$ or $\pi/6$. However, for the simulation results, the azimuth is uniformly distributed within $(-\pi/6, \pi/6]$. The exact cross-correlations between different spreading codes are calculated to determine the total CCI. We observed that both these two methods showed good accuracy regardless of the values of M and φ. Although there are other approximation methods providing more accurate results, SGA method can satisfy the accuracy requirement in the interesting situations. Therefore, to simplify the numerical complexity, SGA method with Fenton-Wilkinson moment matching, i.e., (9), (11) and (12), will be applied in the following numerical results. The azimuth is set to $\varphi = \pi/6$ corresponding to the worse case.

Fig. 3 shows the mean and variance of the CIR in each path versus r/R, with $U = 1$ and $j = 3$. The total interference includes the MAI from the serving cell and the CCI from six neighboring cells. When the desired user approaches the serving BS, i.e., r/R is very small, the total interference is dominated by the MAI, since the propagation loss between the serving BS and the MS is small. Moreover, the desired signal and the MAI experience the same propagation loss. Hence, we have the mean being $(1/j)/(1 - 1/j) = -3$dB and the variance almost equal to zero for $r/R \approx 0$. On the other hand, when r/R increases, the mean dramatically degrades and the variance significantly rises as well. This is due to the increase both in the propagation loss of desired signal and the CCIs from the neighboring cells.

Fig. 4 shows the maximum available encoded bit rate R_b^{max} versus r/R with $\alpha = 1$, $P_e = 0.1$ and different values of U. It shows that the power ratio U strongly impacts on the valid cell coverage under a fixed encoded bit rate, and R_b^{max} is inversely proportional to r/R under a fixed value of U. Hence, when an MS travels away from the serving BS, it experiences the degra-



(a) Mean of CIR versus r/R

(b) Variance of CIR versus r/R

Fig. 3 Mean and variance of CIR

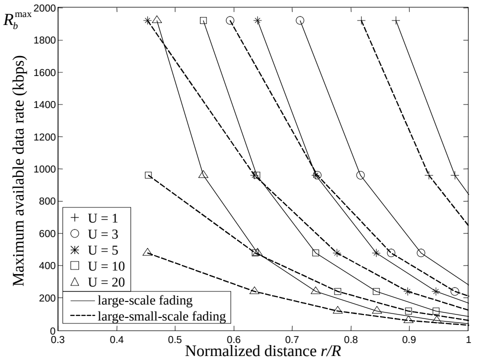


Fig. 4 The maximum available data rate R_b^{max} versus r/R with $\alpha = 1$

dation of QoS on the available data rate. In general, the results of a large-small-scale fading channel are more realistic; thus, the maximum valid cell coverage shrinks from R to about $0.8R$ for $R_b^{max} = 1920$kbps can be easily maintained with $U \leq 5$ for $r \leq 0.4R$. Moreover, $R_b^{max} = 1920$kbps is not available for a large-small-scale fading channel with $U = 10$ or $U = 20$, since the MAI from the serving cell makes $P_e = 0.1$ unachievable under the power ratio. However, for a large-scale fading channel with the multipath number assumed to be $j = 1$, there is no MAI from the serving cell, and thus $R_b^{max} = 1920$kbps is achievable for $U = 10$ and $U = 20$.

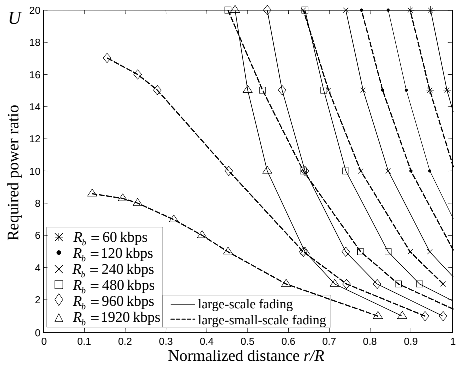


Fig. 5 The required power ratio U versus r/R with $\alpha = 1$

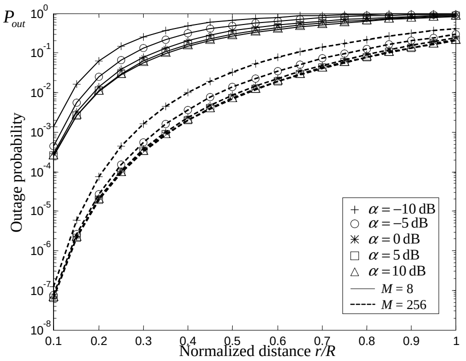
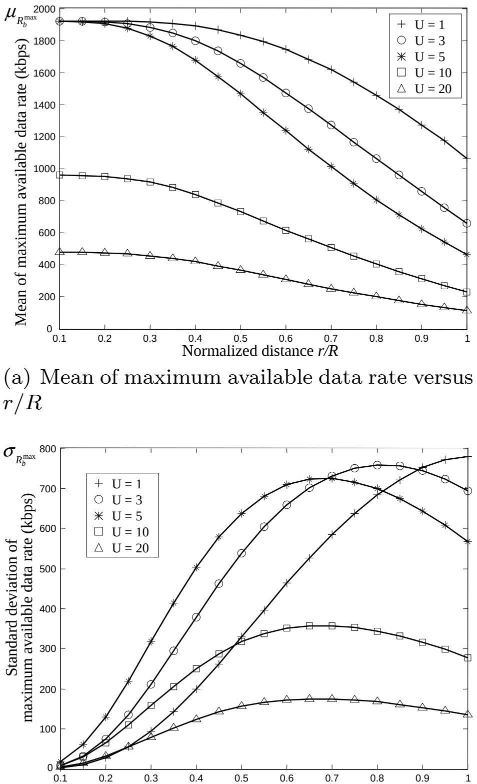


Fig. 6 The outage probability P_{out} versus r/R for different values of α

Fig. 5 shows the required power ratio U versus r/R with $\alpha = 1$, $P_e = 0.1$ and different values of R_b. The results show that the required U is inversely proportional to r/R under a fixed value of R_b. We find that the saturated phenomena occur at $U = 8.5$ and 18 for a large-small-scale fading channel with $R_b = 1920$kbps and 960kbps, respectively. Again, this is due to the MAI from the serving cell. According to this figure, the minimum required transmission power for a specific R_b can be obtained. For example, if $R_b = 1920$kbps is desired, the minimum required transmission power is $0.12P_{tot}$, even the MS is close to the serving BS.

In CDMA systems, the increase of P_{tot} will overcome the AWGN, while it is in vain for the CCI. Seeing that the transmission power is a limited resource, it is important to find out the proper value of P_{tot} equivalent to the parameter α, based on system performance. Fig. 6 shows the outage probability P_{out} versus r/R, with $M = 8$ and 256, $U = 10$ and different values of α. For $\alpha < 0$dB, the AWGN is larger than or comparable to the total CCI at the cell border, and the increase in α will decrease the outage probability. On the other hand, for $\alpha > 0$dB, the total CCI dominates system performance, and the increase in α results in almost no improvement. So, we may conclude that $\alpha = 0$dB$= 1$ is a proper value of providing good system performance yet with moderate total transmission power. Recalling



(a) Mean of maximum available data rate versus r/R

(b) Standard deviation of maximum available data rate versus r/R

Fig. 7 Mean and standard deviation of maximum available data rate

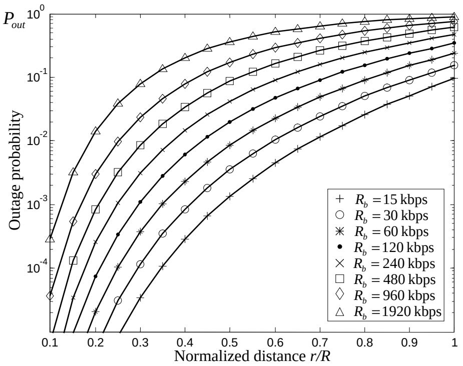


Fig. 8 The outage probability P_{out} versus r/R for different values of R_b

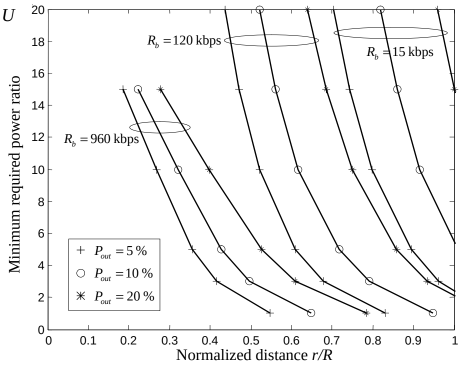


Fig. 9 The required power ratio U versus r/R for different values of P_{out}

4, the proper value of P_{tot}, depending on the cell radius R and the propagation environment, can be obtained by setting $\alpha = 1$.

Fig. 7 shows the mean $\mu_{R_b^{max}}$ and the standard deviation $\sigma_{R_b^{max}}$ versus r/R with different values of U. Recalling Fig. 3, the CIR is with a decrease in mean and an increase in variance as r/R increases. Thus, as we can see, $\mu_{R_b^{max}}$ and $\sigma_{R_b^{max}}$ increases dramatically when r/R increases. However, when r/R approaches 1, the bad CIR forces R_b^{max} varying only among small values, and causes $\sigma_{R_b^{max}}$ to decrease again. Furthermore, we found that the maximum value of $\sigma_{R_b^{max}}$ is about 800kbps for $U = 1$, 3 and 5. Since the variations in R_b^{max} are significant, service adaptation schemes are essential and important for high data rate services in 3G CDMA systems. Our results give a deep insight view into the variations of R_b^{max}.

In some situations, the outage probability P_{out} may be the major concern of service quality. Fig. 8 shows P_{out} versus r/R with $U = 5$ and different values of R_b. We find that P_{out} is proportional to r/R, and a lower encoded bit rate performs better outage probabilities. Fig. 9 shows the required power ratio U versus r/R with different values of R_b and P_{out}. From these results, we can determine the outage probability as well as the minimum required transmission power for a guaranteed QoS. For example, if $R_b^{max} = 1920$kbps and $U = 5$, the outage probability is $P_{out} = 10\%$ for an MS locating at $r = 0.32R$, and is $P_{out} = 50\%$ for $r = 0.6R$. Furthermore, if $R_b^{max} = 120$kbps and $P_{out} = 20\%$, the minimum required transmission power is $P_{tot}/14$ for an MS locating at $r = 0.7R$, and is $P_{tot}/2$ for $r = R$.

In the viewpoint of overall system performance, the percentage of cell coverage, for a specific encoded bit rate, reflects the serviceability of high data rate services. Fig. 10 shows $P_{Coverage}$ versus R_b with different values of U. Under specific values of U and R_b, the result shows the serviceability in the whole cell coverage area. For example, if $U = 3$, the coverage area is about 58% of whole cell for $R_b = 1920$kbps, and 80% for $R_b = 480$kbps. Moreover, the coverage area of $R_b = 1920$kbps is zero for $U = 10$ and 20, since the desired signal power cannot overcome the MAI from the serving cell as shown in Fig. 4.

Finally, in the viewpoint of service adaptation, it is interesting to investigate the probability that a radio link stays in a specific encoded bit rate. Table 1 shows the staying probabilities, derived from (20), of different available encoded bit rates $\{R_{bk}\}_{k=1}^{8}$ at $r/R = 0.2$, 0.4, 0.6, 0.8 and 1.0, with $U = 1$ or 10. It is of note that $R_{bk} = 1920/2^{k-1}$kbps. In the case of $U = 10$, the

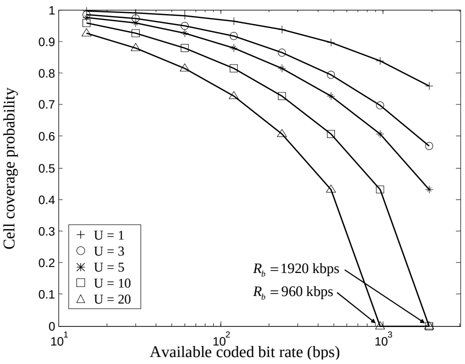


Fig. 10 The cell coverage probability $P_{Coverage}$ versus R_b

Table 1 The staying probabilities of available encoded bit rates.

			$U = 1$		
r/R	0.2	0.4	0.6	0.8	1.0
R_{b1}	9.99E-01	9.74E-01	8.64E-01	6.72E-01	4.38E-01
R_{b2}	4.06E-04	1.51E-02	6.25E-02	1.15E-01	1.36E-01
R_{b3}	1.21E-04	6.59E-03	3.60E-02	8.26E-02	1.20E-01
R_{b4}	3.61E-05	2.77E-03	1.95E-02	5.55E-02	9.85E-02
R_{b5}	1.02E-05	1.09E-03	9.89E-03	3.46E-02	7.47E-02
R_{b6}	2.69E-06	4.01E-04	4.63E-03	2.02E-02	5.29E-02
R_{b7}	6.39E-07	1.37E-04	2.04E-03	1.08E-02	3.46E-02
R_{b8}	1.52E-07	4.39E-05	8.26E-04	5.42E-03	2.13E-02
			$U = 10$		
r/R	0.2	0.4	0.6	0.8	1.0
R_{b1}	0	0	0	0	0
R_{b2}	9.86E-01	7.96E-01	4.87E-01	2.48E-01	1.03E-01
R_{b3}	1.11E-02	1.25E-01	2.25E-01	2.15E-01	1.44E-01
R_{b4}	2.16E-03	4.52E-02	1.24E-01	1.64E-01	1.45E-01
R_{b5}	5.96E-04	1.96E-02	7.40E-02	1.26E-01	1.39E-01
R_{b6}	1.80E-04	8.62E-03	4.32E-02	9.25E-02	1.26E-01
R_{b7}	5.38E-05	3.67E-03	2.39E-02	6.36E-02	1.06E-01
R_{b8}	1.53E-05	1.48E-03	1.24E-02	4.07E-02	8.24E-02

staying probability of $R_b = 1920$kbps is zero, which means that (16) cannot be satisfied for this rate. We also observe that the staying probabilities, for small value of r/R, concentrated only on few data rates. On the contrary, the staying probabilities disperse to all data rates for r/R approaching 1. According to the location of an MS, one can design a service adaptation algorithm to adapt the possible variation of data rates shown in Table 1, and thus the service continuity and reliability can be improved.

