ISSN: 1945-3078 (Print), 1945-3086 (Online) Volume 1, Number 2, July 2009

Wireless Sensor Network

Chief Editor : Kosai Raoof

www.scirp.org/journal/wsn/

Scientific Research

Journal Editorial Board

ISSN 1945-3078 (Print) ISSN 1945-3086 (Online)

http://www.scirp.org/journal/wsn/

Wireless Sensor Network, **2009, 2, 61-121**
Published Online July 2009 in SciRes (http://www.SciRP.org/journal/wsn/).

TABLE OF CONTENTS

Volume 1 Number 2 **July 2009**

Wireless Sensor Network (WSN)

Journal Information

SUBSCRIPTIONS

The *Wireless Sensor Network* (Online at Scientific Research Publishing, www.SciRP.org) is published quarterly by Scientific Research Publishing, Inc., USA.

E-mail: wsn@scirp.org

Subscription rates: Volume 1 2009

Print: $50 per copy.

Electronic: free, available on www.SciRP.org.

To subscribe, please contact Journals Subscriptions Department, E-mail: wsn@scirp.org

Sample copies: If you are interested in subscribing, you may obtain a free sample copy by contacting Scientific Research Publishing, Inc at the above address.

SERVICES

Advertisements

Advertisement Sales Department, E-mail: wsn@scirp.org

Reprints (minimum quantity 100 copies)

Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.

E-mail: wsn@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:

E-mail: wsn@scirp.org

Wireless Sensor Network, 2009, 2, 61-121
doi:10.4236/wsn.2009.12010 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

The Application of Optical CDMA-Based Fiber Radio Networks in Wireless Sensor Networks

Chao-Chin YANG
Department of Electronic Engineering, Kun Shan University, Tainan, Taiwan, China
Email: *ccyang@mail.ksu.edu.tw*
Received April 29, 2009; *revised May* 8, 2009; *accepted May* 9, 2009

Abstract

One fiber radio scheme using shifted prime codes for interference elimination is proposed for optical code-division multiple-access (OCDMA) network. By taking advantage of the cyclic property of the shifted prime codes in the same code groups, the proposed compact decoder is low cost and suitable to be used in the task manager node in the applications of wireless sensor networks. The performance comparison for several OCDMA-based fiber radio networks is also given to clarify the advantage of the proposed one.

Keywords: Optical Code-Division Multiple-Access (OCDMA), Passive Optical Network (PON), Arrayed Waveguide Grating (AWG), Prime Code

1. Introduction

In today's mobile radio networks, the provision of broadband services between a large number of remote base stations (RBSs) is required. To develop simple and low-cost RBSs for high rate transmission, the combination of mobile radio and optical fiber network is one possible solution since the components for complicated operation of signal processing can be moved to control base station (CBS) [1]. In the case of wireless sensor networks (WSNs) nowadays, the condition may be similar [2]. The sensor nodes need to route the collected data back to the sinks, and the sinks may communicate with the task manager node via networks for long distance transmission. Therefore, fiber radio networks can also provide the communication between the sinks (e.g. RBS) and the task manager node (e.g. CBS) to simplify the design of the sinks.

Code division multiple access (CDMA) techniques were investigated for optical network applications during the last 20 to 25 years. These techniques allow many users to access the common channel asynchronously and securely, and, since dedicated time or wavelength slots do not have to be allocated, high statistical multiplexing gain can be offered even in bursty traffic. These characteristics distinguish CDMA from other multiplexing schemes such as time division multiple access (TDMA) and wavelength division multiple access (WDMA). Several OCDMA-based fiber radio schemes were proposed [3−6]. Among these, fiber radio schemes based on spectral-amplitude-coding (SAC) had the advantages of multiple access interference (MAI) elimination and the avoidance of sampling before the optical encoding process [4−6].

A code that can be used in the SAC scheme is denoted as (N, w, λ), where N is code length, w is code weight, and λ is the in-phase cross correlation [6]. Originally unipolar m-sequences were used in the SAC-based networks [7], and, to enhance the network performance, code sequences with complementary code keying ability such as Hadamard codes [7] could be used. However, since these two code families had code weights equivalent to half of the corresponding codelengths, they induced serious phase-induced intensity noise (PIIN) in the photodiodes of the decoders. To suppress the phase-induced intensity noise in the photodiodes of the decoders, several codes with ideal λ ($\lambda = 1$) were proposed, such as Balanced Incomplete Block Design (BIBD) codes [8]. The former had fully cyclic property and $N \times N$ arrayed waveguide grating (AWG) can be used as compact coding devices [9,10]. However, when the total number of

* Phone: 886-6-2050521 ext 3509, Fax: 886-6-2050250.

users in the network increased, a large number of AWG ports were required and the effect of crosstalk was more serious.

To enlarge the network capacity further, new coding schemes using codewords with small λ and having the two-code keying ability can be developed. In [11], one coding scheme was proposed and code sequences with λ=0 or 1 can be used for two-code keying of information bits. By the use of the cyclic property for shifted prime (SP) codes, compact encoder structure based on cascaded AWGs and simplified decoder based on fiber Bragg grating (FBG) were also proposed for the application of passive optical networks (PONs). The optical line terminal (OLT) encoder based on cascaded AWGs had the advantage that only $p \times p$ and $1 \times p$ AWGs were needed in the encoder for SP codes with codelength p^2.

In this paper, the application of SAC schemes using SP codes in radio fiber network is explored. By the modification of the encoder/decoder pair for the downstream transmission in the PON in [11], the encoder/decoder implementation in the upstream transmission (from RBSs to CBS) for the fiber radio network can be obtained. Performance analysis takes thermal noise and PIIN into account. It shows that, as compared to several fiber radio schemes based merely on SAC, the use of SP codes allows larger number of RBSs to access the network simultaneously under a given carrier-to-noise ratio (CNR).

2. Fiber Radio Network Based on SP Code

The proposed SAC OCDMA-based fiber radio network consists of one CBS, the RBSs connected , and one $1 \times M$ splitter connected to the CBS and RBSs by optical fibers, where M is the total number of RBSs in the network. Each RBS is assigned two SP codewords for optical transmission. The RBSs receive the radio signals from the mobile terminals or wireless sensor networks, and, after simply down-conversion process, encode these radio signals with the assigned SP codewords. The resulting encoded optical signals are combined in the $1 \times M$ coupler and transmitted to the CBS through one fiber. Finally the decoders in the CBS receive the signals from the $1 \times M$ coupler and extract the radio signal of each RBS for further processing.

Table 1. The SP codewords for p=3.

e	f	$\mathbf{C}_{e,f}$	$\mathbf{X}_{e,f}$
0	0	0 0 0	100 100 100
0	1	1 1 1	010 010 010
0	2	2 2 2	001 001 001
1	0	0 1 2	100 010 001
1	1	1 2 0	010 001 100
1	2	2 0 1	001 100 010
2	0	0 2 1	100 001 010
2	1	1 0 2	010 100 001
2	2	2 1 0	001 010 100
3	0		111 000 000
3	1		000 111 000
3	2		000 000 111

In the following, the property of SP codes is briefly reviewed: One SP code $X_{e,f}$'s with length $L = p^2$ used in this paper can be obtained from one modified stuffed shifted prime (MSSP) code with the same parameter p in [12] by eliminating the extra codeword and the last p+1 chips of each remaining MSSP codewords. For example, Table 1 shows the SP codewords for p = 3, and ,from [12], the cross-correlation between $X_{e,f}$ and $X_{e',f'}$ is

$$X_{e,f} \odot X_{e',f'} = \begin{cases} p, e = e', f = f', \\ 0, e = e', f \neq f', \\ 1, \ e \neq e', \end{cases} \quad \textbf{(1)}$$

and thus

$$X_{e,f_1} \odot X_{e',f'} - X_{e,f_2} \odot X_{e',f'} = \begin{cases} p, e=e', f_1=f', \\ -p, \ e=e', f_2=f', \\ 0, \text{ otherwise.} \end{cases} \quad \textbf{(2)}$$

where $\odot$ is the dot-product of two vectors. Since the RBSs can use the code property in Equation (2) for MAI elimination, any two distinct SP codewords with the same value of e can be assigned to one specific RBS for the encoding of radio signals, and the value of M is $(p+1)(p-1)/2$. However, the two codewords used for decoding in conventional SAC schemes are always complement to each other. Therefore, this coding scheme has the advantage of security enhancement in the SAC-based fiber radio networks since the eavesdroppers need to guess two codewords used for encoding.

3. Implementation of Coder/Decoder

The RBSs of the proposed fiber radio network use the encoder shown in Figure 2, which is designed for the first SP codewords X_{e,f_1} = [1 0 0 1 0 0 1 0 0] and the second SP X_{e,f_2} = [0 1 0 0 1 0 0 1 0] of this RBS. Light from the broadband light source is incident to one FBG-based encoder cascading with another FBG-based encoder. Suppose that each wavelength chip has spectral width Δλ, the 3dB bandwidths of the FBGs in these two FBG-based encoders should also be Δλ. The center wavelengths of the FBGs in these two FBG-based encoders are { λ_0, λ_3, λ_6 }(corresponding to X_{e,f_1}) and { λ_1, λ_4, λ_7 }(corresponding to X_{e,f_2}), respectively. Therefore, the first (or second) codeword of this RBS is generated in

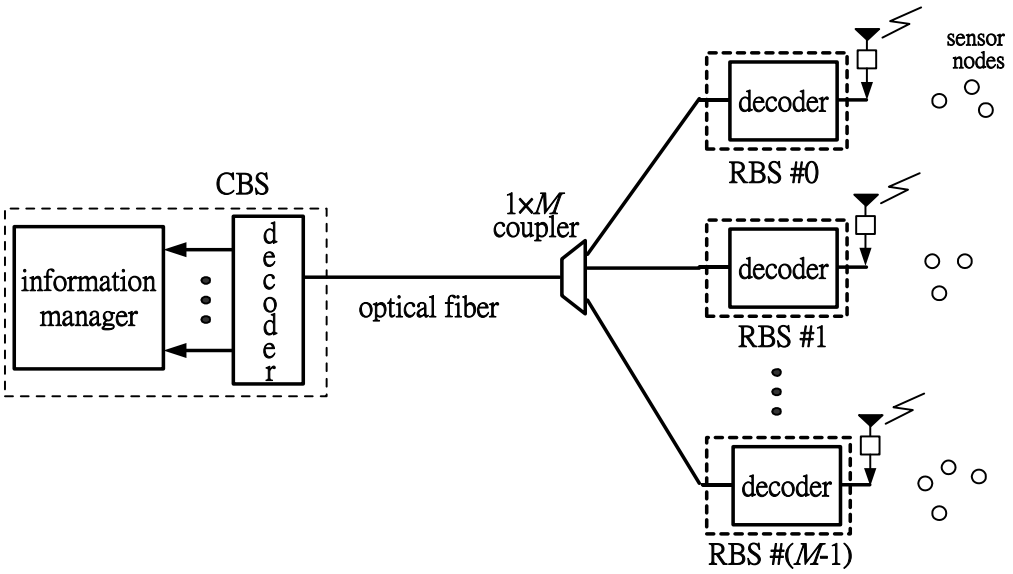


Figure 1. The proposed fiber radio network.

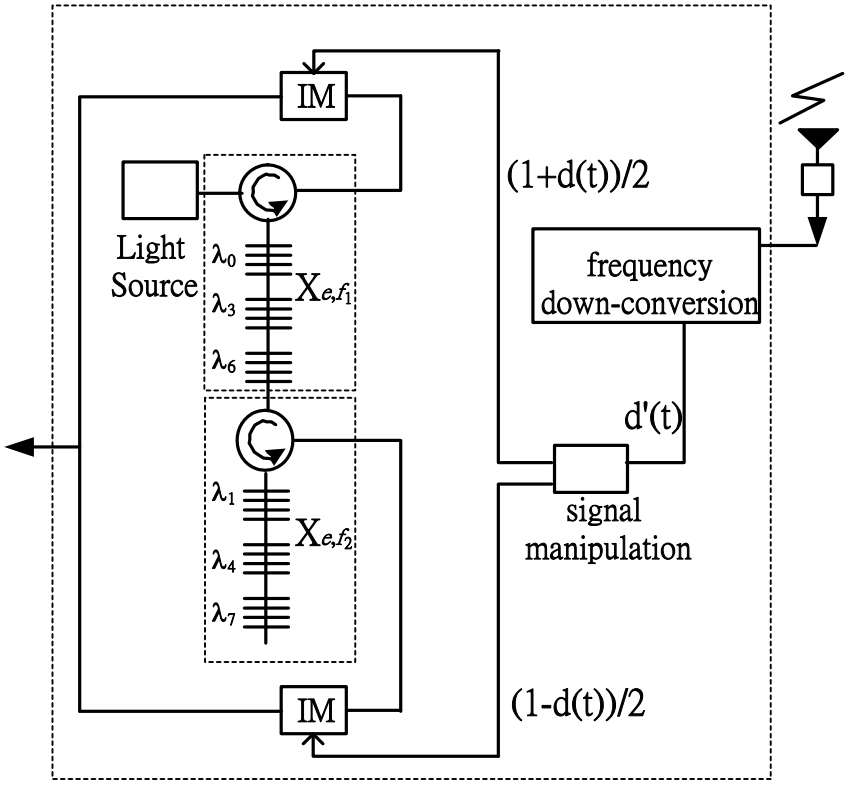


Figure 2. The encoder in the RBS.

the input port of the upper (or lower) intensity modulator (IM) and is intensity-modulated by the signal (1+ $d(t)$)/2 (or (1-$d(t)$)/2). Here $d(t)$ is the radio signal normalized from the radio signal $d'(t)$ such that max $|d(t)| = 1$. To make the receiver bandwidth B as small as possible, the frequency of the received radio signal from the antenna can be first converted to lower frequency before the signal manipulation process.

The decoder used to decode all the SP codewords from the RBSs with $p = 3$ is shown in Figure 3, which is simi-

lar to the AWG-based encoder in [11]. Two kinds of AWGs are used in this encoder: One is coarse AWG with free spectral range $FSR_c = p^2 \Delta\lambda$ and channels bandwidth $p\Delta\lambda$, and the other is fine AWG with $FSR_f = p\Delta\lambda$ and channels bandwidth $\Delta\lambda$ [9]. Note that the center wavelengths of the first output port at the coarse AWG are $\lambda_c - FSR_f \pm r\ FSR_c$ (r =0,1,…), thus there is only one coarse AWG channel at each coarse AWG output port within the encoded spectrum. However, for the fine AWG, the center wavelengths of the AWG channels that route from the first input port to the first output port are $\lambda_0 = \lambda_c - (p\text{-}1)\ \Delta\lambda/2 \pm sFSR_f$ (s = 0,1,…), thus there are p wavelength channels within the encoded spectrum that communicates between any input-output port pair.

The general principle to construct the decoder for one SP code family $X_{e,f}$'s with code length $L=p^2$ is described as follow: the single light source is connected to one 1× (p + 1) splitter. The output port # e (e = 0 ~ p - 1) of the splitter is connected to one coarse AWG cascaded with one fine AWG and the output port #i of the coarse AWG is connected to the input port # $C_{e,0}(i)$ of the fine AWG. Take the encoder (p = 3) in Figure 1(a) as example, light emitted from the filtered light source contains wavelength components $\lambda_0 \sim \lambda_8$ ($\Delta\lambda=\lambda_i-\lambda_{i+1}$) and is split into four. At these four encoder branches, wavelength chips λ_{0+bp}, λ_{1+bp} and λ_{2+bp} are demultiplexed and appear at the output port # b of the coarse AWGs. Due to the connections between the coarse and fine AWG ports described above, $\lambda_{f\oplus 0}$, $\lambda_{p+\ (f\oplus 1)}$ and $\lambda_{2p+(f\oplus 1)}$ (f = 0, 1, ….,p-1) corresponding to $X_{0,f}$ appears at the fine AWG output port # f at encoder branch #0 and $\lambda_{f\oplus 1}$, $\lambda_{p+(f\oplus 0)}$ and $\lambda_{2p+(f\oplus 1)}$ corresponding to $X_{1,f}$ appears at the fine AWG output port # f at encoder branch #1($\oplus$ denotes the modulo-p addition). Let's take the generation of $X_{1,0}$ as example. Wavelengths λ_0, λ_3 and λ_6 from the light source are demultiplexed by the coarse AWG at encoder branch #1 and they appear at the output port #0, #1 and #2, respectively. The output ports # 0, #1 and #2 of this coarse AWG are connected to the input ports #0, #1 and #2 of one fine AWG according to $C_{1,0}$ (= [0 1 2]), respectively, and thus λ_0, λ_3 and λ_6 appear at the input ports #0, #1 and #2 of this fine AWG, respectively. Since wavelength λ_g from the input port # i appears at output port # ($g \oplus (-i)$) for each fine AWG [9], λ_0, λ_3 and λ_6 all appears at output port #0 of this fine AWG which corresponds to $X_{1,0}$ = [100100100].

In this way, the first p^2 codewords can be obtained at the fine AWG output ports of the encoder. Since the wavelengths appeared at the coarse AWG output port # f of the last encoder branch correspond to $X_{3,f}$, the remaining p additional codewords in group #3 are obtained. Balanced detection for each RBS's signal can be accomplished with one balanced detector shown in Figure 3. The upper and lower photo-diodes of the balanced detector for one RBS are connected to the fine AWG output ports in the decoder according to the first and second codewords assigned to that RBS, respectively. Thus each balanced photo-detector can generate electrical current proportional to the value of $X_{e,f_1} \odot S - X_{e,f_2} \odot S$ and the MAI elimination scheme in Equation (2) is realized, where S is the summation signals from all RBSs.

4. Performance Analysis

Assume the light source in the RBS is unpolarized and has flat spectrum for encoding with bandwidth $\Delta\nu$ and magnitude Psr/$\Delta\nu$, where Psr is the effective power received by the photo-diodes in the decoder of the CBS. The average carrier power in the balanced detector is $P_c=R^2 P_{sr}^2/(2p^2)$, where R is the responsivity of the photo-diodes. For the fiber radio network adopting SP codes, the PSD of the upstream optical signal arrived at the input port of the decoder in the CBS can be written as [5,10]

$$s(\nu)=\frac{P_{sr}}{\Delta\nu}\sum_{k=0}^{K-1}\{\frac{1}{2}(1+d_k(t))\sum_{i=0}^{L-1}x_{k,i}^{(0)}\Pi(\nu,i)+\frac{1}{2}(1-d_k(t))\sum_{i=0}^{L-1}x_{k,i}^{(1)}\Pi(\nu,i)\} \quad \textbf{(3)}$$

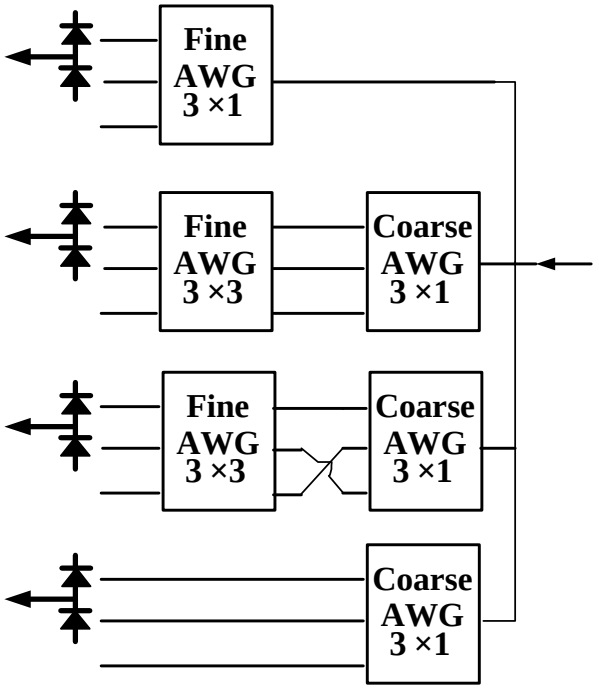


Figure 3. The decoder for SP codes in the CBS.

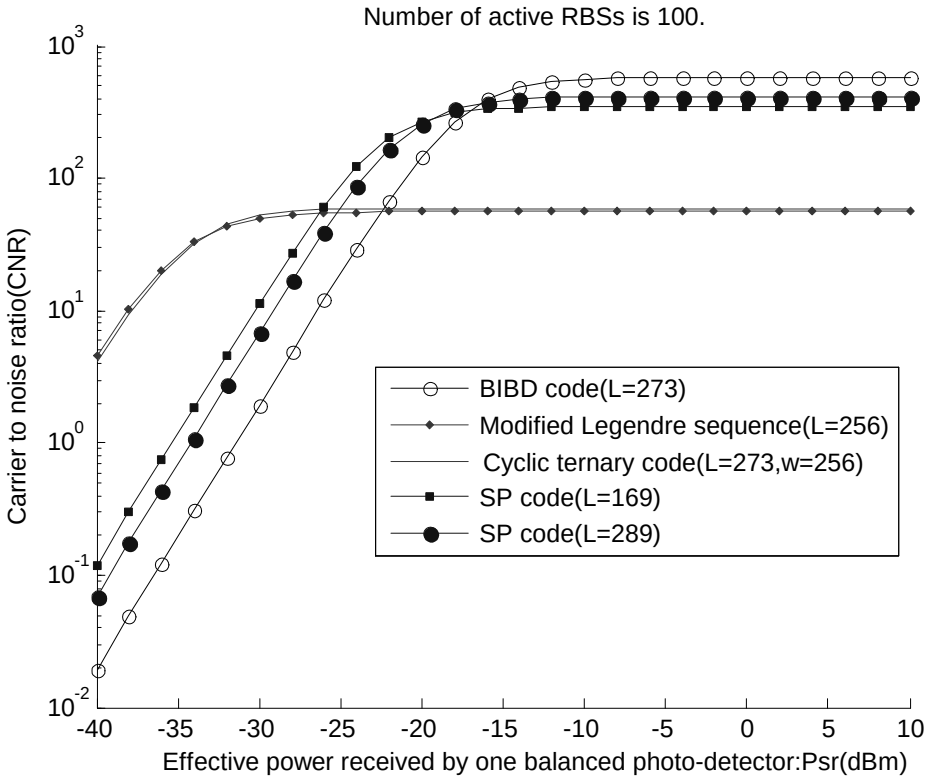


Figure 4. CNR vs. Psr.

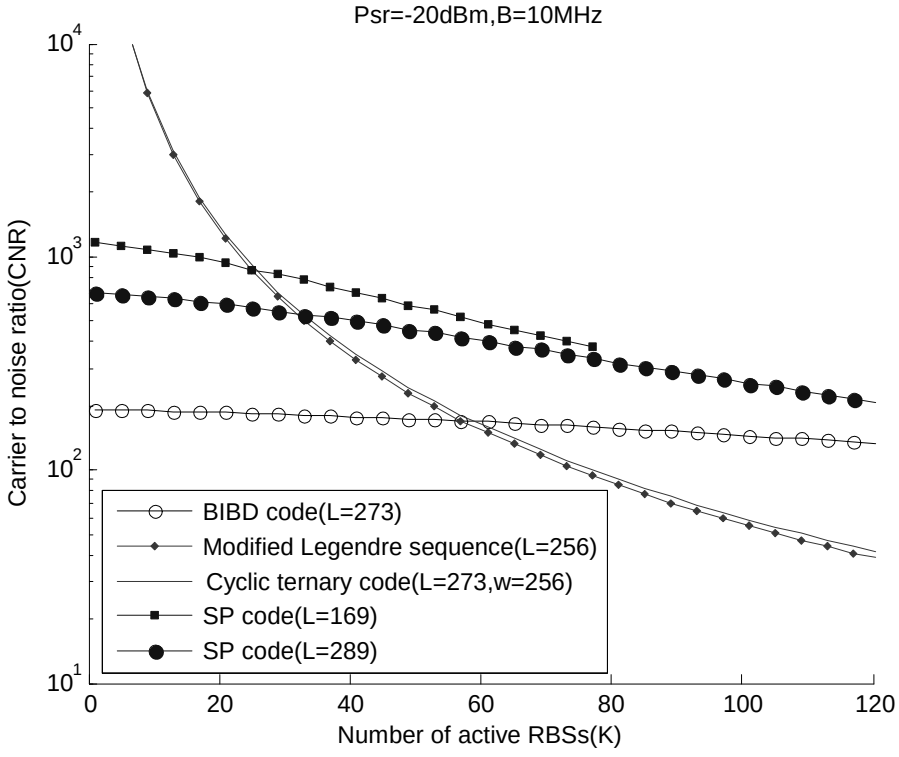


Figure 5. CNR vs. *K*.

where

$$\Pi(\nu,i)=\begin{cases}1, i \le L(\dfrac{\nu-\nu_0}{\Delta\nu}+\dfrac{1}{2}) \le i+1\\ 0, \text{ otherwise.}\end{cases} \quad \textbf{(4)}$$

Here K is the number of active RBSs and $d_k(t)$ is the normalized radio signal of k-th RBS. To simplify the notation, the i-th chip of the first and second codewords of k-th RBS are denoted as $c_{k,i}^{(0)}$ and $c_{k,i}^{(1)}$, respectively. Assume the sensor nodes use modulation with constant envelope such as phase shift keying. If at this time only one sensor node transmit signal to the 0-th RBS, the radio signal is approximately a sine wave and the assumption of time average amplitude $< d_0(t) > = 0$ and time average power $< d_0(t)^2> = 0.5$ can be used [5,6]. Thus the power of phase-induced intensity noise (PIIN) can be obtained [5,10]:

$$\begin{aligned}< I_{\text{PIIN}}{}^2 >&= BR^2\left[\int_0^\infty G_0^2(\nu)d\nu+\int_0^\infty G_1^2(\nu)d\nu\right]\\ &=\frac{BR^2}{p^2\Delta\nu}P_{sr}{}^2(\frac{3}{4}p+\frac{7}{2}K_{eff}+\frac{2}{p}K_{eff}(K_{eff}-1))\end{aligned} \quad \textbf{(5)}$$

where B is noise-equivalent electrical bandwidth of the receiver. The effective number of interfering active RBS is [11]

$$K_{eff}=K-1-\left\lfloor\frac{K-1}{p}\right\rfloor \quad \textbf{(6)}$$

For wireless sensor networks using the 915MHz industrial, scientific, and medical (ISM) band for radio frequency transmission, the data rate may be just about 10k Hz. [1] Therefore, the carrier frequency of the radio signals received by the RBSs can be converted to about 10MHz by the frequency down-conversion circuit in the Figure 2, and B is set to 10MHz here. Other parameters used in the performance analysis are R = 0.8A/W, $\Delta\nu$ = 7.5 THz, and ν_0 = 193.1THz and the power of thermal noise is $<I_{th}{}^2> = 16*10^{-17}A^2$.

The relationship between CNR and K are shown in Figure 4. Here the results for other three codes (e.g. balanced incomplete block design (BIBD) codes, modified Legendre sequences mentioned in [6] and cyclic ternary sequences with weight w =256 in [6] are also shown for comparison and their code lengths are approximately the same. (Note that Hadamard codes in [5,7] obtain the same performance as modified Legendre sequences, and thus only the results for modified Legendre sequences are shown here.) When P_{sr} is low, the influence of PIIN is less and the CNR for each code increases with P_{sr}. This increase gradually disappears due to the domination of PIIN. The CNR curves for modified Legendre sequences and cyclic ternary sequences are approximately the same and these two codes obtain better CNRs then BIBD and SP codes when P_{sr} is low. However, when P_{sr} is high, the situation is reversed. Though the BIBD code obtains slightly better CNRs than SP code with similar length when P_{sr} is larger than -15dBm, it requires very high transmitted power from the light sources in the RBSs when the insertion losses in the encoder/ decoder are taken into consideration. Since the insertion losses of FBG-based encoder in one RBS and the AWG-based decoder in the CBS are approximately 30dB [11], the CNR performance for P_{sr} approximately -20dBm seems more important when the transmitted power from the light source in the RBS is about 10dBm [1], and SP code with L=289 obtains better CNR at P_{sr} = -20dBm for K = 100.

The relationship between CNR and K for the codes in Figure 4 is shown in Figure 5, where P_{sr} is set at -20dBm. It is found that all four codes obtain lower CNR when K increases. When K is relatively small, modified Legendre sequences and cyclic ternary sequences obtain better CNRs then BIBD and SP codes. However, the situation is reversed when K is relatively large. Though SP codes obtain better CNRs than BIBD codes for P_{sr} = -20dBm, the corresponding cardinality is about half of that for BIBD codes with similar code lengths. Since SP codes can use AWG routers with smaller sizes for decoder implementation in the CBS as compared to BIBD codes, it is necessary to consider the tradeoffs between these codes for better network design.

5. Conclusions

One fiber radio scheme for the application of wireless sensor networks is proposed for optical code-division multiple-access (OCDMA) network. By the use of this fiber radio scheme, the design of sinks in the wireless sensor networks can be simplified, which has advantages for temporary deployment of the sensor networks. In addition, the proposed decoder in the CBS is compact and these features make the proposed fiber radio scheme more attractive for further research.

6. References

[1] H. Kim, J. M. Cheong, C. H. Lee, and Y. C. Chung, "Passive optical network for microcellular CDMA personal communication service," IEEE Photonics Technology Letters, Vol. 10, No. 11, pp. 1641–1643, November 1998.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubmmaiam, and E. Cayirci, "A Survey on sensor networks," IEEE Communication Magazine, Vol. 91, No. 8, pp. 102–114, August 2002.

[3] K. Tsukamoto, T. Higashino, T. Nakanishi, and S. Komaki, "Direct optical switching code-division multi-

ple-access system for fiber-optic radio highway networks," IEEE Journal of Lightwave Technology, Vol. 21, No. 12, pp. 3209–3220, December. 2003.

[4] B. K. Kim, S. Park, Y. Yeon, and B. W. Kim, "Radio-over-fiber system using fiber-grating-based optical CDMA with modified PN codes," IEEE Photonics Technology Letters, Vol. 15, No. 10, pp. 1485–1487, October 2003.

[5] T. Demeechai, "Noise-limited performance of spectral-amplitude-coding optical CDMA in fibre-optic radio highway networks," IEE Proceedings of Optoelectronics, Vol. 152, No. 5, pp. 269–273, October 2005.

[6] C. C. Yang, "Optical CDMA fiber radio networks using cyclic ternary sequences," IEEE Communications Letters, Vol. 12, No. 1, pp. 41–43, January 2008.

[7] M. Kavehrad and D. Zaccarin, "Optical code-division-multiplexed systems based on spectral encoding of noncoherent sources," IEEE Journal of Lightwave Technology, Vol. 13, No. 3, pp. 534–545, March 1995.

[8] X. Zhou, H. M. H. Shalaby, C. Lu, and T. Cheng, "Code for spectral amplitude coding optical CDMA systems," Electronics Letters, Vol. 36, pp. 728–729, April 13, 2000.

[9] B. Glance, I. P. Kaminow, and R. W. Wilson, "Applications of the integrated waveguide grating router," IEEE/OSA Journal of Lightwave Technology, Vol. 12, pp. 957–962, June 1994.

[10] C. C. Yang, "The application of spectral-amplitude-coding optical CDMA in passive optical networks," Optical Fiber Technology, Vol. 14, No. 2. pp. 134–142, April 2008.

[11] C. C. Yang, "Optical CDMA passive optical network using prime code with interference elimination," IEEE Photonics Technology Letters, Vol. 19, No. 7, pp. 516–518, April 2007.

[12] C. C. Yang, J. F. Huang, and T. C. Hsu, "Differentiated service provision in optical CDMA network using power control," IEEE Photonics Technology Letters, Vol. 20, No. 20, pp. 1664–1666, October 2008.