5. Conclusions

In this work, the encoded bit error probabilities of 3G CDMA mobile communication systems, over large-scale and large-small-scale fading channels, have been analyzed based on SGA and SIGA methods. Furthermore, analytic methods have been proposed to investigate the issues of coverage shrinking and service data rate variations. The outage probability, cell coverage percentage and the staying probabilities of available data rates were also well examined. From the numerical results, we have the following conclusions:

- Both SGA and SIGA methods provide good accuracy, and SGA method should be applied in considering the computational complexity.
- The noise-raised factor $\alpha = 0$dB$= 1$ is a proper value for providing good system performance yet with moderate total transmission power.
- As the data rate increases, the valid cell coverage shrinks severely. When a user approaches the cell border, a decrease in the mean and an increase in the variance of the available data rate are experienced.
- High data rate services should be provided only on limited area. Maintaining a high data rate service for an MS near the cell border consumes too much power resource, and thus degrades the power efficiency and system capacity.
- For streaming services, if the QoS regarding the service data rate should be maintained, power reservation is essential to overcome the variations of available data rate.
- Since maintaining a high data rate over the whole cell coverage is impractical, a service adaptation mechanism should be proposed to adapt the source data rate to the available encoded bit rate, under a specific value of transmission power.

The proposed analytic methods can be applied, as a preliminary research, to the design of cellular-system-related techniques, such as QoS control, available data rate prediction, power reservation, and service adaptation.

References

[1] R. L. Peterson, R. E. Ziemer, and D. E. Borth, Introduction to Spread-Spectrum Communications. New Jersey: Prentice-Hall, 1995.

[2] A. J. Viterbi, CDMA-Principles of Spread Spectrum Communication. New York: Addison-Wesley, 1995.

[3] Physical layer standard for cdma2000 spread spectrum systems, 3GPP2 C.P0002-A, 3rd Generation Partnership Project 2, October 1999.

[4] Technical Specification Group Radio Access Networks, 3G TS 25.XXX (Release 1999), 3rd Generation Partnership Project, June 2000.

[5] Physical layer specification for LAS-2000, China Wireless Telecommunication Standards, LAS-CDMA Sub-Working Group, 25 April 2001.

[6] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication—Part I: System analysis," IEEE Trans. Commun., vol. COM-25, pp. 759-799, Aug. 1977.

[7] J. S. Lehnert and M. B. Pursley, "Error probabilities for binary direct-sequence spread spectrum communications with random signature sequences," IEEE Trans. Commun., vol. COM-35, pp. 87-98, Jan. 1987.

[8] R. K. Morrow and J. S. Lehnert, "Bit-to-bit error dependence in slotted DS/SSMA packet systems with random signature sequences," IEEE Trans. Commun., vol. 37, pp. 1052-1061, Oct. 1989.

[9] J. M. Holtzman, "A simple, accurate method to calculate spread-spectrum multiple-access error probabilities," IEEE Trans. Commun., vol. 40, pp. 461-464, Mar. 1992.

[10] J. M. Holtzman, ”On using perturbation analysis to do sensitivity analysis: Derivatives vs. differences,” Proc. IEEE Conf. Decision Contr., Tampa, FL, Dec. 1989, pp.2018-2023; also in IEEE Trans. Automat. Contr., vol. 37, pp. 243-247, Feb. 1992.

[11] A. J. Viterbi, ”Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels,” IEEE J. Select. Areas Commun., vol. 8, pp. 641-649, May 1990.

[12] J. E. Salt and S.Kumar, ”Effects of filtering on the performance of QPSK and MSK modulation in D-S spread spectrum systems using RAKE receivers,” IEEE J. Select. Areas Commun., vol. 12, pp. 707-715, May 1994.

[13] Y. Asano, Y. Daido, and J. M. Holtzman, ”Performance evaluation for band-limited DS-CDMA communication system,” in Proc. IEEE 43rd Vehicular Technology Conf., Secaucus, NJ, May 1993, pp. 464-468. 554-558.

[14] J. H. Cho and J. S. Lehnert, ”An optimal signal design for band-limited asynchronous DS-CDMA communications,” IEEE Trans. Inform. Theory, vol. 48, pp. 1172-1185, May 2002.

[15] Y. C. Yoon, ”A simple and accurate method of probability of bit error analysis for asynchronous band-limited DS-CDMA systems,” IEEE Trans. Commun., vol. 50, pp. 656-663, Apr. 2002.

[16] Y. C. Yoon, ”An improved Gaussian approximation for probability of bit-error analysis of asynchronous band-limited DS-CDMA systems with BPSK spreading,” IEEE Trans. Wireless Commun., vol. 1, pp. 373-382, July 2002.

[17] J. H. Cho, Y. K. Jeong, and J. S. Lehnert, ”Average bit-error-rate performance of band-limited DS/SSMA communications,” IEEE Trans. Commun., vol. 50, pp. 1150-1159, Jul. 2002.

[18] Y. C. Yoon, ”Quadriphase DS-CDMA with pulse shaping and the accuracy of the Gaussian approximation for matched filter receiver performance analysis,” IEEE Trans. Wireless Commun., vol. 1, pp. 761-768, Oct. 2002.

[19] G. Zang and C. Ling, ”Performance evaluation for band-limited DS-CDMA systems based on simplified improved Gaussian approximation,” IEEE Trans. Commun., vol. 51, pp. 1204-1213, Jul. 2003.

[20] S. Kumar and S. Nanda, ”High Data-Rate Packet Communications for Cellular Networks Using CDMA: Algorithms and Performance,” IEEE J. Select. Areas Commun,, vol. 17, pp. 472-492, March 1999.

[21] J. D. Lim, ”Air interface capacity and area coverage analyses for cdma2000 voice and packet data services,” in Proc. IEEE Veh. Technol. Conf., pp. 1770-1774, Oct. 2001.

[22] D. Ayyagari and A. Ephremides, ”Cellular Multicode CDMA Capacity for Integrated (Voice and Data) Services,” IEEE J. Select. Areas Commun,, vol. 17, pp. 928-938, May 1999.

[23] A. Abrardo, G. Giambene and D. Sennati, ”Capacity Evaluation of a Mixed-Traffic WCDMA System in the Presence of Load Control,” IEEE Trans. Veh. Technol., vol. VT-52, pp. 490-501, May 2003.

[24] Telecommunications Industry Association (TIA), EIA/TIA-95 Rev B: Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System, Oct. 1998.

[25] G. L. Stuber, Principles of Mobile Communication, KAP, 1996.

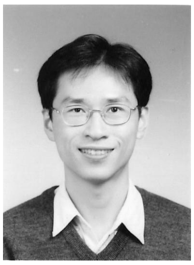
[26] L. F. Fenton, ”The sum of log-normal probability distributions in scatter transmission systems,” IRE Trans. Commun., vol. 8, pp. 57-67, Mar. 1960.

[27] S. C. Schwartz and Y. S. Yeh, ”On the distribution function and moments of power sums with lognormal components,” Bell Systems Technical Journal, Vol. 61, pp.1441-1462, Sep. 1982.

[28] A. Safak, ”Statistical analysis of the power sum of multiple correlated lognormal components,” IEEE transactions on vehicular technology, vol. 42, pp.58-61, Feb. 1993.

[29] N. C. Beaulieu, A. A. Abu-Dayya, and P. J. MacLane, ”Estimating the distribution of a sum of independent lognormal random variables,” IEEE Trans. Commun., vol. 43, pp. 2869-2873, Dec. 1995.

[30] J. G. Proakis, Digital Communications, McGraw-Hill, 2001.



Yuh-Ren Tsai received the B.S. degree in electrical engineering from National Tsing Hua University, Hsing-Chu, Taiwan, in 1989, and the Ph.D. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1994. From 1994 to 2001, he was a Researcher in Telecommunication Laboratories of Chunghwa Telecom Co., Ltd., Taiwan. Since 2001, he has been with the Department of Electrical Engineering and the Institute of Communications Engineering at National Tsing Hua University, Taiwan, and is currently an Assistant Professor. His research interests include sensor networks, mobile cellular systems, CDMA technology and cryptography.

Kai-Jie Yang received the B.S. degree in electrical engineering from National Chung Cheng University, Chiayi, Taiwan, in 2000, and the M.S. degree in communications engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2002. He has been pursuing the Ph.D. degree in communications engineering at National Tsing Hua University,Hsinchu, Taiwan, since 2002. His research interests include mobile ad hoc networks, sensor networks and wireless mobile communication systems.

Yuh-Ren Tsai and Li-Cheng Lin, “Quality Based OVSF Code Assignment and Reassignment Strategies for WCDMA Systems,”

Proc. of *IEEE 2005 International Conference on Wireless Networks, Communications, and Mobile Computing* (*WirelessCom 2005*), June 2005.

Quality Based OVSF Code Assignment and Reassignment Strategies for WCDMA Systems

Yuh-Ren Tsai and Li-Cheng Lin
Institute of Communications Engineering,
National Tsing Hua University
101, Sec. 2, Kuang-Fu Rd., Hsinchu 300, Taiwan
yrtsai@ee.nthu.edu.tw

Abstract—**For the downlink of 3G mobile communication system, OVSF codes are used as channelization codes to support applications with different bandwidth requirements. The code assignment and reassignment problem focuses on the subject of minimizing the number of code relocations. Previous works have been based on some unrealistic assumptions, including fixed service data rates and code-limited system capacities. In this work, we investigated this issue from a realistic perspective, by including variable service data rates and interference-limited system capacities. To achieving the best service quality, we have proposed three OVSF code assignment and reassignment strategies: *sparse-first*; *sparse-first/rightmost*; and *modified sparse-first/rightmost*. The simulation results, based on the evaluation of number of code reassignments, service quality and throughput, were well investigated and compared with those of other strategies. It was found that our proposed strategies outperformed others, with a significant improvement being guaranteed.**

Keywords: **Mobile Cellular Systems; Mobile computing; OVSF; WCDMA; Code Assignment; 3G.**

I. INTRODUCTION

In CDMA mobile cellular systems, all downlink channels, transmitted from the same base station (BS), are spread by different orthogonal codes to maintain the orthogonality. In the second-generation (2G) CDMA system, i.e. IS-95, all orthogonal codes have the same code length (64 chips), and the services are generally limited to low data rate (9.6 kbps or 14.4 kbps) applications, such as voice, facsimile and low-rate data transmission. In 3G CDMA systems, high data rate services are essential for many key applications, such as web browsing, file transfer, and multimedia applications. Since the occupied transmission bandwidth is held to be invariable for different data rates, the variations in transmission data rates are accomplished by applying different lengths of orthogonal codes, referred to as Orthogonal Variable Spreading Factor (OVSF) codes.

The number of available orthogonal codes is restricted to the code length. Being a limited radio resource, efficiently utilizing OVSF codes becomes an important issue. In [1], Minn and Siu proposed the Dynamic Code Assignment (DCA) algorithm to eliminate the code blocking problem. Code assignment and reassignment problems also have been intensively investigated in literature [2]-[8]. All of these works were based on the assumption that the arrival of a new call requested a fixed service data rate for call admission and retained this rate until call termination, while the system capacity has generally been assumed to be code-limited.

In mobile cellular systems, link performance depends on desired signal strength and the amount of total co-channel interference (CCI). Due to the uncertainty of radio propagation channels, link quality, defined as the received carrier-to-total interference power ratio (CIR), varies with time; therefore, link adaptation techniques have been proposed to prevail upon channel conditions, and to maximize the transmission data rate and spectral efficiency. In 3G CDMA systems, the link adaptation schemes adapt the spreading factor (equivalent to processing gain), transmission power (via the power control mechanism) and coding rate to overcome channel conditions and co-channel interferences. Since the available transmission power for a specific channel is limited, if the channel conditions are unfavorable, an OVSF code with a larger code length, corresponding to a smaller data rate, should be used in order to maintain the performance. On the other hand, if the channel conditions are favorable, an OVSF code with a smaller code length, corresponding to a larger data rate, will be used in order to provide the best service effort, as well as to maximize the transmission data rate and spectral efficiency. Therefore, the transmission data rate of a specific channel may vary with time, opening a quite different and more realistic scenario for the study of code assignment and reassignment problems. Under this realistic scenario, we have proposed new code assignment and reassignment strategies, based on service quality.

The remainder of this paper is organized as follows: Section II states the problem of OVSF code utilization and presents the system models; Section III proposes the quality based code assignment and reassignment strategies; in Section IV, the simulation results of system performance are presented; and finally, the conclusions are offered in Section V.

II. PROBLEM STATEMENT AND SYSTEM MODELS

A. *OVSF Code*

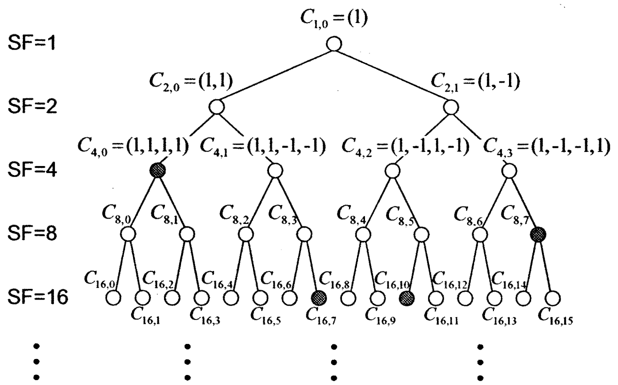
In CDMA cellular systems, all downlink channels transmitted from the same BS are spread by different channelization codes and followed by a common scrambling code. The purpose of using different channelization codes is to preserve orthogonality among downlink channels; therefore, Walsh-Hadamard codes, which are a set of orthogonal sequences and only available for a code length equal to a power of two, are employed [9]-[10]. In WCDMA systems, the channelization codes are also referred to as OVSF codes, which can be represented by a code tree [11], as shown in Fig. 1. Each channeliza-

This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC 93-2752-E-007-003-PAE.

0-7803-9305-8/05/$20.00 ©2005 IEEE

tion code is uniquely described as C_{SF}, where SF is the spreading factor of the code, i.e. the code length, and k, $0 \leq SF \leq k-1$, is the code number. Any two codes in the OVSF code tree are mutually orthogonal, except that they have an ancestor-descendant relationship. If there are two codes with an ancestor-descendant relationship, they are fully correlated and cannot be used simultaneously. For example, in Fig. 1, if the code $C_{4,0}$ is occupied, its descendant codes, $C_{8,0}$, $C_{8,1}$, $C_{16,0}$, $C_{16,1}$, $C_{16,2}$ and $C_{16,3}$, and its ancestor codes, $C_{1,0}$ and $C_{2,0}$, are also regarded as being occupied.

Since the chip rate of OVSF codes is constant, the spreading factor, defined as the number of chips per symbol duration, varies with the transmission data rate. The data rate is inversely proportional to the value of SF, and is represented as kR, where R is the basic data rate of the leaf codes, i.e. the lowest-layer codes, and k is a power of two. If the data rate supported by $C_{16,i}$ is defined as $\tilde{R}$, we have the data rates supported by $C_{8,i}$ and $C_{4,i}$ being $2\tilde{R}$ and $4\tilde{R}$, respectively.


$C_{1,0}=(1)$
SF=1
$C_{2,0}=(1,1)$
$C_{2,1}=(1,-1)$
SF=2
$C_{4,0}=(1,1,1,1)$
$C_{4,1}=(1,1,-1,-1)$
$C_{4,2}=(1,-1,1,-1)$
$C_{4,3}=(1,-1,-1,1)$
SF=4
$C_{8,0}$ $C_{8,1}$ $C_{8,2}$ $C_{8,3}$ $C_{8,4}$ $C_{8,5}$ $C_{8,6}$ $C_{8,7}$
SF=8
$C_{16,0}$ $C_{16,2}$ $C_{16,4}$ $C_{16,6}$ $C_{16,8}$ $C_{16,10}$ $C_{16,12}$ $C_{16,14}$
SF=16
$C_{16,1}$ $C_{16,3}$ $C_{16,5}$ $C_{16,7}$ $C_{16,9}$ $C_{16,11}$ $C_{16,13}$ $C_{16,15}$

*The occupied codes are marked by gray

Fig. 1. The OVSF code tree for WCDMA systems.

B Capacity Limitation in CDMA System

Let the code tree consist of $N_{\max}$ leaf codes, each of which corresponds to a data rate R bps. If a CDMA system is ideally code-limited, the system capacity for a single cell is equal to $N_{\max} \times R$ bps, referred to as code capacity. In general, the existence of an ideal code-limited system is based on the assumption that all assigned codes are mutually orthogonal, and thus there are no mutual interferences among different downlink channels. However, in actual radio environments, CDMA systems are generally not code-limited. Because the frequency reuse factor in CDMA systems is equal to one, co-channel interferences from neighboring cells will degrade system performance as well as system capacity. Moreover, even in the case of a single cell system, mutual interferences among different downlink channels arise from the multipath propagation, and thus the orthogonality between different downlink channels cannot be maintained. Hence, CDMA systems are generally interference-limited, and the real system capacity, referred to as radio capacity, is far beneath code capacity [12]-[15].

In this work, the system capacity was assumed to be interference-limited. Since radio capacity, depending on the propagation environments, user distribution and many other factors, is generally not a determinate value, we introduced a capacity ratio ρ to represent the percentage of code capacity that can actually be supported in a cell. Let Φ_R and $\Phi_c = N_{\max} \times R$ denote radio capacity and code capacity, respectively; the capacity ratio is defined as

$$\rho = \frac{\Phi_R}{\Phi_c} = \frac{\Phi_R}{N_{\max} \times R}, \tag{1}$$

i.e. $\Phi_R = \rho \times N_{\max} \times R$. For example, if $\Phi_c = 256R$ and $\Phi_R = 128R$, the capacity ratio is $\rho = 0.5$.

C. Call Blocking and Code Blocking

When a user requests a data rate, kR, the BS should find a free code corresponding to the rate kR to accommodate the user. There are two possible situations, in which this user will be blocked: *call blocking* and *code blocking*. Call blocking happens when the remaining code capacity of the BS is less than the request rate; whereas code blocking is another situation, where there is no free code available for this user, although the BS has excess code capacity to support the request [1]. To overcome the code blocking problem, the OVSF code relocation for existing users is performed to vacate a branch with rate kR for the requesting user. This code relocation will consume radio resources, and degrade the service quality; thus the number of code relocations must be minimized. To find the optimal branch to be vacated, the DCA algorithm, which minimizes the number of OVSF code relocations, has been proposed [1]. On the other hand, code assignment and reassignment strategies have also been proposed, to prevent the occurrence of code blocking and consequent code relocations [2]-[8]. The code assignment problem addresses the allocation policy when multiple free codes exist in the code tree; whereas the code reassignment problem is devoted to relocating some codes already occupied.

D. Channel and System Models

The channel model plays an important role in link performance. It is assumed to be with large-scale fading effects, including path loss and shadowing effects. Path loss is inversely proportional to the distance between the transmitter and the receiver. Shadowing effects, also known as slow fading, are caused by obstructions in the propagation path, and induce a large variation in the received signal strength for mobile radio environments. As the mobile station (MS) travels, the propagation environment changes accordingly, and so does the link quality, i.e. CIR. Since the available transmission power is limited, and a predetermined performance must be maintained, the transmission data rate should adapt to the link quality accordingly. If the required bit energy-to-total noise density ratio γ_b cannot be maintained, the link adaptation mechanism will magnify the spreading factor, i.e. diminish the transmission data rate, to maintain performance; while, if γ_b has exceeded a threshold, the link adaptation mechanism may diminish the spreading factor to raise the transmission data rate. The data rate control proceeds via a feedback control mechanism, between the MS and the BS, and the data rate may vary on a frame-by-frame basis. For each adaptation time interval, the MS estimates the CIR, based on the received signals, and sends feedback to the BS to decide the transmission data rate.

In this work, the data rate variation was modeled as a three-state Markov chain model [16], as shown in Fig. 2. State 1, State 2 and State 3 represent the data rates $\tilde{R}/2$, $\tilde{R}$ and $2\tilde{R}$, respectively, and $P_{i,j}$ represents the transition probability from State i to State j which depends on the propagation environments. Furthermore, since shadowing effects vary slowly, it is reasonable to assume that the transition from State 1 to State 3 is prohibitive, and vice versa. It should be noted that the same data rate variation model can be applied to different service data rates, since state transitions only depend on the variations of CIR.

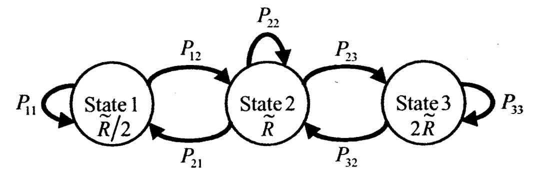


Fig. 2. The data rate variation model – a three-state Markov chain model.

We assumed four types of service in our system: one type of constant-data-rate (CDR) service, and three types of variable-data-rate (VDR) services. The CDR service, assumed to have the basic data rate R, is for the traffic requiring a fixed bandwidth throughout the call duration, such as a voice service that does not change the data rate. The VDR services, assumed to change the data rate within three rates $\{\tilde{R}/2, \tilde{R}, 2\tilde{R}\}$, are intended for real-time applications, like video streaming, multimedia applications or high data rate services. For VDR services, the available data rate, regarded as the measure of service quality, is the major concern. The four types of services are defined as follows. Type-I is a CDR service, with the data rate being the basic rate R. Type-II, Type-III, and Type-IV are all VDR services, with the data rates varying within $\{R, 2R, 4R\}$, $\{2R, 4R, 8R\}$ and $\{4R, 8R, 16R\}$, respectively.

III. CODE ASSIGNMENT AND REASSIGNMENT STRATEGIES

In [4], three code assignment and reassignment strategies, including Random, Leftmost and Crowded-first, were proposed and discussed. The crowded-first strategy has been shown to have better performance than leftmost and random strategies for an ideal code-limited system and a fixed service data rate. In this work, according to the realistic scenario discussed in Section II, we proposed three new code assignment and reassignment strategies, named *sparse-first*, *sparse-first/rightmost* and *modified sparse-first/rightmost*.

A. Sparse-First Strategy

We defined $\phi_n(i, j)$ as the free capacity of the n-th level up ancestor of $C_{i,j}$. In the sparse-first strategy, if there is more than one candidate in the code tree, the one whose ancestor code has the most free capacity (i.e. more sparse) is chosen and assigned to the call. In other words, we search all candidates $C_{i,j} \in$ code set S_0, for codes $C_{i^1,j^1} \in$ code set S_1, where

$$\left(i^1, j^1\right) = \underset{C_{i,j} \in S_0}{\arg\max}\left(\phi_1(i, j)\right). \tag{2}$$

If C_{i^1,j^1} is not unique, we go two levels up and search all codes in S_1 for $C_{i^2,j^2} \in$ code set S_2, where

$$\left(i^2, j^2\right) = \underset{C_{i^1,j^1} \in S_1}{\arg\max}\left(\phi_2\left(i^1, j^1\right)\right). \tag{3}$$

This procedure is repeated until a unique code with the most free capacity is found, or the upper bounded rate of this service is reached. If candidates end up in the same ancestor code, the leftmost one is chosen. Take Fig. 3 as an example. Supposing that a user requests a type-III service and the initial data rate is $2R$, one of the seven candidates $S_0 = \{C_{16,2}, C_{16,4}, C_{16,5}, C_{16,8}, C_{16,9}, C_{16,11}, C_{16,15}\}$, will be chosen to serve this call. According to (2), we found that both $C_{8,2}$ (the ancestor of $C_{16,4}$ and $C_{16,5}$) and $C_{8,4}$ (the ancestor of $C_{16,8}$ and $C_{16,9}$) had the maximum free capacity $4R$. Hence, we had the code set $S_1 = \{C_{16,4}, C_{16,5}, C_{16,8}, C_{16,9}\}$, and went two levels up to compare the free capacities of codes $C_{4,1}$ and $C_{4,2}$ again. According to (3), we had the code set $S_2 = \{C_{16,8}, C_{16,9}\}$, and the leftmost one $C_{16,8}$ was chosen. If the BS intends to raise the data rate of the above-mentioned user from $2R$ to $4R$, the applied channelization code will change from $C_{16,8}$ to $C_{8,4}$ and no reassignment is required for this change.

Under the same scenario, the crowded-first strategy choose the code $C_{16,15}$. If the BS intends to raise the data rate to $4R$, a new free code with capacity $4R$ must be sought, or the user on $C_{16,14}$ must be reassigned. Similarly, if the leftmost strategy is used, $C_{16,2}$ will be chosen for the newly arrived call. When the data rate changes to $4R$, reassigning this user or the user on $C_{16,3}$ is required, and this may lead to other reassignments. Therefore, the number of code reassignments, caused by the rate change, can be reduced by using the sparse-first strategy.

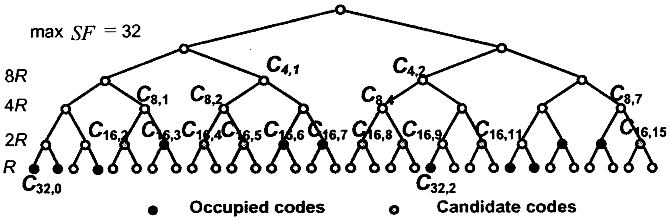


Fig. 3. Code assignment for the sparse-first strategy.

B. Sparse-First/Rightmost Strategy

Since the data rate for CDR services is fixed, the sparse-first strategy is not suitable for these services. Hence, we modified the strategy for CDR services to enhance system performance. When a user requests a CDR service and there is more than one candidate in the code tree, the rightmost one will be chosen. In other words, we search all candidates $C_{i,j} \in$ code set S_0, for the code $C_{i,\hat{j}}$, where

$$\hat{j} = \max(j). \tag{4}$$

If a VDR service is requested, the code assignment will follow the sparse-first strategy. Using this *sparse-first/rightmost* strategy makes the codes serving the CDR services more crowded, and vacates a larger capacity in the left-hand side of the code tree for the requests of high-rate VDR services. Compared with the sparse-first strategy, this can reduce the number of code reassignments due to code blocking.

C. Modified Sparse-First/Rightmost Strategy

In the previous two strategies, if ties occur in capacity comparisons of a VDR service assignment decision, the leftmost candidate will be chosen. However, it may not be the best choice. A new mechanism was introduced to consider the impact of the existing users, which may reassign candidate codes due to potential rate changes. We defined $\eta(i,j)$ as the number of occupied codes, each of which may raise its rate by a factor of 2 and has the ancestor-descendant relationship with candidate $C_{i,j}$. For example, in Fig. 4, if $C_{16,5}$ is a candidate and $C_{8,3}$ is already occupied by a Type-III service, the first level up ancestor of $C_{8,3}$, i.e. $C_{4,1}$, is the second level up ancestor of $C_{16,5}$. So, we have $\eta(16,5)$ equal to one. The value of $\eta(i,j)$ corresponds to the possibility that $C_{i,j}$ will be reassigned by an existing user. After the capacity comparisons among all candidates have been completed, some candidates with tied comparisons are selected and contained in the code set $\tilde{S}$. In this strategy, if there is more than one candidate in the code set $\tilde{S}$, the one with the smallest $\eta(i,j)$ is chosen. In other words, we search all candidates $C_{i,j} \in \tilde{S}$, for the code $C_{\tilde{i},\tilde{j}}$, where

$$(\tilde{i},\tilde{j}) = \underset{C_{i,j} \in \tilde{S}}{\arg\min}(\eta(i,j)). \tag{5}$$

In Fig. 4, if a type-II service with the initial data rate $2R$ is requested, the candidates are in the code set $S_0 = \{C_{16,2}, C_{16,4}, C_{16,5}, C_{16,8}, C_{16,9}, C_{16,11}, C_{16,15}\}$. The types of existing users are as follows: $C_{8,0}$ and $C_{8,3}$ are type-III services; $C_{16,3}$ and $C_{16,14}$ are type-II services; and the others are type-I ($\{R\}$) services. According to the sparse-first strategy, all candidates in $\tilde{S} = \{C_{16,4}, C_{16,5}, C_{16,8}, C_{16,9}\}$ have the same free capacity. Hence, the parameters $\eta(i,j)$ are calculated, and we have $\eta(16,4) = \eta(16,5) = 1$ and $\eta(16,8) = \eta(16,9) = 0$. Finally, the leftmost one, $C_{16,8}$, is chosen.