Wireless Sensor Network, **2009, 2, 61-121**
doi:10.4236/wsn.2009.12011 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

On the Performance of Blind Chip Rate Estimation in Multi-Rate CDMA Transmissions Using Multi-Rate Sampling in Slow Flat Fading Channels

Siavash GHAVAMI, Bahman ABOLHASSANI
School of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran
Email: *sghavami@ee.iust.ac.ir, abolhassani@iust.ac.ir*
Received March 7, 2009; *revised March* 19, 2009; *accepted March* 20, 2009

Abstract

This paper considers blind chip rate estimation of DS-SS signals in multi-rate and multi-user DS-CDMA systems over channels having slow flat Rayleigh fading plus additive white Gaussian noise. Channel impulse response is estimated by a subspace method, and then the chip rate of each signal is estimated using zero crossing of estimated differential channel impulse response. For chip rate estimation of each user, an algorithm which uses weighted zero-crossing ratio is proposed. Maximum value of the weighted zero crossing ratio takes place in the Nyquist rate sampling frequency, which equals to the twice of the chip rate. Furthermore, bit time of each user is estimated using fluctuations of autocorrelation estimators. Since code length of each user can be obtained using bit time and chip time ratio. Fading channels reduce reliability factor of the proposed algorithm. To overcome this problem, a receiver with multiple antennas is proposed, and the reliability factor of the proposed algorithm is analyzed over both spatially correlated and independent fading channels.

Keywords: Multi-Rate Sampling, Chip Time, Multi-Rate CDMA, Blind Estimation

1. Introduction

Direct-sequence code division multiple accesses (DS-CDMA) systems are nowadays of increasing importance in wireless cellular communications because of their inclusion in most of the proposals about both terrestrial and satellite based standards for third-generation (3G) wireless networks [1,2]. On the other hand, direct sequence spread spectrum (DS-SS) signals are well-known and are used in secure communication for their low probability of interception, their statistics are similar to those of noise; furthermore, they are usually transmitted below the noise level.

One of the salient features of 3G cellular systems is the capability of supporting transmission data as diverse as voice, packet data, low-resolution video, and compressed audio. Since these heterogeneous services produce digital information streams with different data rates, their implementation requires the use of multi-rate CDMA systems where each user may transmit his data at one among a set of available data rates. An easy way to view the multi-rate CDMA transmission is to consider the variable spreading length (VSL) technique where all users employ sequences with the same chip period; moreover, the data rate is tied to the length of the spreading code of each user. Another way to view a multi-rate CDMA transmission is to consider a constant spreading length where users employ sequences with different chip periods. In general, chip time and spreading sequence length of users can be selected variable.

Two recent systems are often applied in military systems based on spread spectrum. In the literature, different methods have been presented for chip time estimation, many number of those methods are based on cyclic cumulant method, which have been presented in litera-

tures [3−7]. Cyclic cumulant method in multi-rate and multi-user system doesn't exhibit good performance because cyclic frequencies of different users overlap and chip time estimation of different users is difficult. In [8] we proposed a blind chip time estimation algorithm based on multi-rate signal processing.

In this paper, blind chip time estimation technique is considered, which is based on channel impulse response estimation using singular value decomposition of estimated received signal covariance matrix. For spreading sequence estimation in slow flat Rayleigh fading channel plus additive white Gaussian noise (AWGN), length of code must be determined. Bit time and chip time must be estimated, for length of code estimation. In the proposed method in this paper, chip time is determined using multi-rate sampling of channel impulse response. Maximum number of zero crossing in differential of channel impulse response take places in sampling frequency twice of the chip rate. Therefore, number of zero crossing is obtained as a function of sampling rate. Then bit time is determined using fluctuations of correlation estimator. Code length of each user is determined using their corresponding chip times and bit times. Finally, performance of the proposed method is analyzed in fading channels and a receiver with multiple antennas is proposed for performance improvement in fading channels. Estimated parameter using this method is useful and applicable in noted systems in pervious paragraph in very low signal to noise ratio (SNR) (negative SNR in dB). This needs no prior knowledge about transmitter in the receiver side; it is typically the case in blind signals interception in the military field or in spectrum surveillance.

The remaining of this paper is organized as follows. In Section 2, System model are introduced. In Section 3, subspace methods for channel impulse response estimation are reviewed. Section 4 proposed blind chip time estimation based on multi-rate sampling. Blind bit time estimation is considered in Section 5. In Section 6, the performance of the proposed chip rate estimation algorithm will be analyzed over spatially correlated and non correlated flat fading channels. Simulation results are expressed in Section 7. Finally conclusions are performed in Section 8.

2. System Model

We consider the down link scenario of a multi-rate DS-CDMA network. The base station transmits signals and a mobile station receives a combined signal of K active users, which is given by

$$r(t)=\sum_{i=1}^{S}\sum_{k=1}^{K_i}\sum_{j=-\infty}^{+\infty}A_k d_k[j]h_{k,i}(t-jT_{s_i}-\tau_{k,i})+n(t), \quad \textbf{(1)}$$

where A_k, $d_k[j]$, $h_k(t)$, T_{s_i} and $\tau_{k,i}$ are the received amplitude, j^{th} data symbol, convolution of channel impulse response and spreading sequence waveform, symbol time and delay of k^{th} user in i^{th} rate respectively. Also, K_i is number of active users in i^{th} rate, S is the number of available rates, and $n(t)$ is additive white Gaussian noise. In the down link scenario, $\tau_{k,i}$ s are equal, and the $h_{k,i}(t)$ is expressed by

$$h_{k,i}(t)=L_i^{-1/2}\sum_{j=0}^{L_i-1}c_{k,i}\,p_i(t-jT_{c_i}) \quad \textbf{(2)}$$

where L_i is the spreading sequence length, $T_{c_i}=T_{s_i}/L_i$ is the chip time of i^{th} rate, $p_i(t)$ is explained in the next paragraph, $c_{k,i}[m]$ is the value of the m^{th} chip and i^{th} rate with $|c_{k,i}[m]|=1$. Data symbols $\{d_k[j]\}$ of different users have independent identically distributions (i.i.d.). In this paper, channel model is considered slow flat Rayleigh fading plus AWGN. We assume that the channel in the duration of one processing window remains approximately constant.

In (2), $p_i(t)$ is the total channel impulse response of the k^{th} user in i^{th} rate, it is convolution of the transmitter filter $e(t)$, channel filter $s(t)$ and receiver filter $g(t)$ i.e.

$$p_i(t)=e(t)*s(t)*g(t) \quad \textbf{(3)}$$

where these filters have unit energy. $s(t)=\beta_k(t)e^{j\varphi_k(t)}$ $\delta(t)$, β_k s have Rayleigh distribution and φ_k s have uniform distribution, * operation denotes convolution operation. We assume that

1) Channel model is slow flat Rayleigh fading plus AWGN.

2) Signal power is lower than noise power (SNR < 0 in dB).

3) A symbol time is equal to the spreading sequence length, i.e. $T_{c_i}=T_{s_i}/L_i$ (i.e. short spreading code).

4) Symbols have zero mean and are uncorrelated.

3. Estimation of Channel Impulse Response

Covariance matrix of the received signal can be written as

$$\mathbf{R}_r=E\{\mathbf{r}\mathbf{r}^H\}=E\{\mathbf{x}\mathbf{x}^H\}+E\{\mathbf{n}\mathbf{n}^H\}=\mathbf{R}_x+\mathbf{R}_n \quad \textbf{(4)}$$

where $\mathbf{r}$ is vector of received signal, $\mathbf{x}$ is vector of received signal when noise is absent , H denotes hermitian operation, and $\mathbf{n}$ is vector of additive white Gaussian

noise. Furthermore, $\mathbf{R}_x$ and $\mathbf{R}_n$ show covariance matrix of the received signal when noise is absent and noise covariance matrix, respectively. Covariance matrix of the received signal can be written as [9]

$$R = \sigma_n^2 \left\{ \sum_{k=0}^{K-1} \beta_k \left\{ (1-\alpha_k)\mathbf{v}_k^0 (\mathbf{v}_k^0)^* + \alpha_k \mathbf{v}_k^{-1} (\mathbf{v}_k^{-1})^* \right\} + \mathbf{I} \right\} \quad \textbf{(5)}$$

where, σ_n^2 is noise variance, $\alpha_k = \tau_k / T$, $\beta_k = \sigma_{sig_{(k,i)}}^2 T \Big/ \left(\sigma_n^2 T_{s_i} \right)$, in which $\sigma_{sig_{(k,i)}}^2$ is power of signal of k^{th} user in i^{th} rate. $\mathbf{v}_k^0$ and $\mathbf{v}_k^{-1}$ are normalized eigenvector of the estimated covariance matrix of received signal, they correspond to k^{th} asynchronous user in i^{th} rate, $\boldsymbol{T}$ is the sampling period, and $\mathbf{I}$ is the identity matrix.

In the downlink scenario, when signal of one user are synchronized all user are synchronized (α_k s =0) and covariance matrix of the received signal is given by

$$R = \sigma_n^2 \{ \sum_{k=1}^{K-1} \beta_k \left(\mathbf{v}_k \mathbf{v}_k^* \right) + \mathbf{I} \} \quad \textbf{(6)}$$

4. Blind Chip Rate Estimation

In singular value decomposition of the received signal covariance matrix, if the chip rate is less than half of the sampling frequency, estimated eigenvector of the received signal covariance matrix (corresponds to maximum eigenvalue) is the convolution of $e(t)$, $s(t)$, $g(t)$ and spreading sequence of synchronized user. If the impulse response of transmitter and receiver filters have been assumed ideal, and channel is constant in the duration of one *processing window*, estimated eigenvector will be the multiplication of spreading sequence of synchronized user in constant value. Hence for estimating spreading sequence from estimated impulse response of channel by subspace method i.e. $\mathbf{v}_k$, chip rate must be known exactly.

Sampling ratio is defined as

$$m = \frac{F_s}{R_c} \quad \textbf{(7)}$$

The total impulse response of the channel related to each user, (2), can be written as follows

$$h_{k,i}(t) = \sum_{j=0}^{mL_i - m} c_{k,i}\, p(t - j\frac{T_c}{m}). \quad \textbf{(8)}$$

If the sampling frequency is known, by estimation $h_{k,i}(t)$ using eigenvector of synchronized signal and relation between sampling time and chip time (i.e. $T_e = T_c / m$), we can obtain chip rate of the signal. Hence we focus on the estimation of sampling ratio (*SR*).

First, we calculate sign of $h_{k,i}(t)$ to obtain the normalized spreading sequence, it is performed due to binary phase shift keying (BPSK) spreading assumption of transmit signal. Sign of $h_{k,i}(t)$ is given by

$$\operatorname{sgn}(h_{k,i}(t)) = \sum_{j=0}^{mL_i - m} c_{k,i} \left[u(t - (j+1)\frac{T_c}{m}) - u(t - j\frac{T_c}{m}) \right] \quad \textbf{(9)}$$

where $u(t)$ is the unit step function. Then, the differential of $\operatorname{sgn}(h_{k,i}(t))$ is calculated with respect to t to estimate the number of zero crossings of the spreading sequence, which is given by

$$\frac{d \operatorname{sgn}(h_{k,i}(t))}{dt} = N^{-1/2} \sum_{j=0}^{mL_i - m} c_{k,i}[m] \delta(t - j\frac{T_c}{m}) \quad \textbf{(10)}$$

$d \operatorname{sgn}(h_{k,i}(t)) / dt$ includes impulse sequences by zero crossings at times $t = j T_c / m$. It is obvious that in sampling frequencies greater than the Nyquist rate, by increasing the sampling frequency, number of zero crossings on differential of spreading sequence remains constant. Figure 1(a) shows a sequence with length of 63, and Figure 1(b) shows the same sequence with the *SR* of 4. Figure 1(c) shows the differential of the sequence shown in Figure 1(a) and Figure 1(d) is similar to Figure 1(c) for differential of sequence which has been shown in Figure 1(b). It can be seen in this figure, the number of zero crossings remains constant in differential of each sequence by over sampling greater than the Nyquist rate. Figure 2(a) shows the number of zero crossings of the differential of sequence in terms of different sampling frequencies. Bandwidth of signal is considered 10 *MHz* and it can be seen the number of zero crossings remains constant for sampling frequency greater than 20 *MHz*. Zero crossing ratio is defined as the following

$$ZCR = \frac{\text{Number of zero crossings}}{\text{Length of diffrentiated Sequence}} \quad \textbf{(11)}$$

Figure 2(b) shows the number of zero crossings ratio of the differential sequence in terms of sampling frequency. It is obvious that for sampling frequencies greater than the Nyquist rate, *ZCR* reduces by increasing the sampling frequency and the maximum value of *ZCR* occurs at the sampling frequency equal to the chip rate of the sequence (i.e. length of 63). But for sampling frequencies less than the Nyquist rate, for loss of some data samples and decreasing length of differential

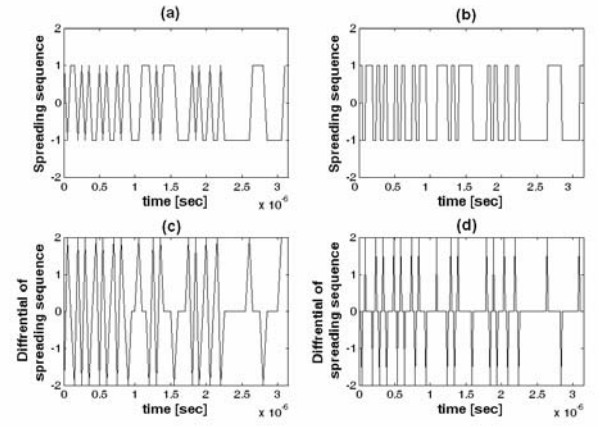


Figure 1. (a) Spreading sequence (length of 63). (b) Over sampled spreading sequence (length of 63) and *SR* = 4. (c) Differential of Spreading sequence (length of 63). (d) Differential of Over sampled spreading sequence (length of 63) and *SR* = 4.

sequence, *ZCR* may be greater than *ZCR* corresponding to the Nyquist rate. In Figure 2(b) an example of this situation can be seen. Sampling frequencies less than Nyquist rate losses some of data samples and decrease length of differential of sequence, it causes *ZCR* increases for sampling frequency lower than the Nyquist rate. Since, by weighting *ZCR*, it is possible to reduce *ZCR* for sampling frequencies less than the Nyquist rate. Therefore, we propose weighted zero crossing ratio, which is defined as

$$WZCR = \text{The Number of zero crossings} \times ZCR \quad \textbf{(12)}$$

Figure 2(c) shows weighted zero crossing ratio in terms of frequency sampling, it can be seen the maximum of WZCR corresponds to the sampling frequency is equal to the chip rate.

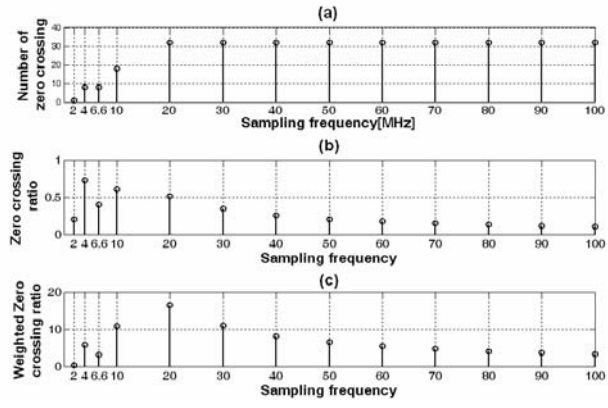


Figure 2. (a) Number of Zero crossing in terms of sampling frequency in MHz. (b) Zero crossing ratio in terms of sampling frequency. (c) Weighted zero crossing ratio in terms of sampling frequency.

Increasing the code length increases the computational complexity of subspace decomposition of the received signal covariance matrix due to increasing the dimension of covariance matrix, instead, output SNR in output of covariance matrix estimator increases by the factor of N_2/N_1, as,

$$\frac{SNR_{out_2}}{SNR_{out_1}} = \frac{N_2}{N_1} \quad \textbf{(13)}$$

where N_2 and N_1 are spreading sequence lengths corresponds to SNR_{out_2} and SNR_{out_1}, respectively. Therefore by increasing code length it is possible to estimate spreading sequence of active users in lower SNR values.

5. Blind Bit Rate Estimation

For blind bit rate estimation, we use fluctuation of correlation estimator [10]. To compute the fluctuations, we divide the received signal into *M* temporal windows with duration of T_F for each window. Then, a correlation estimator is applied to each window prior to the computation. For the m^{th} window, an estimation of the autocorrelation for any signal $r(t)$ is given by

$$\hat{R}_r^m(\tau) = \frac{1}{T_F}\int_0^{T_F} r_m(t) r_m^*(t-\tau)dt \quad \textbf{(14)}$$

where $r_m(t)$ is the signal sample over the m^{th} window. Replacing $r_m(t)$ in (14) with the m^{th} window of $r(t)$, and taking expectation of $\hat{R}_r^m(\tau)$ for *M* windows, lead to the following equation, which can be used for both the uplink (asynchronous user) and downlink (synchronous) transmission,

$$\hat{R}_r(\tau) = \hat{R}_s(\tau) + \hat{R}_n(\tau) \quad \textbf{(15)}$$

where $\hat{R}_n(\tau)$ is autocorrelation of noise and

$$\widehat{R}_s(\tau) = \sum_{i=0}^{S-1}\sum_{k=0}^{K_i} \hat{R}_{s_{k,i}}(\tau). \quad \textbf{(16)}$$

In (16), $\hat{R}_{s_{k,i}}(\tau)$ is the noise-unaffected estimation of the autocorrelation for the $(k, i)^{\text{th}}$ signal. Indeed, since the fluctuations are computed from many randomly-selected windows, they do not depend on the relative delays of signals. The use of *M* windows allows us to estimate the second-order moment of the estimated correlation $\hat{R}_r^m(\tau)$ as

$$\Phi(\tau) = \hat{E}\left\{\left|\hat{R}_r(\tau)\right|^2\right\} = \frac{1}{M}\sum_{m=0}^{M-1}\left|\hat{R}_r(\tau)\right|^2 \quad \textbf{(17)}$$

where $\hat{E}\{\cdot\}$ is the estimated expectation of $(\cdot)$. Hence, $\Phi(\tau)$ is a measure of the fluctuations of $\hat{R}_r^m(\tau)$. The difference between the successive and equal amplitude peaks in (16) determines symbol time for each rate [10].

6. Performance of the Proposed Chip Rate Estimation Algorithm

In the proposed chip rate estimation algorithm, if the eavesdropping receiver fallen in the deep fading; the estimated chip rate is not valid. In this section, effect of fading channel is analyzed on the performance of the proposed chip rate estimation algorithm. This analysis is first done for a receiver using a single antenna. Second, the analysis is done for a receiver using multiple antennas to overcome fading.

The threshold value of SNR for chip rate estimation algorithm is defined as $SNR_{\min}$, which is the minimum SNR for achieving a reliability factor (*RF*) of chip rate estimation greater than 0.99. $SNR_{\min}$ is obtained using computer simulation in non-fading channels and its value depends directly on the spreading factor.

By assuming Rayleigh distribution for the channel, the probability that the received signal voltage r is less than a given the threshold *R*, which is corresponding value of $SNR_{\min}$, found from the following equation

$$\Pr(r \le R) = \int_0^R p(r)dr = 1 - \exp(-\rho^2) \qquad \textbf{(18)}$$

where $p(r)$ is the probability density function (PDF) of the Rayleigh distribution, and $\rho = R/R_{rms}$ is the value of the specified level *R*, normalized to the local *rms* amplitude of the fading duration envelope, hence $\rho^2 = SNR_{\min}/SNR$. The chip rate estimation algorithm is failed, when SNR in *M* processing blocks of signal is less than $SNR_{\min}$.

So, the P_d for AWGN channels is changed to

$$RF_{fading} = RF \times \left(1 - \Pr(r \le R)\right) = RF \times \left(\exp(-\rho^2)\right). \qquad \textbf{(19)}$$

From the above equation, it is obvious that the reliability factor of chip rate estimation algorithm reduces by the factor of $\exp(-\rho^2)$ when all *M* processing windows have been faded.

Table 1 shows the $SNR_{\min}$ for achieveing a *RF* greater than 0.99 in terms of spreading factor (*SF*) over AWGN channels. These results are extracted using computer simulations. It can be seen, by 3 dB increase in *SF*, the $SNR_{\min}$ increases by 3 dB.

Table 1. SNR_{min} **for achieving a** *RF* **greater than 0.99 in terms of** *SF* **over AWGN channels.**

SF [dB]	15	18	21	24	27	30
$SNR_{\min}$ [dB]	-5	-8	-11	-14	-17	-20

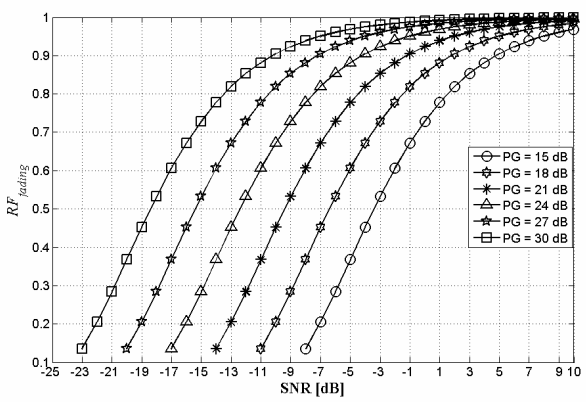


Figure 3. RF_{fading} **in terms of received SNR for different values of** *SF*.

Figure 3 shows *RF* in terms of SNR in flat fading channels with Rayleigh distribution for different values of *SF*. It can be seen that the *RF* is greater than 0.95 for SNR = 8 dB in *SF* = 15 dB. It is obvious that the performance of the proposed algorithm degraded 13 dB with respect to that of AWGN channels.

Now, we propose a solution to overcome effect of fading channels on the performance of the proposed algorithm. If the distance between two neighboring antennas in eavesdropping receiver is large enough and communication environment is rich scattering, fading channels experienced by these antennas are independent. So, the reliability factor of the proposed chip rate estimation algorithm improves. So that

$$RF_{fading_m} = RF \times \left(1 - \left(\Pr(r \le R)\right)^L\right), \qquad \textbf{(20)}$$

where RF_{fading_m} denotes reliability factor of the proposed chip rate estimation algorithm with *L* receiving antennas in eavesdropping receivers over independent fading channels. Figure 4 shows RF_{fading_m} in terms of SNR for spreading factor of 15 dB and different numbers for receiving antennas in eavesdropping receiver. It can be seen from this Figure that by increasing the number of receiving antennas from 1 to 4, RF_{fading_m} increases from 0.53 to 0.95 at SNR = -3 dB.

In practice, assuming independent channel gains for receiving antennas is not realistic due to limitation size

of the eavesdropping receiver. The correlation matrix of an antenna array depends not only on the array configuration but also on the incident angle of the incoming signal. Different parameters affect the correlation in different ways. Branch correlation depends on the antenna height and antenna separation through their ratio [11]. For a given antenna array, correlation between signals received by two neighbouring antennas varies with their corresponding angle of incidence so that it reaches its maximum value when the signal comes from the endfire direction and reduces gradually as the signal direction moves towards the broadside [12]. The covariance matrix among channel power gains e.g. $\left[\delta_{i,1} \quad \cdots \quad \delta_{i,L}\right]$ in the receiver is denoted by

$$\Omega = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,L} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{L,1} & \omega_{L,2} & \cdots & \omega_{L,L} \end{bmatrix}_{L\times L} . \qquad \textbf{(21)}$$

$RF_{corr-fading_m}$, which is the reliability factor of the proposed chip rate estimation algorithm for correlated fading channel, is obtained using Bayes' theorem for correlated fading channels as follows

$$RF_{corr-fading_{\mathrm{m}}} = RF \times \left[1 - \Pr\left(r \le R\right)^{1+\sum_{j=2}^{L}\prod_{i=1}^{j-1}\left(1-\omega_{i,j}\right)}\right] \qquad \textbf{(22)}$$

In the above equation, $\omega_{i,j}$ s for $i \neq j$ are less than one. So $\left(1-\omega_{i,j}\right)$ and $\prod_{i=1}^{j-1}\left(1-\omega_{i,j}\right)$ is less than 1. Therefore,

$$\sum_{j=2}^{L}\prod_{i=1}^{j-1}\left(1-\omega_{i,j}\right) \le L-1. \qquad \textbf{(23)}$$

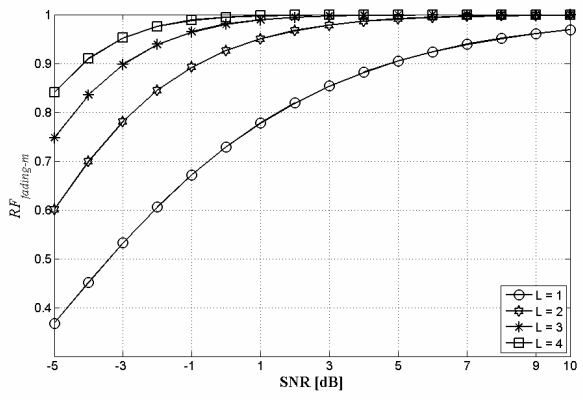


Figure 4. RF_{fading_m} **in terms of received SNR for different number of received antennas and** SF **= 15 dB .**

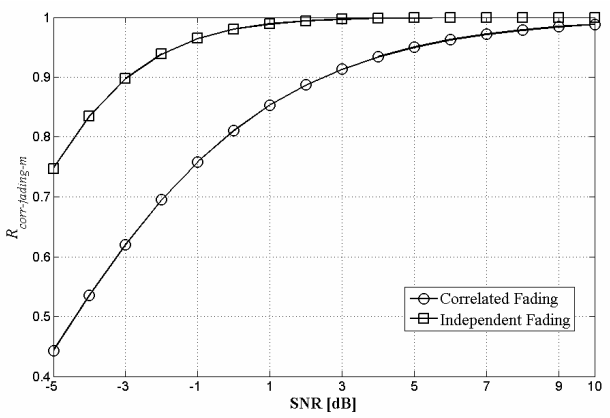


Figure 5. $RF_{corr-fading_m}$ **in terms of received SNR for a 3 antenna array with correlated and independent fading channels and SF = 15 dB.**

Therefore, the power of $\Pr\left(r \le R\right)$ in equation (22) is less than L. In comparison to (20), it is obvious that, $RF_{corr-fading_{\mathrm{m}}}$ is less than $RF_{fading_{\mathrm{m}}}$.

Figure 5 shows $RF_{corr-fading_{\mathrm{m}}}$ in terms of received SNR for a 3 antenna array receiving correlated signals. Similar to that of [12], we assume triangular array at the eavesdropping receiver, located at the height of 30.43 m, and its correlation matrix is given by

$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & 0.727 & 0.913 \\ 0.727 & 1 & 0.913 \\ 0.913 & 0.913 & 1 \end{bmatrix}$$

It can be seen from Figure 5, that $RF_{corr-fading_{\mathrm{m}}}$ for triangular and antenna arrays is 0.62 at SNR = -3 dB. It is obvious that $RF_{corr-fading_{\mathrm{m}}}$ decreases by 0.27 compared with $RF_{fading_{\mathrm{m}}}$ of an antenna array consisting of 3 antennas with independent fading channels.

7. Simulation Results

Computer simulations have been performed using two streams of spread spectrum signals with both BPSK spreading and data modulation. Gold sequences with lengths of 63 (user 1) and 31 (user 2) are considered as spreading sequences. Their chip rates are 20 *Mega chips/sec* and 40 *Mega chips/sec*, respectively. Processing window size is considered to be 160 μsec , and M = 10 is the number of windows, which is used for averaging. Chip rates of each signal have been obtained using estimated impulse response of the channel. Bit rate of each signal has been estimated using the time difference

between the successive equal amplitude peaks in the output of the correlation estimator.

Figure 6 shows fluctuations of the correlation estimator at SNR of -5 dB. The time differences between two successive equal amplitude peaks have been estimated 0.775 μsec and 3.175 μsec.

Figure 7(a) shows the estimated channel impulse response for the spreading sequence with length of 63 and sampling frequency of $F_S = 20MHz$. This signal has more SNR relative to the other user and produces the greatest eigenvalue in the estimated received signal covariance matrix; therefore, its first corresponding estimated eigenvector is selected for the chip rate estimation. A window of estimated channel impulse response as long as bit time = 3.175 μsec is considered for the chip rate estimation. Figure 7(b) shows the estimated, which is given by (9), and exact spreading sequence, it is obvious they are matched together. Figure 7(c) shows $d\,\mathrm{sgn}(h_{k,i}(t))/dt$, it can be seen from this figure, zero crossing occurs when sign of signal changes.

Figure 8(a) shows that the distribution of zero crossing number for the $d\,\mathrm{sgn}(h_{k,i}(t))/dt$ in terms of sampling frequency. Figure 8(b) shows the *ZCR* of the $(h_{k,i}(t))/dt$, which is defined in (11), in terms of sampling frequency. Figure 8(c) shows *WZCR* for $d\,\mathrm{sgn}(h_{k,i}(t))/dt$, which is defined in (12), in terms of sampling frequency. It obvious maximum number of *WZCR* corresponds to the sampling frequency equal to Nyquist rate, which is equivalent to a chip rate 20 *Mega chips/sec*. Hence, $T_c = 1/(20\times10^6) = 50n\,\mathrm{sec}$, code length is obtained as $T_b/T_c = 3.15\mu\,\mathrm{sec}/\,50n\,\mathrm{sec} = 63$.

Figure 9(a) shows the probability of chip rate detection related to first user for 10000 times Monte Carlo test. It can be seen in this Figure, sampling frequency of 40 *MHz* is estimated by probability of more than 97.5% as a

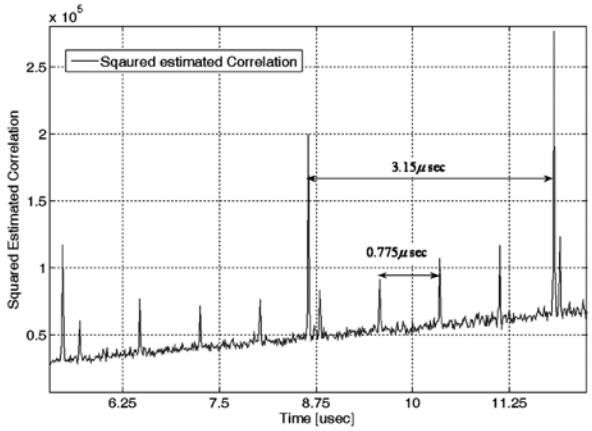


Figure 6. Fluctuations of correlation estimator output.

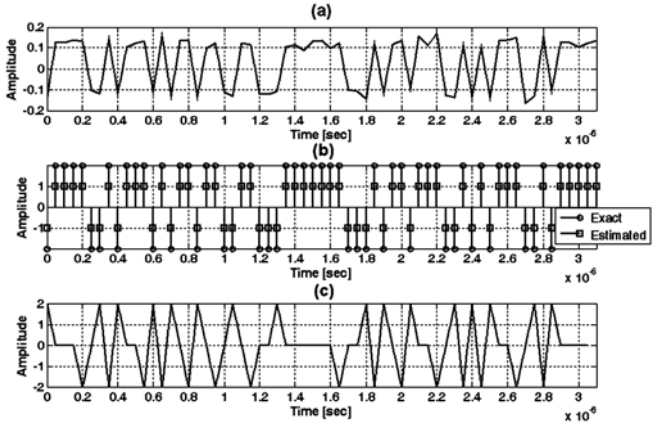


Figure 7. (a) Estimated eigenvector corresponding to the maximum eigenvalue of the received signal covariance matrix. (b) Exact and estimated spreading sequence of user 1. (c) The differential of the estimated eigenvector.

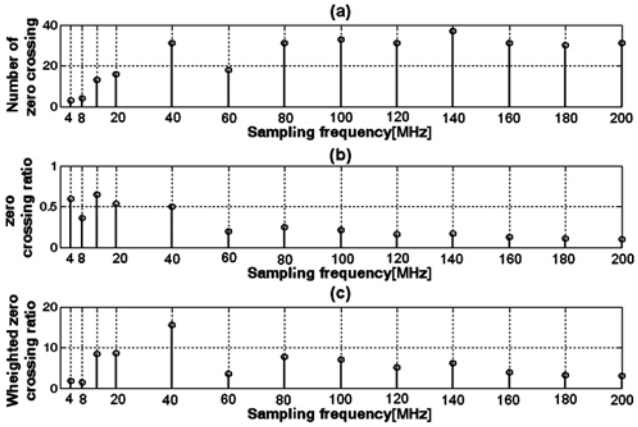


Figure 8. (a) Number of Zero crossings in terms of sampling frequency. (b) Zero crossing ratio in terms of sampling frequency. (c) weighted zero crossing ratio in terms of sampling frequency, all of them are related to user 1.