Under the same scenario, for both crowded-first and leftmost strategies, $C_{16,2}$ is chosen and assigned to the requesting user. If this user intends to raise the service rate in the future, the code $C_{16,3}$ will be reassigned. Moreover, if $C_{16,3}$ or $C_{8,0}$ intends to raise the service rate, $C_{16,2}$ will also be reassigned. On the other hand, both sparse-first and sparse-first/rightmost strategies choose the code $C_{16,4}$. Although $C_{16,4}$ does not reassign other codes when rate change occurs, it will be reassigned when the user on $C_{8,3}$ raises the service rate. If the modified sparse-first/rightmost strategy is applied and accordingly $C_{16,8}$ is chosen, no codes will be reassigned by the rate change of this user, and no rate changes require $C_{16,8}$ to be reassigned.

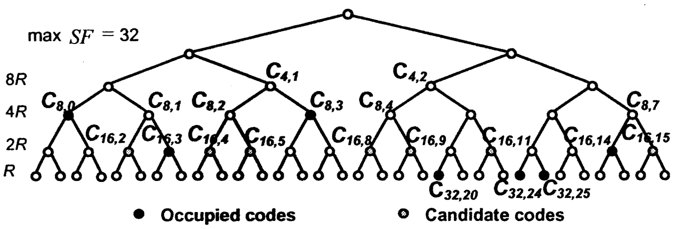


Fig. 4. Code assignment for the modified sparse-first/rightmost strategy.

IV. SIMULATION MODEL AND RESULTS

A. Simulation Model

In this work, the system performance of six code assignment and reassignment strategies, including random (RAN), leftmost (LM), crowded-first (CRF), sparse-first (SF), sparse-first/rightmost (SFR) and modified sparse-first/rightmost (MSFR), were evaluated via simulations. Some system parameters used in the simulations are shown as follows:

- Maximum spreading factor: max $SF = 256$. Therefore, the code capacity is $256R$.
- The call arrival process is a Poisson process, and the call holding times are exponentially distributed, with a mean $\mu = 4$ units of time.
- Input traffic pattern is defined as the ratio of different types of users. For example, Type-I : Type-II : Type-III : Type-IV $= a : b : c : d$. The traffic load G is defined as the average total request data rate in a unit of time.

$$G = \frac{a + 2b + 4c + 8d}{a + b + c + d} \times \lambda \mu R. \tag{6}$$

- Data rate variation model is defined in Fig. 2. For simplicity, the transition probabilities between different states are assumed to be the same, i.e. $P_{i,j} = P$ for $i \neq j$. The step interval is assumed to be 0.1 unit of time. Two models, with transition probabilities $P = 1/3$ and $P = 0.1$, are applied.
- Capacity ratio is defined in (1) for an interference-limited system, and is assumed to be $\rho = 0.6$.

The performance comparison is based on number of code reassignments NOR, service quality Q and throughput. NOR is defined as the number of code reassignments, due to new call arrivals and rate changes, during the simulation interval. For some type of VDR service, if the available rate is held at the highest rate, the service quality is regarded as 100%. If the available rate is held at the middle rate or the lowest rate, the service quality is regarded as 60% or 20 %, respectively. We also define T_1, T_2 and T_3 as the time durations correspond to highest, middle and lowest data rates. Hence, the quality indicator of service can be defined as

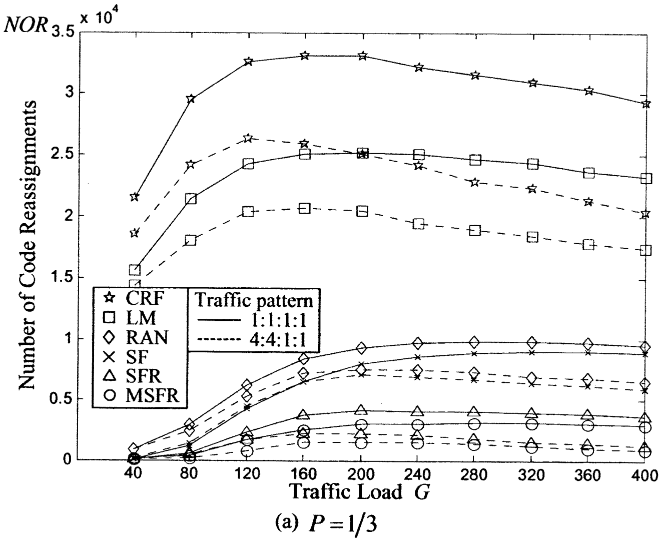
$$Q = \frac{0.2 \times T_1 + 0.6 \times T_2 + 1 \times T_3}{T_1 + T_2 + T_3}. \qquad (7)$$

However, other definitions can be applied to evaluate service quality of VDR services. Throughput is defined as the transmitted bits per unit of time, normalized to R. We also assumed that the maximum number of code reassignments in a unit of time is limited, since the computational power and radio resources of a BS are limited. We defined M_{NOR} as the maximum number of code reassignments that can be performed in a unit of time.

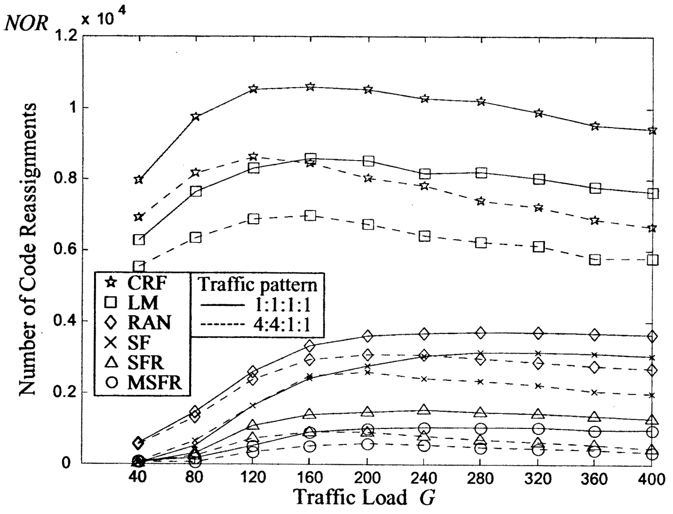
B. *Number of Code Reassignments*

In the evaluation of number of code reassignments, the results were obtained by averaging 10 simulation runs, each of which was terminated when 4000 calls had been accepted. Fig. 5(a) and 5(b) shows the number of code reassignments NOR versus traffic load G for different strategies, with traffic pattern 1:1:1:1 or 4:4:1:1, $\rho = 0.6$, $P = 1/3$ and an unlimited number of code reassignments, i.e. $M_{NOR} = \infty$.

It shows that NOR is proportional to G for a small value of G; however, NOR saturates and slightly decreases for $G > 240$, since the code tree is almost filled up. Furthermore, MSFR has the best performance and CRF has the worst performance. Comparing the results of different input traffic patterns, the pattern $4:4:1:1$ implies more CDR users than pattern 1:1:1:1; thus, NOR decreases for the same traffic load. Fig. 5(b) shows NOR versus G for different strategies, under the same scenario except for the transition probability $P = 0.1$. Compared with the results in Fig. 5(a), smaller values of NOR were observed, since the probability of a rate change decreases from $P = 1/3$ to $P = 0.1$.



(a) $P = 1/3$



(b) $P = 0.1$

Fig. 5. The performance of NOR versus G, with $\rho = 0.6$ and $M_{NOR} = \infty$.

C. *Service Quality and Throughput*

For the performance of service quality and throughput, the results were obtained by averaging 5 simulation runs, each of which was terminated when 10000 calls had been accepted. In the following results, we considered practical limitations and assumed that the maximum number of code reassignments, in a unit of time, was limited to M_{NOR}. Fig. 6 shows the average service quality Q of VDR services versus G for different strategies, with traffic pattern 1:1:1:1 or 4:4:1:1, $\rho = 0.6$, $P = 1/3$ and $M_{NOR} = 30$. Our proposed strategies show far better service quality than that of the others, for different traffic patterns.

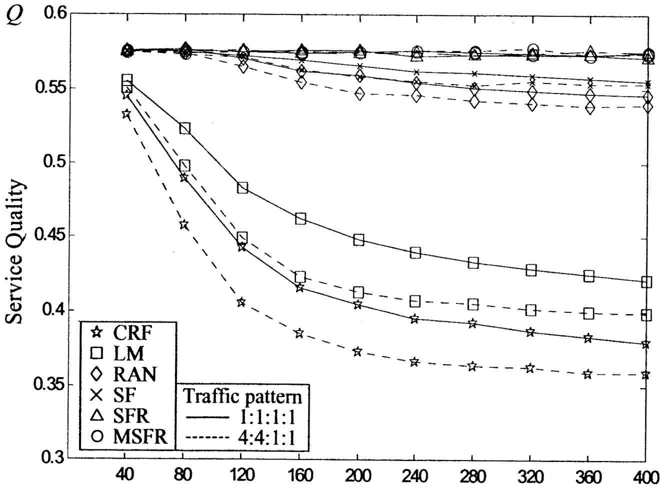


Fig. 6. The performance of service quality Q of VDR services versus G, with $\rho = 0.6$, $P = 1/3$ and $M_{NOR} = 30$.

Fig. 7 shows throughput versus G for different strategies, with traffic pattern 1:1:1:1 or 4:4:1:1, $\rho = 0.6$, $P = 1/3$ and $M_{NOR} = 30$. It shows that the three proposed strategies shows larger throughput than that of the others. When the channel condition is favorable, our proposed strategies maximized the transmission rate with a minimum number of code reassignments; hence, system throughput was greatly improved. The MSFR and SFR improved system throughput over 40 %, when compared to CRF for $M_{NOR} = 30$.

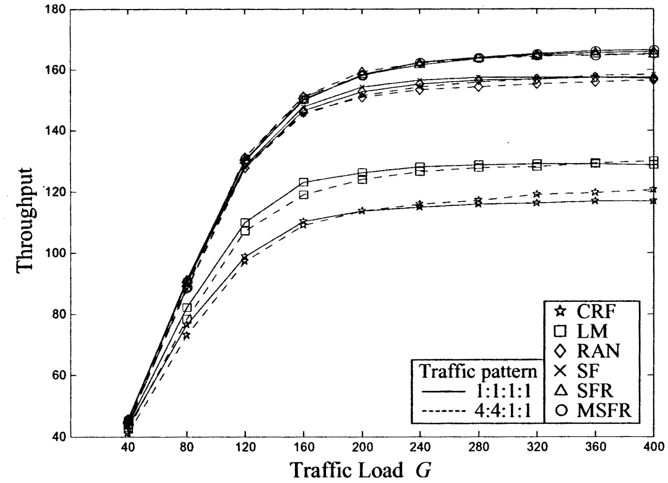


Fig. 7. The performance of throughput versus G, with $\rho = 0.6$, $P = 1/3$ and $M_{NOR} = 30$.

D. Impact of Limitation in NOR

The maximum number of code reassignments in a unit of time, M_{NOR}, is an important system parameter that influences service quality and throughput. Fig. 8 shows service quality Q of VDR services versus M_{NOR} for different strategies, with traffic pattern 1:1:1:1, $\rho = 0.6$, $G = 320$ and $P = 1/3$ or 0.1. It shows that Q is proportional to M_{NOR}, and that the three proposed strategies outperformed the others, especially for small values of M_{NOR}. Notice that the required M_{NOR} in our proposed strategies is much smaller than the requirements of others for achieving the same quality. However, the quality of CRF or LM is limited by M_{NOR}, since more NOR is required for the same conditions.

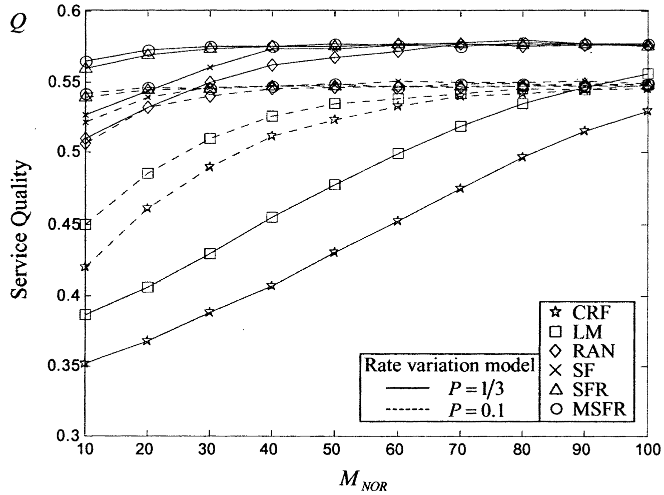


Fig. 8. The performance of service quality versus M_{NOR}, with traffic pattern 1:1:1:1, $\rho = 0.6$ and $G = 320$.

V. CONCLUSION

In this work, we have investigated the OVSF code assignment and reassignment problem for 3G CDMA systems, under a more realistic scenario. The system capacity was assumed to be interference-limited, not code-limited, and the service data rate was assumed to vary with time. We proposed a data rate variation model, with a three-state Markov chain, to represent the variations of service data rates. Under these realistic scenarios, and in view of achieving the best service quality, three OVSF code assignment and reassignments strategies were proposed: sparse-first; sparse-first/rightmost; and modified sparse-first/rightmost. The results showed that our proposed strategies outperformed the others, guaranteeing a significant improvement, especially in the cases with higher rate change probabilities.

REFERENCES

[1] T. Minn and K.-Y. Siu, "Dynamic Assignment of Orthogonal Variable-Spreading-Factor Codes in W-CDMA," IEEE J. Selected Areas in Comm., vol. 18, no. 8, pp. 1429-1440, Aug. 2000.

[2] M. Dell'Amico, F. Maffioli and M.L. Merani, "A tree partitioning dynamic policy for OVSF codes assignment in wideband CDMA," IEEE Trans. Wireless Comm., vol. 3, pp. 1013 - 1017, July 2004.

[3] Y. Yang and T.-S.P. Yum, "Maximally flexible assignment of orthogonal variable spreading factor codes for multirate traffic," IEEE Trans. Wireless Comm., vol. 3, pp. 781 - 792, May 2004.

[4] Y.-C. Tseng and C.-M. Chao, "Code Placement and Replacement Strategies for Wideband CDMA OVSF Code Tree Management", IEEE Trans. on Mobile Computing, vol. 1, pp. 293-302, Oct.-Dec. 2002.

[5] A. N. Rouskas and D. N. Skoutas, "OVSF Code Assignment and Reassignment as the Forward Link of W-CDMA 3G systems," in Proc. IEEE PIMRC, vol. 5, pp. 2404-2408, 2002.

[6] W. T. Chen, Y. P. Wu and H.C. Hsiao, "A Novel Code Assignment Scheme for W-CDMA Systems," in Proc. IEEE Vehicular Technology Conference, VTC-Fall, pp. 1182 -1186, 2001.

[7] R. Assarut, K. Kawanishi, U. Yamamoto, Y. Onozato, and M. Masahiko, "Region Division Assignment of Orthogonal Variable-Spreading-Factor Codes in W-CDMA", in Proc. IEEE Vehicular Technology Conference, VTC-Fall, pp. 1884-1888, 2001.

[8] F. Shueh and W. S. E. Chen, "Code Assignment for IMT-2000 on Forward Radio Link," in Proc. IEEE Vehicular Technology Conference, VTC-Spring, pp. 906-910, 2001.

[9] E. H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and Wideband CDMA cellular networks," *IEEE Communication magazine*, vol. 9, pp. 48–54, Sep. 1998.

[10] 3GPP TS 25.213, V6.0.0, Spreading and modulation (FDD), December 2003

[11] F. Adachi, M. Sawahashi, and K. Okawa, "Tree-Structured Generation of Orthogonal Spreading Codes with Different Lengths for Forward Link of DS-CDMA Mobile Radio," Electronic Letters, vol. 33, pp. 27-28, Jan. 1997.

[12] W. Choi and J. Y. Kim, "Forward-link capacity of a DS/CDMA system with mixed multirate sources," IEEE Trans. Veh. Technol., vol. 50, pp. 737-749, May 2003.

[13] M. M. Matalgah, J. Qaddour, A. Sharma and K. Sheikh, "Throughput and spectral efficiency analysis in 3G FDD WCDMA cellular systems," in Proc. IEEE Globecom Conference, pp. 3423 - 3426, Dec. 2003.

[14] D. Staehle, K. Leibnitz, K. Heck, P. Tran-Gia, B. Schroder and A. Weller, "Analytic approximation of the effective bandwidth for best-effort services in UMTS networks," in Proc. IEEE Vehicular Technology Conference, VTC-Spring, pp. 1153 - 1157, 2003.

[15] S. Dehghan, D. Lister, R. Owen and P. Jones, "W-CDMA capacity and planning issues," Journal of Electronics & Communication Engineering, vol. 12, pp. 101 - 118, June 2000.

[16] H. Stark and J. W. Woods, Probability and Random Processes with Application to Signal Processing, 3rd Prentice Hill, 2002.

Yuh-Ren Tsai and S.J. Wang, “Routing Security and Authentication Mechanism for Mobile Ad Hoc Networks,” Proc. of *IEEE 2004 Vehicular Technology Conference* (*VTC-2004 Fall*), Sept. 2004.

Routing Security and Authentication Mechanism for Mobile Ad Hoc Networks

Yuh-Ren Tsai
Institute of Communications Engineering,
National Tsing Hua University
101, Sec. 2, Kuang-Fu Rd., Hsinchu 300, Taiwan
yrtsai@ee.nthu.edu.tw

Shiuh-Jeng Wang
Department of Information Management,
Central Police University
Taoyuan, Taiwan 3333
sjwang@sun4.cpu.edu.tw

***Abstract*—Mobile ad hoc networks (MANETs) is proposed as an extremely flexible technology for establishing wireless communications. In comparison with fixed networks or traditional mobile cellular networks, MANETs introduce some new security issues. Especially, the routing security is the most important and complicated one. In this work, we propose a two-tier authentication mechanism for MANETs. The first tier, based on hash function and the concept of MAC, provides fast message verification and group identification. The second tier, based on secret sharing technology, provides secure user identification. This two-tier authentication mechanism can prevent internal and external attacks, including black hole, impersonation, routing table overflow and energy consummation attacks.**

Keywords—Mobile ad hoc networks (MANET); Authentication; Routing security; Secret sharing; Hash function.

I. Introduction

An ad hoc network is formed by a collection of self-governing nodes which can communicate with each other without any fixed infrastructure, such as base stations or mobile switching centers, or any centralized administration. Due to limited radio propagation coverage, each node can only connect to a few neighboring nodes. If the communication link between any two nodes which are not in the same radio coverage area is required, a multiple-hop radio connection relying on other intermediate nodes to relay messages is established to provide services. If the nodes of ad hoc networks are mobile and with wireless communication to maintain the connectivity, it is known as mobile ad hoc network (MANET) and is proposed as an extremely flexible technology for establishing communications in situations which demand a fully decentralized network without any fixed base stations, such as battlefields, military applications, and other emergency and disaster situations. Since all nodes are mobile, the network topology of a MANET is generally dynamic and may change frequently.

In comparison with fixed networks or traditional mobile cellular networks, MANETs introduce some new security issues; especially, the routing security is the most important and complicated one [1]-[5]. In this work, we propose a two-tier authentication mechanism according to the characteristics of MANETs. The first tier is the cluster authentication and the second tier is the individual authentication. The rest of this paper is organized as follows. In section II we will illustrate the routing security problems in MANETs. Section III concentrates on the proposed authentication mechanism. In Section IV, the discussions for our proposal are presented. Finally, this paper is concluded in section V.

II. Routing Security in MANETs

A. Routing Protocols of MANETs

In MANETs, each node functions as a host as well as a router and will forward packets for other mobile nodes in the network which is not within direct wireless transmission range of each other. Since the network topology is dynamic, each node participates in the network management and routing mechanism. According to the routing protocols, each node can discover multiple-hop paths through the network to any other node. Many different routing protocols, such as Destination-sequenced distance-vector (DSDV) [6], ad hoc on-demand distance vector (AODV) [7] and Dynamic Source Routing (DSR) [8] protocols, have been developed for MANETs and are generally classified into three categories: proactive, reactive and hybrid protocols.

Proactive protocols, also known as table-driven protocols, require each node to possess one or more tables to store routing information and attempt to maintain consistent and up-to-date routing information for each node; thus periodically refreshing or updating of the existing routing information is essential. On the contrary, reactive protocols, also known as source-initiated on-demand driven protocols, do not periodically update the routing information and routes are created only when desired by source node. When a source node requires a route to a specific destination, it initiates a route discovery process within the network. Since neither the pure proactive nor the reactive approach is sufficient, hybrid protocols make use of both reactive and proactive approaches by adapting the protocols to the specific conditions and is in general the optimal choice.

B. Routing Security Problems of MANETs

MANETs suffer from some new weaknesses, such as no trustworthy administration, easy theft of nodes, vulnerability of tampering, limited computational abilities, battery powered operation and transient nature of services and devices; and thus

This work was supported in part by the Chung-Shan Institute of Science and Technology, Taiwan, R.O.C., under Grant BC-93-B14P and the National Science Council, Taiwan, R.O.C., under Grant NSC 93-2752-E-007-003-PAE.

0-7803-8521-7/04/$20.00 © 2004 IEEE

MANETs introduce some new and critical security issues. Especially, the operation of MANETs is based on the assumption of routing protocol functioning normally; thus routing security becomes the Achilles' heel of MANETs in an insecure environment. To defend against malicious attacks, the security mechanisms should be able to protect the network management information (such as routing information), user traffic and any other vulnerable information from any adversaries, and to maintain the routing protocol functioning normally. Furthermore, the design of security mechanisms should also consider the availability, confidentiality, integrity, authenticity and non-repudiation similar to the traditional networks.

Attacks against MANETs can be classified as passive attacks and active attacks. Passive attacks typically involve only eavesdropping and the attacker attempt to discover valuable information by listening to the routing traffic. Active attacks involve actions performed by adversaries; generally the attacker must be able to inject arbitrary packets into the network and attempts to improperly modify data, gain authentication, replicate or delete data, insert false packets or modify packets transition through the network. Furthermore, active attacks can be divided into external attacks and internal attacks. External attacks are caused by nodes that do not belong to the network while internal attacks are from compromised or hijacked nodes that belong to the network. External attacks can typically be prevented by using standard security mechanisms; while internal attacks are typically more severe, since malicious nodes already belong to the network as an authorized party and thus the attacks are protected with the security mechanisms offered by the network.

Some types of active attacks, including black hole, wormhole, denial of service (DOS), routing table overflow, impersonation, energy consummation and Information or location disclosure, have been addressed in literature [1]-[5]. These attacks can be easily performed against a MANET in the network layer and may paralysis the operation of part or whole network. Details of these attacks will be explained in the final version.

III. Authentication Mechanism

Due to the special characteristics of MANETs, including no centralized administration, dynamic network topology, limited computational power, limited transmission bandwidth and high user mobility, the authentication mechanism suitable for MANETs should be feasible for highly changing network topology and be with low computational complexity and low bandwidth consumption. Thus we propose a two-tier authentication mechanism for message verification and user identification. The first tier is the cluster authentication which provides a fast messages verification and basic user identification mechanism with low computation complexity and low bandwidth consumption. The second tier is the individual authentication which provides a secure user identification mechanism with moderate computation complexity and low bandwidth consumption.

A. Basic Assumptions and Definitions

Followings are the basic assumptions and definitions for this work. It is assumed that each node holds the following items:

- The plaintext M : which is the original message sent by a node.
- The ciphertext C : which is the output of message M encrypted by the cryptosystem.
- Time synchronization: time synchronization is imposed in this system so that each node is able to synchronize the same time.
- Each node holds a common secret key K_c, which is the same for all nodes and is undisclosed to any outsiders.
- Each node holds a symmetric cryptosystem with the property of $M = D_{K_c}(E_{K_c}(M))$, such as 3-DES system or AES system, where $E_{K_c}(\bullet)$ and $D_{K_c}(\bullet)$ are the encryption and decryption functions, respectively; thus all the transmitted messages are well protected from the outsiders.
- Each node holds a collision-free hash function H, such as SHA/MD5.
- The message authentication code (MAC): which is defined by $MAC = H(K_c;\vartheta)$, where K_c is the protection key and ϑ denotes the encoded message.
- Each node holds a set of secret shadows for all other nodes. For example, $K_{A,E}$, kept inside Node $\mathcal{A}$, is the secret shadow associated with Node $\mathcal{E}$. The secret shadows are assumed to be symmetric between any pair of nodes, i.e. $K_{i,j} = K_{j,i}$, and are undisclosed to any other nodes in the same cluster or any outsiders.

B. Two-tier Authentication Mechanism

1) *First tier: cluster authentication*

The goal of cluster authentication is to verify whether a user (node) belongs to the same group or to verify whether a message is come from a (any) node of this group. For example, a soldier should make sure the received message is come from a partner, not come from any adversary. The concept of message authentication code (MAC) is commonly used for message verification [9]. Also it was proposed to play a role in packet leashes for the defense against wormhole attacks [10]. By applying the hash operation, the MAC of a specific message can be easily obtained. In this work, we apply the concept of MAC and hash function for cluster authentication. For each transmitted packet, the original message accompanied with the cluster signature, which is obtained from the hash function by using the time stamp and the original message as inputs, will be sent by the originator. For a node having received a packet, it verifies the cluster signature and then determines the validity of this packet. Thus, the false packets can be detected, and the attacks from any outsider can also be prevented.

2) *Second tier: individual authentication*

The goal of individual authentication is to verify the identity of a specific user (node) of this group or to verify whether the message is come from a specific node or not. For example, a soldier should make sure the received message is come from an officer not come from a soldier which might be a hijacked node. Generally, the public-key cryptosystems [10] can be utilized to fulfill the above-mentioned requirements; however it suffers from high computational complexity and seems unsuitable for MANETs applications. In this work, we utilize the secret sharing concept and propose a user authentication mechanism to verify the identity of a node. In cryptography, secret sharing is usually applied to the key management. With this, a secret (specific key) can be well protected from the unauthorized access and the intentional destruction. The concept of secret sharing was first due to Shamir's interpolating polynomial construction [11]. Afterwards, there are a lot of reports developed on the area explorations of secret sharing [12]-[14]. Among these proposals, the secret key construction (or management) designed in a low-complexity computation is entirely satisfied with the need of MANETs. Accordingly, the low power-consuming authentication procedure based on the secret sharing on low-computation purpose is proposed in our scheme.

C. Proposed Authentication Mechanism

Shown in Fig. 1 is a possible network topology for MANET. The routing protocol is assumed to be an on-demand driven routing protocol, such as AODV. As shown in Fig. 1, when a source node, say Node $\mathcal{A}$, requires a route to a destination node, say Node $\mathcal{E}$, it initiates a *route discovery* process and broadcasts a route request (*RREQ*) packet to its neighbors. All intermediate nodes forward the received *RREQ* to its neighbors, and so on, until either the destination or an intermediate node with a fresh route to the destination is located. As shown in Fig. 2, according to the received *RREQ*, the destination node or an intermediate node with a fresh route will unicast a route reply (*RREP*) packet following the shortest path to inform the source node about route information. After establishing a route, it is maintained by a *route maintenance* process until either the destination becomes inaccessible or the route is no longer desired.

1) *First tier: cluster authentication*

For any message M transmitted from a source node, say Node $\mathcal{A}$, following steps will be progressed before the message is sent.

Procedure 1.

Step 1: According to the system time, Node $\mathcal{A}$ generates a time stamp T_s.

Step 2: By applying M and T_s as inputs, Node $\mathcal{A}$ uses the hash function $H(\bullet)$ to generate MAC for this message, i.e.

$$MAC_M = H(K_c; T_s, M). \tag{1}$$

Step 3: By applying T_s as input, Node $\mathcal{A}$ uses $H(\bullet)$ to generate the cluster signature, i.e.

$$MAC_T = H(K_c; T_s). \tag{2}$$

Step 4: By applying the global symmetric cryptosystem, Node $\mathcal{A}$ generates the encrypted message body

$$E_{K_c}(MAC_M, T_s, M). \tag{3}$$

Step 5: Node $\mathcal{A}$ forms and transmits the output packet

$$PKT_M = \{MAC_T; T_s; E_{K_c}(MAC_M, T_s, M)\}. \tag{4}$$

■

For each intermediate node, if a packet is received, it checks the following two conditions:

Condition 1: $MAC_T = H(K_c; T_s)$, and

Condition 2: T_s is in a reasonable time delay range.