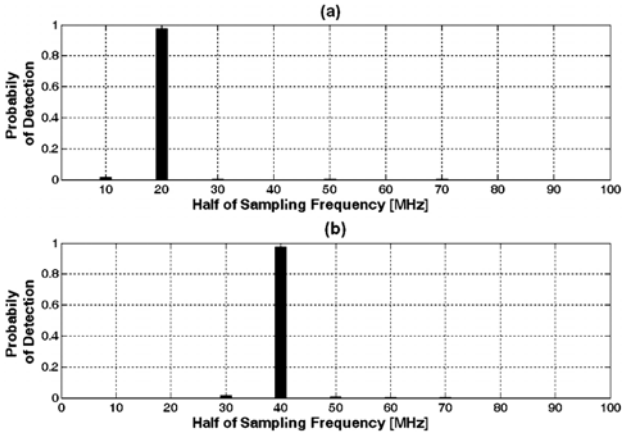


Figure 9. (a) Probability of chip rate detection related to the first user for 10000 times Monte Carlo Test. (b) Probability of chip rate detection related to the second user for 10000 times Monte Carlo Test.

sampling frequency with maximum *WZCR*, which is equivalent to chip rate of 20 *Mega chips/sec*. Figure 9(b) is similar to Figure 9(a) except it is corresponds to second user. It can be seen sampling frequency of 80 *MHz* is estimated by probability greater than 97% as a sampling frequency with maximum *WZCR*, which is equivalent to a chip rate 40 *Mega chips/sec*.

8. Conclusions

This paper considered the problem of blind chip rate estimation of direct sequence spread spectrum (DS-SS) signals in multi-rate and multi-user direct-sequence code division multiple accesses (DS-CDMA). The estimation is based on the multi-rate sampling of the estimated differential channel impulse response. Simulation results showed that the chip rate and bit rate can be determined exactly in very low SNR (-5 dB) and in multi-rate and multi-user field. Therefore, it is possible to blindly estimate the spreading factor of each user. In fading channels, the reliability factor of the proposed chip rate estimation algorithm is analyzed and a receiver with multiple antennas is proposed to improve the reliability factor of the proposed algorithm.

9. References

[1] E. Dahlman, P. Beming, J. Knutsson, F. Ovesjo, M. Persson, and C. Roobol, "WCDMA-the radio interface for future mobile multimedia communications," IEEE Transactions on Vehicular Technology, Vol. 47, pp. 1105–1118, November 1998.

[2] P. Taaghol, B. G. Evans, E. Buracchini, G. De Gaudinaro, Joon Ho Lee, and Chung Gu Kang, "Satellite UMTS/ IMT2000 W-CDMA air interfaces," IEEE Communications Magazine, Vol. 37, pp. 116–126, September 1999.

[3] Z. Zhang, L. Li, and X. Xiao; "Carrier frequency and chip rate estimation based on cyclic spectral density of MPSK signals," in Proceedings of Communications, Circuits and Systems Conference, Vol. 2, pp. 859 –862, June 2004.

[4] W. J. Ma, S. M. Yang, W. Ren, Z. H. Xue, and W. M. Li, "Spectral correlation function in low SNR environment," in Proceedings of Radio Science Asia-Pacific Conference, pp. 197–200, August 2004.

[5] Y. X. Li, M. Yi, Q. Yang, X. C. Xiao, and H. M. Tai, "Low SNR BPSK signal chip rate estimation using a wavelet based spectral correlation algorithm," in Proceedings of Circuit and Systems Conference, Vol. 3, pp. 247–249, August 2002.

[6] P. G. Turner, R. Hewitt, and D. J. Purle, "Chip-rate processing for software defined radios," in Proceedings of 3G Mobile Communication Technologies, pp. 292–296, March 2001.

[7] Y. Jin and H. B. Ji, "A cyclic-cumulant based method for DS-SS signal detection and parameter estimation," in Proceedings of Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications Conference, Vol. 2, pp. 966–969, August 2005.

[8] S. Ghavami and B. Abolhassani, "Blind chip time estimation in multirate CDMA transmissions using multirate sampling in slow flat fading channels," in the Proceedings of IEEE-ICCS08, pp. 1344–1348, November 2008.

[9] C. N. Nzéza, R. Gautier, and G. Burel, "Parallel blind multiuser synchronization and sequence estimation in multirate CDMA transmissions," in Proceedings of IEEE - ACSSC, pp. 2157–2161, November 2006.

[10] C. N. Nzéza, R. Gautier, and G. Burel, "Blind multiuser detection in multirate CDMA transmissions using fluctuations of correlation estimators," in Proceedings of IEEE - Globecom, pp. 1–5, November 2006.

[11] Q. T. Zhang, "Maximal-ratio combining over Nakagami fading channels with an arbitrary branch covariance matrix," IEEE Transactions on Vehicular Technology, Vol. 48, No. 4, pp.1141–1150, July 1999.

[12] W. C. Lee, "Mobile communications: Design fundamentals," 2nd Edition, New York, Wiley, pp. 202–211, 1993.

Wireless Sensor Network, 2009, 2, 61-121
doi:10.4236/wsn.2009.12012 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

Micro Controller Based Ac Power Controller

S. A. HARI PRASAD[1], B. S. KARIYAPPA[1], R. NAGARAJ[2], S. K.THAKUR[3]
[1]*Department of Electronics & Communication Engineering, R.V. College of Engineering, Bangalore, India*
[2]*Director, Center for Cognitive Technologies,R.V.C.E. Campus, Bangalore, India*
[3]*Deputy Director, Naval Research Board, Defence Research Development Organization, New-Delhi, India*
Email: *harivat*2002*@yahoo.co.in*
Received April 25, 2009; *revised May* 8, 2009; *accepted May* 9, 2009

Abstract

This paper discusses the design and implementation of single phase PWM inverter using 8051 microcontroller. The main features of 8051 based PWM inverter are simpler design, low cost, maximum range of voltage control and compact in size. The designed PWM inverter is tested on various AC loads like AC motor and intensity control of incandescent lamp in a closed loop environment.

Keywords: Gate Signals Generation, Micro Controller, Pulse Width Modulation, PWM Generation

1. Introduction

The pulse width inverters can be broadly classified as

- Analog bridge PWM inverter [1]
- Digital bridge PWM inverters [2]

The advantage of Analog based PWM inverter controller is that, the level of inverter output voltage can be adjusted in a continuous range and the throughput delay is negligible. The disadvantages of Analog based PWM inverters are:

Analog component output characteristics changes with the temperature and time. They are prone to external disturbances. Analog controller circuitry is complex and bulky. They are non-programmable, hence not flexible.

On the other hand Microcontroller based PWM inverter controller (Digital bridge PWM inverter) makes the controller free from disturbances and drift, but the performance is not very much high due to its speed limitation. However to minimize throughput delay, some microcontroller based PWM inverters, retrieves switching patterns directly from memory so that calculation can be minimized, but this technique demands more memory. This drawback can be eliminated if switching patterns are generated by executing simple control algorithms [3]. Even after using simple control algorithms, sometimes throughput delay may be substantial.

With the availability of advanced microcontrollers and DSP [Digital signal processor] controllers [4], with many advanced features like inbuilt PWM generator, event managers, time capture unit, dead time delay generators, watch dog timers along with high clock frequency, the limitation of speed, associated with microcontroller based PWM inverters [5] can be neglected to some extent.

This paper presents a simple and cost effective technique of implementing single-phase AC [alternating current] voltage controller, used to control AC loads ,which doesn't demand very high precisions, using 8051 microcontroller.

The paper is organized as follows.

- Review of PWM inverters.
- Block diagram of controller.
- Controller implementation (software and hardware).
- Results and Conclusion.

2. PWM Bridge Inverter Review

Inverters can be classified as single-phase and three phase inverters [6] and they are further classified as Voltage fed inverter [VSI.], current fed inverter [CFI], and variable DC [direct current] linked inverter. In Voltage fed inverter, input voltage remains constant, in current fed inverter [CFI], input current remains constant and in variable DC [direct current] linked inverter, input voltage is controllable.

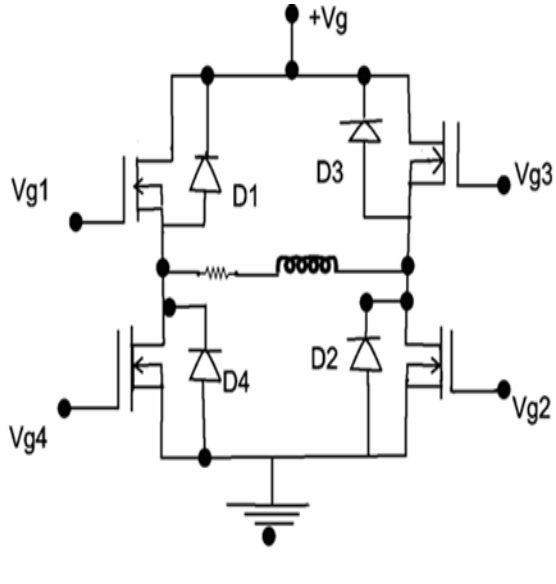


Figure 1. Single phase inverter.

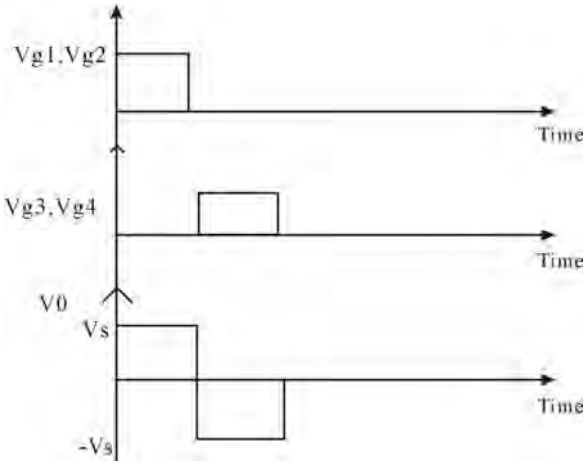


Figure 2. O/P voltage/gate signals.

Figure 1 shows single phase bridge inverter with MOSFET switches [6]. In spite of MOSFET switches having high ON state resistance and conduction losses [7], in this paper MOSFET switches are used because of the following reasons. MOSFET being a voltage controlled device, it can be driven directly from CMOS or TTL logic and the same gate signal can be applied to diagonally opposite switches. Also the gate drive current required is very low [7].

The working principle of Single-phase bridge inverter can be explained as follows.

Positive Voltage 'Vs' appears across the load, when MOSFET Q1 and Q2 conduct simultaneously. Negative voltage 'Vs' appears across the load, when Q3 and Q4 conduct simultaneously.

To overcome the effect of back emf in case of inductive load diodes, D1-D4 are used. Diode D1 and D2 are called feedback diodes, because when they conduct the energy is feedback to the DC source. The RMS output voltage is given by

$$V_o = V_s\sqrt{p/\Pi}$$

where P is pulse width. The O/P voltage and gate signals are as shown in Figure 2.

3. Controller Block Diagram

The block diagram of microcontroller based bridge PWM inverter is as shown in Figure 3. The required four digit speed in RPM [Rotation per Minute] is entered through the keyboard and corresponding to the key pressed, digital equivalent of that RPM is stored in memory.

Current running speed of the AC motor is sensed through speed sensor, and the analog output given by the sensor is converted to digital data using Analog to Digital converter [ADC].

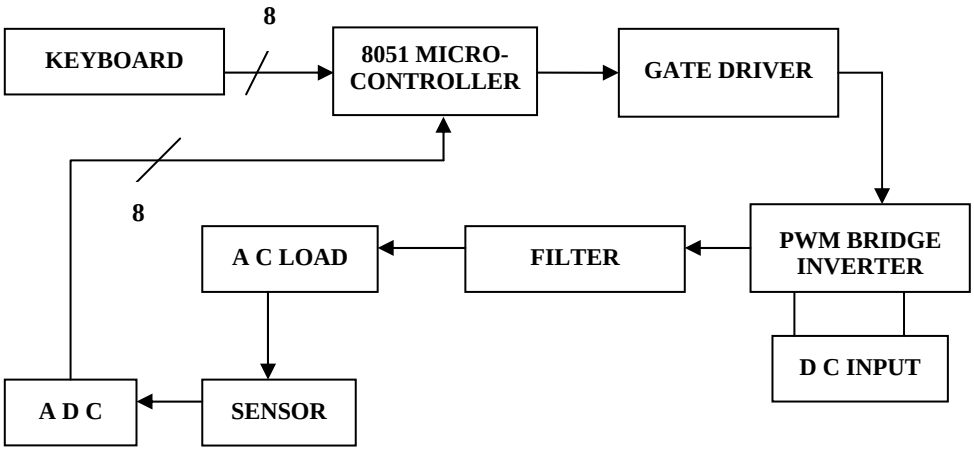


Figure 3. Block diagram of controller.

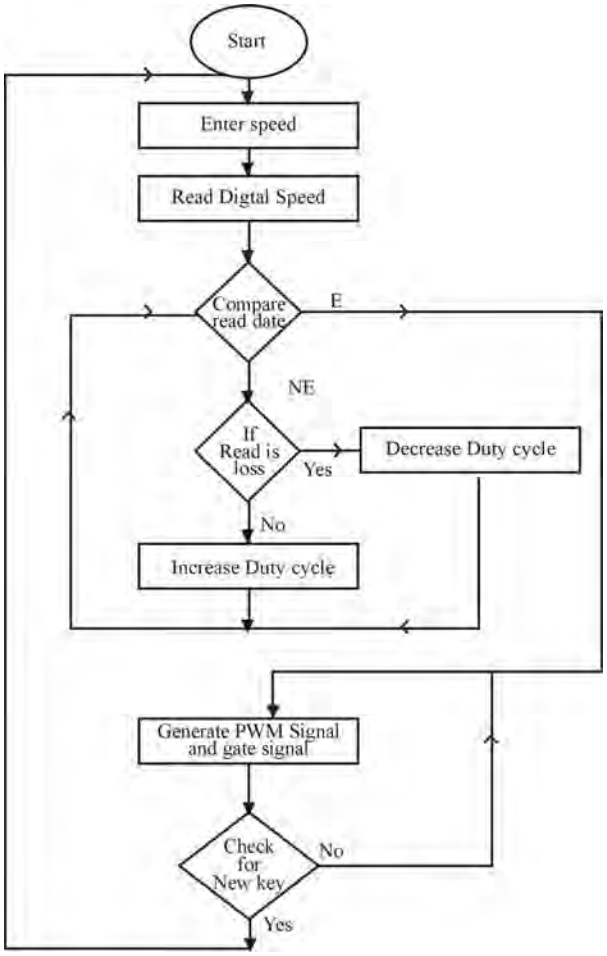


Figure 4. Flowchart of basic operation.

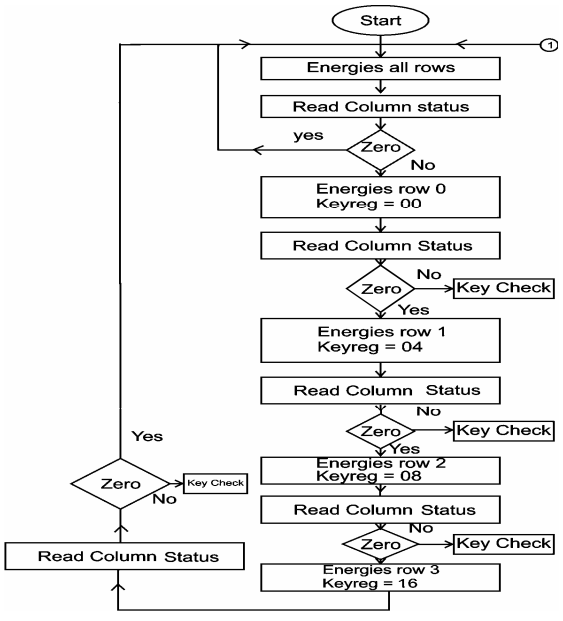


Figure 5. Flow chart of keyboard logic.

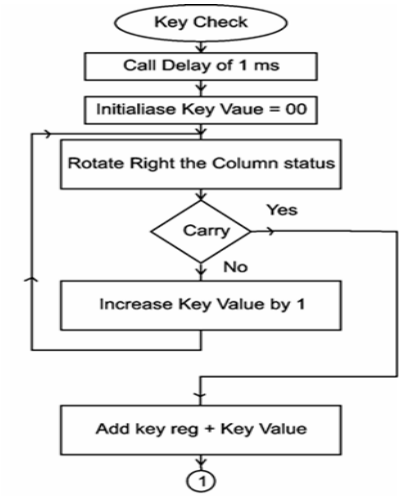


Figure 6. Flowchart of keyboard logic.

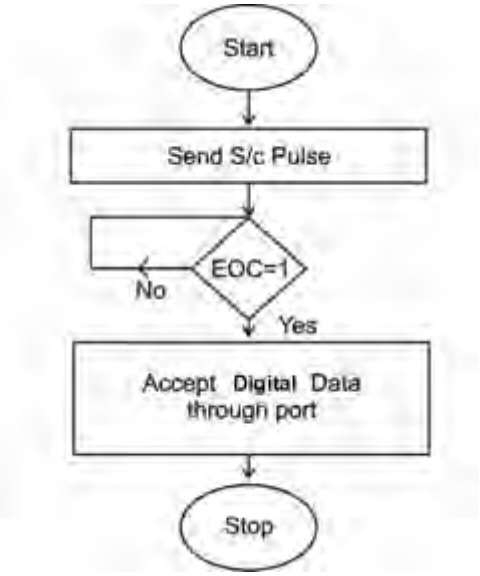


Figure 7. Flowchart of A/D converter.

The digital data is accepted through 8051 microcontroller ports and is compared with required speed's equivalent digital data. In accordance with the error signal, the width (duty cycle) of PWM signal is varied, which in turn controls the AC voltage.

From the generated PWM signal, required two gate signals are generated using external interrupt to drive the bridge inverter circuit.

Gate signals are boosted up to a sufficient voltage level by using gate driver circuit, so that it can drive the MOSFET switches of bridge inverter to the ON state. User can alter the speed at any instant of time in accordance to his requirements. Many additional features can be further added like sensing the temperature of room and automatically controlling either the speed of

the fan or the level of air conditioning required. Figure 4 explains the logic flow of the basic operation.

4. Controller Design

Controller is designed by using simpler low cost components like 8051 microcontroller, 8 or 12 bit Analog to Digital Converter (ADC), 4×4 keypad, 4 chopper MOSFET switches (IRFZ48) and speed/Intensity sensor.

The controller design can be explained under 4 sections as:

- Keypad interface with 8051 μc.
- ADC interface with μc.
- Generating PWM signals and gate signals using 8051 microcontroller.
- Gate driver circuit implementation.

4.1. Keypad Interface

A 4×4 keypad is interface with 8051 microcontroller as shown in Figure 5, through which four keys are accepted. After accepting the four keys they are combined to represent four digit required RPM, which actually represents the external memory address, in which digital equivalent of speed is stored.

For example if the keys entered are 1 (01), 2 (02), 3 (03), 4 (04), then they are combined as 1234 (RPM), which represents External memory address, in which 8 bit digital equivalent of that speed is stored. Higher byte of the memory address is stored in DPH [data pointer high byte]. Lower byte of the memory address is stored in DPL [data pointer low byte]. This method saves time since it doesn't require any program execution to convert the entered speed in RPM into its digital equivalent. The other method is to enter equivalent digital data of RPM directly, provided a conversion chart is available [external look-up table]. This technique will save some memory access time, since communication with memory is avoided.

4.2. ADC Interfacing

Whenever speed varies from zero to maximum, the speed sensor O/P varies from zero to five volts respectively. An 8-bit ADC with resolution $1/2^8$ is used to convert the analog voltage to digital data. Minimum of 19.5 mv change in voltage (corresponding change in RPM) is required to change the digital state of ADC. This limits the accuracy of the application. The logic of interfacing ADC is as explained in the flowchart given in the Figure 7.

4.3. PWM Generation

8051 microcontroller do not have on-chip PWM generator. It is implemented using 'A' register and any other register (R0-R7) as shown in Figure 8.

A count (ON period time) is loaded onto one of the GPR (General purpose register), which can be called as Duty cycle register and accumulator ('A') is loaded with zero. Register 'A' is incremented in steps of one and continuously compared with duty cycle register.

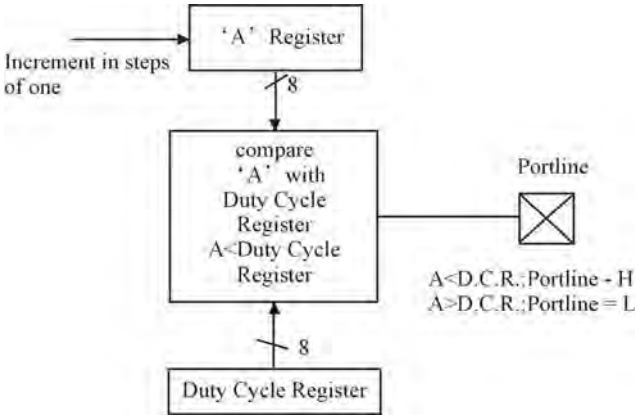


Figure 8. PWM generation.

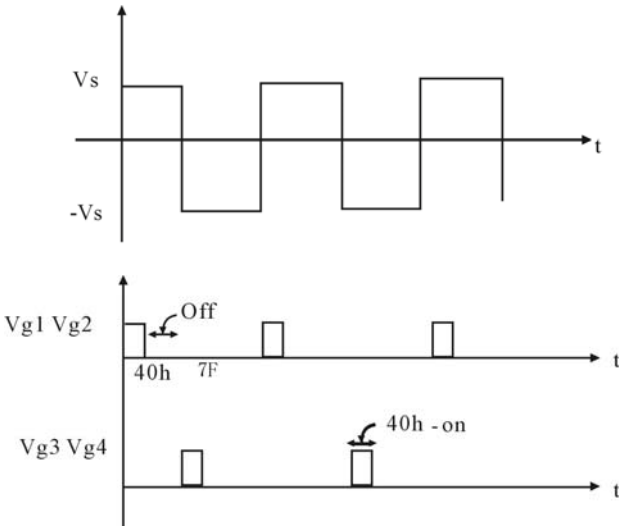


Figure 9. Gate signal generation using interrupt.

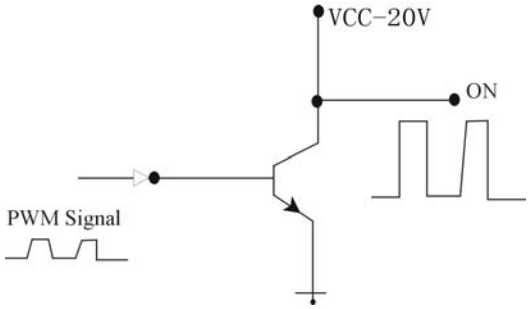


Figure 10. Gate signal booster circuit.

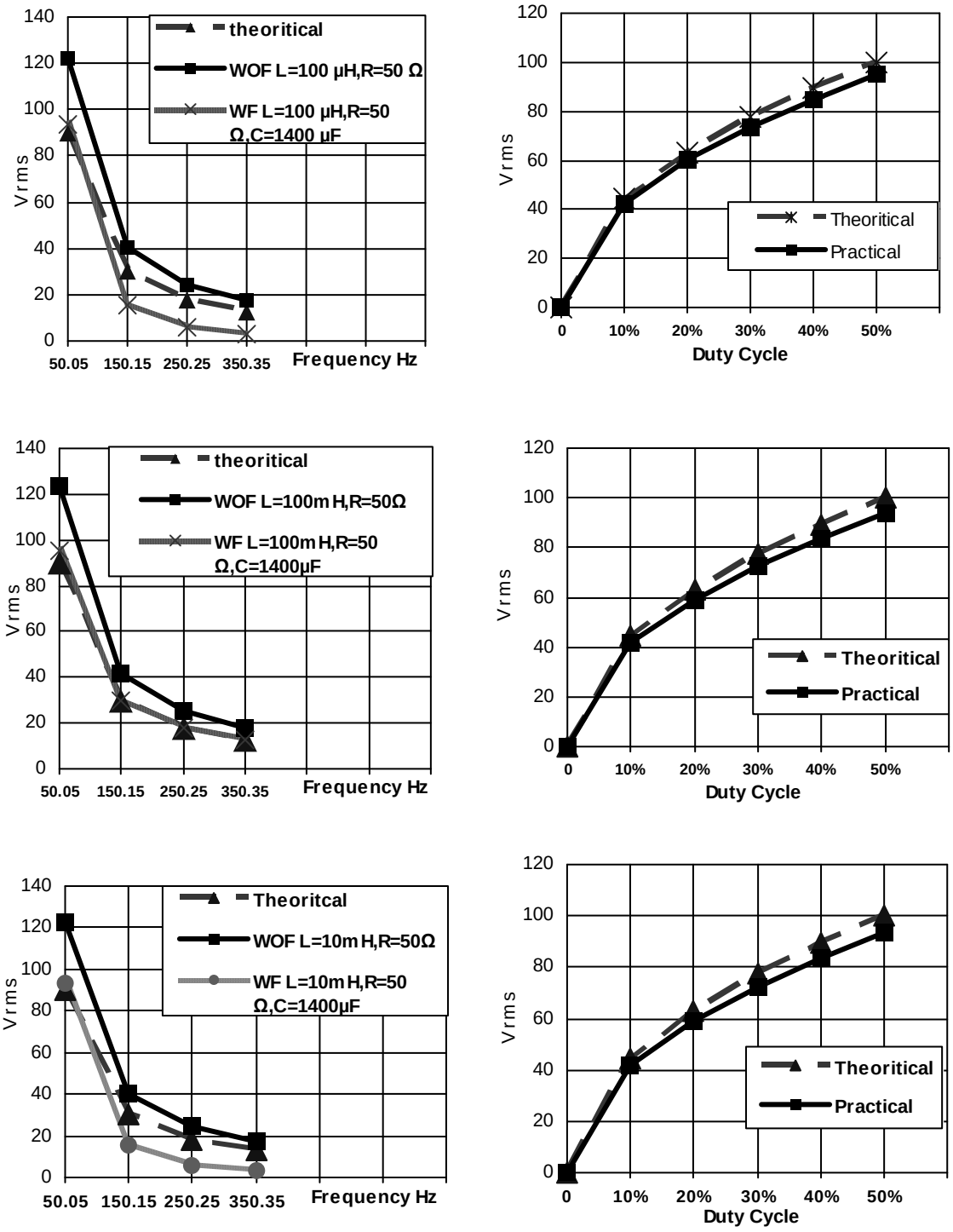


Figure 11. Response for various loads with corresponding duty cycles.

If the 'A' contents are less than duty cycle register, high level is maintained at port line P1.1. When 'A' is higher than duty cycle register content a low level is maintained on port line. The alternate technique is to use Timer as Counter by applying clock pulses externally and comparing the count present in the counter with 'A' register (duty cycle register). This demands external clock source, since 8051 do not have any clock out pin.

Since the maximum time period is limited to 256 microseconds, the minimum frequency of PWM signal will

be 4 KHz, but this can be changed using software delays. The AC signal frequency generated by PWM bridge inverter depends on PWM signal frequency. The error signal is generated by comparing the required speed with accepted digital equivalent speed divided by two. In proportionate with the error signal, PWM duty cycle is varied. When the required speed value is less than the accepted one, duty cycle register value and accepted value is decremented by one continuously till accepted value is equal to the required speed's digital value. When the required speed value is more than the accepted one, duty cycle register values and accepted value is incremented by one continuously till accepted value is equal to the required speed digital values.

4.4. Gate Signal Generation

The generated controlled PWM signal itself will be one set of gate signal (g_1, g_2) and other set of gate signals (g_3, g_4) is generated using interrupt technique. The controlled PWM signal generated is given to the external interrupts, which is initialized as falling edge sensitive interrupt type. When controlled PWM signal's falling edge occurs, an interrupt service routine meant for that particular external interrupt is executed.

In the interrupt service routine, a delay is created equal to the time, 7FH minus duty cycle register content, after which, the port line is made high and is retained high for the time duration decided by the contents of duty cycle register (Figure 9).

The gate signal (vg_1 vg_2, vg_3, vg_4) are boosted to a sufficient voltage level by Gate drive circuitry as shown in Figure 10, so that they are capable of driving MOSFET'S to the ON state, when the gate signals are high.

A transistor switch (with inverted gate signals as input) is made used to boost the gate signal. The same DC supply, which is used for inverter is also used to drive the transistor by reducing the DC level using voltage dividers. The other technique is to use opto-isolators. Both of these techniques use the same inverter DC source to boost up the gate signals, thus avoiding more usage of DC sources.

5. Results and Conclusions

The designed application is tested by designing 60V MOSFET bridge inverter.

Harmonics are removed by using simple capacitor filter and the AC voltage is stepped up to 220 V using step-up transformer. The performance of application is tested on various A.C loads and the plots of the same are as shown in Figure 10. The design exhibits good results for the load values of 50 ohm and 100 mH/ 10mH. A simple PWM technique with 100% duty cycle variation, which reduces hardware and software complexity, is used rather than using the most often used complex sinusoidal PWM technique (For Single-phase inverters). Required dead time is generated through interrupt, which avoids the usage of dead time delay generators. With minor modifications the same work can be used to control light intensity, temperature etc., The accuracy can be further improved by using high resolution ADC's and the delay involved in the software can be overcome using higher versions of controllers.

6. References

[1] H. Parasuram and B. Ramaswami, "A three phase sine wave reference generator for thyristorized motor controllers," IEEE Transactions on Industrial Electronics, Vol. IE-23, pp. 270–276, August 1976.

[2] J. M. D. Murphy, L. S. Howard, and R. G. Hoft, "Microprocessor control of PWM inverter induction motor drive," in Record of the 1979 IEEE Power Electron Specialist Conference, pp. 344–348.

[3] G. S. Buja and P. Fiorini, "Microcomputer control of PWM inverters," IEEE Transactions on Industrial Electronics, Vol. IE-29, pp. 212–216, August 1982.

[4] G. S. Buja and P. De Nardi, "Application of a signal processor in PWM inverter control," IEEE Transactions on Industrial Electronics, Vol. IE-32, No. 1, February 1985.

[5] Y. K. Peng, et al., "A novel PWM technique in digital control," IEEE Transactions on Industrial Electronics, Vol. 54, February 2007.

[6] M. H. Rashid, "Power Electronics Circuits, Devices and Applications," 3rd Edition, Prentice-Hall of India, Private limited, New-Delhi, 2004.

[7] V. Jagannathan, "Introduction to power electronics," Prentice-Hall of India, Private limited, New-Delhi, 2006.

Wireless Sensor Network, 2009, 2, 61-121
doi:10.4236/wsn.2009.12013 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

Target Detection in Three-Dimension Sensor Networks Based on Clifford Algebra

Tiancheng HE, Weixin XIE, Wenming CAO
Institute of Intelligent Information System, Shenzhen University, Shenzhen, China
Email: *sd*76130201@*szu.edu.cn*
Received April 27, 2009; *revised May* 8, 2009; *accepted May* 9, 2009

Abstract

The three-dimensional sensor networks are supposed to be deployed for many applications. So it is significant to do research on the problems of coverage and target detection in three-dimensional sensor networks. In this paper, we introduced Clifford algebra in 3D Euclidean space, developed the coverage model of 3D sensor networks based on Clifford algebra, and proposed a method for detecting target moving. With Clifford Spinor, calculating the target moving formulation is easier than traditional methods in sensor node's coverage area.

Keywords: 3D Sensor Networks, Clifford Algebra, Spinor, Target Detection, Coverage

1. Introduction

Three-dimensional sensor networks [1] could enable a broad range of applications: Video Surveillance, Ocean Sampling, Environmental Monitoring, Assisted Navigation, and etc. As an emerging technology which can put the information field to a new stage, the theory and applications of three-dimensional intelligent sensor networks have become a key research aspect. With the coming availability of low cost, short range radios along with advances in wireless networking, it is expected that wireless ad hoc sensor networks will become commonly deployed. So it is useful to study three-dimensional intelligent sensor network systems. The coverage problem and target detection problem are the fundamental issues in 3D intelligent sensor network systems. Studies on sensor networks include distributed network, distributed information acquisition, distributed intelligent information fusion and so on. Xie [2,3] proposed a coverage analysis approach for sensor networks based on Clifford algebra. In a 2-dimensional plane, a homogeneous computational method of distance measures has been provided for points, lines and areas, and a homogeneous coverage analysis model also has been proposed for targets with hybrid types and different dimensions. Thus, an analysis framework has been established for sensor networks in Clifford geometric space. To evaluate the quality of network coverage, Megerian [4] used Voronoi diagram and Delaunay triangulation respectively to define the worst and best-case coverage in sensor networks. There are also a lot of improved methods to solve these problems [5–8].

The target detection problem in sensor networks also has been a topic of extensive study under different metrics and assumptions [9–11]. There are already some related theories and algorithms proposed for solving the problems of target detection in sensor networks. The work of target detection in sensor networks includes many aspects, as the following:

The Traversing Path Detection [12]: A traversing path without being detected should not intersect the sensing areas of any sensors [13]. The detection rate of a sensor network is interested in application scenarios such as vehicles crossing a battlefield in military operations. Meguerdichian mentioned a novel coverage model for the target detection [14], and proposed an approximate value algorithm for calculating the traversing path [15]. There are also some distributed algorithms to calculate the efficient value in the sensor networks for detecting target [16]. Another way to solve the problem is the localized algorithm with lower computational complexity [17]. It uses polar coordinates to detect target and calcu-

late the path, and refers the Euler equation to calculate the minimal exposure traversing path.

The Exposure of Target [18]: The exposure of target detection is another aspect of the related work. The most popular definition is the information exposure to the estimation of target parameters [18]. With the information exposure formulation and grid graph, the minimal exposure traversing path in detecting target could be achieved. Meanwhile, some heuristic algorithms were also proposed for nodes deployment according to the degree of information exposure [8,15].

The Deployment of Nodes Density [19]: The nodes density is used for ensuring the probability of target detection[19]. It is assumed that the motive target could traverse the deployed area of sensor networks with a fixed velocity and a line path. The studies for deployment of nodes density include the probability sensor model and exposure model [20], grid deployment and random deployment in wireless sensor networks [21], and The critical nodes density based on continuum percolation [22]. The all-sensor field intensity can be modeled as a two dimensional Poisson shot noise process for large-scale sensor networks under the general sensing model [23].

Barrier Coverage [24]: Barrier coverage was proposed by Kumar [24] who mentioned it as an appropriate notion of coverage when a sensor network is deployed to detect targets traversing a protected region, which represents a promising and popular class of applications for wireless sensor networks. There are also some studies for the problem such as minimum segment barrier coverage [25], and double barrier coverage [11].

The coverage and target detection problem can be solved optimally in 2D plane by dividing the polygon into non-overlapping triangles, but it becomes NP-hard in 3D space. In this paper, we present a method for calculating target moving in 3D sensor networks with the coverage analysis approach based on Clifford algebra, which establishes a coordinate-free, homogeneous coverage model for different dimensional spaces and targets with hybrid types. This approach gives the intact relative geometric description between sensor node and target. With the Clifford Spinor, calculating target moving formulation will be more simply and effectively than traditional method for sensor networks.

The paper is organized as follows. In Section 2, we state relevant background of Clifford algebra. In Section 3, we present our model and target moving formulation. In Section 4, the algorithm based on Clifford algebra is proposed. In Section 5, we present our conclusions.

2. Clifford Algebra in 3D Euclidean Space

Rectangular Cartesian coordinate system is adopted in 2D Euclidean space, and any vector $\mathbf{a}$ can be written as $\mathbf{a} = a_1 e_1 + a_2 e_2$, where e_1, e_2 are the unit vectors of x, y directions, respectively. Hereby the geometric product between vector $\mathbf{a}$ and $\mathbf{b}$ can be calculated as:

$$\begin{aligned}\mathbf{ab} &= (a_1 e_1 + a_2 e_2)(b_1 e_1 + b_2 e_2) \\ &= a_1 b_1 + a_2 b_2 + a_1 b_2 e_1 e_2 + a_2 b_1 e_2 e_1\end{aligned}$$

Here $e_1 e_2 = e_2 e_1 = 1$. If we appoint $e_1 e_2$ as the area unit with direction, and let $e_1 e_2 = -e_2 e_1$, so the geometric product is

$$\begin{aligned}\mathbf{ab} &= a_1 b_1 + a_2 b_2 + (a_1 b_2 - a_2 b_1) e_1 e_2 \\ &= \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \wedge \mathbf{b}\end{aligned} \quad \textbf{(1)}$$

$\mathbf{a} \bullet \mathbf{b}$ is inner product, and $\mathbf{a} \wedge \mathbf{b}$ is outer product. The inner product is always coincident with dot product in vector algebra, while outer product is the measurement of parallelogram area composed by adjacent borders $\mathbf{a}$, $\mathbf{b}$. If $\mathbf{a}$ is rotated counter-clockwise with an angle which is no more than π to overlap $\mathbf{b}$, the measurement of parallelogram area is larger than zero. Otherwise it is less than zero. The area uses $e_1 e_2$ as a measurement unit, and its absolute value is $\mathbf{ab}\sin\theta$, where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$. Here $\mathbf{a} \bullet \mathbf{b} = \mathbf{ab}\cos\theta$.

Suppose that $e_1 e_2 = i$, then $e_2 e_1 = -i$, and

$$i^2 = e_1 e_2 e_1 e_2 = -e_1 e_1 e_2 e_2 = -1$$

Because $e_1 e_2 = -e_2 e_1$, exchanging the position of $e_1 e_2$ will appear a minus, and exchanging the position n times should multiply (-1)n. We call it negative exchange, and should take care of it in Clifford algebra.

After confirming the $e_1 e_2 = i$, we will have

$$e_1 i = e_1 e_1 e_2 = e_2, \quad e_2 i = e_2 e_1 e_2 = -e_1 e_2 e_2 = -e_1,$$

$$\mathbf{a}i = (a_1 e_1 + a_2 e_2) e_1 e_2 = a_1 e_2 - e_2 e_1$$

So each vector multiplies i on right side to let the vector be rotated counter-clockwise with $\pi/2$. And

$$\begin{aligned}\mathbf{a}e^{i\theta} &= (a_1 e_1 + a_2 e_2)(\cos\theta + i\sin\theta) \\ &= a_1 e_1 \cos\theta + a_1 e_1 i \sin\theta + a_2 e_2 \cos\theta + a_2 e_2 i \sin\theta \\ &= (a_1 \cos\theta - a_2 i \sin\theta) e_1 + (a_1 \sin\theta + a_2 \cos\theta) e_2\end{aligned}$$

It definitely describes that the ordinate unit vectors $e_1 e_2$ are fixed, vector $\mathbf{a}$ is rotated counter-clockwise with θ. It is obvious that i is not only just as -1, but also has specific geometric significant. The $\mathbf{a}ze^{i\theta}$ also denotes that $\mathbf{a}$ is rotated with θ, and magnified z times. z is the module of complex number.

The nodes in sensor networks need scalar and directed quantity to be described together, so we use Equation (1) for calculating. The quantity $\mathbf{a} \wedge \mathbf{b}$ is called dual vector

or bivector, its unit is $e_1e_2 = i = e_1 \wedge e_2$. According to (1), $e_1e_2 = e_1 \bullet e_2 + e_1 \wedge e_2 = e_1 \wedge e_2$ due to $e_1 \bullet e_2 = 0$.

The vector $\mathbf{a}$ in 3D Euclidean space $\Re^3$ can be written as $\mathbf{a} = a_1e_1 + a_2e_2 + a_3e_3$, where e_1, e_2, e_3 are orthonormal vectors. So

Definition 1: The geometric product of vector $\mathbf{a}$ and $\mathbf{b}$ in 3D space is

$$\begin{aligned}\mathbf{ab} &= (a_1e_1 + a_2e_2 + a_3e_3)(b_1e_1 + b_2e_2 + b_3e_3) \\ &= a_1b_1 + a_2b_2 + a_3b_3 + (a_1b_2 - a_2b_1)e_1e_2 \\ &\quad + (a_2b_3 - a_3b_2)e_2e_3 + (a_3b_1 - a_1b_3)e_3e_1 \\ &= \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \wedge \mathbf{b}\end{aligned} \tag{2}$$

Let $e_1e_2e_3 = i$, and $e_1e_2 = e_1e_2e_3e_3 = ie_3$, $e_2e_3 = ie_1$, $e_3e_2 = ie_2$. So

$$\begin{aligned}\mathbf{a} \wedge \mathbf{b} &= (a_1b_2 - a_2b_1)ie_3 + (a_2b_3 - a_3b_2)ie_1 \\ &\quad + (a_3b_1 - a_1b_3)ie_2 \\ &= i(\mathbf{a} \times \mathbf{b})\end{aligned}$$

where $\mathbf{a} \times \mathbf{b}$ is the scalar product in vector algebra.

Suppose that $ie_k = i_k (k = 1,2,3)$, $i_1 = e_2e_3$, $i_2 = e_3e_1$, $i_3 = e_1e_2$, and the 3D vector i_k is the product of the pseudo-scalar i and the three basis vectors e_k. Here $i_1e_2 = e_2e_3e_2 = -e_3$, and $e_2i_1 = e_2e_2e_3 = e_3 = -i_1e_2$. The i is direct volume unit in 3D space.

In 3D space, any rotation is denoted as the result of two vector reflections because the vectors are not always coplanar. The vector x_1 is from the reflection of vector x with the plane I whose normal is basis vector u, depicted in figure 1. And the vector x' is from the reflection of vector x_1 with the plane II whose normal is basis vector v. So vector x rotates to x' with the angle θ, and $\theta = 2\alpha$ where α is the angle between u and v. With Clifford algebra, it is shown as

$$x' = -v(-uxu)v = vuxuv$$

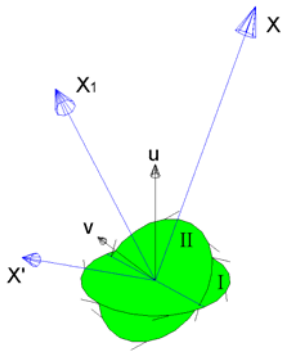


Figure 1. Vector rotation in 3D space.

Let $R = uv$, so

$$x' = R^T xR \tag{3}$$

where R^T is the reversion of R. R is called Spinor [26] which is composed of scalar $u \cdot v$ and bivector $u \wedge v$. In general

$$\begin{aligned}R &= u \cdot v + u \wedge v = u \cdot v + i(u \times v) \\ &= \cos\alpha + in\sin\alpha \\ &= e^{in\alpha}\end{aligned} \tag{4}$$

where n is the basis vector whose direction is decided by $u \times v$. α is the angle between u and v. $R = e^{in\alpha}$ can be written as $R = \alpha + ib$, where α is scalar and b is vector, here $\alpha^2 + b^2 = 1$. Any bivector can be written as the product of i and a vector due to $e_1e_2 = ie_3$, so ib is a bivector. $R^T = e^{-in\alpha}$ if $R = e^{in\alpha}$ while $R^T = \alpha - ib$ if $R = \alpha + ib$ and $R^T = vu$ if $R = uv$. That means $RR^T = R^TR = 1$. $R = e^{in\theta/2}$ if the angle between x and x' is θ.

3. Target Detection in 3D Sensor Networks

Coverage analysis in sensor networks is essentially to determine whether an arbitrary point in a space is perceived by sensor nodes. Previous methods use models of different dimensional geometric targets to determine whether the interested points covered or not in different dimensional space, respectively. For sensor networks with hybrid types of targets, those methods cannot provide a homogeneous and effective coverage analysis in different dimensional subspaces. The geometric operation in Clifford algebra is independent of coordinates with a determinate dimension. Hence, Xie [2] proposed a coverage model based on the rotation operator in 3D space for sensor networks.

Definition 2: Set an omnidirectional sensor node $S = (\mathbf{v}_s, \rho e_1 \wedge e_2 \wedge e_3)$ with perceived radius ρ, the coverage range of S is $\mathbf{D} = \{x \mid \|x - S\| \le \rho\}$, here vector $\mathbf{v}_s \in S$, so the coverage discrimination model for task region $\mathbf{B}$ in space $\mathcal{G}_3$ is given by

$$C(\mathbf{B}) = \begin{cases} 1, & \forall x \in \mathbf{B}, \|\mathbf{R}\|_x \le 1 \\ 0, & \forall x \in \mathbf{B}, \|\mathbf{R}\|_x > 1 \end{cases} \tag{5}$$

where the Spinor $\mathbf{R} = x / \mathbf{v}_s$. Figure 2 illustrates the coverage model.

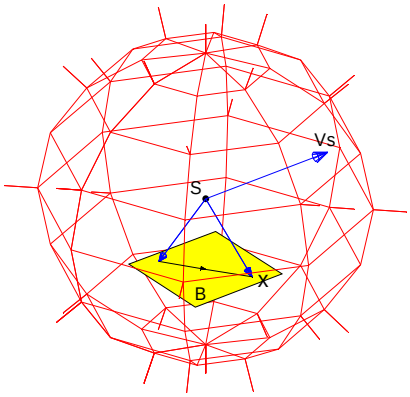


Figure 2. Coverage model based on Clifford Spinor in 3D space.

The relationship between sensor node and target is always depicted by Euler angle, and it would be illustrated by Clifford algebra without coordinates. So the Clifford algebra can present 3×3 Spinor matrix, and the angle transformation is

$$\begin{aligned}\mathbf{R}x &= R_\phi^T R_\theta^T R_\varphi^T x R_\varphi R_\theta R_\phi \\ &= e^{-ie_3\phi/2} e^{-ie_3\theta/2} e^{-ie_3\varphi/2} x e^{ie_3\varphi/2} e^{ie_3\theta/2} e^{ie_3\phi/2}\end{aligned} \quad (6)$$

Let $f_k = \mathbf{R}\,\mathrm{e}_k$, and f_k is the new vector that is isomorphic with e_k in new position after transformation.

Theorem 1: The vector ξ which is perpendicular with axis rotates to ξ', so

$$\xi' = \mathbf{R}\xi = R^T \xi R = \xi R^2 \quad (7)$$

Proof: ξ is perpendicular with vector b in $R = \alpha + ib$ whose direction is same as axis, so $b\xi = b\cdot\xi + b\wedge\xi = b\wedge\xi = -\xi\wedge b = -\xi b$, that means the negative exchange between vectors which are perpendicular with each other. So

$$\begin{aligned}\xi' &= R^T \xi R = (\alpha - ib)\xi(\alpha + ib) \\ &= (\alpha - ib)(\xi\alpha + i\xi b) \\ &= \alpha^2\xi - i\alpha b\xi + i\alpha\xi b + b\xi b \\ &= \xi(\alpha^2 + i\alpha b + i\alpha b - b^2) \\ &= \xi(\alpha + ib)(\alpha + ib) = \xi R^2\end{aligned}$$

Here the exchange among vector ξ, pseudo-scalar i and α is positive, but is negative between ξ and vector b.

With $\xi' = \xi R^2$, consider $R_\phi^T R_\theta^T R_\varphi^T e_1 R_\varphi R_\theta R_\phi$, here

$$\begin{aligned}R_\varphi^T e_1 R_\varphi &= e_1 R_\varphi{}^2 = e_1 e^{ie_3\varphi} \\ &= e_1(\cos\varphi + ie_3\sin\varphi) \\ &= e_1\cos\varphi + e_2\sin\varphi\end{aligned} \quad (8)$$

And

$$\begin{aligned}&R_\theta^T(e_1\cos\varphi + e_2\sin\varphi)R_\theta \\ &= e_1\cos\varphi + e_2\sin\varphi e^{ie_1\theta} \\ &= e_1\cos\varphi + e_2\sin\varphi\cos\theta + e_3\sin\varphi\sin\theta\end{aligned} \quad (9)$$

So

$$\begin{aligned}f_1 &= R_\phi^T(e_1\cos\varphi + e_2\sin\varphi\cos\theta + e_3\sin\varphi\sin\theta)R_\phi \\ &= e_1\cos\varphi(e^{ie_3\phi}) + e_2\sin\varphi\cos\theta(e^{ie_3\phi}) + e_3\sin\varphi\sin\theta \\ &= e_1\cos\varphi\cos\phi + e_2\cos\varphi\sin\phi + e_2\sin\varphi\cos\theta\cos\phi \\ &\quad - e_1\sin\varphi\cos\theta\sin\phi + e_3\sin\varphi\sin\theta\end{aligned} \quad (10)$$

Meanwhile

$$\begin{aligned}f_2 &= e_2\cos\varphi\cos\theta\cos\phi - e_1\cos\varphi\sin\theta\sin\phi - e_1\sin\varphi\cos\phi \\ &\quad - e_2\sin\varphi\sin\phi + e_3\cos\varphi\sin\theta\end{aligned} \quad (11)$$

$$f_3 = e_3\cos\theta - e_2\sin\theta\cos\phi + e_1\sin\theta\sin\phi \quad (12)$$

The Spinor matrix $R_{jk} = (f_j\cdot e_k)$ is

$$R_{jk} = \begin{bmatrix} \cos\varphi\cos\phi - \sin\varphi\cos\theta\sin\phi & \cos\varphi\sin\phi + \sin\varphi\cos\theta\cos\phi & \sin\varphi\sin\theta \\ -\cos\varphi\sin\theta\sin\phi - \sin\varphi\cos\phi & \cos\varphi\cos\theta\cos\phi - \sin\varphi\sin\phi & \cos\varphi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{bmatrix} \quad (13)$$

With the method of Clifford algebra, we can easily calculate the 3×3 Spinor matrix of arbitrary angle θ rotating around any vector n. The element of this matrix is

$$R_{jk} = f_j \bullet e_k = (R' e_j R)\bullet e_k \quad (14)$$

Generally, it is more suitable that the element under this circumstance can be written as $R_{jk} = <(R^T e_j R)\bullet e_k>_0$. Here $<\ >_0$ denotes the vector of grade zero which is scalar part, because the result will normally be $<\ >_0 + <\ >_1 + <\ >_2 + <\ >_3$. We have known that the established R_{jk} just is a simple scalar quantity, so we don't have to denote its vector of grade zero, where

$$R^T e_j R = (\alpha - ib)e_j(\alpha + ib)$$
$$= (\alpha - ib)(\alpha e_j + ie_j b)$$
$$= \alpha^2 e_j - i\alpha b e_j + i\alpha e_j b + be_j b$$

With

$$i\alpha(e_j b - be_j) = 2i\alpha(e_j \wedge b) = 2i\alpha i(e_j \times b) = -2\alpha(e_j \times b),$$

$$be_j b = b \bullet (e_j \bullet b + e_j \wedge b) + b \wedge (e_j \bullet b + e_j \wedge b)$$
$$= b \bullet e_j b - b \bullet be_j + e_j \bullet bb = 2b \bullet e_j b + b^2 e_j$$

Hence

$$R^T e_j R = (\alpha^2 - b^2)e_j - 2\alpha(e_j \times b) + 2b \bullet e_j b,$$

And

$$(R^T e_j R) \bullet e_j = (\alpha^2 - b^2)\delta_{jk} - 2\alpha(e_j \times b) \bullet e_j + 2(b \bullet e_j)(b \bullet e_k)$$

where $\alpha = \cos\frac{\theta}{2}$, $b = n\sin\frac{\theta}{2}$, θ is the angle, n is the unit vector for rotating. (3.10) can be written as

$$R_{jk} = \cos\theta \bullet \delta_{jk} - \sin\theta(e_j \times n) \bullet e_k + (1 - \cos\theta)n_j n_k \quad \textbf{(15)}$$

Here δ_{jk} is Kronecker symbol. When $j = k$, $\delta_{jk} = 1$, otherwise $\delta_{jk} = 0$. The second part in (15) is a general vector operation which can be easily calculated by specific j, k because $(e_j \times n) \bullet e_k = -(e_j \times e_k) \bullet n$. For example, $j = 2$, $k = 3$, and $e_2 \times e_3 = e_1$, so $(e_j \times e_k) \bullet n = e_1 \bullet n = n_1$. Therefore, the second part is $n_1 \sin\theta$ here which can infer the other parts. The final result is

$$R_{jk} = \begin{bmatrix} \cos\theta + n_1^2(1-\cos\theta) & n_1 n_2(1-\cos\theta) + n_3\sin\theta & n_1 n_3(1-\cos\theta) - n_2\sin\theta \\ n_1 n_2(1-\cos\theta) - n_3\sin\theta & \cos\theta + n_2^2(1-\cos\theta) & n_2 n_3(1-\cos\theta) + n_1\sin\theta \\ n_3 n_1(1-\cos\theta) + n_2\sin\theta & n_3 n_1(1-\cos\theta) - n_1\sin\theta & \cos\theta + n_3^2(1-\cos\theta) \end{bmatrix}$$

The conventional methods can also achieve this result, but this method is straight and do not need the graph for help. Hence the target moving can be detected in coverage area of sensor networks by Clifford algebra, as shown in figure 3. If the target moves from x to x', that is

$$x' = R^T xR + a = \mathbf{R}x + a \quad \textbf{(16)}$$

where x is the position of target.

Each element in (16) would be time function, so the target moving formulation is

$$\dot{x}' = \dot{\mathbf{R}}x + \mathbf{R}\dot{x} + \dot{a} \quad \textbf{(17)}$$

Theorem 2: In (17), $\dot{\mathbf{R}}x = \dot{R}^T xR + R^T x\dot{R}$ can be calculated by (18):

$$\dot{\mathbf{R}}x = R^T(\omega \times x)R \quad \textbf{(18)}$$

where ω is target's angular velocity to sensor node.

Proof: Suppose that

$$\dot{R}x = \frac{1}{2}R\Omega$$

And

$$\dot{R}^T R + R^T \dot{R} = 0$$
$$\frac{1}{2}\Omega^T R^T R + \frac{1}{2}R^T \dot{R}\Omega = 0$$
$$\Omega^T = -\Omega, \dot{R}^T = -\frac{1}{2}\Omega R^T$$

Here $R^T R = 1$, $\dot{R}^T = (\dot{R})^T$. Ω is bivector, and can be denoted as $\Omega = i\omega$. So $\Omega = 2R^T\dot{R}$ or $i\omega = 2R^T\dot{R}$.

Thus

$$\dot{\mathbf{R}}x = \dot{R}^T xR + R^T x\dot{R} = -\frac{1}{2}\Omega R^T xR + \frac{1}{2}R^T xR\Omega = \frac{1}{2}R^T(-R\Omega R^T x + xR\Omega R^T)R$$

and

$$\dot{\mathbf{R}}x = R^T\frac{1}{2}(x\Omega - \Omega x)R = \mathbf{R}x \cdot \Omega = R^T(x \cdot \Omega)R$$

Or

$$\dot{\mathbf{R}}x = R^T\frac{i}{2}(x\omega - \omega x)R = R^T(\omega \times x)R$$

With $\omega \wedge x = i(\omega \times x)$ and $\omega \times x = -i(\omega \wedge x)$,

$$\dot{x}' = R^T[(\omega \times x) + \dot{x}]R + \dot{a} \quad \textbf{(19)}$$

The target moving formulation can further be denoted as

$$\ddot{x}' = \dot{\mathbf{R}}[(\omega \times x) + \dot{x}] + \mathbf{R}[(\dot{\omega} \times x) + (\omega \times \dot{x}) + \ddot{x}] + \ddot{a} \quad \textbf{(20)}$$

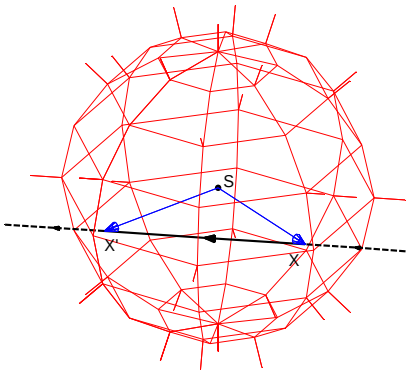


Figure 3. The target moving path detected by sensor node.

where

$$\begin{aligned}\dot{\mathbf{R}}(\omega\times x) &= \dot{R}^T(\omega\times x)R + R^T(\omega\times x)\dot{R}\\ &= -\frac{1}{2}i\omega' R^T(\omega\times x)R + \frac{1}{2}R^T(\omega\times x)Ri\omega'\\ &= R^T\frac{i}{2}[(\omega\times x)R\omega' R^T - R\omega' R^T(\omega\times x)]R\\ &= R^T\frac{i}{2}[(\omega\times x)\omega - \omega(\omega\times x)]R\\ &= R^T\omega\times(\omega\times x)R\\ &= R^T(x\cdot\Omega)\cdot\Omega R\end{aligned}$$

And

$$\dot{\mathbf{R}}\dot{x} = \dot{R}^T\dot{x}R + R^T\dot{x}\dot{R} = R^T(\omega\times\dot{x})R = R^T\dot{x}\cdot\Omega R\,.$$

Therefore,

$$\ddot{x}' = R^T[\ddot{x} + 2(\omega\times\dot{x}) + \omega\times(\omega\times x) + \dot{\omega}\times x]R + \ddot{a} \quad \textbf{(21)}$$

In the 3D space, the movement formulation of the target m is

$$m\ddot{x}' = g'\,,$$

where is the power of the target so

$$R^T m[\ddot{x} + 2(\omega\times\dot{x}) + \omega\times(\omega\times x) + \dot{\omega}\times x]R + m\ddot{a} = g'$$

$$m[\ddot{x} + 2(\omega\times\dot{x}) + \omega\times(\omega\times x) + \dot{\omega}\times x] + Rm\ddot{a}R^T = Rg'R^T = g$$

$$m\ddot{x} = g - m[2(\omega\times\dot{x}) + \omega\times(\omega\times x) + \dot{\omega}\times x] - Rm\ddot{a}R^T$$

We can achieve the movement formulation of the target in sensor networks, which will help us to analyze the state of the detected target.

4. Algorithm

With the movement formulation of the target in 3D sensor networks, we propose a motive target detection algorithm based on Clifford Spinor. We assume that the motive target in sensor networks can be detected by some nodes, and its positions in the range of these nodes also can be calculated. Because the surveillance range of a node is always small, the track of the target in the range will be considered as a line approximately. Our algorithm for tracking the motive target is following:

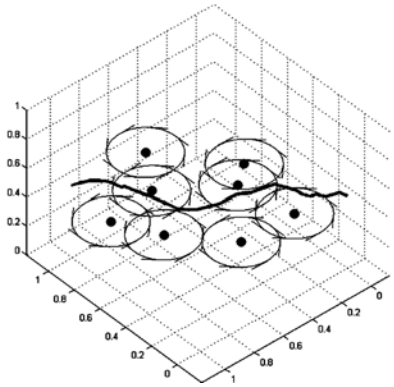
1) Detect the two vertexes of the track when the target is in the range of node;

2) Use (17) or (20) to calculate the movement formulation of the target;

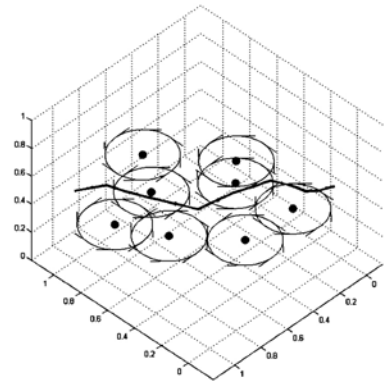
3) Use the movement states of target to estimate the direction and velocity so as to inform the next correlative node turning on and starting to detect the target;

4) Collect all information from each node, and achieve the track of the motive target in the sensor networks.

Figure 4 shows the motive target detection using moving formulation calculated by Clifford Spinor in 3D sensor networks, where (a) is the target's actual path to traverse sensor networks, and (b) is the target's traversing path achieved by moving formulation. The target positions in sensor's coverage area can be used to calculate the moving formulation to get the traversing path in this area. Connecting all of these paths which are achieved from each node can get the target's track in sensor networks. It is helpful for target tracking and forecasting.



(a)



(b)

Figure 4. Target detection in 3D sensor networks.

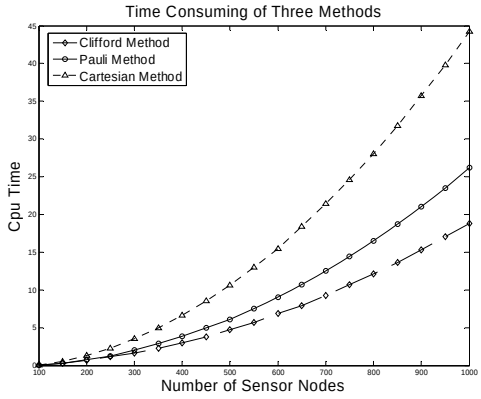


Figure 5. The comparison of time consuming among three methods.

Figure 5 denotes the time consuming to calculate target moving formulation using Cartesian method, polar coordinates method and Clifford method, respectively. It is obvious that the time consuming using Clifford method is lower than that of the other two methods when the number of sensor nodes increases. This is because the Cartesian method should use the information of three axes in 3D space and the polar coordinates method would calculate three 3×3 Pauli matrices. The quantity of calculation is decreased in Clifford method due to just using Spinor equation to get the moving formulation without axes information in 3D space.

5. Conclusions

In this paper, we developed the coverage model of 3D sensor networks with the coverage analysis approach based on Clifford algebra, which establishes a coordinate-free, homogeneous coverage model for different dimensional spaces and targets with hybrid types, and proposed the method for detecting target moving. With Clifford Spinor, calculating the target moving formulation is easier than traditional methods in sensor node's coverage area. It is helpful to the research of coverage analysis in sensor networks.

6. Acknowledgements

This research is supported by National Natural Science Foundation (No.60872126) and Guangdong Province Natural Science Foundation (No.8151806001000002).

7. References

[1] C. Tian, W. Liu, J. Jin, Y. Wang, and Y. Mo, "Localization and synchronization for 3D underwater acoustic sensor network," Proceedings of UIC 2007, pp. 622–631, July 2007.

[2] W. X. Xie, W. M. Cao, and S. Meng, "Coverage analysis for sensor networks based on Clifford algebra," Science in China Series F-Information Sciences, Vol. 51, No. 5, pp. 460–475, 2008.

[3] W. X. Xie, T. C. He, and W. M. Cao, "Path analyses of sensor networks based on Clifford algebra," Acta Electronica Sinica, Vol. 35, No. 12A, pp. 27–31, 2007.

[4] S. Megerian, F. Koushanfar, M. Potkonjak, M. B. Srivastava, "Worst and best-case coverage in sensor networks," IEEE Transactions on Mobile Computing, Vol. 4, No. 1, pp. 84–92, 2005.

[5] X. Li, P. Wan, and O. Frieder, "Coverage problem in wireless ad-hoc sensor networks," IEEE Transactions on Computers, Vol. 52, No. 6, pp. 753–763, 2003.

[6] T. C. He, W. M. Cao, and W. X. Xie, "Research of breach issue for sensor networks based on Clifford geometrical algebra," Chinese Journal of Scientific Instrument, Vol. 28, No. 8, pp. 161–165, 2007.

[7] A. D. Gupta, A. Bishnu, and I. Sengupta, "Optimisation problems based on the maximal breach path measure for wireless sensor network coverage," ICDCIT, pp. 27–40, 2006.

[8] G. Veltri, Q. Huang, G. Qu, and M. Potkonjak, "Minimal and maximal exposure path algorithms for wireless embedded sensor networks," Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys'03), Los Angels, CA, USA, pp. 40–50, 2003.

[9] L. Lazos, R. Poovendran, and J. A. Ritcey, "Probabilistic detection of mobile targets in heterogeneous sensor networks," in Proceedings of IPSN'07, pp. 519–528, 2007.

[10] W. Z. Zhang, M. L. Li, and M. Y. Wu, "An algorithm for target traversing based on local Voronoi diagram," Journal of Software, Vol. 18, No, 5, pp. 1246–1253, 2007.

[11] C. D. Jiang and G. L. Chen, "Double barrier coverage in dense sensor networks," Journal of Computer Science and Technology, Vol. 23, No. 1, pp. 154–165, 2008.

[12] M. Cardei, M. T. Thai, Y. Li, and W. Wu, "Energy-efficient target coverage in wireless sensor networks," Proceedings of IEEE INFOCOM'05, Miami, FL, pp. 1976–1983, 2005.

[13] M. Cardei, J. Wu, M. Lu, and M. O. Pervaiz, "Maximum network lifetime in wireless sensor networks with adjustable sensing ranges," Proceedings of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'05), Montreal, Canada, pp. 1–8, 2005.

[14] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava, "Coverage problems in wireless ad-hoc sensor networks," in IEEE INFOCOM, Vol. 3, pp. 1380–1387, 2001.