If the two above-mentioned conditions are all satisfied, this intermediate node forwards this packet to the next node; otherwise this packet is discarded. For the destination node, without loss of generality, it will follow the **Condition 1** and **Condition 2** to check the status of an arriving packet. Furthermore, it decrypts $E_{K_c}(MAC_M, T_s, M)$ and checks the following two conditions:

Condition 3: $MAC_M = H(K_c; T_s, M)$ and

Condition 4: The decrypted T_s is the same as the one in the packet without encryption.

If all conditions are satisfied, this packet is regarded as a valid packet; otherwise this packet is discarded.

2) *Second tier: individual authentication*

Assume that system has chosen a large prime number, *p*, and a primitive root *g*. As in Fig. 1, Node $\mathcal{A}$ is the source node and Node $\mathcal{E}$ is the destination node. The individual authentication mechanism is incorporated into the *route discovery* process as shown in the following steps.

Procedure 2.

Step 1: Node $\mathcal{A}$ generates a random number a_0 and a random challenge number $RAND$.

Step 2: Node $\mathcal{A}$ finds a linear polynomial $f_1(x) = a_1 x + a_0$ mod *p*-1 with the number a_1 obtained by solving

$$K_{A,E} = f_1(ID_E) = a_1 \times ID_E + a_0 \bmod (p-1), \tag{5}$$

where ID_E denotes the identity of Node $\mathcal{E}$ in this system.

Step 3: Node $\mathcal{A}$ generates

$$\Gamma = g^{f(1)} \bmod p = g^{a_1 + a_0} \bmod p. \tag{6}$$

Step 4: Node $\mathcal{A}$ broadcasts the *RREQ* packet containing Γ and $RAND$ to discover Node $\mathcal{E}$.

Step 5: Node $\mathcal{E}$ receives *RREQ*, and then generates

$$\Lambda = g^{K_{E,A}} \bmod p = g^{a_1 ID_E + a_0} \bmod p\,, \tag{7}$$

according to the secret shadow $K_{E,A}$.

Step 6: Node $\mathcal{E}$ finds Z satisfying the relation of

$$(ID_E - 1) \times Z = 1 \bmod (p-1)\,. \tag{8}$$

Step 7: According to the Γ in the received *RREQ*, Node $\mathcal{E}$ calculates

$$K_S = \left(\Gamma^{ID_E} / \Lambda\right)^Z \bmod p\,. \tag{9}$$

Step 8: By using K_S as key, Node $\mathcal{E}$ performs an encryption process for the received challenge number $RAND$ and obtains the authentication reply code $AUTHR$, i.e.

$$AUTHR = E_{K_S}(RAND)\,. \tag{10}$$

Step 9: According to the shortest path, Node $\mathcal{E}$ replies a *RREP* packet containing the authentication reply code $AUTHR$ to Node $\mathcal{A}$.

Step 10: Node $\mathcal{A}$ computes

$$V = g^{a_0} \bmod p\,. \tag{11}$$

and verifies whether the received $AUTHR$ is equal to

$$AUTHR' = E_V(RAND)\,, \tag{12}$$

where $RAND$ is the challenge number generated in Step 1. If yes, the responding node is treated as the intended recipient and a secure routing path to Node $\mathcal{E}$ is guaranteed. Furthermore, a common session key $K_S = V = g^{a_0} \bmod p$ is obtained which can be used for message verification and protection in further data transactions. ■

IV. Discussion

In this section, we discuss the applicability and feasibility of our schemes for MANETs.

1) *First tier: cluster authentication*

For the first tier authentication, all messages are encrypted by a global symmetric cryptosystem. No outsider can eavesdrop and acquire the message contents. Furthermore, each intermediate node can verify the validity of the received packets. If any outsider intends to inject a false packet, such as a false routing request, into the network, the **Condition 1** or **Condition 2** will be violated and no intermediate node will forward this false packet. If any outsider intends to modify the encrypted message body or replace it by a replay version of a valid packet, the **Condition 3** or **Condition 4** will be violated and the destination node will discard this packet. According to the foregoing discussion, our scheme can adequately prevent external attacks from any outsider, including routing table overflow and energy consummation attacks.

To be a feasible scheme for MANETs, the computational complexity, bandwidth consumption and power consumption are major concerns. Furthermore, long latency is not acceptable for many real-time applications. The cryptosystem applied in ***Procedure 1*** is a symmetric one with very low computational complexity and power consumption. Moreover, the message verification in any intermediate node involves only the computation of hash function, which is with very low computational complexity. Thus, the message verification procedure will not burden the network with high power consumption and long latency. As considering the bandwidth consumption, the hash output contains only limited number of bits which just increases the required bandwidth slightly.

2) *Second tier: individual authentication*

In MANETs, the internal attacks are typically more severe, since malicious nodes already belong to the network. To prevent such attacks, authenticating the unique identity of any node is necessary. Our individual authentication scheme provides an efficient way to verify the identity of a node. According to ***Procedure 2***, only the destination node can acquire the secret $K_S = g^{a_0} \bmod p$ by the concept of secret sharing. The source node can verify the validity of authentication reply code $AUTHR$ to authenticate the identity of destination node. It is noteworthy that the secret shadow won't be disclosed to other intermediate nodes during the *routing discovery-authentication* procedure since it is subject to the hard problem of discrete logarithms when breaking the secret is desired. Furthermore, the *authentication* manner can be reiterated in later time. In the meantime, a different secret can be imposed to obtain a different common session key. Since the identity of the destination node is well verified, the internal attacks from member of this network, including black hole and impersonation attacks, can be adequately prevented. Consider the crucial damage with the attack of black hole in MANETs. If a malicious node receives a *RREQ* packet requesting a route to the destination node, it may reply a *RREP* packet claiming that it has the shortest path to the destination node. Therefore the malicious node can intercept the packets to the destination node and disrupts the correct functioning of the routing protocol. However, in our proposal, the *RREP* packet replied by the malicious node will be ignored since the authentication reply code is invalid. On the contrary, the source node will confide the *RREQ* packet under the correct comparison in (12). Thus, a secure routing path is obtained and the black hole attack can be definitely prevented.

In the following, we further show that the source and destination nodes share a common secret key, K_S.

Theorem 1. *The destination Node $\mathcal{E}$ and the source Node $\mathcal{A}$ share a common secret key, K_S, in **Procedure 2**.*

Proof.

In ***Procedure 2***, $K_S = g^{a_0} \bmod p$.

Consider the computation in the **Step 7** at the side of Node $\mathcal{E}$.

$$\begin{aligned} K_S &= \left(\Gamma^{ID_E} / \Lambda\right)^Z \bmod p \\ &= (g^{(ID_E - 1)a_0})^Z \bmod p \\ &= g^{(ID_E - 1) \times Z \times a_0} \bmod p \end{aligned}$$

Since $(ID_E - 1) \times Z = 1 \bmod (p-1)$ is sustained in **Step 6** of

Procedure 2, Node $\mathcal{E}$ can obtain the result of $g^{a_0} \bmod p$, which is identical to K_S generated at the side of Node $\mathcal{A}$.

Q.E.D.

For ***Procedure 2***, the computational efforts dominated at the source node are **Step 3** and **10**, and the major computation required in destination note is at Steps **Step 5** and **7**. By carefully choosing the parameters, only limited number of multiplication and modulo operations are needed. As compared to existing exponent-computation-like public-key cryptosystem, our scheme is capable of satisfying the requirements of low computational complexity and less power consumption. As a result, it is suitable for MANETs.

V. CONCLUSIONS

In this work, we have proposed a two-tier authentication mechanism. The first tier, based on hash function and the concept of MAC, is the cluster authentication. This tier can verify whether a user (node) belongs to the same group or whether the message is come from a (any) node of this group, and prevent external attacks, such as routing table overflow and energy consummation attacks. The second tier, based on secret sharing technology, is the individual authentication. This tier can verify the identity of a specific user (node) or verify whether the message is come from a specific node, and prevent the internal attacks from member of this network, including black hole and impersonation attacks. In sum, the first tier mechanism provides limited secure ability with very low complexity; while the second tier mechanism provides a high degree of secure ability with moderate complexity.

REFERENCES

[1] L. Zhou and Z. J. Haas, “Securing Ad Hoc Networks,” IEEE Network., vol. 13, pp. 24-30, Nov./Dec. 1999.

[2] H. Deng, W. Li and D. P. Agrawal, “Routing Security in Wireless Ad Hoc Networks,” IEEE Communications Magazine pp. 70-75, Oct. 2002.

[3] F. Hu and N. K. Sharma, "Security Considerations in Ad Hoc Networks," to be appeared in Ad Hon Network, 2004.

[4] R.D. Pietro, L.V. Mancini, and S. Jajodia, "Providing Secrecy in Key Management Protocols for Large Wireless Sensors Networks," Ad Hoc Networks, Vol. 1, 2003, pp. 455-468.

[5] H.Deng, W. Li, and D.P. Agrawal, "Routing Security in Wireless Ad Hoc Networks,” IEEE Communication Magazine, Oct. 2002.

[6] C. E. Perkins, and P. Bhagwat, “Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers,” Proc. of SIGCOMM, pp. 234–244, 1994.

[7] C. E. Perkins, E. M. Royer, and S. R. Das, “Ad Hoc On-Demand Distance Vector (AODV) Routing,” IETF Mobile Ad Hoc Networks Working Group, Internet Draft, work in progress, 17 Feb. 2003.

[8] D. B. Johnson, D. A. Maltz, and Y-C. Hu, “The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR),” IETF Mobile Ad Hoc Networks Working Group, Internet Draft, work in progress, 15 Apr. 2003.

[9] W. Stallings, Cryptography and Network Security, 3/e, Prentice-Hall, 2003.

[10] Y.-C. Hu, A. Perrig and D. B. Johnson, “Packet Leashes: A Defense against Wormhole Attacks in Wireless Networks,” Proc. of IEEE INFOCOM 2003.

[11] A. Shamir, “How to Share a Secret,” Comm. ACM, Vol. 22, No. 11, 1979, pp. 612-613.

[12] C. Asmuth and J. Bloom, “A Modular Approach to Key Safeguarding,” IEEE Trans. Information Theory, vol. IT-29, no. 2, 1983, pp. 208-210.

[13] E.D. Karnin, J.W. Greene, and M.E. Hellman, “On Sharing Secret Systems,” IEEE Tran. Information Theory, vol. IT-29, 1983, pp. 35-41.

[14] G.J. Simmons, An Introduction to Shared Secret and/or Shared Control Schemes and Their Application, in G. Simmons, Editor, Contemporary Cryptology: The Science of Information Integrity, IEEE Press, 1992.

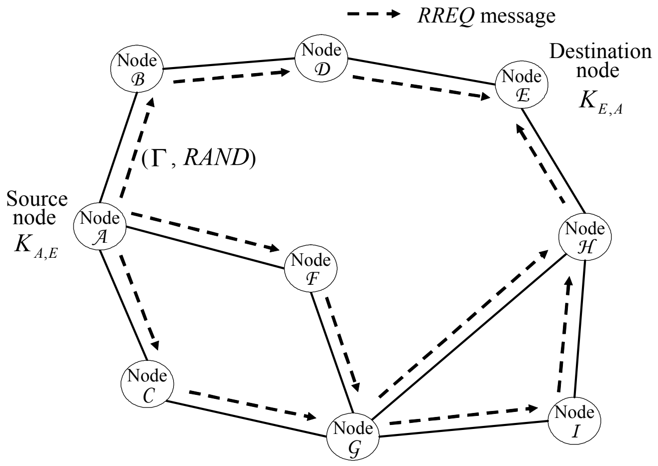


Fig. 1. MANET topology and *RREQ* packet forwarding.

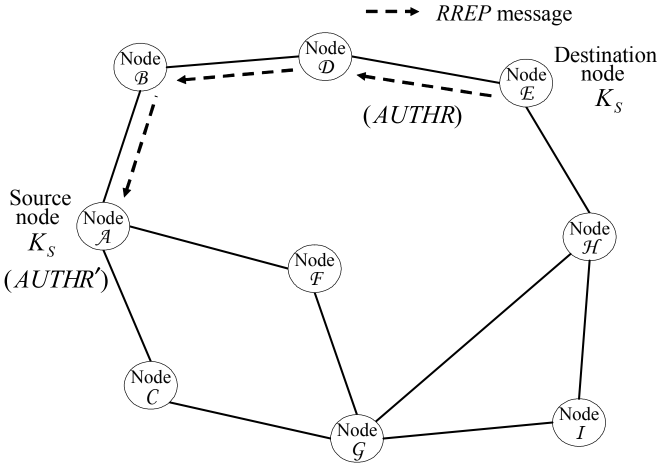


Fig. 2. The reply of *RREP* for shortest path.

Yuh-Ren Tsai, “Coherent *M*-ary Spreading-Code-Phase-Shift-Keying Modulation for Direct-Sequence Spread Spectrum Systems,”

Proc. of *IEEE 2004 Vehicular Technology Conference* (*VTC-2004 Fall*), Sept. 2004.

Coherent *M*-ary Spreading-Code-Phase-Shift-Keying Modulation for Direct-Sequence Spread Spectrum Systems

Yuh-Ren Tsai
Institute of Communications Engineering
National Tsing Hua University
101, Sec. 2, Kuang-Fu Rd.,
Hsinchu 300, Taiwan
yrtsai@ee.nthu.edu.tw

Abstract – **In this work, a generalized form of Spreading-Code-Phase-Shift-Keying (SCPSK) modulation is proposed. The performance of *M*-SCPSK modulation with coherent receiver in an AWGN channel exposed to multi-user interference is investigated. Two types of *M*-ary SCPSK modulation—one employing short spreading codes and the other employing long spreading codes—are examined and the union bounds of symbol error probabilities are derived. Additionally, feasible designs of the *M*-CPSK transmitter and receiver are proposed. The *M*-SCPSK system with long spreading codes employed is found to be more flexible and therefore more suitable for packet data service applications.**

Keywords: Spreading-Code-Phase-Shift-Keying; Code Division Multiple Access; Spread Spectrum Communication.