[15] S. Meguedrichian, S. Slijepcevic, V. Karayan, and M. Potkonjak, "Localized algorithms in wireless ad-hoc networks: Location discovery and sensor exposure," Proceedings of ACM MobiHoc'01, Long Beach, CA, USA, pp. 106–116, 2001.

[16] M. Lu, J. Wu, M. Cardei, and M. Li, "Energy-efficient connected coverage of discrete targets in wireless sensor networks," Proceedings of 2005 International Conference on Computer Networks and Mobile Computing (ICCN-MC'05), Zhangjiajie, China, pp. 1–10, 2005.

[17] X. Li, P. Wan, and O. Frieder, "Coverage in wireless ad-hoc sensor networks," Proceedings of IEEE International Conference on Communications (ICC'02), New York, NY, USA, pp. 1–5, 2002.

[18] S. Meguerdichian, F. Koushanfar, G. Qu, and M. Potkonjak, "Exposure in wireless ad hoc sensor networks," Proceedings of ACM MobiCom'01, pp. 139–150, 2001.

[19] S. Adlakha and M. Srivastava, "Critical density thresholds for coverage in wireless sensor networks," Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'03), NewOrleans, Louisiana, USA, pp. l6–20, 2003.

[20] B. Wang, et al., "Worst and best information exposure paths in wireless sensor networks," Proceedings of MSN 2005, pp. 52–62, December 2005.

[21] B. Liu and D. Towsley, "On the coverage and detectability of large-scale wireless sensor networks," Proceedings of WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, INRIA Sophia-Antipolis, France, pp. 1–3, 2003.

[22] R. Meester and R. Roy, "Continuum percolation," Balletin of the American Mathematical Society, Vol. 34, No. 4, pp. 447–448, 1997.

[23] B. Liu and D. Towsley, "A study of the coverage of large-scale sensor networks," Proceedings of the 1st IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS'04), Florida, USA, pp. 475–481, 2004.

[24] S. Kumar, T. Lai, and A. Arora, "Barrier coverage with wireless sensors," Wireless Network, Vol. 13, pp. 817–834, 2007.

[25] S. Kloder and S. Hutchinson, "Barrier coverage for variable bounded-range line-of-sight guards," in Proceedings of IEEE International Conference on Robotics and Automation 2007, pp. 391–396, April 2007.

[26] J. O'Rourke, "Computational geometry: Column 15," International Journal of Computational Geometry and Applications, Vol. 2, No. 2, pp. 215–217, 1992.

Wireless Sensor Network, **2009, 2, 61-121**
doi:10.4236/wsn.2009.12014 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

Gaussian Convolution Filter and its Application to Tracking

Qing LIN, Jianjun YIN, Jianqiu ZHANG, Bo HU
Electronic Engineering Department, Fudan University, Shanghai, China
E-mail: *qlin98@fudan.edu.cn*
Received April 29, 2009; *revised May* 11, 2009; *accepted May* 12, 2009

Abstract

A new recursive algorithm, called the Gaussian convolution filter (GCF), is proposed for nonlinear dynamic state space models. Based on the convolution filter (CF) and similar to the Gaussian filters, the GCF approximates the posterior density of the states by Gaussian distribution. The analytical results show the ability to deal with complex observation model and small observation noise of the GCF over the Gaussian particle filter (GPF) and the lower complexity, more amenable for parallel implementation than the CF. The Simulation in the Tracking domain demonstrates the good performance of the GCF.

Keywords: Signal Processing, Tracking, Nonlinear Estimation

1. Introduction

To estimate the dynamic state from the history of noisy observations is the main objective of filtering, which arise in many fields including statistics, economics and engineering such as tracking and navigation. Based on the difference of the dynamic state space models, usually filtering can be divided into two categories: linear and nonlinear, which correspond to the linear Gaussian models and nonlinear and/or non-Gaussian models respectively. For linear filtering, Kalman filter (KF) [1] usually gives the optimal results. For nonlinear filtering, the extended Kalman filter (EKF) [2] is most popular. However, the linearization process of the EKF is liable to large errors threatening the convergence of the algorithm, particularly for models with high nonlinearity. A recently-popularized technique for numerical approximation, termed as the particle filter (PF) [3−5], offers a general tool for the state estimation of nonlinear non-Gaussian systems. The core idea behind the PF is to use samples (particles) to approximate the concerned distributions. Usually the PF gives better results than the EKF and the unscented Kalman filter (UKF) [6]. However, it also has drawbacks. Firstly, the algorithm is complex and difficult to parallel implementation [7]. Secondly it is prone to divergence when the observation noise is too small [8]. Thirdly, it does not work when the likelihood function can not be obtained analytically [8]. The first drawback has been overcome by the Gaussian particle filter (GPF) [7], which approximates the posterior distributions by single Gaussians, and avoid the resampling step, which reduces the complexity and is more amenable for fully parallel implementation in VLSI. However the second and third shortages are still within the GPF. The convolution filter (CF) [8] has circumvented the second and third drawbacks by using convolution kernels, however, the first one still remains.

In this paper, we propose a new algorithm, namely the Gaussian convolution filter (GCF), which can overcome all the three drawbacks above. The GCF is based on the convolution filtering concept, and it approximates the posterior distributions by single Gaussians. It is shown that the GCF is asymptotically optimal in the number of particle under the Gaussianity assumption. The Simulation results in the Tracking demonstrate the performance of the GCF when the observation noise is too small and the GPF fails.

2. Background

In this section, we first describe the problem formulation. The convolution filter is then recalled.

2.1. Problem Formulation

Let the following general model of a state space system

be considered [3]:

$$x_t = f_t(x_{t-1}, \omega_t) \quad \textbf{(1)}$$

$$y_t = h_t(x_t, \upsilon_t) \quad \textbf{(2)}$$

where ω_t and υ_t denote the known independent process and measurement noise respectively, $\{x_t\}$ the state of the system complying with the Markov process, $\{y_t\}$ the measurements of the system, and both f_t and h_t are known nonlinear or linear functions. Again, let $X_t = \{x_1, \cdots x_t\}$, $Y_t = \{y_1, \cdots y_t\}$, and $p(x_0|x_{-1}) = p(x_0)$ be the initial density. Here, the purpose is supposed to estimate the posterior probability density function (pdf) $p(x_t|Y_t)$ by the Bayes recursion

$$p(x_t \mid Y_{t-1}) = \int p(x_t \mid x_{t-1})\, p(x_{t-1} \mid Y_{t-1})\, dx_{t-1} \quad \textbf{(3)}$$

$$p(x_t|Y_t) = p(y_t|x_t)p(x_t|Y_{t-1}) \Big/ \left(\int p(y_t|x_t)p(x_t|Y_{t-1})dx_t\right) \quad \textbf{(4)}$$

2.2. The Convolution Filter

Usually in PF scheme, the weight of each particle is given by the likelihood function. The observations thus operate the filter through the likelihood function which is assumed to exist and to be known. This assumption is rather restricting in practice. Moreover it rules out the non-noisy case and will also cause trouble when the observation noise is too small and also when the noise is non-additive as in the general system (1) and (2). These drawbacks can be circumvented by using convolution kernels to weight the particles in the CF [8]. We can approximate the weights consistently by simulating the observations, and use this approximation in place of the true function in the PF, i.e.

$$w_t^{(i)} \propto p\left(y_t \mid x_t^{(i)}\right) \approx K_{hn}^z\left(y_t - y_t^{(i)}\right) \quad \textbf{(5)}$$

where $w_t^{(i)}$ denote particle weights, K_{hn}^z is Parzen-Rosenblatt kernel of appropriate dimensions (A Parzen-Rosenblatt kernel is a bounded positive and symmetric function for which $\int K d\lambda = 1$, where λ is the Lebesgue measure, and $\lim \|x\|^d K(x) = 0$ as $\|x\| \to \infty$, where d is the dimension of variable y and $\|\cdot\|$ denotes the squared norm), h_n is called the kernel bandwidth, and $\hat{y}_t^{(i)}$ are the samples from observation. Then we can get the estimation as

$$p(x_t \mid Y_t) = \sum_{i=1}^n w_t^{(i)} K_{hn}^x\left(x_t - \hat{x}_t^{(i)}\right) \Big/ \sum_{i=1}^n w_t^{(i)} \quad \textbf{(6)}$$

A brief description of the resampled CF (RCF) is given in Table 1.

Table 1. The RCF algorithm.

For $t = 0$

Given p_0 be the probability density of the initial state distribution

For $t \geq 1$

(1) resample: $\left\{x_{t-1}^{(i)}\right\}_{i=1}^n \sim p_{t-1}$

(2) evolving: $x_t^{(i)} \sim f_t(x_{t-1}^{(i)}, \cdot)$, $y_t^{(i)} \sim h_t(x_t^{(i)}, \cdot)$, $i = 1, \cdots, n$

(3) estimation:

$$p_t(x_t \mid Y_t) = \frac{\sum_{i=1}^n K_{hn}^y\left(y_t - y_t^{(i)}\right) K_{hn}^x\left(x_t - x_t^{(i)}\right)}{\sum_{i=1}^n K_{hn}^y\left(y_t - y_t^{(i)}\right)}.$$

3. The Gaussian Convolution Filter

In this section, we present the main results of the paper, the GCF recursion. The main idea behind the GCF is to estimate the posterior distributions by CF, and then approximate them by Gaussians.

3.1. The Measurement Update

Assume the density of prediction is approximated by a Gaussian [7], i.e.,

$$p(x_t \mid Y_{t-1}) \approx \mathrm{N}\left(x_t; \bar{\mu}_t, \bar{\Sigma}_t\right) \quad \textbf{(7)}$$

where $\mathrm{N}(x; \mu, \Sigma)$ denotes the Gaussian distribution with the mean μ and covariance Σ. Take (7) as the importance density and get samples from it, i.e.,

$$x_t^{(i)} \sim \mathrm{N}\left(x_t; \bar{\mu}_t, \bar{\Sigma}_t\right), \quad i = 1, \cdots, N \quad \textbf{(8)}$$

Obtain the observations

$$y_t^{(i)} \sim h_t(x_t^{(i)}, \cdot), \quad i = 1, \cdots, N \quad \textbf{(9)}$$

and the particle weights

$$w_t^{(i)} \propto \hat{p}\left(y_t \mid x_t^{(i)}\right) = K_{hn}^z\left(y_t - y_t^{(i)}\right) \quad \textbf{(10)}$$

By substituting (7) into (4) we get

$$p(x_t \mid Y_t) \approx m_t \cdot p(y_t \mid x_t) \cdot \mathrm{N}\left(x_t; \bar{\mu}_t, \bar{\Sigma}_t\right) \quad \textbf{(11)}$$

where

$$m_t = 1 \Big/ \int p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t \quad \textbf{(12)}$$

We approximate (11) by a Gaussian, i.e.,

$$p(x_t \mid Y_t) \approx \mathrm{N}\left(x_t; \mu_t, \Sigma_t\right) \quad \textbf{(13)}$$

where

$$\begin{aligned} \mu_t &= \sum_{i=1}^N w_t^{(i)} x_t^{(i)} \Big/ \sum_{i=1}^N w_t^{(i)} \\ \Sigma_t &= \sum_{i=1}^N w_t^{(i)} \left(x_t^{(i)} - \mu_t\right)\left(x_t^{(i)} - \mu_t\right)^T \Big/ w_t^{(i)} \end{aligned} \quad \textbf{(14)}$$

where A^T denotes the transpose of matrix *A*. We now give the corollary to verify the convergence of the algorithm above. Let us note that the form of corollary here is similar to that in GPF [7], however the difference is that the one here is based on the CF.

Corollary 1: Assume that at time t, the prediction distribution is Gaussian, i.e. $p(x_t \mid Y_{t-1}) \approx \mathrm{N}(x_t; \bar{\mu}_t, \bar{\Sigma}_t)$. The GCF measurement updates the filtering distribution by the algorithm above. Then, μ_t computed in (14) converges to the MMSE estimate of x_t almost surely as $N \to \infty$, and the MMSE estimate given by Σ_t in (14) converges to the true MMSE estimate almost surely as $N \to \infty$.

Proof: 1) mean

$$\begin{aligned}
\mu_t &= \hat{E}(x_t \mid Y_t) = \frac{\sum_{i=1}^N x_t^{(i)} w_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}} = \frac{\sum_{i=1}^N x_t^{(i)} \hat{p}(y_t \mid x_t^{(i)})}{\sum_{i=1}^N \hat{p}(y_t \mid x_t^{(i)})} \\
&\xrightarrow{x_t^{(i)} \sim p(x_t \mid Y_{t-1})} \frac{\int x_t \hat{p}(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t}{\int \hat{p}(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t} \\
&\approx \frac{\int x_t p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t}{\int p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t} \\
&= \frac{\int x_t p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t}{p(y_t \mid Y_{t-1})} \\
&= \int x_t p(x_t \mid Y_t)\, dx_t = E(x_t \mid Y_t)
\end{aligned} \tag{15}$$

2) covariance

$$\begin{aligned}
\Sigma_t &= \hat{E}\left(\left(x_t - \hat{E}(x_t \mid Y_t)\right)\left(x_t - \hat{E}(x_t \mid Y_t)\right)^T \mid Y_t \right) \\
&= \frac{\sum_{i=1}^N \left(x_t^{(i)} - \mu_t\right)\left(x_t^{(i)} - \mu_t\right)^T w_t^{(i)}}{\sum_{i=1}^N w_t^{(i)}} \\
&= \frac{\sum_{i=1}^N \left(x_t^{(i)} - \mu_t\right)\left(x_t^{(i)} - \mu_t\right)^T \hat{p}(y_t \mid x_t^{(i)})}{\sum_{i=1}^N \hat{p}(y_t \mid x_t^{(i)})} \\
&\xrightarrow{x_t^{(i)} \sim p(x_t \mid Y_{t-1})} \frac{\int \left(x_t - E(x_t \mid Y_t)\right)\left(x_t - E(x_t \mid Y_t)\right)^T \cdot \hat{p}(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t}{\int \hat{p}(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t} \\
&\approx \frac{\int \left(x_t - E(x_t \mid Y_t)\right)\left(x_t - E(x_t \mid Y_t)\right)^T p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t}{\int p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t} \\
&= \frac{\int \left(x_t - E(x_t \mid Y_t)\right)\left(x_t - E(x_t \mid Y_t)\right)^T p(y_t \mid x_t)\, p(x_t \mid Y_{t-1})\, dx_t}{p(y_t \mid Y_{t-1})} \\
&= \int \left(x_t - E(x_t \mid Y_t)\right)\left(x_t - E(x_t \mid Y_t)\right)^T p(x_t \mid Y_t)\, dx_t \\
&= E\left(\left(x_t - E(x_t \mid Y_t)\right)\left(x_t - E(x_t \mid Y_t)\right)^T \mid Y_t \right)
\end{aligned} \tag{16}$$

3.2. The Time Update

By substituting (13) to (3) we have

$$\begin{aligned}
p(x_{t+1} \mid Y_t) &= \int p(x_{t+1} \mid x_t)\, p(x_t \mid Y_t)\, dx_t \\
&\approx \int p(x_{t+1} \mid x_t)\, \mathrm{N}(x_t; \mu_t, \Sigma_t)\, dx_t
\end{aligned} \tag{17}$$

Draw samples

$$x_t^{(i)} \sim \mathrm{N}(x_t; \mu_t, \Sigma_t), \quad i = 1, \cdots, N \tag{18}$$

and then a Monte Carlo approximation of the predictive distribution is given by

$$p(x_{t+1} \mid Y_t) \approx \frac{1}{N} \sum_{i=1}^N p\left(x_{t+1} \mid x_t^{(i)}\right) \tag{19}$$

Obtain samples at time t+1 from the process model by

$$x_{t+1}^{(i)} \sim f_{t+1}(x_t^{(i)}, \cdot) \tag{20}$$

from which the mean and covariance of $p(x_{t+1} \mid Y_t)$ are computed as

$$\begin{aligned}
\bar{\mu}_{t+1} &= \frac{1}{M} \sum_{i=1}^N x_{t+1}^{(i)} \\
\bar{\Sigma}_{t+1} &= \frac{1}{M} \sum_{i=1}^N \left(x_{t+1}^{(i)} - \bar{\mu}_{t+1}\right)\left(x_{t+1}^{(i)} - \bar{\mu}_{t+1}\right)
\end{aligned} \tag{21}$$

Recall that the prediction distribution is approximated as a Gaussian, we have

$$p(x_{t+1} \mid Y_t) \approx \mathrm{N}(x_{t+1}; \bar{\mu}_{t+1}, \bar{\Sigma}_{t+1}) \tag{22}$$

3.3. Summary of the GCF

We summarize the algorithm above in Table 2.

The GCF does away with the need of resampling step, this means that the GCF is more amenable for fully parallel implementation in VLSI than the CF. Moreover, because of the use of convolution kernels the GCF can deal with scenarios that the observation noise is too small or the likelihood function can not be obtained analytically.

4. Tracking Simulation Results

An example: Consider the problem of tracking a maneuvering target [9], whose position and velocity at instant t are given by a continuous random vector $x_{2t} \in R^{n-1}$, and where the maneuver/regime of the target is represented by the discrete random variable $x_{1k} \in R$. The state to be estimated is $x_t = \{x_{1t}; x_{2t}\}$. The model is as follows:

$$X_{2t} = Fx_{2t-1} + Bx_{1t} + w_t;$$

$$Z_t = Cx_{2t} + e_t$$

Table 2. The GCF algorithm.

For $t=0$

Given $p(x_0)=\mathrm{N}(x_0;\mu_0,\Sigma_0)$

For $t\geq 1$

(1) Time update

(i) Draw samples from posterior density

$x_{t-1}^{(i)}\sim \mathrm{N}(x_{t-1};\mu_{t-1},\Sigma_{t-1}),\ \ i=1,\cdots,N$

(ii) Draw samples from the process model

$x_t^{(i)}\sim f_t(x_{t-1}^{(i)},\cdot),\ \ i=1,\cdots,N$

(iii) Compute the mean and covariance

$$\bar{\mu}_t=\frac{1}{M}\sum_{i=1}^{N}x_{t1}^{(i)}$$

$$\bar{\Sigma}_t=\frac{1}{M}\sum_{i=1}^{N}\left(x_t^{(i)}-\bar{\mu}_t\right)\left(x_t^{(i)}-\bar{\mu}_t\right)^T$$

(2) Measurement update

(i) Draw samples from importance density

$x_t^{(i)}\sim \mathrm{N}(x_t;\bar{\mu}_t,\bar{\Sigma}_t),\ \ i=1,\cdots,N$

(ii) Draw samples from the observation model

$y_t^{(i)}\sim h_t(x_t^{(i)},\cdot)$

(iii) Compute and normalized the weights

$w_t^{(i)}\propto K_{hn}^{z}\left(y_t-\hat{y}_t^{(i)}\right)$

$w_t^{(i)}=w_t^{(i)}/\sum_{i=1}^{N}w_t^{(i)}$

(iv) Compute the mean and covariance

$\mu_t=\sum_{i=1}^{N}w_t^{(i)}x_t^{(i)}$

$\Sigma_t=\sum_{i=1}^{N}w_t^{(i)}\left(x_t^{(i)}-\mu_t\right)\left(x_t^{(i)}-\mu_t\right)$

Additionally, given

$$p(x_{1t}\mid x_{1t-1})=\sqrt{{x_{1t-1}}^2+0.01t}+\theta_t,\quad \theta_t\sim N(0,10).$$

w_t and e_t are zero mean Gaussian noises, with covariance matrices Q and R. Since given X_{1t}, the dynamics of x_{2t} are linear-Gaussian. In our model, we use $x_{2t}=[x_t\ y_t\ x_t'\ y_t']^T$ where (x; y) is the position of the target in a cartesian plane. We take :

F=[1,0,0.3,0;0,1,0,0.3;0,0,1,0;0,0,0,1]T,

B=[1.25;-1.25;0.25;-0.25]T,

C=[1,0,0,0;0,1,0,0].

We use the simulation data as follows:

$\omega_t\sim N\left(0,\left(0.01,0;0,0.01\right)\right)$, $e^t\sim N(0,R)$, $\hat{x}_{0|1}=$ (20, 30, 0.5, 0.5)^T where the measurement noise variance varies during simulation. Both GCF and GPF are adopted in the simulation. Figure 1 shows the absolute value of errors of the state estimates given by the GPF (denoted by asterisk-solid line) and GCF (denoted by dashed line) respectively when the observation noise variance R=5. In this case both the GPF and the GCF works well, also shown in Table 3. However, when R is reduced, e.g. R=0.1, the GPF diverges while the proposed GCF still works well, as shown in Table 3, where ϕ means the divergence of the results as shown in [8], and the RMSE

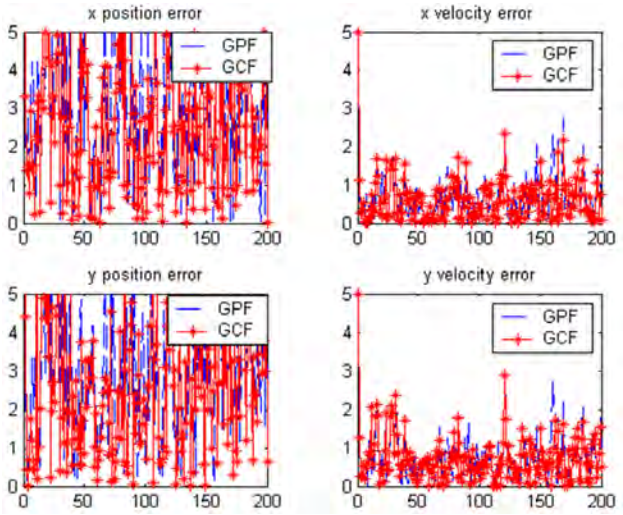


Figure 1. Absolute value of estimated error (R=5).

Table 3. Detailed simulation results.

meth-ods	*R*	x position RMSE	y position RMSE	x velocity RMSE	y velocity RMSE
GPF	10	8.691727	9.132639	43.681761	42.659641
	1	0.936217	0.898894	4.448537	5.993423
	0.1	ϕ	ϕ	ϕ	ϕ
GCF	10	8.777454	9.123733	46.752717	47.554381
	1	0.951737	0.901504	3.882222	5.185661
	0.1	0.189300	0.198540	1.274176	1.247551

is calculated according to

$$\sqrt{\sum_{i=1}^{t}\left(\hat{x}_i-x_i\right)^2\Big/N},$$

where $\hat{x}_i$ and x_i are the estimated and true values respectively, N denotes the total estimation times.

5. Conclusions

The proposed Gaussian convolution filter (GCF) overcomes the drawbacks of the existing GPF, which is limited in the applications to scenarios that have non-noisy or near non-noisy observations or lack the knowledge of the likelihood function. Moreover the parallelizability of the GCF and the absence of resampling step makes it more convenient for VLSI implementation and, hence, feasible for practical real-time applications than the existing CF. Simulation results are also presented to demonstrate the performance of the GCF when the observation noise is small and the GPF fails.

6. References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," Transactions of the ASME–Journal

of Basic Engineering, Vol. 82-D, pp. 35–45, 1960.

[2] M. K. Steven, "Fundamentals of statistical signal processing," PTR Prentice-Hall, Englewood Cliffs, N.J., 1993.

[3] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," IEE Proceedings-F, Vol. 140, No. 2, pp. 107–113, 1993.

[4] A. Doucet, F. Nando, and N. J. Gordon, "Sequential Monte Carlo in practice," Springer, New York, 2001.

[5] T. Schon, F. Gustafsson, and P. J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," IEEE Transactions on Signal Processing 2005, Vol. 53, No. 7, pp. 2279–2289.

[6] J. J. Simon and K. U. Jeffrey, "Unscented filtering and nonlinear estimation," Proceedings of the IEEE, Vol. 92, No. 3, pp. 401–422, 2004.

[7] J. H. Kotecha and P. M. Djuric, "Gaussian particle filtering," IEEE Transactions on Signal Processing, Vol. 51, No. 10, pp. 2592–2601, 2003.

[8] V. Rossi and J.-P. Vila, "Nonlinear filter in discreet time: A particle convolution approach," Biostatic Group of Monetepellier, Technical Report 04-03, 2004. http://vrossi.free.fr/recherche.html.

[9] F. Mustiere, M. Bolic, and M. Bouchard, "A modified Rao-blackwellised particle filter[C]," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006, Toulouse, France, Vol. 3, pp. 21–24, May 2006.

Wireless Sensor Network, 2009, 2, 61-121
doi:10.4236/wsn.2009.12015 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

Recurrent Polynomial Neural Networks for Enhancing Performance of GPS in Electric Systems

Mohammad Reza MOSAVI
Associate Professor, Department of Electrical Engineering,
Iran University of Science and Technology, Narmak, Iran
Email: *M_Mosavi@iust.ac.ir*
Received April 2, 2009; *revised April* 28, 2009; *accepted April* 30, 2009

Abstract

Global Positioning System (GPS) is a worldwide satellite system that provides navigation, positioning, and timing for both military and civilian applications. GPS based time reference provides inexpensive but highly-accurate timing and synchronization capability and meets requirements in power system fault location, monitoring, and control. In the present era of restructuring and modernization of electric power utilities, the applications of GIS/GPS technology in power industry are growing and covering several technical and management activities. Because of GPS receiver's error sources are time variant, it is necessary to remove the GPS measurement noise. This paper presents novel recurrent neural networks called the Recurrent Pi-Sigma Neural Network (RPSNN) and Recurrent Sigma-Pi Neural Network (RSPNN). The proposed NNs have been used as predictor in GPS receivers timing errors. The NNs were trained using the dynamic Back Propagation (BP) algorithm. The actual data collection was used to test the performance of the proposed NNs. The experimental results obtained from a Coarse Acquisition (C/A)-code single-frequency GPS receiver strongly support the potential of the method using RPSNN to give high accurate timing. The GPS timing RMS error reduces from 200 to less than 40 nanoseconds.

Keywords: Accurate Timing, GPS, Electric Systems, Neural Networks

1. Introduction

Most descriptions of Global Positioning System (GPS) focus on its use as a system to provide precise latitude, longitude and altitude information. Often it is used to determine speed as well. GPS is depicted as a dynamic positioning system which provides the raw information needed to navigate, that is, to find where we are and to figure how to get from there to some desired place (or, perhaps, to avoid some undesired place). This is a fundamental use for GPS but it is far from the only use of the system [1].

Continuous access to precise time and frequency, at low cost and anywhere it is needed, is a revolutionary development. It allows, for example, improved synchronization and timing of both wired and wireless communications systems. Users see higher quality (fewer dropped calls), increased capacity (no delays getting on), improved data transmission (low error rates) and new services (lifetime phone number). Or, consider timing electrical transients arriving at substations in a geographically dispersed power delivery system. A fault (a downed line, for instance) can be precisely located and crews can be transported to the precise geographic spot without delay. Similar statements can be made for wide-area computer networks. GPS allows precise transfer of time between the world's timing centers ensuring we all tick on the same clock. In general, wide availability of precise time and frequency at low cost will improve many scientific, manufacturing, business, R&D and just plain fun activities [2,3].

GPS provides services for two levels of users. These

* Tehran 16846-13114
Tel.: 0098-21-77240492,3, Fax.: 0098-21-77240490.

are referred to as the Standard Positioning Service (SPS) and the Precise Positioning Service (PPS). The latter is reserved, almost entirely, for the exclusive use of the DoD. The U.S. DoD states very clearly in the Federal Radio navigation Plan (FRP) what SPS and PPS provide: (1) SPS is a positioning and timing service which will be available to all GPS users on a continuous, worldwide basis with no direct charge. SPS will be provided on the GPS L1 frequency which contains a Coarse Acquisition (C/A) code and a navigation data message. SPS is planned to provide, on a daily basis, the capability to obtain timing accuracy within 340nsec (95 percent probability). The GPS L1 frequency also contains a Precision (P) code that is reserved for military use and is not a part of the SPS. Although available during GPS constellation build-up, the P code will be altered without notice and will not be available to users that do not have valid cryptographic keys. (2) PPS is a highly accurate military positioning, velocity, and timing service which will be available on a continuous, worldwide basis to users authorized by the DoD. PPS will be the data transmitted on GPS L1 and L2 frequencies. PPS was designed primarily for U.S. military use and will be denied to unauthorized users by use of cryptography. PPS will be made available to U.S. Federal and Allied Government (civilian and military) users through special agreements with the DoD. Limited, non-Federal Government, civilian use of PPS, both domestic and foreign, will be considered upon request and authorized on a case-by-case basis. PPS has timing accuracy with 200nsec [4,5].

For effective use of GPS timing information in power systems, it is essential to model and predict these errors. The better the prediction, the smaller the error values become. Hence, the efficiency of the predictive system depends highly on the predictor. Linear predictors have been widely used because of their simple implementation. In this case, the predicted value is the linear combination of the previous data elements. Nonlinear predictors provide better results than the linear predictor; however their use is limited due to the mathematical complexity of such predictor structures. NNs provide an alternative to this problem. The nonlinear nature and the simplicity of the learning algorithm of the NNs attracted many researchers to use NNs as predictors for GPS receivers timing errors [6]. This paper is organized as follows. Section 2 presents GPS applications in power systems. The proposed prediction methods using Recurrent Pi-Sigma Neural Network (RPSNN) and Recurrent Sigma-Pi Neural Network (RSPNN) are described in Section 3. In Section 4, the experimental tests results are reported with collected real data. Conclusions are presented in Section 5.

2. Precise Timing Applications in Power Systems

Precise timing in power systems is one of the key technologies that will enable the development of new control systems and the monitoring required to maintain them. Some of these areas of potential development are described in the following paragraphs [7−10].

2.1. GPS Traveling Wave Fault Locator Systems

An important monitoring device is a fault locator. A short circuit or fault usually can be cleared by momentarily disconnecting the line. Occasionally equipment is damaged and repair is required. Automatic fault location is much faster and cheaper than patrolling the entire line. When a line fault occurs, such as and insulator flashover or fallen conductor, the abrupt change in voltage at the point of the fault generates a high frequency electromagnetic impulse called a traveling wave which propagates along the line in both directions away from the fault point at velocities close to that of light (The velocity is determined by the distributed parameters of the line and it varies in the range 295-296m/μs). The fault location is determined by accurately time-tagging the arrival of the traveling wave at each end of the line, and comparing the related time difference ΔT to the total propagation time of the line T_p. The equation for calculating the distance $L1$ between the fault and the nearest terminal is as follows [10]:

$$L1 = 0.5(T_p - \Delta T)\frac{L}{T_p} \quad \textbf{(1)}$$

where L is total length of the line. Precise detection of the arrival time of the traveling wave is critical to the accuracy of the fault locator. A specially developed Fault Transient Interface Unit (FTIU) is used for this purpose. This device couples to the transmission lines by tapping off an inductive drain coil that is connected in series to the ground lead of a capacitive voltage transformer. The FTIU discriminates for a valid traveling wave by measuring the rise time and amplitude of the incoming signal. A signal whose rise time falls within two predetermined values (0.7-8.3μs, which corresponds to frequencies in the range 30-350KHz) and is of sufficient amplitude is considered to be a valid traveling wave and will cause the FTIU to produce a trigger pulse that is coincident with the leading edge of the detected traveling wave. The trigger pulse is fed to a GPS time code receiver which then timestamps the arrival of the traveling wave. The design goal of the fault location system is an accuracy of ±300 meters (one tower span) which translates to a time-tagging accuracy of better than 1μs (assuming that the velocity of the traveling wave is about 300m/μs). GPS receivers easily fulfill this requirement by providing a timestamp to within ±0.3μs of Universal Coordinate Time (UTC).

The marriage of a stable atomic clock and GPS satellite constellation makes possible an instrument with great accuracy and stability. It is possible to discipline a local clock with clocks of greater accuracy resident in the GPS satellite resulting in both short term and long accuracy improvements. The timing generator is a very stable and accurate instrument that has been designed to produce an accurate time standard with an absolute time accuracy of ±200nsec [11]. It has the capability of maintaining this accuracy for 24 hours after the total loss of the GPS signal. Each GPS satellite contains four stable atomic clocks that are traceable to the National Institute of Standards and Technology (NIST). There are two cesium beam clocks and two rubidium clocks aboard each satellite. It is these atomic clocks which give the GPS satellite the accuracy to be used as a continual calibration source for the timing generator rubidium clock.

2.2. Sources of Synchronization

Synchronization signal could be distributed over any of the traditional communication media currently in use in power systems. Most communication systems, such as leased lines, microwave, or AM radio broadcasts, place a limit on the achievable accuracy of synchronization, which is too coarse to be of practical use. Fiberoptic links, where available, could be used to provide high precision synchronization signals, if a dedicated fiber is available for this purpose. If a multiplexed fiber channel is used, synchronization errors of the order of 100 microseconds are possible, and are not acceptable for power system measurements. GOES satellite systems have also been used for synchronization purposes, but their performance is not sufficiently accurate.

The technique of choice at present is the Navstar GPS satellite transmissions. This systems is designed primarily for navigational purposes, but it furnishes a common-access timing pulse, which is accurate to within 1 microsecond at any location on earth. The system uses transmissions from a constellation of satellites in nonstationary orbits at about 10,000 miles above the earth's surface. For accurate acquisition of the timing pulse, only one of the satellites need be visible to the antenna. The antenna is small (about the size of a water pitcher), and can be easily mounted on roof of a substation control house. The experience with the availability and dependability of the GPS satellite transmission has been exceptionally good.

2.3. Phasor Measuring Units

Phasor Measuring Units (PMUs) using synchronization signals from the GPS satellite system have evolved into mature tools and are now being manufactured commercially. The GPS receiver provides the 1 Pulse-Per-Second (PPS) signal, and a time tag, which consist of the year, day, hour, minute, and second. The time could be the local time, or the Universal Time Coordinated (UTC). The 1-PPS signal is usually divided by a phase-locked oscillator into the required number of pulses per second for sampling of the analog signals. In most systems being used at present, this is 12 times per cycle of the fundamental frequency. The analog signals are derived from the voltage and current transformer secondaries, with appropriate anti-aliasing and surge filtering.

The microprocessor determines the positive sequence phasors according to the recursive algorithm described previously, and the timing message from the GPS, along with the sample number at the beginning of a window, is assigned to the phasor as its identifying tag. The computed string of phasors, one for each of the positive sequence measurements, is assembled in a message stream to be communicated to a remote site. The messages are transmitted over a dedicated communication line through the modems. A 4800-baud communication line can support the transmission of the phasor stream at the rate of about every 2-5 cycles of the fundamental frequency, depending upon the number of positive sequence phasors being transmitted.

2.4. State Estimation

Modern electric utility centers use state estimators to monitor the state of the power system. The state estimator uses various measurements (such as complex powers and voltage and current magnitudes) received from different substations, and, through an iterative nonlinear estimation procedure, calculates the power system state. The sate (vector) is a collection of all the positive sequence voltage phasors of network, and, from the time the first measurement is taken to the time when the state estimate is available, several seconds or minutes may have elapsed. Because of the time skew in the data acquisition process, as well as the time it takes to converge to a state estimate, the available state vector is at best an averaged quasi-steady-state description of the power system. Consequently, the state estimators available in present-day control centers are restricted to steady-state applications only.

Now, consider the positive sequence voltages measured by the synchronized phasor measurement units. If voltages at all system substations are measured, one would have a true simultaneous measurement of the power system state. No estimation of the state vector is necessary. From a practical point of view, it is sensible to use the positive sequence currents also, which provide data redundancy. This leads to a linear estimator of the power system state, which uses both current and voltage measurements. The estimate results from the multiplication, of a constant matrix by the measurement vector, and is extremely fast.

In addition to a much simplified static state estimator, synchronized phasor measurements also provide the first

real possibility of providing a dynamic state estimator. By maintaining a continuous stream of phasor data from the substations to the control center, a state vector that can follow the system dynamics can be constructed. With normal dedicated communication circuits operating at 4800 or 9600 baud, a continuous data stream of one phasor measurement every 2-5 cycles (33.3-83.33msec) can be sustained. Considering that the usual power system dynamic phenomena fall in the range of 0-2Hz, it is possible to observe in real-time the power system dynamic phenomena with high fidelity at the center.

Another application of directly measured dynamic phenomena is to validate power system models used in transient stability studies. For the first time in history, synchronized phasor measurements have made possible the direct observation of system oscillations following system disturbances. By trying to simulate these events, one can learn a great deal about the models of major system components, and correct them as needed until the simulations and observed phenomena match well.

2.5. Improved Control

Power system control elements, such as generation excitation systems, HVDC terminals, variable series capacitors, SVCs, etc., use local feedback to achieve the control objective. However, often the control objective may be defined in terms of a remote occurrence. As an example, consider the task of damping power swings between two areas by controlling (modulating) the flow on a dc line. Such a controller must have a built-in mathematical description (model), which must relate the dc power to the angle between the two regions. To the extent that the assumed model is not valid under the prevailing system conditions, the controller does not do the job for which it was intended.

With synchronized phasor measurements being brought to the controller location, it becomes possible to provide direct feed-back from the angular difference between the two systems. Studies of this nature have shown that improved control performance is achieved when a model-based controller is replaced by one based upon feedback provided by the phasor measurement system.

2.6. Quasi-Traveling Wave Schemes

Quasi-traveling wave schemes compare only the relative phase of the charge in impedance at the inception of fault at the local end with a signal representing the relative change at the remote end. When a fault occurs, the instantaneous voltage will usually fall and the instantaneous current will rise; either quantity may be positive or negative at that time. The relative change between the two represents the change in impedance and the direction of the fault. The relay is triggered by a rate of change in the voltage and current and sends a directional signal. Trip decision times are short but must allow for transmission time of the carrier system, relative end-to-end phasing of the voltage/current is not normally critical.

3. GPS Receivers Timing Errors Modeling Using Neural Network

The NN concept is used in forecasting, by considering historical data to be the input to a black box, which contains hidden layers of neurons. These neurons compare and structure the inputs and known outputs by nonlinear weightings, which are determined by a continuous learning process (Back-Propagation (BP)). The learning process continues until forecast outputs are reasonably close to known actual outputs. The structure of the black box is then used for forecasting actual future outputs. For time series forecasting the inputs are the past observations of the data series and the output is the future value. The GPS receiver's time errors $x(t)$ is difference between the two sequence time at time t, i.e., $x(t) = UTOD(t) - UTOD(t-1)$ [12]. The NNs estimate $x(t)$ at future time $t+1$.

3.1. Modeling Using Recurrent Pi-Sigma Neural Network

Similarly to feed-forward PSNN, the RPSNN consists of three layers, the input layer, the summing units layer and the output layer. In the output layer, the NN calculates the product of the weighted sum of the inputs and passes the result to a nonlinear transfer function. Then the output of the network is fed back to the inputs. The NN has a regular structure and requires a smaller number of weights than conventional single-layers, High-Order Neural Networks (HONNs). The weights from the summing units to the product unit are fixed at unity, which implies that the summing units layer is not hidden. The adoption of the smaller number of weights results in faster training. The order of the NN corresponds to the number of Σ units connected to the Π unit. Figure 1 shows the architecture of the proposed RPSNN.

Let the number of external inputs to the NN to be $M+2$ and the number of outputs to be 1. Let $y(n+1)$ be the output of the network at time $n+1$ and $x_j(n)$ be the $j-th$ input to the NN at time n. The overall inputs to the NN are the concatenation of $x(n)$ and $y(n)$ and is referred to as $z(n)$:

$$z_j(n) = \begin{cases} x_j(n) & ; \quad 1 \le j \le M \\ 1 & ; \quad j = M+1 \\ y(n) & ; \quad j = M+2 \end{cases} \qquad \textbf{(2)}$$

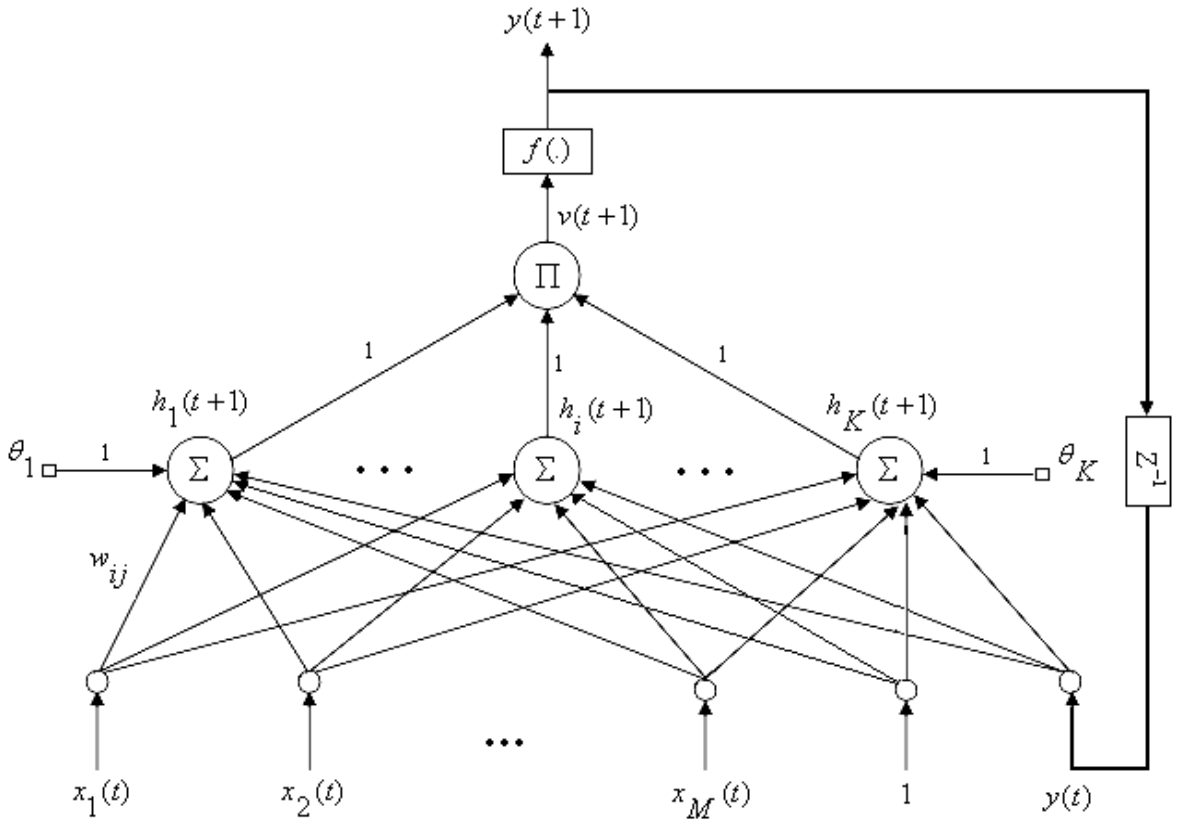


Figure 1. RPSNN architecture with $M+2, K, 1$ structure.

The bias is incorporated into the structure of the NN by adding an extra input line of value 1. The dynamic equations of a $k-th$ order network are as follows:

$$h_i(n+1) = \sum_{j=1}^{j=M+2} w_{ij}(n) z_j(n) + \theta_i(n) \qquad \textbf{(3)}$$

where $h_i(n+1)$ represents the net sum of the $i-th$ sigma unit and w_{ij} is the interconnection weight between the $i-th$ hidden neuron and the $j-th$ input node. The size of the weights matrix is $K \times (M+2)$. $\theta_i(n)$ is an adjustable threshold of the $i-th$ summing unit.

$$v(n+1) = \prod_{i=1}^{i=K} h_i(n+1) \qquad \textbf{(4)}$$

where $v(n)$ is activation function for the output neuron.

$$y(n+1) = f[v(n+1)] = f[\prod_{i=1}^{i=K} h_i(n+1)] \qquad \textbf{(5)}$$

where $y(n+1)$ is the NN output. The transfer function of the output is the logistic sigmoid which is defined as follows:

$$f(x) = \frac{1}{1+e^{-x}} \qquad \textbf{(6)}$$

The RPSNN is trained using dynamic BP. This is a gradient descent learning algorithm with the assumption that the initial state of the NN is independent of the initial weights. Let $d(n+1)$ represent the desired response at time *n*. The error of the NN at time $n+1$ is defined as [13]:

$$e(n+1) = d(n+1) - y(n+1) \qquad \textbf{(7)}$$

The cost function of the NN is the squared total error where:

$$J(n+1) = \frac{1}{2}[e(n+1)]^2 \qquad \textbf{(8)}$$

The aim of the dynamic BP learning algorithm is to minimize the total error by a gradient descent procedure. Therefore, the change for any specified weight w_{ij} is determined as follows:

$$\Delta w_{ij}(n+1) = -\eta \frac{\partial J(n+1)}{\partial w_{ij}} + \alpha \Delta w_{ij}(n) \quad \textbf{(9)}$$

where η is a positive real number representing the learning rate and α is the momentum term.

$$\frac{\partial J(n+1)}{\partial w_{ij}} = e(n+1)\frac{\partial e(n+1)}{\partial w_{ij}} = -e(n+1)\frac{\partial y(n+1)}{\partial w_{ij}} \quad \textbf{(10)}$$

In this case, $\frac{\partial y(n+1)}{\partial w_{ij}}$ is found by using the chain rule:

$$\frac{\partial y(n+1)}{\partial w_{ij}} = \frac{\partial y(n+1)}{\partial h_i(n+1)} \cdot \frac{\partial h_i(n+1)}{\partial w_{ij}} \quad \textbf{(11)}$$

By differentiating the dynamic equations of the NN, $\frac{\partial y(n+1)}{\partial h_i(n+1)}$ can be obtained as follows:

$$\frac{\partial y(n+1)}{\partial h_i(n+1)} = f'\Big[\prod_{l=1}^{l=K} h_l(n+1)\Big].\Big[\prod_{l=1\ ;\ l\#i}^{l=K} h_l(n+1)\Big] \quad \textbf{(12)}$$

and the value $\frac{\partial h_i(n+1)}{\partial w_{ij}}$ is calculated as follows:

$$\frac{\partial h_i(n+1)}{\partial w_{ij}} = z_j(n) + w_{iM+2}\frac{\partial y(n)}{\partial w_{ij}} \quad \textbf{(13)}$$

Let $P_{ij}(n+1) = \frac{\partial y(n+1)}{\partial w_{ij}}$, then weights updating rule is:

$$\Delta w_{ij}(n+1) = \eta e(n+1) P_{ij}(n+1) + \alpha \Delta w_{ij}(n) \quad \textbf{(14)}$$

With:

$$P_{ij}(n+1) = f'\Big[\prod_{l=1}^{l=K} h_l(n+1)\Big].\Big[\prod_{l=1\ ;\ l\#i}^{l=K} h_l(n+1)\Big].[z_j(n) + w_{iM+2}P_{ij}(n)] \quad \textbf{(15)}$$

where $f'(.)$ is the derivative of the nonlinear transfer function and is determined as follows:

$$f'(x) = f(x)[1 - f(x)] \quad \textbf{(16)}$$

The change for any specified weight θ_i is determined using BP learning algorithm as follows:

$$\Delta\theta_i(n+1) = -\eta \frac{\partial J(n+1)}{\partial \theta_i} + \alpha \Delta\theta_i(n) \quad \textbf{(17)}$$

where $\frac{\partial J(n+1)}{\partial \theta_i}$ is obtained as:

$$\begin{aligned}\frac{\partial J(n+1)}{\partial \theta_i} &= e(n+1)\frac{\partial e(n+1)}{\partial \theta_i} \\ &= -e(n+1)\frac{\partial y(n+1)}{\partial \theta_i}\end{aligned} \quad \textbf{(18)}$$

$\frac{\partial y(n+1)}{\partial \theta_i}$ is found by using the chain rule:

$$\frac{\partial y(n+1)}{\partial \theta_i} = \frac{\partial y(n+1)}{\partial h_i(n+1)} \cdot \frac{\partial h_i(n+1)}{\partial \theta_i} \quad \textbf{(19)}$$

The value $\frac{\partial h_i(n+1)}{\partial \theta_i}$ is calculated as follows:

$$\frac{\partial h_i(n+1)}{\partial \theta_i} = 1 + w_{iM+2}\frac{\partial y(n)}{\partial \theta_i} \quad \textbf{(20)}$$

Let $Q_i(n+1) = \frac{\partial y(n+1)}{\partial \theta_i}$, then weights updating rule is:

$$\Delta\theta_i(n+1) = \eta e(n+1) Q_i(n+1) + \alpha \Delta\theta_i(n) \quad \textbf{(21)}$$

With:

$$Q_i(n+1) = f'\Big[\prod_{l=1}^{l=K} h_l(n+1)\Big].\Big[\prod_{l=1\ ;\ l\#i}^{l=K} h_l(n+1)\Big].[1 + w_{iM+2}Q_i(n)] \quad \textbf{(22)}$$

3.2. Modeling Using Recurrent Sigma-Pi Neural Network

The architecture of a RSPNN with $M+2$ inputs and one output is shown in Figure 2. In this network, hidden layer neurons output is the product of the input terms and the network output is the sum of these products. It also has a single layer of adaptive weights in the second layer.

The RSPNN learning procedure using the BP method can be summarized as follows:

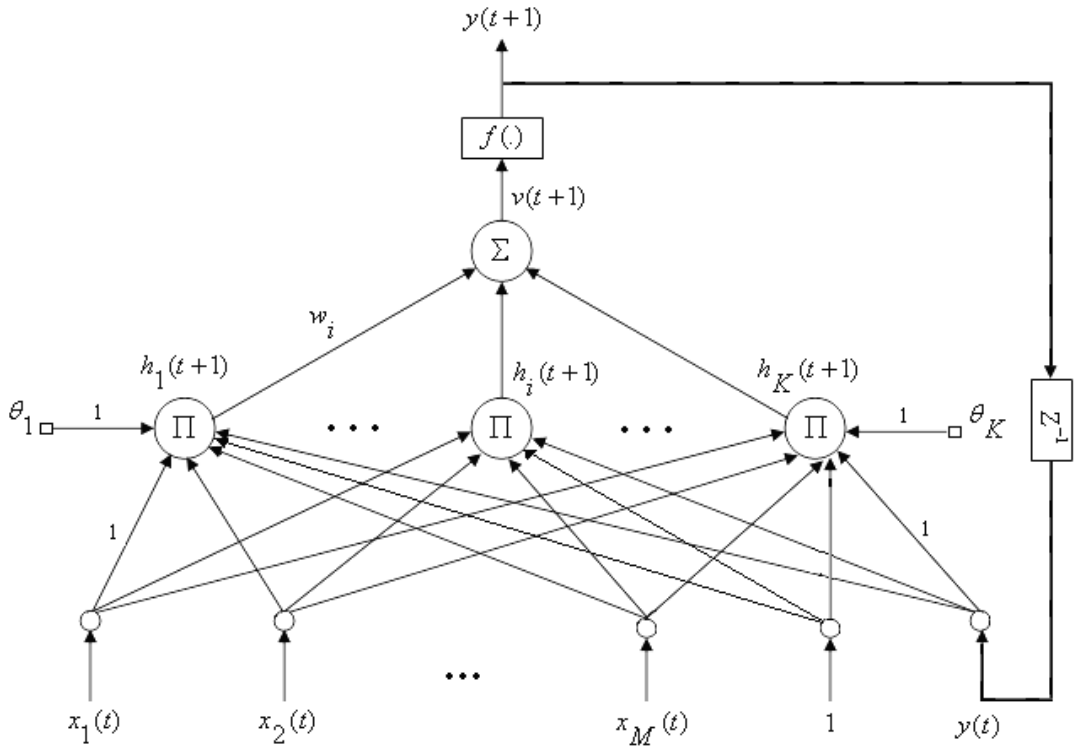


Figure 2. RSPNN architecture with $M+2, K, 1$ structure.

Step1: Weight Vector Initialization

Set all of the weights and thresholds of the network to small random numbers that are uniformly distributed.

Step2: Forward Calculations

$$z_j(n) = \begin{cases} x_j(n) & ; \quad 1 \le j \le M \\ 1 & ; \quad j = M+1 \\ y(n) & ; \quad j = M+2 \end{cases} \tag{23}$$

$$h_i(n+1) = \prod_{j=1}^{j=M+2} z_j(n) + \theta_i(n) \tag{24}$$

$$v(n+1) = \sum_{i=1}^{i=K} w_i h_i(n+1) \tag{25}$$

$$y(n+1) = f[v(n+1)] \tag{26}$$

Step3: Learning Process

$$\Delta w_i(n+1) = \eta e(n+1) P_i(n+1) + \alpha \Delta w_i(n) \ ;$$

$$P_i(n+1) = f'[\sum_{i=1}^{i=K} w_i h_i(n+1)].w_i.[\prod_{j=1}^{j=M+1} z_j(n)].P_i(n) \tag{27}$$

$$\Delta \theta_i(n+1) = \eta e(n+1) Q_i(n+1) + \alpha \Delta \theta_i(n) \ ;$$

$$Q_i(n+1) = f'[\sum_{i=1}^{i=K} w_i h_i(n+1)].w_i.[1+(\prod_{j=1}^{j=M+1} z_j(n)).Q_i(n)] \tag{28}$$

Step4: Iteration

Increment time *n* by one unit and go back to Step2.

4. Experimental Results

To test the proposed NNs for GPS receivers timing errors prediction a system was built. The test setup was implemented and installed on the building of Computer Control and Fuzzy Logic Research Lab in the Iran University of and Technology. The observation data received by a low cost and single frequency GPS receiver manufactured by Rockwell Company. The collected data were processed with developed programs by the paper author. Figure 3 shows the data collection system adopted in this research.

In preparing the training data, all input and output variables are normalized in the range [0,1] to reduce the training time [14]. Observation at time t is applied to NNs inputs and the networks must predict the value of instant $t+1$. The choice of the order for the NNs is very important in on-line prediction. It is more difficult to

formulate the order of a nonlinear model. In this paper, the order selection is based on the experimental results. For optimizing NNs structure, various combinations of input variables were tried. The proposed methods were implemented and developed by author of this paper using Microsoft Visual Basic6. These models were validated using a set of data points.

The RPSNN and RSPNN benefit from both the advantages of feed-forward HONNs and RNNs. The incorporation of higher order terms allows the networks to make use of nonlinear interactions between the inputs, thus functionally expanding the input space into a higher-dimensional space, where linear separability, or reduction in the dimension of the nonlinearity is possible. Furthermore, the adaptation of the small number of weights allows the NNs to be trained faster than HONNs which suffer from the combinational explosion of the high-order terms and demonstrate slow learning, when the order of the NNs becomes excessively high. In contrast to fully RNNs that are trained using the RTRL algorithm, the small structure of the RPSNN and RSPNN accelerates the learning of the NNs using the dynamic BP.

To evaluate the performance of the presented training algorithms, they were tested by collected data sets. Six statistical measures (maximum, minimum, RMS, average, variance, standard deviation), are used to evaluate prediction results. Table 1 presents statistical measures on 500 test data by using the proposed NNs. In order to evaluate the prediction accuracy, we used RMS as a measure of closeness between predicted and observed values [15]. From Table 1 can be seen that time accuracy has improved by factor of about 5.

Table 2 shows the comparison of test results of different models for GPS timing errors prediction. The simulation results demonstrated that RPSNN and RSPNN are efficient than PSNN,RNN and SPNN,RNN, respectively.

5. Conclusions

Accurate timing using GPS can revolutionize the field of monitoring, protection, and control of power systems. It is with great excitement that we look for other applications, not yet thought of, that can advance the state of the art in electric power engineering. The past few years have witnessed increasing interest in synchronized accurate timing and how they may be used for various power system applications. The development of new types of computer-based hardware and the completion of the GPS of satellites provide the components needed for true synchronized monitoring systems. GPS time synchronization enables the accurate time tag of each recorded data sample to better than 1 microsecond accuracy. In this paper, a RPSNN and RSPNN were implemented and used as predictor in GPS system. The proposed NNs were trained using the dynamic BP, which is a gradient descent learning algorithm. The actual data collection was used to train the networks. The trained weights of the NNs were fixed and used to predict GPS receivers timing errors. Extensive tests have shown that third order NNs provide the most promising results. The tests results using the RPSNN predictor have shown an improvement in the GPS timing accuracy over the linear predictor, multilayer perceptorns, HONNs, RNNs. The GPS timing RMS error reduced from 200 to less than 40 nanoseconds.

Figure 3. Data collection and processing system.

Table 1. Performance evaluation of the proposed NNs.

Parameters	Error Values [nsec] (using RPSNN with (3,3,1) Structure)	Error Values [nsec] (using RSPNN with (3,4,1) Structure)
Max	90.364	101.643
Min	-41.784	-20.944
Average	-0.224	0.836
Variance	3.137	3.870
Standard Deviation	1.771	1.967
RMS	**39.562**	**43.955**

Table 2. Comparison of test results of different models for GPS timing errors prediction.

Model Name	RMS
RNN	57.2
PSNN	45.8
RPSNN	39.6
SPNN	50.0
RSPNN	43.9

6. References

[1] K. D. McDonald, "The modernization of GPS: Plans, new capabilities, and the future relationship to Galileo," Journal of Global Positioning System, Vol. 1, No. 1, pp. 1–17, 2002.

[2] W. Lewandowski, J. Azoubib, and W. J. Klepczynski, "GPS: Primary tool for time transfer," Proceedings of the IEEE, Vol. 87, No. 1, pp. 163–172, January 1999.

[3] T. E. Parker and D. Matsakis, "Time and frequency dissemination advances in GPS transfer techniques," GPS World Magazine, pp. 32–38, November 2004.

[4] K. Mohammadi and M. H. Refan, "A new method for improving of GPS receivers time accuracy using Kalman filter," Journal of Engineering Science, Iran University of Science and Technology, Vol. 13, No. 1, pp. 11–24, 2002.

[5] M. R. Mosavi, "GPS receivers timing data processing using neural networks: Optimal estimation and errors modeling," Journal of Neural Systems, Vol. 17, No. 5, pp. 383–393, October 2007.

[6] A. J. Hussain and P. Liatsis, "A new recurrent polynomial neural network for predictive image coding," IEEE Conference on Image Processing and its Applications, Vol. 1, No. 465, pp. 82–86, 1999.

[7] K. E. Martin, "Precise timing in electric power systems," IEEE Conference on Frequency Control, pp. 15–22, 1993.

[8] A. G. Phadke, "Synchronized phasor measurements in power systems," IEEE Computer Applications in Power, pp. 11–15, April 1993.

[9] M. A. Street, I. P. Thurein, and K. E. Martin, "Global positioning system applications for enhancing the performance of large power systems," CIGRE, pp. 1–6, 1994.

[10] H. Lee and A. M. Mousa, "GPS traveling wave fault locator systems: Investigation into the anomalous measurements related to lighting strikes," IEEE Transactions on Power Delivery, Vol. 11, No. 3, pp. 1214–1223, 1996.

[11] M. R. Mosavi, "Modeling of GPS SPS timing error using multilayered neural network," IEEE Conference on Signal Processing, China, November 16–19, 2006.

[12] M. R. Mosavi, "Real time prediction of GPS receivers timing errors using parallel-structure neural networks," Journal of Geoinformatics, Vol. 3, No. 3, pp. 53–61, September 2007.

[13] M. R. Mosavi, "A practical approach for accurate positioning with L1 GPS receivers using neural networks," Journal of Intelligent and Fuzzy Systems, Vol. 17, No. 2, pp. 159–171, March 2006.

[14] M. R. Mosavi, "Precise real-time positioning with a low cost GPS engine using neural networks," Journal of Survey Review, Vol. 39, No. 306, pp. 316–327, October 2007.

[15] M. R. Mosavi, "Comparing DGPS correction prediction using neural network, fuzzy neural network, and Kalman filter," Journal of GPS Solutions, Vol. 10, No. 2, pp. 97–107, May 2006.

Wireless Sensor Network, 2009, 2, 61-121
doi:10.4236/wsn.2009.12016 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

Centralized Quasi-Static Channel Assignment for Multi-Radio Multi-Channel Wireless Mesh Networks

Juan REN, Zhengding QIU
Institute of Information Science, Beijing Jiaotong University, Beijing, China
Email: {04112047, *zdqiu*}*@bjtu.edu.cn*
Received February 19, 2009; *revised March* 19, 2009; *accepted March* 20, 2009

Abstract

Employing multiple channels in wireless multihop networks is regarded as an effective approach to increasing network capacity. This paper presents a centralized quasi-static channel assignment for multi-radio multi-channel Wireless Mesh Networks (WMNs). The proposed channel assignment can efficiently utilize multiple channels with only 2 radios equipped on each mesh router. In the scheme, the network end-to-end traffics are first modeled by probing data at wireless access points, and then the traffic load between each pair of neighboring routers is further estimated using an interference-aware estimation algorithm. Having knowledge of the expected link load, the scheme assigns channels to each radio with the objective of minimizing network interference, which as a result greatly improves network capacity. The performance evaluation shows that the proposed scheme is highly responsive to varying traffic conditions, and the network performance under the channel assignment significantly outperforms the single-radio IEEE 802.11 network as well as the 2-radio WMN with static 2 channels.

Keywords: Wireless Mesh Networks, Multihop Network, Channel Assignment, Multi-Radio

1. Introduction

WMN [1] is a promising wireless technology for numerous applications, e.g., broadband home networking, community and neighborhood networks, enterprise networking, building automation, etc. [2,3]. However, interference among wireless links significantly impacts the performance of WMNs. As a multi-hop wireless network, the actual goodput available to WMN applications decreases a lot when forwarding or relaying packets over multiple wireless hops.

Fortunately, the IEEE 802.11 PHY specification permits simultaneous operation of multiple non-overlapping channels. By deploying multi-radio routers in infrastructure-based networks and assigning radios to non-overlapping channels, the routers can communicate simultaneously with little interference in spite of being in direct interference range of each other. Therefore, the capacity of wireless networks can be increased. While due to the limited number of channels available, the interference cannot be completely eliminated. In addition, the channel assignment must be restricted to the number of radios on each wireless node. So it's a challenging problem deserving our research.

In equipping routers with multiple radios, a naive strategy would be to equip each router with the number of radios equal to the number of orthogonal channels. However, this strategy is economically prohibitive due to the significant number of non-overlapping channels. Another channel assignment strategy is to frequently change channel on the interface, for instance, for each packet transmission based on current state of the medium. Such dynamic channel assignment approaches [4–6] require channel switching at a very fast time scale (per packet or a handful of packets). The fast-channel switching requirement makes these approaches unsuitable for use with commodity hardware, where channel switching delays itself can be in the order of milliseconds [4]. Some of the dynamic channel assignment approaches also re-

*This work is supported by National Basic Research Program of China (973 Program) (2007CB307100).

quire specialized MAC protocols or extensions of 802.11 MAC layer, making them further unsuitable for use in commodity 802.11 hardware.

In order to use multiple channels with commodity hardware, several researches [7–9] focused on developing techniques that assign channels statically. Such static assignments can be changed whenever there are significant changes to traffic load or network topology. Since WMN is an infrastructured network and aims to provide reliable broadband services, such changes are infrequent enough that the channel-switching delay and traffic measurement overheads are insignificant. We refer to the above as quasi-static channel assignments. However, most of the existing quasi-static channel assignments are performed offline and bound with routing.

In this paper, we address the problem of quasi-static channel assignment independent of routing. A centralized quasi-static channel assignment algorithm is proposed in the context of networks with multi-radio nodes. In the channel assignment, we use a novel scheme to estimate the traffic load on each wireless link. The estimation considers the traffic on the link itself as well as the interfering traffics introduced by its neighbors. Having knowledge of the expected load on each link, the algorithm can intelligently select different channels for each radio with the objective of minimizing network interference, which as a result efficiently improves the network capacity. To evaluate the algorithm performance, a corresponding channel assignment protocol is implemented in ns-2 simulations [10] and we incorporate the well-known WCETT (Weighted Cumulative Expected Transmission Time) path metric [11], which is tailored for multi-radio multihop wireless networks, into the AODV (Ad Hoc On Demand Distance Vector) routing protocol [12] as our multi-radio routing protocol. The performance evaluation shows that the proposed scheme is highly responsive to varying traffic conditions, and the network performance under the channel assignment significantly outperforms the single-radio IEEE 802.11 network as well as the 2-radio WMN with static 2 channels.

The rest of the paper is organized as follows. Section 2 gives the system architecture of the proposed multi-radio WMN. In Section 3, we describe the centralized quasi-static channel assignment scheme. In Section 4, we evaluate the performance of our channel assignment algorithm using the ns-2 simulations. Section 5 concludes the paper.