I. Introduction

Code division multiple access (CDMA) technologies, based on spread spectrum technologies, provide many advantages such as anti-jamming, anti-multipath interference, anti-fading, low probability of intercept, low frequency reuse pattern in cellular system, etc. [1][2]. The CDMA technology, based on direct-sequence spread spectrum (DSSS), has become the main stream of third-generation mobile communications system standards, including cdma2000 [3], UMTS, and LAS-CDMA. Direct-sequence spread spectrum is achieved by directly multiplying the input data sequences with a wideband spreading code (pseudonoise, PN, code) and then the resultant sequence is modulated. Especially in cdma2000 standards, different reverse link channels in the same cell are distinguished via different code phases of the same spreading code.

The processing gain of a DSSS system is defined as the amount of performance improvement achieved through the use of spread spectrum. It is often approximated as the ratio of the spread chip rate to the information rate [1]. Since the transmission bandwidth is fixed, the increase of information rate will result in decrease of the processing gain. Thus, if a high data rate is desired, the link performance may degrade severely.

Code shift keying (CSK), a real *M*-ary DSSS modulation scheme which combines *M*-ary data modulation with spread spectrum modulation, was proposed to increase the transmission efficiency of spread spectrum systems and to overcome the above-mentioned spreading gain vs. data rate limitation [4]-[9]. In [8]-[9], the CSK modulation is also known as code-phase-shift-keying (CPSK). The *M*-ary CSK modulation employs *M* signaling waveform codes, which are obtained from different code phase shifts of the same PN code, to represent the input data. In [8]-[9], the code phase difference between any two signaling waveform codes is an integer number of chips; moreover, each signaling waveform code is a full period version of the same PN code sequence (i.e. the symbol duration is equal to the PN code period). The *M* signaling waveform codes can be obtained through cyclic shifting the PN code sequence; hence it is also known as cyclic code shift keying (CCSK). If a maximal length sequence (m-sequence) is adopted and the cross-correlations between different signaling waveform codes are ignored, a performance similar to that of coherent *M*-ary FSK can be obtained.

In this work, a generalized form of *M*-ary DSSS modulation scheme is proposed and investigated. This scheme involves switching the code phase of the spreading code in accordance with the incoming data. Thus it is addressed as spreading code phase shift keying (SCPSK) modulation. Each signaling waveform code is assumed to be a partial period version of the same PN code sequence, i.e. the symbol duration is shorter than the PN code period. This *M*-ary SCPSK modulation scheme can improve the performance and flexibility of a DSSS system and is suitable for multi-user applications. The rest of this paper is organized as follows. In section II we will illustrate the *M*-SCPSK modulation and demodulation schemes, and propose the structures of transmitter and receiver. Section III concentrates on the analysis of system performance. Numerical examples and discussions are given in Section IV. Finally, this paper is concluded in Section V.

II. *M*-CPSK Modulation and Demodulation

Different from the conventional DSSS system, the SCPSK modulation scheme integrates the data modulation into spectrum spreading. The *M*-SCPSK transmitter groups the input data stream into *k*-bit ($k = \log_2 M$) symbols with a symbol duration of *T* which is *k* times the bit duration, and there are *M* ($M = 2^k$) different symbols. According to the value of the input symbol, a code sequence, with length *N* (referred to as the spreading factor, *N* >> *k*) chips and chip

This work was supported in part by the Ministry of Education, Taiwan, R.O.C., under Grant 89-E-FA04-1-4 and the National Science Council, Taiwan, R.O.C., under Grant NSC 93-2752-E-007-003-PAE.

duration T_c, is selected from a set of signaling waveform codes $\{\phi_i(t), i = 0, 1, \cdots, M-1\}$, and then it is modulated and transmitted. Thus, the data modulation and spectrum spreading are simultaneously accomplished. Each signaling waveform code $\phi_i(t)$ is a polar code sequence, which takes on the values ± 1, and is a partial version with a distinct code phase of the same PN sequence. It is noteworthy that the signaling waveform code length *N* is less than or equal to the period of the PN sequence *P*, i.e $N \le P$.

There are many choices of PN spreading sequences, including m-sequences, Gold sequences, Kasami sequences, and Walsh-Hadamard sequences. Owing to their good properties [1], m-sequences are widely used in CDMA communications systems, such as cdmaOne and cdma2000 systems [3]. In this work, the employed PN spreading sequence is assumed to be an m-sequence which is generated by a linear feedback shift-register generator using a primitive characteristic polynomial over GF(2). As in [3], the spreading codes with distinct code phases can be extracted from the same m-sequence generator via the employing of a set of "user code masks," which are used for user addressing and are unique to different users. However, in this work, we modify the m-sequence generator to generate the signaling waveform codes for *M*-SCPSK modulation.

Fig. 1(a) shows the signaling waveform codes generator for *M*-SCPSK modulation. The lower block is the conventional m-sequence generator, and the upper block carries out the masking operation on the shift-register contents. Assuming that the degree of the characteristic polynomial is m, the code mask consists of m bits, including k bits for data modulation and $m-k$ bits for user addressing as shown in Fig. 1(b). For a specific user, the $m-k$ bits for user addressing are fixed; the other k bits are determined by the input symbol. It should be noted that the code mask is used to determine the code phase of the output signaling waveform, not the actually transmitted signaling waveform. By using the shift-and-add property of m-sequence [1], the output sequence $\phi(t)$ is the same as that defined by the characteristic polynomial; however there is a certain code-phase offset between $\phi(t)$ and $c(t)$. Different input data symbols imply different code masks which inspire different code-phase offsets. During a specific symbol duration, the output signal is an $N-$chip code sequence determined by the input symbol. As shown in Fig. 2, if the input symbol of a specific user is i ($0 \le i \le M-1$), the corresponding k bits of code mask will be assigned by i; therefore, the inspired code-phase offset is O_i, and the unique output signaling waveform code will be $\phi_i(t)$. Similarly, if the input symbol of a specific user is j, the output signaling waveform code will be $\phi_j(t)$.

By employing the modified m-sequence generator, the transmitter of an *M*-SCPSK modulation system can be easily implemented, as shown in Fig. 3. For a multi-user transmitter such as the base station in mobile cellular system, only one conventional m-sequence generator is required. Every user can share the same conventional m-sequence generator and generate the unique signaling waveform code by the masking operation against its own code mask.

For the *M*-SCPSK demodulation, the architecture similar to the modulator is imposed. We assume that the carrier phase of desired signal is available and thus a coherent detection can be realized. Fig. 4 shows the receiver architecture for coherent *M*-SCPSK. The conventional m-sequence generator is assumed to be well synchronized to the received signal which can be achieved by an appendant pilot signal or conventional synchronization and tracking techniques. The receiver consists of a bank of M correlators matched to the M specific signaling waveform codes. Each correlator can be obtained via masking operation against the code mask, which corresponds to the user address and a possible symbol. Subsequently, the received signal is correlated with these M correlators and then a decision that picks the largest value will be made.

III. Performance Analysis

A. System Model

In this work, an additive white Gaussian noise (AWGN) channel with multi-user interference is considered. Assuming that the input data symbol of a desired user is i, the i^{th} signaling waveform code $\phi_i(t)$ is selected and transmitted. The received signal can be denoted as

$$r(t) = A\phi_i(t)\cos(2\pi f_c t + \theta) + n(t) + I(t) \qquad (1)$$

where A is the amplitude of the desired signal, f_c is the carrier frequency, θ is the random carrier phase uniformly distributed over $[0, 2\pi)$, $n(t)$ is the received AWGN with two-sided power spectral density $N_0/2$, and $I(t)$ is the received multi-user interference from other $K-1$ co-channel users.

The multi-user signals are also assumed to be SCPSK signals and have equal signal power as desired signal at the front end of the receiver. For an asynchronous system, $I(t)$ can be represented as

$$I(t) = \sum_{u=2}^{K} A\phi'_u(t-\tau_u)\cos(2\pi f_c t + \theta_u), \qquad (2)$$

where θ_u is the random phase, $\phi'_u(t)$ is the equivalent transmitted signaling waveform code of the u^{th} user in the desired symbol interval, and τ_u is the relative random time delay of the u^{th} user. When $K = 1$, there is no multi-user interference and it is the special case of single-user transmission.

B. Performance of Long Spreading Code

In general, we assume that the period of m-sequence P is much larger than the signaling waveform code length N, i.e. $P >> N$. It is referred to as the case of using long spreading code. In virtue of the properties of m-sequence, *M*-SCPSK modulation scheme can be regarded as a quasi-orthogonal modulation; however, as $P >> N$, the cross-correlations between signaling waveform codes cannot be ignored. For simplicity, the signaling waveform codes are regarded as purely random code for $P >> N$.

For the desired signaling waveform code $\phi_i(t)$ and an arbitrary receiving signaling waveform code $\phi_j(t)$, it is assumed that there are S ($S \in \{0,1,2,\cdots,N\}$) chips with the identical value (1 or -1), i.e. there are $N-S$ chips with opposite values, during a symbol duration *T*. Then the partial auto-correlation of an m-sequence becomes

$$y_{i,j} = \frac{1}{T}\int_0^T \phi_i(t)\phi_j(t)dt = (2S-N)/N \cdot \tag{3}$$

By ignoring the high frequency components, the decision variable at the j^{th} integrator output can be expressed as

$$R_j = \int_0^T r(t)\cos(2\pi f_c t + \theta)\phi_j(t)\, dt = \frac{AT}{2} y_{i,j} + W_j + I_j, \tag{4}$$

where

$$I_j = \frac{AT}{2}\sum_{u=2}^{K} \cos\theta'_u y'_{u,j}, \tag{5}$$

is the received multi-user interference, and W_j is a zero mean Gaussian random variable with variance $Var[W_j] = N_0 T/4$. For $j = i$, the decision variable corresponding to the transmitted signaling waveform code can be written as

$$R_i = \frac{AT}{2} + W_i + I_i \cdot \tag{6}$$

For simplicity, we apply the standard Gaussian approximation (SGA) method to approximate the multi-user interference. Thus I_j is approximated as a Gaussian random variable with mean zero and variance $Var[I_j] = A^2T^2(K-1)/3N$.

It is noted that the random variables $R_0, R_1, \cdots, R_{M-1}$ are correlated. We have the probability of error decision from symbol i to symbol j conditioning on y_{ij} as

$$P_{i,j|y_{i,j}} = Q\left(\sqrt{\frac{A^2T^2(1-y_{i,j})/2}{\frac{N_0 T}{4} + \frac{A^2T^2(K-1)}{3N}}}\right) = Q\left(\sqrt{(1-y_{i,j})\times\gamma}\right), \tag{7}$$

where $\gamma = 1\Big/\left[\frac{N_0}{E_b \log_2 M} + \frac{2(K-1)}{3N}\right]$ is the equivalent symbol energy to total interference power density ratio, and $E_b = A^2T/2\log_2 M$ is the energy per bit. Consequently, the probability of error decision from symbol i to symbol j, conditioning on a specific value of *S*, becomes

$$P_{i,j|S=s} = Q\left(\sqrt{\frac{N-s}{(\frac{N}{2\log_2 M})(\frac{N_0}{E_b}) + \frac{K-1}{3}}}\right) \cdot \tag{8}$$

Averaging over all possible values of S, $P_{i,j}$ becomes

$$P_{i,j} = \frac{1}{2^N}\sum_{s=0}^{N} C_s^N Q\left(\sqrt{\frac{N-s}{(\frac{N}{2\log_2 M})(\frac{N_0}{E_b}) + \frac{K-1}{3}}}\right) \cdot \tag{9}$$

Finally, the actual symbol error probability will be upper-bounded by the union bound, i.e.

$$P_{error} \le \frac{M-1}{2^N}\sum_{s=0}^{N} C_s^N Q\left(\sqrt{\frac{N-s}{(\frac{N}{2\log_2 M})(\frac{N_0}{E_b}) + \frac{K-1}{3}}}\right) \cdot \tag{10}$$

C. Performance of Short Spreading Code

In contrast to the case of $P >> N$, if $P = N$ we refer to it as the case of using short spreading code. In such a case, the signaling waveform code is a full period of the m-sequence and the cross-correlation between different signaling waveform codes becomes a determinate value, i.e. $y_{i,j} = -1/N = -1/P$. It is noted that the multi-user signals will contribute the same amount of interference to each decision variable; thus the multi-user interference term in (8) will vanish. From (8), with setting $S = (N-1)/2$ and K = 1, we have

$$P_{i,j} = Q\left(\sqrt{(\frac{N+1}{N})(\log_2 M)(\frac{E_b}{N_0})}\right), \tag{11}$$

Finally, the union bound of symbol error probability for using short spreading code becomes

$$P_{error} \le (M-1)Q\left(\sqrt{(\log_2 M)(\frac{E_b}{N_0})}\right) \cdot \tag{12}$$

IV. Results and Discussions

In the simulation results, two practical m-sequences are employed. For the case of using short spreading code, the code period is $N = P = 2^7 - 1 = 127$ and the chosen characteristic polynomial is

$$p_1(x) = 1 + x^3 + x^7 \cdot \tag{15}$$

While, for the case of using long spreading code, an m-sequence with period $P = 2^{42} - 1$ is selected and the chosen characteristic polynomial is [3]

$$p_2(x) = 1 + x^1 + x^2 + x^3 + x^5 + x^6 + x^7 + x^{10} + x^{16} + x^{17} + x^{18} + x^{19} + x^{21} + x^{22} + x^{25} + x^{26} + x^{27} + x^{31} + x^{33} + x^{35} + x^{42} \cdot \tag{16}$$

Accordingly, if $M = 32$ (i.e. $k = 5$) and $p_1(x)$ is employed, the number of bits for user addressing in code mask is 2 and only $2^2 - 1 = 3$ users can be accommodated in this system; however, if $p_2(x)$ is employed, the number of bits for user addressing in code mask is 37 and $2^{37} - 1$ users can be accommodated, which is much larger than that in the case of using short spreading code.