2. System Architecture

In this section, we formulate the interference problem involved in wireless multihop networks and present the architecture of multi-radio multi-channel WMNs to resolve this problem.

2.1. Interference Problem

Traditional 802.11-based wireless networks can't transmit data simultaneously as wired networks because of the intra-path and inter-path interference. For example in Figure 1, although the two flows transmit separately on path 1 and path 2, nodes must compete with each other for a common channel, which reduces network throughput hardly. If node 3 is in transmission, all the nodes in interference range of node 3 should keep silence, or a collision will occur. In contrast, if we assign interfering hops to different channels, then one collision domain can be broken into several collision domains with each operating in a different frequency range. When the ingress-egress node pairs that originally pass through the collision domain now take different paths to route their traffic, hops using different channels can transmit simultaneously and the network throughput will increase.

2.2. Multi-Radio Multi-Channel WMN Architecture

Figure 2 gives the architecture of multi-radio multi-channel WMN. The wireless mesh backbone network consists of mesh routers (MR), mesh access routers (MAR) and the gateway. Mesh routers provide purely wireless routing services. Mesh access routers provide not only wireless routing services but also wireless access services. Each WMN has at least one gateway, which can also be served as the access point for wireless users. The integration of WMN with other networks such as the Internet, cellular, IEEE 802.11, IEEE 802.15, IEEE 802.16, sensor networks, etc., can be accomplished through the gateway and bridging functions in the mesh routers.

In WMN only end users may frequently move and mesh backbone facilities are almost static once been settled. In addition, both the gateway and the mesh access routers have aggregation capability. So we use them to measure the ingress-egress network traffic in our load estimation algorithm. Although in our network each router is equipped with only 2 radios, the overall network can utilize more channels with intelligent channel assignment to every link. This is the fundamental reason for non-linear improvement in throughput with respect to the increase in number of radios per node.

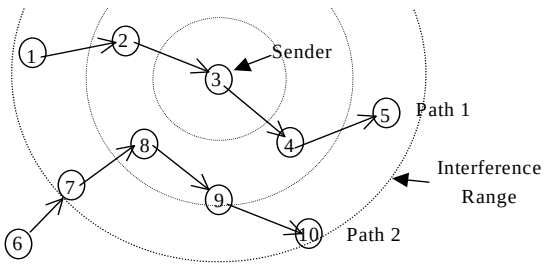


Figure 1. Interference in wireless communication.

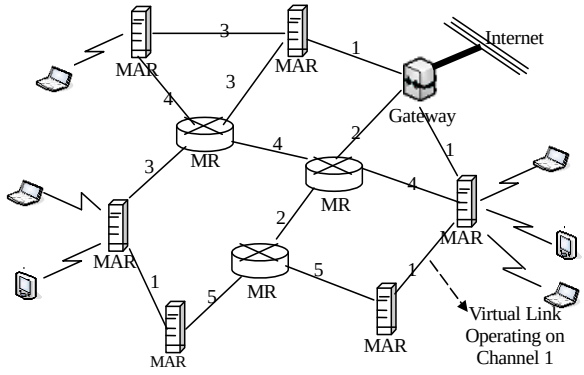


Figure 2. Architecture of multi-radio multi-channel WMN.

3. Design of Channel Assignment

The goal of channel assignment in multi-radio WMNs is to bind each radio to a channel in such a way that the available bandwidth on each link is proportional to its expected load. The link here means direct communication on a channel between the routing pair. In this section, we describe the load estimation and channel assignment algorithms.

3.1. Traffic Measurement

Optimizations of channel assignment using load estimation require knowledge of network traffic information, so we propose to measure the end-to-end traffics between mesh access routers. The traffic measurement procedure is described as follows:

At first, each mesh access router (including the gateway) measures its ingress-egress flows by probing data periodically (the interval is set to 10s in simulations). For convenient expression, we use the access router to indicate both the mesh access router and the gateway in the rest of this paper. Then each access router aggregates its ingress-egress flows and sends the information to the gateway. (The gateway is used as the computation center since it owns the most powerful capacity.)

After receiving the flow information, the gateway calculates the end-to-end traffic value between every pair of access routers by further aggregate the flow information. In this way, the real time value of the end-to-end traffic between each pair of access routers is measured at each echo interval. However, since what the quasi-static algorithm needed is a long-term measured traffic, the gateway performs an exponentially weighted moving average (EWMA) of each end-to-end traffic load to get an approximate long-term traffic. That is:

$$T(s,d) = \alpha T_{old}(s,d) + (1-\alpha)T(s,d) \qquad \textbf{(1)}$$

where $T(s,d)$ denotes the end-to-end traffic between access router pair *s* and *d*. In simulations, the smoothing factor $\alpha = 0.7$.

3.2. Initial Expected Load

Having the knowledge of the end-to-end traffics, the gateway estimates the expected load on each wireless link and assigns channels to links in order of the expected loads. The gateway is required to perform a new estimation of the expected load either when it receives the traffic information for the first time, or when the difference between the new traffic and the last one is large enough.

To initial the load estimation, we assume that there is a link between each pair of routers in direct communication range, and each end-to-end traffic load is equally divided among all the paths with the least hops between the pair of access routers. (Note that this link won't really exist if there is no common channel assigned to the pair of routers on this link.)

If the number of all shortest paths between node *s* and *d* is $P(s,d)$, and in those paths there are $P_i(s,d)$ paths passes link *i*, then the initial expected load for link *i* to carry is calculated as follow:

$$\varphi(i) = \sum_{s,d} \frac{P_i(s,d)}{P(s,d)} T(s,d) \qquad \textbf{(2)}$$

Here we only count the shortest paths because the path with less hops always have much better performance compared with longer paths in multi-hop wireless networks if they all have enough bandwidth.

3.3. Channel Assignment

Having knowledge of the expected loads on all network links, we start to assign channels to links as follows:

At first, all links are sorted by their expected loads. Since links expected to carry higher traffic load should be given more bandwidth, the link with the most expected load is prior to other links in choosing channels. Assume every node is equipped with *q* radios and node *a* and *b* is connected by link *i*. There are three conditions for choosing a channel to link *i*:

1) If both of node *a* and *b* have used less than *q* radios, then choose a channel with the least interference with link *i*. (The chosen channel should also be different from the used channels of *a* and *b*.) Meanwhile, update the channel lists of node *a* and *b*.

2) If one of the two nodes (for example *a*) has used *q* channels, then choose a channel from the channel assignment list of node a with the least interference with link *i*. Meanwhile update the channel list of node *b*.

3) When both of node *a* and *b* have used *q* channels: if there exists a common channel between node *a* and *b*, then assign this channel to link *i*; else, choose one channel from the list of node *a* and *b* respectively and merge the two channels to one. Meanwhile update the channel

lists of all nodes which connect to node *a* or *b* using the two channels directly or indirectly.

When link *i* has been assigned a channel, we remove it from the unassigned link list. As a result, the link with lower expected load now owns the priority to choose channel next. This continues until all network links have been visited, which is denoted as a cycle of channel assignment.

3.4. End or Feedback

After a cycle of channel assignment for all links, we need to judge whether current channel assignment satisfies all bandwidth requirements. If so, we terminate the whole procedure and output the channel assignment results, or we feedback the expected link loads under new channel assignment to find a better channel assignment.

It's easy to see that, if the available bandwidth on each link is more than the traffic load it's expected to carry, no congestion will occur. So at first we need to estimate the capacity for each link in the network. However, in wireless networks, channels are shared by all links in the same interference range. So when estimating the usable capacity of a link, we should consider all traffic loads in its interference range. According to the channel assignment rules, the higher load a link is expected to carry, the more bandwidth it should get. On the other side, the higher loads its interfering links are expected to carry, the less bandwidth it could obtain. Thus, the link capacity should be proportional to its traffic load, and be inversely proportional to all other interfering loads. So the capacity for a link *i* is given by:

$$C(i) = \frac{B * \varphi(i)}{\sum_{j \in Intf(i)} \varphi(j)} \quad \textbf{(3)}$$

where B is the channel bandwidth and $Intf(i)$ stands for the set of links in the interference range of *i* including itself.

Then the residual capacity of link i can be obtained as below:

$$RC(i) = C(i) - \varphi(i) \quad \textbf{(4)}$$

We use the minimal residual capacity of all links on a path as the available bandwidth for this path. Then we consider the bandwidth requirement of an end-to-end traffic is satisfied if there exists a path with its available bandwidth more than the required traffic value.

At the start of the traffic allocation, the initial $RC(i)$ of each link is set to the expected capacity $C(i)$, and the expected load $\phi(i)$ of each link is set to 0. When a path is chosen to carry an end-to-end traffic, all links on the path reduce their residual capacities and increase their expected loads by the allocated traffic value.

Assume there are totally N end-to-end traffics with respectively the value of $T_n (n = 1..., N)$. We choose path P with the maximal available bandwidth among all the paths with both the least hop and enough available bandwidth for traffic T_n, which is because the WCETT path metric used in multi-channel routing protocol prefers to choose high throughput paths in multiple channels. When a path is chosen for traffic T_n, there are two conditions when allocating traffic to this path:

1) If $T_n \le \min_{i \in P}(RC(i))$, we can allocate the whole traffic T_n to path P. So we decrease the residual capacity and add the expected load of each link i on path P by T_n. We also decrease the total unallocated traffic by T_n for a later comparison of each iteration.

2) If $T_n > \min_{i \in P}(RC(i))$, we can only allocate $\min_{i \in P}(RC(i))$ traffic to path P. So we decrease the residual capacity and add the expected load of each link *i* in path P by $\min_{i \in P}(RC(i))$. We also decrease the total unallocated traffic value by $\min_{i \in P}(RC(i))$.

When all the N traffics have been checked, we terminate the whole algorithm if the total unallocated traffic equals to 0. Or, we compare the total unallocated traffic of this cycle to the last one. If the unallocated traffic of this cycle is no less than the last one, it means no improvement is made and we also terminate the whole process. Otherwise, we feedback the new expected link loads to the channel assignment for a better scheme.

After the whole algorithm ends, the gateway broadcasts the channel assignment results. Then each mesh router adjusts its radios to the assigned channels when they receive the results.

If we combine the traffic measurement, load estimation and channel assignment, the whole algorithm process can be depicted as in Figure 3.

4. Simulation Analysis

To evaluate the performance of our channel assignment, we run simulations using ns-2. We use the IEEE 802.11 MAC protocol with RTS/CTS enabled. The AODV protocol is used as the routing protocol for the single-radio IEEE 802.11 network simulations. We modify the AODV protocol using the WCETT metric, a prevailing path metric, as our multi-radio routing protocol. The topology of all networks in the simulations is a 25-node square grid. To simplify and generalize the simulation, we configure the center node as the gateway and all the other 24 nodes as mesh access routers. The bandwidth of each channel is 2 Mbps. The ratio between interference and communication range is 2 and all nodes in multi-channel networks are equipped with 2 radios.

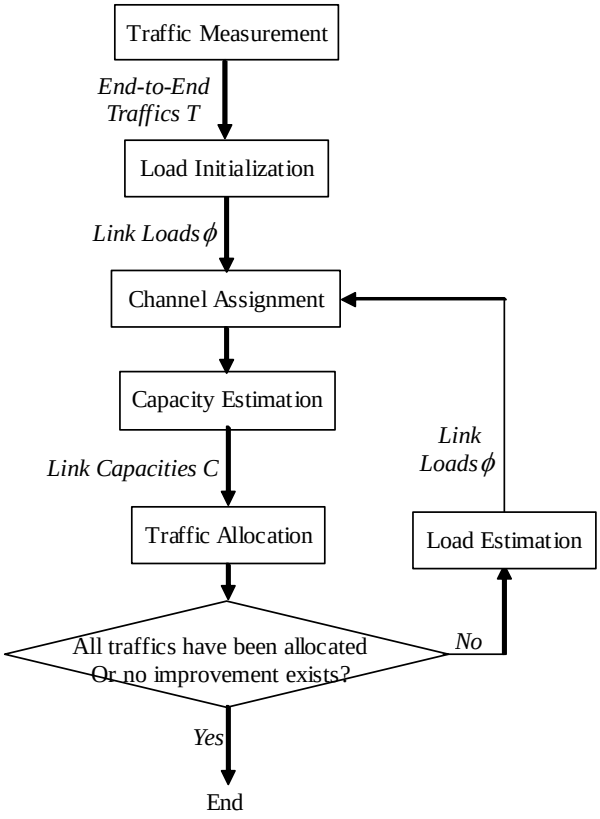


Figure 3. The whole algorithm process.

4.1. Static Traffic Simulation

4.1.1. 10 Flows Simulation

At first, we randomly choose 10 pair of nodes from the 25 nodes and assign each pair with a different CBR UDP flow. The rate of each flow is chosen randomly between 0 and 0.8Mbps. The packet size is 1000 bytes and flows run for 100 seconds. We test our channel assignment algorithm in a 5-channel network with the 10 flows. To have a good comparison, we also run the same flows in a standard IEEE 802.11 network and a 2-radio 2-channel network. The two networks both needn't any channel assignment and the 2-channel network simulation is just the original WCETT multi-radio routing simulation. We evaluate the improvement of network performance by comparing the aggregate throughputs, the average packet delays and the average packet drop probabilities of the 3 networks.

The aggregate throughputs of the 3 networks are shown in Figure 4. The standard 802.11 network throughput is quite low and the 2-channel network throughput is about twice of the standard 802.11 network. Although the improvement is significant, we can see both the two network throughputs suffer sharp vibration. This is because all network nodes using common channels is quite easy to bring great interference. While in the 5-channel network, we can efficiently limit interference to several small areas using our channel assignment. So the network throughput is quite stable and the value of the throughput also increases a lot.

The average packet delay of each flow is shown in Figure 5. We draw the delays of flows that haven't received any data packet successfully in the whole simulation to the max value of the y axis. We can see the 1-channel network performs badly with 1 flow receiving no packet. Although all flows in the 2-channel network can transport data, the average packet delays for most flows are quite large and exceed 0.5s. The maximal average packet delay of the 2-channel network is up to 1.83s. In the 5-channel network, the average packet delays are much smaller. The maximal delay of all flows is 1.15s and there are 7 flow delays are below 0.5s. This further proves that under our channel assignment, the mesh network can efficiently utilize 5 or even more channels with only 2 radios.

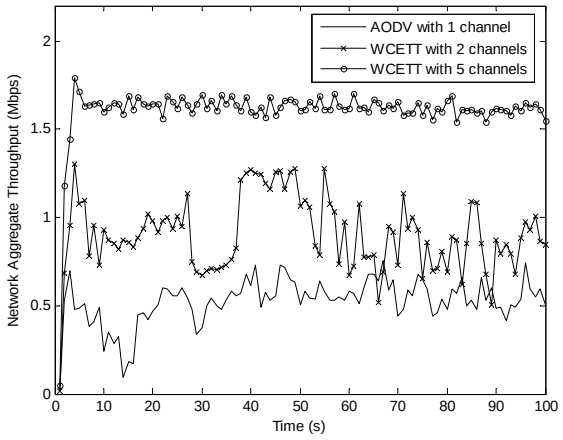


Figure 4. The aggregate network throughput for 10 flows.

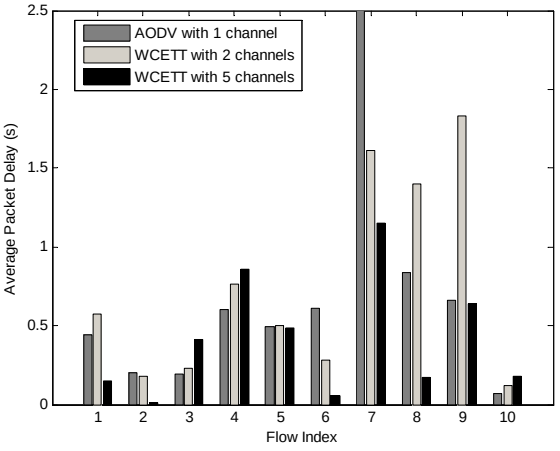


Figure 5. The average packet delay for 10 flows.

We also compare the average packet drop probability of each flow in Figure 6. The probabilities for the previous two networks are all quite large. In the 5-channel network, the probabilities of 4 flows are quite small with values below 4%. Although the ones for the other flows are large, but they have decreased a lot compared to the previous two networks.

4.1.2. 20 Flows Simulation

To derive the network to saturation, we randomly choose 20 pair of nodes from the 25 nodes and run the same simulations in the above 3 networks. The aggregate throughput of the 3 networks is shown in Figure 7. From Figure 7 we can see the throughput gain for 2-channel network is not significant under the heavy traffic. While both the value and the stability of throughput for 5-channel network get further significant increase, which is due to the sufficient bandwidth brought by efficient utilization of multiple channels.

The average packet delay of each flow is shown in Figure 8. We can see that in the 1-channel network, 4 flows received no packet in the whole simulation. In the 2-channel network, the average packet delays for most flows are also quite large with the maximal average delay up to 2.32s for the 19th flow. While in the 5-channel network, the average packet delays for most flows are below 0.5s and the maximal one is only 0.96s.

We compare the average packet drop probability of each flow in Figure 9. The average packet drop probabilities for all the three networks are all quite large because of the heavy traffic. While compared to the two networks without channel assignment, the average packet drop probability for the 5-channel network is still much smaller and the average drop probability of 9th flow is nearly 0, which means it has almost transmitted all the data packets successfully.

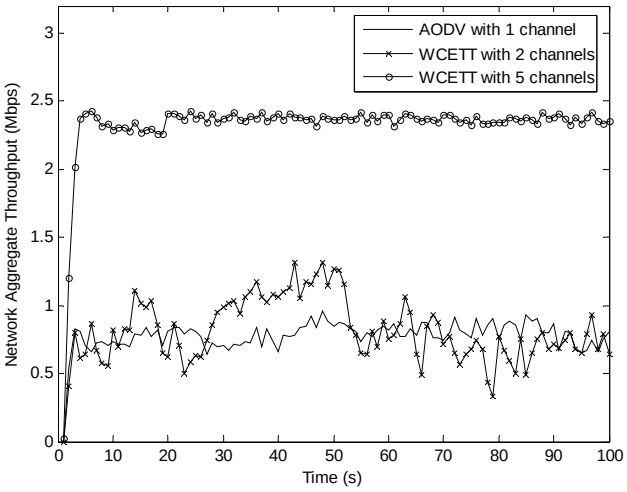


Figure 7. The aggregate network throughput for 20 flows.

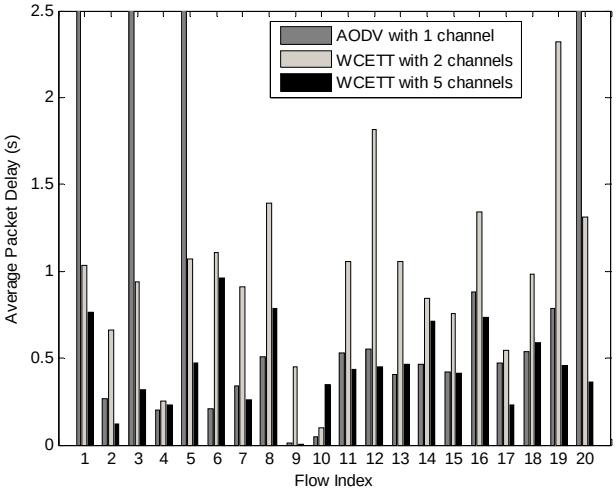


Figure 8. The average packet delay for 20 flows.

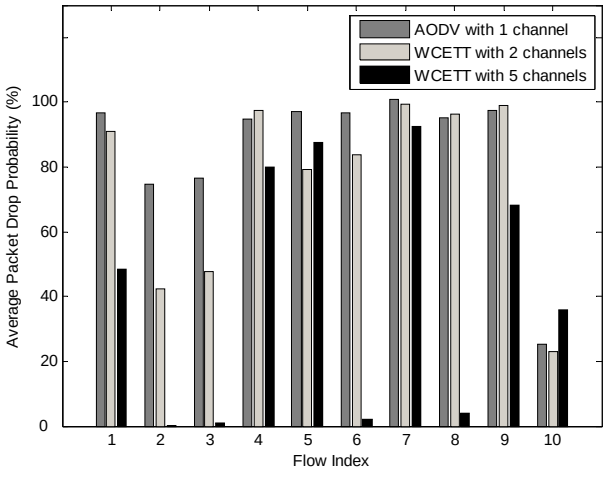


Figure 6. The average packet drop probability for 10 flows.

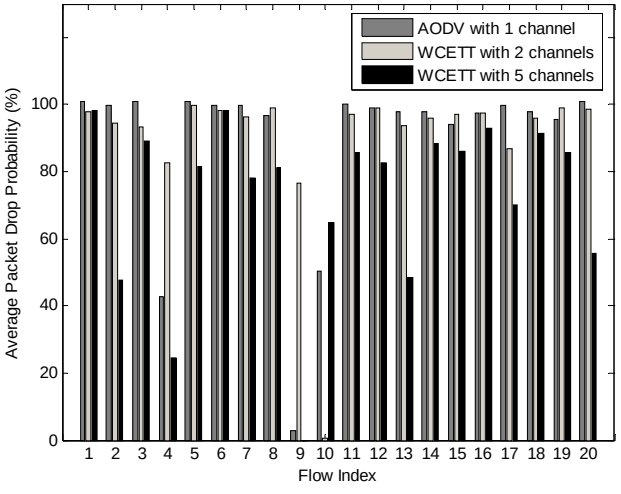


Figure 9. The average packet drop probability for 20 flows.

Table 1. Comparison of the average aggregate throughput.

Average Aggregate Througput (Mbps)	AODV	WCETT	WCETT + Channel Assignment		
	1 channels	*2 channels*	*3 channels*	*4 channels*	*5 channels*
10 flows	0.525	0.903	1.231	1.437	1.608
20 flows	0.777	0.814	1.448	2.080	2.323

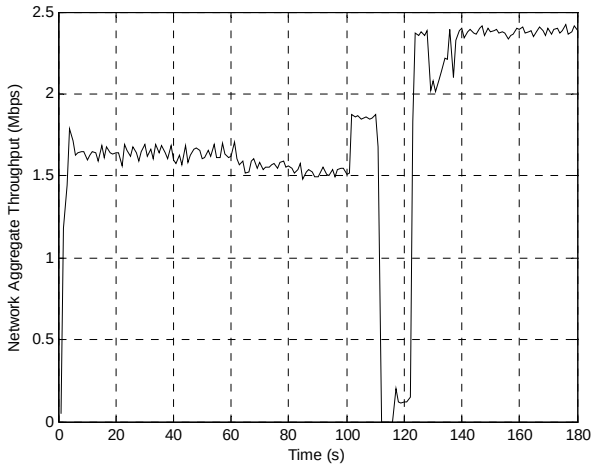


Figure 10. Impact of traffic change on aggregate network throughput.

4.1.3. The Average Aggregate Throughput Comparison

Without loss of generality, we also run the 10 and 20 flows respectively in the networks with 3 and 4 channels, using the proposed channel assignment. We list the average aggregate throughputs of the 5 networks with 2 different traffics in Table 1. We see that the network throughput increases with more channels assigned to the network. While the improvement can't be achieved as high times as the number of channels, due to the unneglectable management packets and probing packets for both the WCETT path metric and traffic measurement.

4.2. Varying Traffic Simulation

To evaluate the impact of traffic change on network performance, we randomly choose 3 flows in the 10-flow scene and vary their sending rates in the simulation. Flows run for 180 seconds. At the 60th second, we increase the sending rate of a flow by 0.2Mbps. Then at the 80th second, we increase the sending rate of another flow by 0.4Mbps. At the 100th second, we decrease the sending rate of the third flow by 0.4Mbps. After this, the traffic changes are large enough for the gateway to reassign channels. Because the traffic measurement interval is 10 seconds, the gateway should detect the traffic change and start to reassign channels at the 110th second. Besides, each node need to fresh its routing table since the network channel assignment has changed. So we purge the history information of the WCETT path metric in each node after every new channel assignment.

Figure 10 shows the serial changes of aggregate network throughput by varying the flow sending rates. We can see from the 60th second to the 110th second, the aggregate network throughput changes slightly after each variation of flow sending rate and at last stays around 1.85Mbps. Then at the 110th second, the aggregate throughput suddenly decreases to 0, which means the network begins to reassign channels. After about 15 seconds vibration, the aggregate throughput comes back to near 2.4Mbps which means the new channel assignment as well as the WCETT path metric routing have terminated.

Compared to the last 1.85Mbps throughput before channel reassignment, the network now owns a much better bandwidth assignment. In addition, the channel reassignment and routing together cost only about 15 seconds. So we can say our channel assignment combining with the WCETT path metric routing provides a good choice for the multi-radio WMNs.

5. Conclusions

This paper formulated the channel assignment problem in multi-radio multi-channel WMNs. Since the backbone of WMN is an infrastructured network, we assume the mesh routers are stationary and each of them is equipped with 2 radios. Based on network traffic measurement and the load estimation of wireless links, we present a centralized quasi-static channel assignment with the objective of minimizing network interference, which as a result greatly improves network capacity. Extensive simulations show that the proposed scheme is highly responsive to varying traffic conditions, and the network performance under the channel assignment significantly outperforms the single-radio IEEE 802.11 network as well as the 2-radio WMN with static 2 channels.

6. References

[1] Mesh Networks Inc. http://www.meshnetworks.com.

[2] I. F. Akyildiz, X. D. Wang, and W. L. Wang, "Wireless mesh networks: A survey," Computer Networks, Vol. 47, No. 4, pp. 445–487, 2005.

[3] Mesh Networking Forum, "Building the business case for implementation of wireless mesh networks," Mesh Networking Forum 2004, San Francisco, CA, October 2004.

[4] R. Chandra and P. Bahl, "MultiNet: Connecting to multiple IEEE 802.11 networks using a single wireless card," INFOCOM, Vol. 2, pp. 882–893, 2004.

[5] I. Wormsbecker and C. Williamson, "On channel selection strategies for multi-channel MAC protocols in wireless ad hoc networks," IEEE Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'2006), pp. 212–220, 2006.

[6] J. So and N. Vaidya, "Multi-channel MAC for ad hoc networks: Handling multi-channel hidden terminals using a single transceiver," MobiHoc'04, May 24–26, 2004.

[7] A. Raniwala, K. Gopalan, and T. Chiueh, "Centralized channel assignment and routing algorithms for multi-channel wireless mesh networks," ACM Mobile Computing and Communications Review, Vol. 8, No. 2, pp. 50–65, 2004.

[8] J. Tang, G. Xue, and W. Zhang, "Interference-aware topology control and QoS routing in multi-channel wireless mesh networks," ACM SIGMOBILE, Urbana-Champaign, IL, pp. 68–77, 2005.

[9] A. Subramanian, H. Gupta, and S. R. Das, "Minimum-interference channel assignment in multi-radio wireless mesh networks," Proceedings of 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON'07), pp. 481–490, June 18–21, 2007.

[10] "ns-2 simulator," http://www.isi.edu/nsnam/ns.

[11] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in Proceedings of ACM MOBICOM, pp. 114–128, September 2004.

[12] C. Perkins, E. Royer, and S. Das, "Ad hoc on demand distance vector (AODV) routing," IETF Internet Draft, draft-ietf-manet-aodv2-10.txt, January 24, 2002.

Wireless Sensor Network, 2009, 2, 61-121
doi:10.4236/wsn.2009.12017 Published Online July 2009 (http://www.SciRP.org/journal/wsn/).

A Novel DSA-Driven MAC Protocol for Cognitive Radio Networks

Hua SONG, Xiaola LIN
School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China
Email: *songhua@mail2.sysu.edu.cn, linxl@mail.sysu.edu.cn*
Received February 19, 2009; *revised April* 29, 2009*; accepted May* 6, 2009

Abstract

With the deployment of more wireless applications, spectrum scarcity becomes an issue in many countries. Recent reports show that the reason for this spectrum shortage is the underutilization of some spectrum resources. Fortunately, the emergence of open spectrum and dynamic spectrum assess (DSA) technology in cognitive radio networks relieves this problem. In this paper, we propose a novel DSA-driven cognitive MAC protocol to achieve highly efficient spectrum usage and QoS provisioning. In the proposed protocol, secondary users are divided into several non-overlapping groups, and all leftover channels are allocated among groups taking the groups' bandwidth requirements into consideration. Moreover, the allocation of vacant channels can be adjusted dynamically when members join/leave groups or primary users return/leave the current network. Simulations show that the proposed MAC protocol greatly improves the quality of service for secondary users and maximizes the utilization ratio of spectrum resources.

Keywords: Cognitive Radio, DSA-Driven MAC Protocol, QoS Provisioning, Dynamic Spectrum Access

1. Introduction

The deployment of wireless services and devices has been increasing rapidly in recent years, but current usable spectrum has almost been allocated to various spectrum-based services, which greatly blokes the development of wireless communication. However, extensive reports indicate that the reason for this spectrum shortage is not the scarcity of the radio spectrum, but the low utilization (only 6%) of the licensed radio spectrum in most of the time [1].

This underutilization of spectrum resources has prompted the emergence of cognitive radio. In 2003, Federal Communications Commission (FCC) suggested a new concept/policy for dynamically allocating the spectrum [2]. Thus, a promising implementation technique called cognitive radio is proposed to alleviate the scarcity of spectrum bandwidth. Based on cognitive radio, open spectrum and dynamic spectrum access (DSA) technologies have shown great interest recently [3]. In this technology, primary users (licensed users) have high priority to use their spectrum; secondary users (unlicensed users) are allowed to opportunistically access the spectrum only when the spectrum is not used by primary users.

Although the research community has proposed several cognitive MAC protocols to address various issues in cognitive network [4–8], commonly, they pay more attention to save the number of transceivers, and improve throughput of the whole system or decrease session delays. However, all these protocols do not lay emphasize on quality of service for secondary users and high usage of leftover spectrum with dynamically adjusting allocation. For instance, in [5], a decentralized protocol, called hardware-constrained cognitive MAC protocol (HC-MAC), for managing and coordinating spectrum access is proposed. Under HC-MAC, a pair of secondary users can use several channels to communicate simultaneously after they have sensed the vacant channels, but if the leftover spectrum allocated to the pair of secondary users is more than they can utilize, this part of surplus spectrum is wasted and cannot be used by other users whose

* This work was supported in part by NSFC under Projects 60773199, U0735001, and 985 II fund under Project 3171310.

bandwidth requirements have not been satisfied.

To effectively provide QoS for secondary users and achieve highly efficient spectrum usage, in this paper, we propose a novel DSA-driven MAC protocol which is of significant importance in ad hoc cognitive network to guarantee the QoS requirements of secondary users. Different from the existing cognitive MAC protocols, the main advantages of the proposed MAC protocol include: 1) Maximizing the utilization ratio of spectrum resources; 2) Using bonding/aggregation and dynamical channel allocation techniques to guarantee the QoS requirements of secondary users; 3) Ensuring the fairness of channel allocation for groups.

The rest of the paper is organized as follows. Related work is discussed in Section 2. The preliminaries and system model is introduced in Section 3. The proposed MAC protocol is presented in Section 4. Performance of the proposed MAC protocol is evaluated by experiments in Section 5. Finally, conclusion is drawn in Section 6.

2. Related Work

Over the past several years there have been increasing interests in cognitive radio. In addition, wireless MAC protocol has a principal part in spectrum reuse and efficiency management. Therefore, various cognitive MAC protocols have been proposed for more flexible and efficient use of spectrum resources [4−8].

Hamdaoui and Shin [4] propose the OS-MAC protocol. This protocol divides the secondary users into several groups, at each Opportunistic Spectrum Period, the channel used by a group can be adjusted dynamically according to the channels' states of the whole system. The OS-MAC protocol also discusses the non-cooperative mode between primary users and secondary users, but it does not give a feasible solution (we will present our solution on this problem in subsequent paper). Furthermore, as only one channel can be used in a group at anytime, the spectrum resources cannot be used to the maximum with taking secondary users' QoS into consideration.

Jia et al. [5] present the HC-MAC protocol. This protocol uses *k-stage* look-ahead method to sense unused channels with high efficiency and takes hardware-constraints into consideration (including sensing constraint and transmission constraint). Besides the flaw mentioned in Section 1, there is a problem called sensing exposed terminal problem in this protocol. That is, if a secondary pair senses unused channels while their neighbors who didn't receive the C-RTS/C-CTS packets are performing their operations freely, this pair of users can not sense the vacant channels accurately.