Fig. 5 shows the union bound of symbol error probability P_{error} versus E_b/N_0 with $N = 127$ and $M = 32$. In the case of using short spreading code, the multi-user number is assumed to be $K = 1$, 2, or 3. The numerical and simulation results show that the multi-user interference brings almost no influence on the *M*-SCPSK symbol error performance; this is owing to that the cross-correlation between signaling waveform codes approaches zero as $N \to \infty$ since full period of an m-sequence is used. The union bound with either considering or ignoring correlations between decision variables shows acceptable accuracy. This implies that ignoring the correlations between decision variables will not miscalculate the symbol error probability. Furthermore, the *M*-SCPSK system with short spreading code employed can be regarded as an orthogonal modulation system if N is large enough. In the case of using long spreading code, the multi-user number is assumed to be $K=1$, 4, or 8. As compared to the case of using short spreading code, the performance is slightly degraded for $K=1$. This is mainly due to the fact that the cross-correlations between different signaling waveform codes are now the partial period auto-correlations of an m-sequence. Furthermore, P_{error} is severely degraded by multi-user interference. Again, it is owing to the cross-correlations between different users' signaling waveform codes. The union bound with considering correlations between decision variables shows acceptable accuracy; however, the union bound with ignoring these correlations reveals a noticeable error even when E_b/N_0 is large.

Fig. 6 shows the union bound of P_{error} versus E_b/N_0 with $M=32$, $K=1$ and several values of N. For the case of using short spreading code, the increase of N makes no significant improvement, since it is an orthogonal-modulation-like system. On the contrary, N makes significant impact on P_{error} for the case of using long spreading code. As $N \to \infty$, the P_{error}, with long spreading code employed, approaches that with short spreading code employed. In other words, the *M*-SCPSK system acts like an orthogonal modulation system for $N \to \infty$.

Shown in Fig. 7 is the plot of the union bound of P_{error} versus E_b/N_0 with $N=127$, $K=1$ and several values of M. The results for coherent *M*-FSK and BPSK are obtained by using the union bound and exact solution, respectively. The results show that the performance and spectral efficiency, which is proportional to $\log_2 M/N$, can be greatly enhanced by increasing M. In brief, power efficiency and bandwidth efficiency of *M*-SCPSK modulation can be improved by using high order modulation. The major cost that should be taken is the increase of system complexity in the receiver.

V. CONCLUSIONS

In this work, we have proposed a generalized form of spreading-code-phase-shift-keying modulation scheme. The performance of *M*-SCPSK modulation with coherent detection is investigated in an AWGN channel exposed to multi-user interference. Two types of *M*-ary SCPSK modulation—one employing short spreading codes and the other employing long spreading codes—were examined, and the union bounds of the corresponding symbol error probabilities have been derived. Additionally, feasible designs of the *M*-SCPSK transmitter and receiver were well presented.

From the numerical results, we found that the performance of coherent *M*-SCPSK with short spreading code employed is comparable to that of coherent *M*-FSK and the correlations between different decision variables can be ignored. However, for the coherent *M*-SCPSK with long spreading code employed, the performance is worse than that of coherent *M*-FSK and the correlations between different decision variables should be taken into consideration; ignoring the correlations will induce unacceptable error. For multi-user applications, we discovered that multi-user interference has only ignorable influence on the performance of the *M*-SCPSK with short spreading code employed. However, for the case of using long spreading code, the symbol error performance is severely degraded by the multi-user interference. This performance degradation can be further improved by other techniques which are left for further study. Furthermore, in regard to system flexibility, the *M*-SCPSK system with long spreading code employed is more flexible and can accommodate a larger number of users in the system.

REFERENCE

[1] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread-Spectrum Communications*. New Jersey: Prentice-Hall, 1995.

[2] A. J. Viterbi, *CDMA-Principles of Spread Spectrum Communication*. New York: Addison-Wesley, 1995.

[3] Physical layer standard for cdma2000 spread spectrum systems, 3GPP2 C.P0002-A, 3rd Generation Partnership Project 2, October 1999.

[4] A. Hammer and D. J. Shaefer, "Performance analysis of *M*-ary code shift keying in code division multiple access systems," in *Proceeding of IEEE International Conference on Communications*, 7E. 2. 1, 1982.

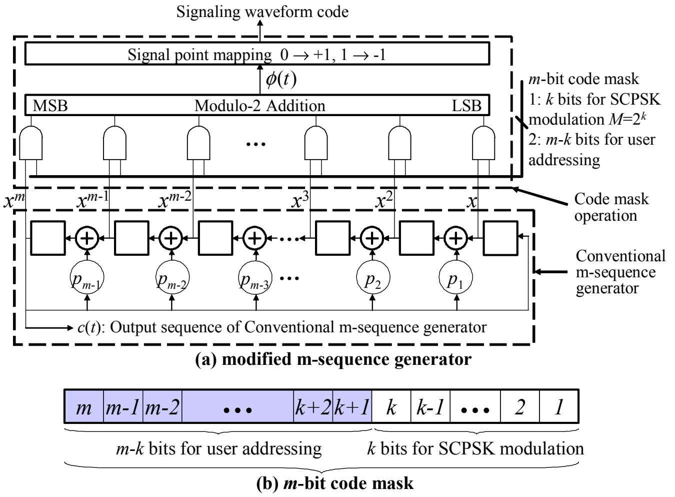
[5] L. Guo, N. Kuroyanagi and N. Suehiro, "Transmission efficiency of code shift keying," in *Proceeding of IEEE Conference on Military Communications*, pp. 35 – 40, Oct. 1993.

[6] A. G. Burr, "Capacity improvement of CDMA systems using *M*-ary code shift keying," in *Proceeding of IEEE Sixth International Conference on Mobile Radio and Personal Communications*, pp. 63 – 67, Dec. 1991.

[7] A. G. Burr, "Coded *M*-CSK for CDMA systems," *IEE Colloquium on Spread Spectrum Techniques for Radio Communication Systems*, pp. 2/1 – 2/6, June 1992.

[8] A. Y.-C Wong and V. C. M. Leung, "Code-phase-shift keying: a power and bandwidth efficient spread spectrum signaling technique for wireless local area network applications," in *Proc. IEEE 1997 Canadian Conference on Electrical and Computer Engineering,* vol. 2, pp. 478 – 481, 1997.

[9] S. K. S. Chan and V. C. M. Leung, "An FPGA receiver for CPSK spread spectrum signaling," *IEEE Trans. Consumer Electronic*, vol. 45, pp. 181-191, Feb. 1999.

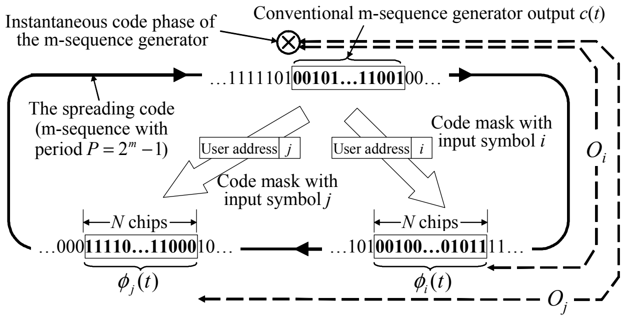


Signaling waveform code
Signal point mapping $0 \to +1$, $1 \to -1$
$\phi(t)$
MSB
Modulo-2 Addition
LSB
m-bit code mask
1: *k* bits for SCPSK modulation $M=2^k$
2: *m-k* bits for user addressing
Code mask operation
x^m
x^{m-1}
x^{m-2}
x^3
x^2
x
p_{m-1}
p_{m-2}
p_{m-3}
p_2
p_1
Conventional m-sequence generator
$c(t)$: Output sequence of Conventional m-sequence generator
m
m-1
m-2
k+2
k+1
k
k-1
2
1
m-k bits for user addressing
k bits for SCPSK modulation

(a) modified m-sequence generator

(b) *m*-bit code mask

Fig. 1. The modified m-sequence generator and *m*-bit code mask.


Instantaneous code phase of the m-sequence generator
Conventional m-sequence generator output $c(t)$
…1111101**00101…11001**00…
The spreading code (m-sequence with period $P = 2^m - 1$)
User address | j
User address | i
Code mask with input symbol i
Code mask with input symbol j
O_i
N chips
N chips
…000**11110…11000**10…
…101**00100…01011**11…
$\phi_j(t)$
$\phi_i(t)$
O_j

Code length of signaling waveform: $N << P$

$\phi_i(t)$: signaling waveform code for input symbol i

$\phi_j(t)$: signaling waveform code for input symbol j

O_i: code-phase offset for code mask with input symbol i

O_j: code-phase offset for code mask with input symbol j

Fig. 2. Signaling waveform codes and code-phase transitions.

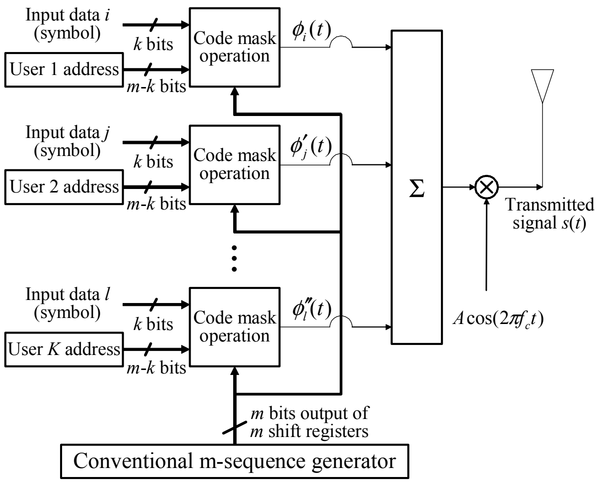


Fig. 3. The M-CPSK multi-user transmitter.

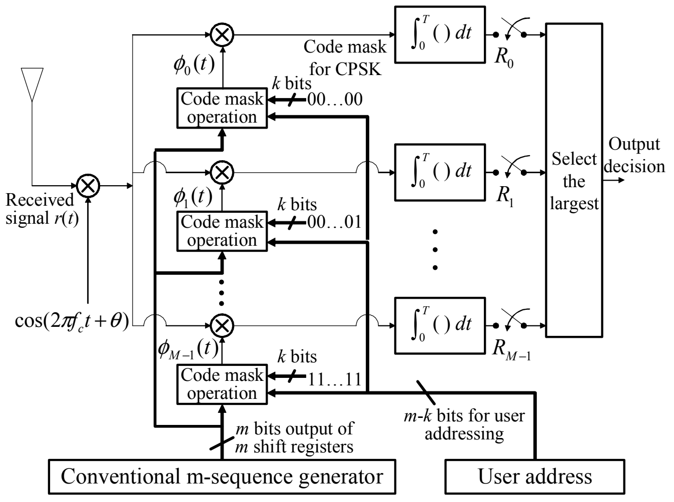


Fig. 4. The M-CPSK coherent receiver.

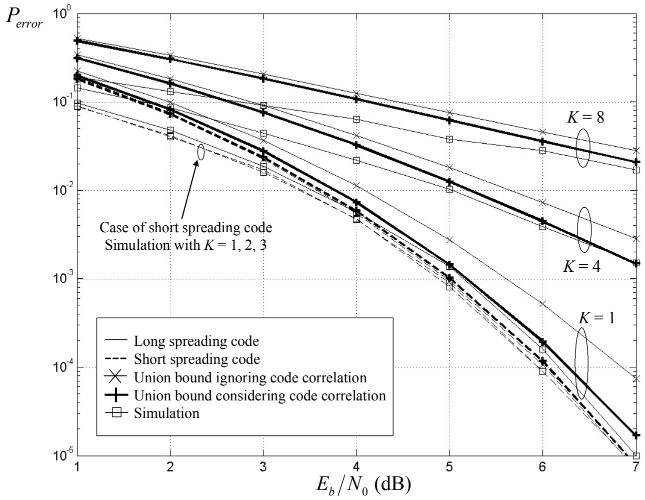


Fig. 5. P_{error} versus E_b/N_0 with N = 127, and M = 32.

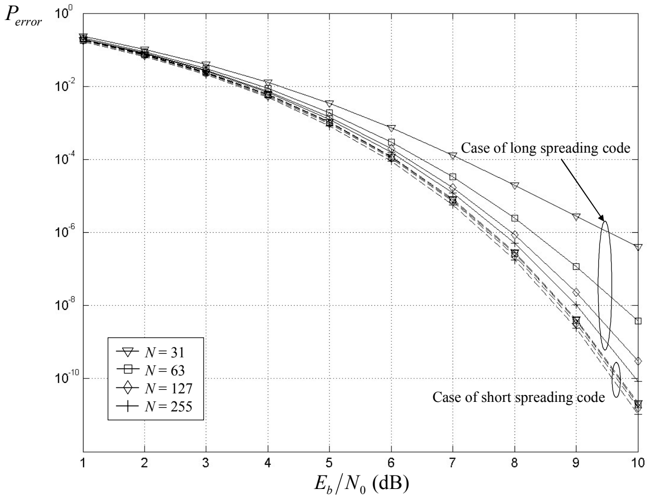


Fig. 6. P_{error} versus E_b/N_0 with M = 32, and K = 1.

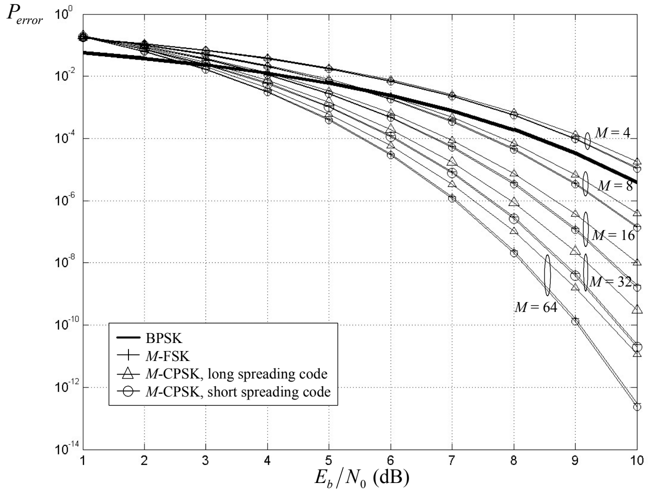


Fig. 7. Performance comparison between M-CPSK, M-FSK and BPSK.