Su and Zhang [6] propose the cross-layer based opportunistic MAC protocols. In the protocols two sensing policies, the random sensing policy and the negotiation-based sensing policy, are presented. Like [5], the protocols also use bonding/aggregation technique to transmit data through several channels. In essence, the main contribution of their work is to reveal the tradeoff between throughput and delay, which provides the guidelines to support the different QoS requirements over cognitive radio based wireless networks. But, at anytime, only a pair of secondary users can use vacant channels, thus the leftover channels cannot be used efficiently.

Thoppian et al. [7] propose a CSMA-Based MAC protocol. In the protocol, each node maintains a list of favorable channels for each of its neighbors based on the previous history of communication on each of the channels, and a secondary pair chooses the most favorable channel for communication. As it does not consider the channels' utility of the whole system, it also cannot use the radio spectrum resources efficiently.

Similarly, the methods in [9−12] all do not address the issues of the QoS of secondary users and the spectrum resources' utility of the entire system.

3. Preliminaries and System Model

We first present an introduction to the channel bonding/aggregation technique and the main framework of the system.

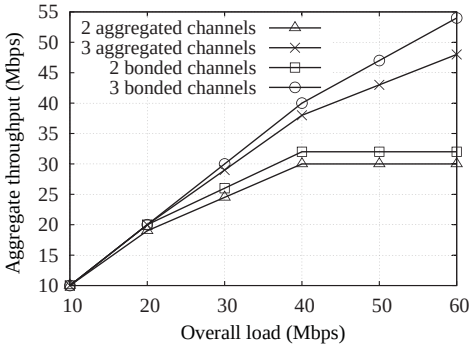
3.1. Channel Bounding/Aggregation Technique

From the definition of channel bonding/aggregation technique in [13] we can see that channel bonding is used for contiguous channels, but channel aggregation is for discrete ones. So, if several contiguous channels can be used, channel bonding is the appropriate technique. Otherwise, we can adopt channel aggregation technique.

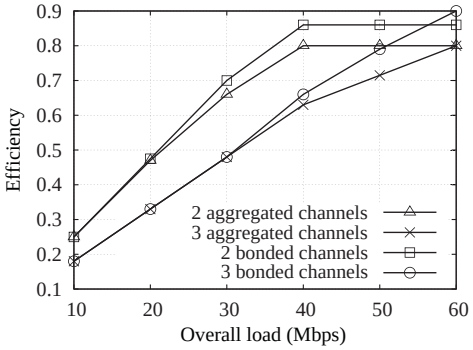
In MAC perspective, channel bonding incurs no additional overhead as all control messages are transmitted only once, and an access point (AP) with channel bonding also has much greater control and more freedom on resource allocation and transmit power. In contrast, with channel aggregation, the overhead increases considerably with the number of channels used, and for an effective channel aggregation solution, features such as sophisticated scheduling, load balancing and channel management are needed.

Figure 1(a) and Figure 1(b) compare the aggregate throughput and efficiency of the two techniques respectively. From Figure 1, we can see that channel aggregation incurs much more overhead than channel bonding and these two techniques are designed for medium-high loads. Also, this figure presents that channel bonding can offer much better channel utilization with less overhead.

In [5], hardware constrains of bonding/aggregation are pointed out: spectrum used by a secondary user has maxi-



(a) Aggregate throughput against overload.



(b) Efficiency against overload.

Figure 1. Channel bonding vs. channel aggregation.

mum bandwidth limit and maximum fragmentation number limit. For convenience, we will not take these constrains into consideration presently, but they are included in our future research plan.

3.2. Secondary Users' Devices

The whole available radio spectrum in a CR network is divided into a number of non-overlapping *data channels* (DCs) and a common *control channel* (CC), for transmitting data information and control information respectively. Further, we assume the CR network is a wireless ad hoc network formed by a great many secondary users, each of which is equipped with two wireless transceivers. One is called *data transceiver* used to sense leftover DCs and exchange data through these channels, but this transceiver is not able to operate in sensing and transmitting mode concurrently. The other transceiver is called *control transceiver* used to send or receive messages from CC. Note that, the transceivers are half-duplex, thus they are not able to send and receive messages simultaneously. We also assume that every secondary user can use aggregated spectrum and has full spectrum sensing ability.

3.3. Control Mechanism

Based on distributed environment, we divide secondary users into several non-overlapping groups, and each group has a leader (the leader can be adjusted dynamically) who is responsible for group members' management, group channel's management and group channels' application. Moreover, there is a manager in the whole wireless ad hoc network who is elected among leaders and is used to manage all the leftover channels and allocate these channels to groups fairly. The manager communicates with leaders through CC, and the control messages in a group are exchanged through DCs. So the common control channel (CC) is light loaded and will not be the bottleneck of the network. As the manager and leaders can be changed dynamically, each of them is impossible to become a single failure point. Thus, the system is provided good scalability and extensibility.

4. Our Proposed DSA-Driven MAC Protocol

In this section, we present the design of our proposed DSA-driven MAC protocol for cognitive radio networks.

4.1. Definitions and Notations

In order to present the proposed MAC protocol more clearly, we would now like to introduce some definitions and notations here.

1) *Tables*: These tables are used for the management of the whole network.

- **GroupTable**: This table is maintained by the manager. It contains the group number, the leader and members in the group.
- **MemberTable**: This table is maintained by the leader. It contains the list of group members. Note that each member has a MemberTable.

2) *Control frames*: These frames are used for the control of protocol's realization. The first five frames are sent through CC, the LeaveReq frame is sent through DC and the last one can be sent through CC or DC.

- **Invite**: It contains the manager's id.
- **JoinReq**: It contains the id of a secondary user whom the request user wants to communicate with.
- **JoinACK**: It contains the leader's id of chosen group.
- **Sense**: It contains the group number and the sensing range of channels.
- **Allocate**: It contains the group number and a list of channels.
- **LeaveReq**: A member uses this frame to apply for leaving the current group.
- **Notify**: It contains ManagerChanged and LeaderChanged fields.

3) *Timers*: These timers are used for the maintenance of the whole system.

- **ManagerHeartBeatTimer**: This timer is started at the manager when it broadcasts the **Invite** frame on CC. In essence, ManagerHeartBeatTimer has two functions, the first one is to let secondary users

know who is the manager currently and the other one is to make leaders assure the manager is still alive.

- **ManagerFailureTimer**: This timer is started at a leader when it receives an **Invite** frame on CC. If this timer expires before the leader receives another **Invite** frame, it knows that the manager is failed.
- **LeaderHeartBeatTimer**: This timer is started at a leader when it broadcasts the LeaderHeartBeat frame on DC in a group. As soon as members receive the LeaderHeartBeat frames they know that the leader is alive.
- **LeaderFailureTimer**: This timer is started at a member when it receives a LeaderHeartBeat frame from DC. If this timer expires before the member receives another LeaderHeartBeat frame, it knows that the leader is failed.

4.2. Overview

There are two kinds of control frames in our proposed MAC protocol: inter-group control frame and intra-group control frame. Inter-group control frames are transmitted between the leaders and manager through CC, while intra-group control frames are exchanged between the leader and members through DCs in a group. If several leaders want to send messages to the manager, they must negotiate with each other via the contention-based algorithms, such as IEEE 802.11 Distributed Coordination Function (DCF) [14] and p-persistent Carrier Sense Multiple Access (CSMA) protocols [15]. In order to guarantee reliable transmission, we use acknowledge mechanism in control messages' exchange.

Figure 2 shows the principle of our proposed MAC protocol. It divides time into Periods and each of which consists of three consecutive phases: Sensing Phase, Allocating Phase and Transmitting Phase. At any time, only one pair of members in a group can exchange information. In order to let all members in the same group use channels fairly, the Transmitting Phase is divided into several time slots according to the number of users. Furthermore, a few of time slots are reserved for intra-group control frames exchange.

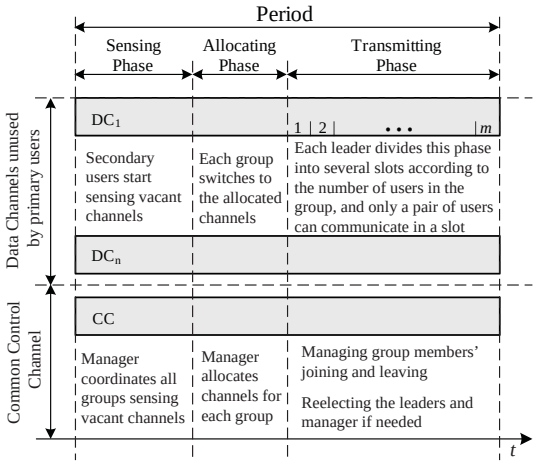


Figure 2. The principle of our proposed MAC protocol.

In Sensing Phase, manager coordinates all groups sensing vacant channels (it's an efficient method to check all unused channels in a short time). Then, based on the feedback of sensing results and bandwidth requirements from all groups, the manager calculates out a best allocation scheme and allocates the available channels to each group in Allocating Phase. When groups gain their new allocated channels, they switch to these channels immediately and begin to transmit information in Transmitting Phase.

4.3. Details of the Proposed MAC Protocol

In the proposed MAC protocol, each secondary user in the network will be in one of the following phases at any given time.

1) *Initialization Phase*: If a secondary user is not involved in any group and seeks channels to transmit information, it will listen to CC.

a) If the secondary user receives an **Invite** frame from CC, it sends a **JoinReq** control frame to apply for joining a group. After receiving the **JoinReq** frame, the manager looks up its GroupTable to decide which group this user should join in according to the user's request. If the manager finds an appropriate group, it updates its GroupTable with a **JoinACK** control frame to send back. When the secondary user receives the **JoinACK** frame, it communicates with the leader of the chosen group through CC and tunes its data transceiver to the according DCs. However, the manager may not find a suitable group for the secondary user to join in, then it checks whether it's possible to establish a new group for this user. If so, the manager creates a new group and makes this secondary user a leader; otherwise a **REJ** control frame is responded.

b) If the ManagerFailureTimer expires before the secondary user receives an **Invite** frame, the user knows that it is the only secondary user in the network. Then it establishes a new group, makes itself a leader and manager, creates a MemberTable and a GroupTable, and broadcasts an **Invite** frame on CC.

2) *Sensing Phase*: all secondary users cooperate to sense the channels unused by primary users.

a) The manager allocates sensing channels among all groups according to GroupTable, and sends a **Sense** frame to each leader via CC.

b) Each group senses vacant channels whose range is indicated by the **Sense** frame.

c) Leaders report sensing results to the manager through CC.

3) *Allocating Phase*: When the manager gets the vacant channel information from all groups, it allocates these channels among groups.

a) Each leader sends bandwidth requirement of its group to manager via CC.

b) The manager determines a channel allocation scheme according to the bandwidth requirement of each group, and sends this scheme to leaders.

c) Each group switches to the allocated channels.

4) *Transmitting Phase*: As soon as groups switch to the allocated channels, they can use bonding/aggregation technique to transmit data.

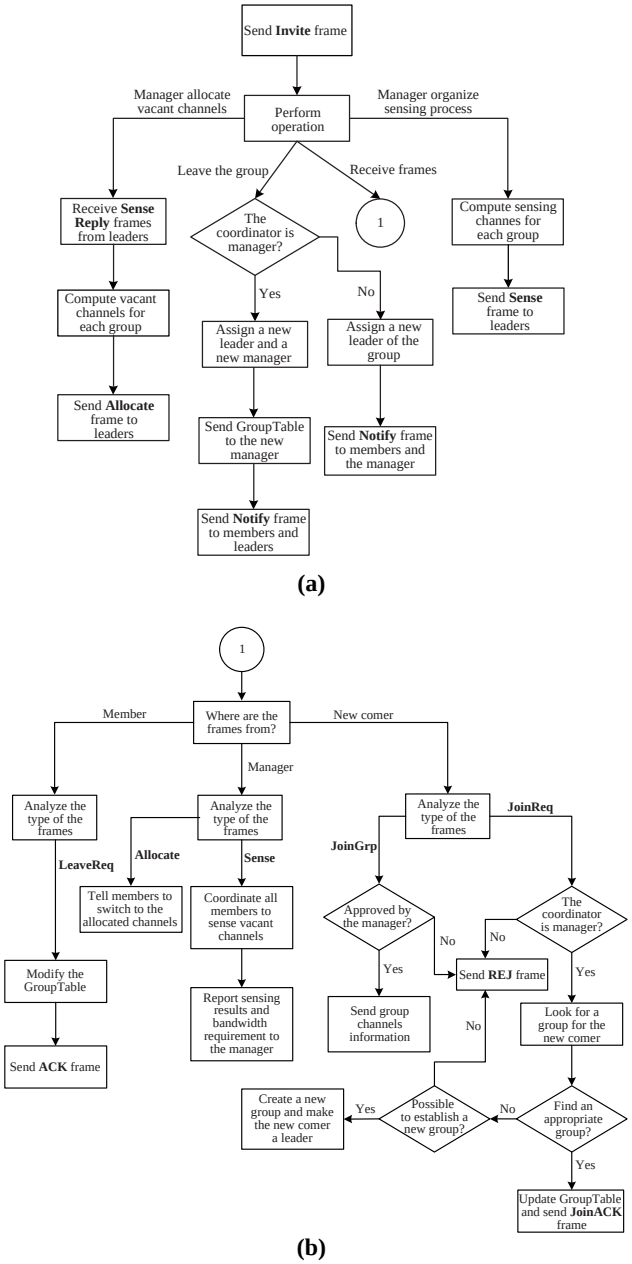


Figure 3. The operation flow chart of a coordinator.

a) As leaders divide the whole Transmitting Phase into several time slots, in each group, the members use channels in turn.

b) At the end of Transmitting Phase, the manager and leaders exchange group information through CC in order to update GroupTable and MemberTable respectively. Also, in each group the leader broadcasts GroupTable to all members.

c) If a member wants to leave a group, the following operations will be done:

- If the member is not a leader or manager, it sends a **LeaveReq** frame to the leader through DC, then the leader modifies the GroupTable with an **ACK** frame is replied.
- If the member is a leader, it assigns a new leader from rest members of this group and broadcasts a **Notify** frame on DCs in the group; it also sends this frame to the manager via CC. In case this leader is the last one in the group, it sets leader's id NULL in the LeaderChanged field of **Notify** frame.
- If the member is a manager, first, it selects a new leader from rest members in its group and a new manager from all leaders; second, it sends GroupTable to the new manager; last, it broadcasts a **Notify** frame on DCs in the group, as well as posts this frame to leaders via CC.

d) If a secondary user wants to join a group, it will do the operation as shown in Initialization Phase.

In the following, we call the manager or leader coordinator. Figure 3 delineate the operation flow of a coordinator according to the protocol's details mentioned above. For conciseness, we only give the foremost control operations.

4.4. Channel Allocation Mechanism

We suppose a radio spectrum system consisting of M groups which apply for n vacant channels. For each group, the allocation optimization problem can be described as:

$$\begin{aligned} &\max \sum_{j=1}^{c_i} b_{ij} \\ &s.t.\ \ c_i \le n \\ &\sum_{j=1}^{c_i} b_{ij} < \min(r_i, (r_i \sum_{i=1}^{n} B_i / \sum_{i=1}^{M} r_i)) \\ &\qquad + (\sum_{i=1}^{n} B_i / \lceil n/M \rceil n) \end{aligned} \quad (1)$$

where $c_i (i = 1, \cdots, M)$ is the number of leftover channels allocated to the i^{th} group, $b_{ij} (j = 1, \cdots, c_i)$ denotes the bandwidth of the j^{th} channel which is allocated to the i^{th} group (In this paper, we assume that all channel's bandwidths are not the same and they are uniformly

distributed), $r_i (i = 1, \cdots, M)$ is the bandwidth requirement of the i^{th} group, and $B_i (i = 1, \cdots, n)$ is the bandwidth of the i^{th} vacant channels. The last constraint indicates that the bandwidth allocated to each group cannot exceed a certain upper bound which is defined as $\min(r_i, (r_i \sum_{i=1}^{n} B_i / \sum_{i=1}^{M} r_i)) + (\sum_{i=1}^{n} B_i / \lceil n / M \rceil n)$. The augend of the upper bound represents each group's deserved bandwidth according to its requirement and the addend is an adjusting factor which guarantees the channel allocation is accurate and even.

The object of problem (1) is to maximize the utility of vacant channels. For this kind of problems, branch and bound algorithm [16, 17] is an effective method to search the optimal solutions. In this paper, we propose a branch and bound algorithm, which is more suitable for the channel allocation optimization problem in cognitive radio networks than the general one. For each group, the detail steps are as follows:

1) *Initialization*: The feasible solution can be presented as $P = \{x_1, \cdots, x_t\}$, Where t $(t \le n)$ is the number of available vacant channels for the current group, and $x_i \in \{0,1\}$ $(i = 1, \cdots, t)$. If the value of x_i is 1, it means that the i^{th} channel is allocated to the current group. We define $P_0 = \{0, 0, \cdots, 0\}$. In this algorithm, the channels are sorted by $B_i (i = 1, \cdots, t)$ which is the bandwidth of the i^{th} channel, that is $B_j \ge B_k, \forall j < k,\ j, k \in I$,; because the channel with higher bandwidth has larger impact to the optimal solution. Let $S_0 = \{P_0\}$ to indicate the initial feasible set and initialize $k = 1$. The initial lower and upper bound of the optimal solution can be represented as

$$\alpha_0 = 0 \tag{2}$$

$$\beta_0 = \sum_{i=1}^{t} B_i \tag{3}$$

2) *Branching*: Based on the characteristics of the channel allocation optimization problem, the branching method puts each $P \in S_{k-1}$ in duplicate to set S_k and set the value of x_k .

$$P^{(1)} = \{P \mid x_k = 0, x_k \in P\} \tag{4}$$

$$P^{(2)} = \{P \mid x_k = 1, x_k \in P\} \tag{5}$$

3) *Bounding*: The lower bound of the optimal solution in a set can achieve by a greedy method. In this method, the channel with high bandwidth has the priority to be chosen in order to achieve the optimal capacity. Thus, the lower and upper bound of the optimal value in *P* is

$$\alpha(P) = \sum_{i=1}^{k} x_i B_i \tag{6}$$

$$\beta(P) = \sum_{i=1}^{k} x_i B_i + \sum_{i=k+1}^{t} B_i \tag{7}$$

The lower and upper bounds of the optimal solution in the entire feasible region *D* are

$$\alpha_k = \max_{P \in S_k} \alpha(P) \tag{8}$$

$$\beta_k = \max_{P \in S_k} \beta(P) \tag{9}$$

4) *Pruning*: For $P \in S_k$, it should be pruned if one of the following conditions is satisfied:

a) The constraints are violated.

b) The upper bound of this set is smaller than the maximum lower bound, which is $\beta(P) < \alpha_k$.

In this step, many subsets are pruned to accelerate the searching speed.

5) *Convergence*: In this step, examine whether it is the t^{th} iteration.

a) If $k < t,\ k+1 \to k$ and go to the branching step.

b) If $k = t$, the optimal solutions of all the unpruned subsets are confirmed.

In the convergence step of the t^{th} iteration, end the iteration and enumerate the optimal solutions of all unpruned subsets in P_t to search the global optimal solution.

6) *Return*: After getting the optimal solution for current group, the algorithm returns to calculate for the next group. Thus, the channels allocated to the current group will be excluded in the next round, and the number of available channels *t* will also be modified.

4.5. Election Mechanism

If the coordinator is crashed, members must elect a new one. In our proposed MAC protocol, we use a simple and high efficient algorithm to realize the election mechanism.

When any member notices the coordinator is not functioning, it sends an **Election** frame on CC/DCs (manager election frame on CC and leader election frame on DCs) to apply for becoming the new coordinator. If several members detect the coordinator's malfunction concurrently, they have to compete for the CC/DCs through the IEEE 802.11 random access scheme. Recall that for a successful transmission, an **ACK** frame will be sent back to the sender. Under our proposed MAC protocol, as all members are listening to CC/DCs at all times, each member is able to receive this **Election** frame, and if anyone replies, the **ACK** frame will be heard by all members. The member who is the first to successfully

deliver an **Election** frame is automatically appointed as the new coordinator. Therefore, upon receiving an **ACK** frame notifying a successful reception, this member considers itself the new leader. Also, any other members who hear the **ACK** frame, and hence would know that someone else is appointed to be the new leader and need not send its own acknowledgement.

In Figure 4, an example of election algorithm is given. The group consists of eight processes, numbered from 0 to 7. Member 7 is the coordinator which has just crashed. Member 4 is the first one to notice this, so it sends an **Election** frame on CC/DCs. After hearing the **Election** frame from Member 4, Member 6 affirms Member 7's malfunction and replies with an **ACK** frame. Upon receiving this **ACK** frame, Member 4 knows that it has been permitted to be the new coordinator. Furthermore, because the rest members in the same group all can hear the **Election** and **ACK** frames, they are also aware of Member 7's crash, as well as the new coordinator's generation.

Note that the election mechanism is very efficient, it exploits the already existing ACK mechanism and does not require any extra message exchange, thus incurring no bandwidth overhead.

5. Simulation and Performance Evaluation

In this section, we present the simulation results for the performance evaluation of the protocol. The simulations are performed using the network simulator ns-2 [18]. The used parameters are presented in Table1.

In the simulation, the proposed MAC protocol is compared with the following existing MACs: OS-MAC and CO-MAC.

1) *OS-MAC*: OS-MAC [4] is an opportunistic spectrum MAC. It divides time into periods each of which is called OSP (Opportunistic Spectrum Period) and consists of three consecutive phases: Select, Delegate, and Update.

- **Select Phase**: Each SUG (Secondary User Group) selects a "best" DC, and uses it for communication until the end of the current OSP.
- **Delegate Phase**: On each DC, a DSU is appointed among members to represent the group during the Update Phase.

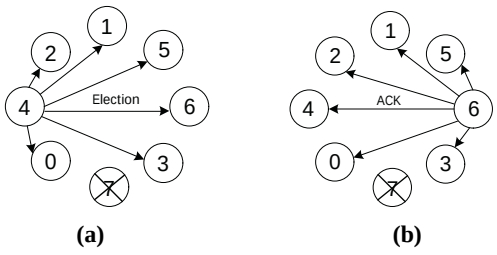


Figure 4. Election algorithm.

Table 1. Parameters used in the simulations.

Parameter	Value
Number of secondary user group	5, 8
Number of DCs	30
Bandwidth of each DC (not the same)	1~1.5Mbps
Required bandwidth of each group	4.5Mbps
Transmit power	0.01W
Transport protocol	UDP

- **Update Phase**: All DSUs switch to CC to update each other about their channel conditions while all non-DSUs continue communicating on their DCs.

2) *CO-MAC*: Like OS-MAC, CO-MAC [6], an opportunistic MAC protocol, uses two transceivers, a dedicated CC, and N DCs. The time is divided into a number of periodical time slots and each slot is divided into two phases, namely, Reporting Phase and Negotiating Phase. Reporting phase can be further divided into n mini-slots, each of them corresponding to one of the n licensed channels.

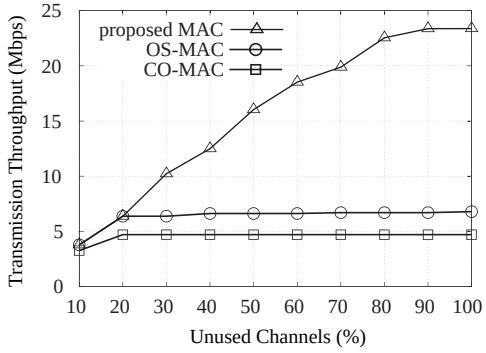
- **Reporting Phase**: Each secondary user is equipped with one SDR transceiver, and by using this transceiver only one of n licensed channels can be sensed. Thus the secondary user is unable to accurately know the states of all the channels by itself. However, the goal of the Reporting Phase is to empower the secondary users to have a large picture of all the channels' states through their cooperation. In particular, each secondary user senses a channel in corresponding mini-slot, if the channel is idle, the user sends a beacon during this mini-slot over the control channel. Otherwise, no beacon is posted.
- **Negotiating Phase**: the secondary users use the control transceivers to negotiate about the data channels among the secondary users by exchanging request-to-send (RTS) and clear-to-send (CTS) packets over the control channel. Meanwhile, the only secondary user which is the winner in contending for the data channels during the last time slot uses the SDR transceiver to transmit data packets over all the unused licensed channels in the current time slot.

For conciseness, we define the unused channels' percentage in the entire system as $p_n = n / N$, where N is the number of channels licensed to primary users.

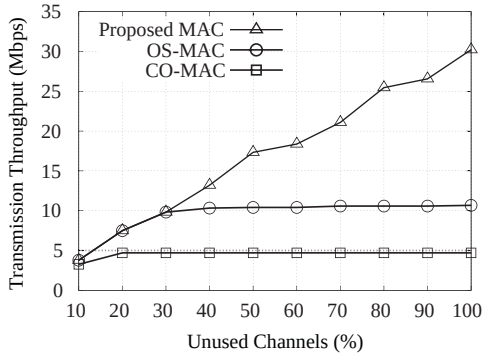
5.1. Throughput Analysis

In this section, we analyze the throughput with different numbers of unused channels under three protocols: the proposed MAC, OS-MAC and CO-MAC. The throughput allows us to evaluate the protocols' performance, that is, the higher throughput, the higher performance.

Figure 5 shows the throughput comparison for our proposed MAC protocol with the other two MAC protocols under each of the two network scenarios: (Figure 5(a)) the number of secondary user groups with *M*=5 and (Figure 5(b)) the number of secondary user groups with *M*=8. First note that with the increase of unused channels' percentage (p_n), the system throughput of these three protocols all increases. Also, observe that the throughput of the proposed MAC protocol increases sharply while CO-MAC attains a bound rapidly and OS-MAC increases little after reaching a fixed value. This demonstrates that our proposed protocol can sufficiently use the vacant channels with the increment of p_n. In OS-MAC, as only one channel can be used by a group and all channel's bandwidths are uniformly distributed between 1*Mbps* and 1.5*Mbps*, when each group monopolizes a channel (i.e., this group does not share the channel with other groups), the overall throughput won't increase much with the increment of p_n but with the number of secondary user groups (*M*). For example, in Figure 5 (a), the system throughput of OS-MAC in $p_n = 40\%$ is a litter higher than which in $p_n = 30\%$, because with the increasing of p_n, although each group cannot gain another channel, it has more opportunity to switch to a

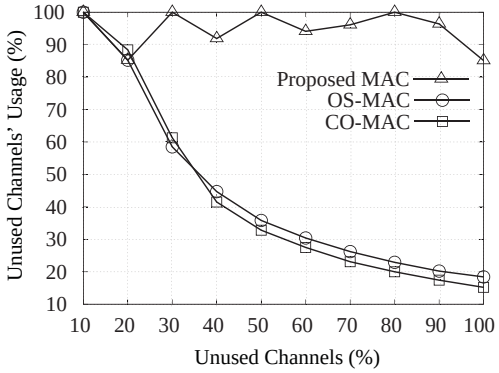


(a) The number of secondary user groups *M* = 5.

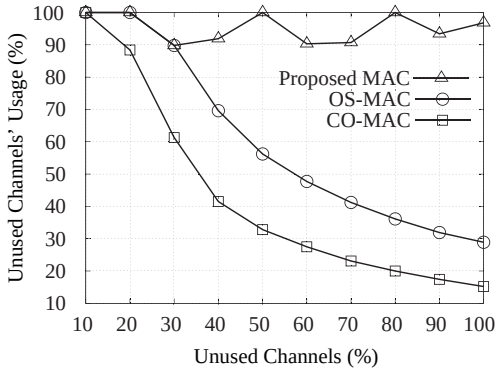


(b) The number of secondary user groups *M* = 8.

Figure 5. The throughput with different number of unused channels.



(a) The number of secondary user groups *M* = 5.



(b) The number of secondary user groups *M* = 8.

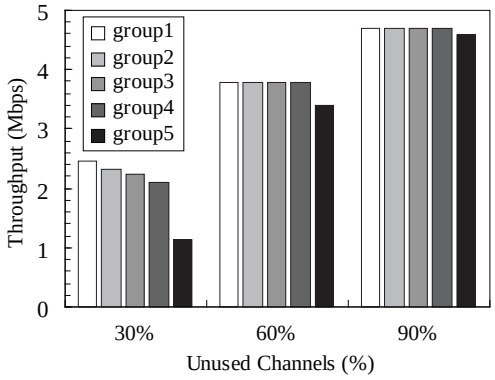
Figure 6. The unused channels' usage with different number of unused channels.

"best" one, Moreover, in Figure 5(b), the throughput of OS-MAC is much higher than the counterpart in Figure 5(a), because more vacant channels are utilized by groups as *M* is increasing. In CO-MAC, since only a pair of secondary users uses multiple vacant channels at anytime and each pair's required bandwidth is the same in our "best" one, Moreover, in Figure 5(b), the throughput of OS-MAC is much higher than the counterpart in Figure 5(a), because more vacant channels are utilized by groups experiment, as soon as the required bandwidth is satisfied, the entire throughput will not change. Another point that requires attention is that, in Figure 5(a), the curve represented for the system throughput of our proposed protocol is closed to a fixed value when $p_n = 80\%$. This is because of the throughput of the entire system maintains a certain value and more vacant channels are not needed when all groups' bandwidth requirements are satisfied.

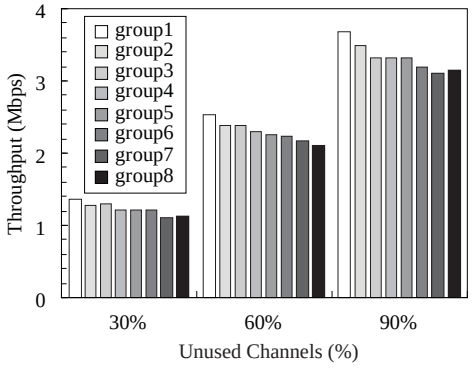
5.2. Ratio of Unused Channels' Utilization Analysis

Figure 6 depicts the variation of unused channels' utilization rate (ρ) of the three MAC protocol with the change of unused channels' quantity (p_n) under each of

the two network scenarios: (Figure 6(a)) the number of secondary user groups with *M*=5 and (Figure 6(b)) the number of secondary user groups with *M*=8. There are three observations to make. First, the unused channels' usage (ρ) of our proposed MAC protocol fluctuates between 0.9 and 1.0, while the other two protocols' unused channels' usage (ρ) decreases with the increase of p_n. Second, we find that, in the proposed protocol, unless the secondary users' bandwidth requirements are satisfied, they utilize the vacant channels to their fullest extent and conduct high channel usage. However, the unused channels usage (ρ) in OS-MAC drops rapidly with p_n due to the utilization of fixed number of vacant channels except for the increment of *M*. Third, in Figure 6 (a), the curve represented for ρ under our proposed MAC protocol declines quickly after $p_n = 80\%$ caused by the satisfaction of all groups' bandwidth requirements. However, unlike CO-MAC, in the proposed protocol, with the increment of p_n, part of unused channels will not be occupied by some secondary users whose bandwidth requirements are already satisfied, thus these channels are able to be allocated to other users who need them. This is an important feature.



(a) The number of secondary user groups *M* = 5.



(b) The number of secondary user groups *M* = 8.

Figure 7. Group throughput with different p_n in our proposed protocol.

5.3. QoS and Fairness Analysis

In this subsection, we evaluate QoS and fairness of the proposed MAC protocol.

Figure 7 presents each group's throughput against the different number of unused channels ($p_n = 30\%$, $p_n = 60\%$ and $p_n = 90\%$) under each of the two network scenarios: (Figure 7(a)) the number of secondary user groups with *M* = 5 and (Figure 7(b)) the number of secondary user groups with *M* = 8. Commonly, the throughput comparison of all groups with certain p_n reflects the fairness of the whole system, and the comparison of certain group's throughput with different p_n shows the QoS support of the network. In Figure 7, we observe that, with certain p_n, the throughputs of all groups are much closer. Especially when p_n = 90%, the groups' throughputs are almost the same as those in Figure 7(a). We also note that the throughput of certain group increases monotonically with the increment of p_n. For instance, in Figure 7(b), the throughput of Group1 raises from 1.365*Mbps* to 3.671*Mbps* as p_n increases from 30% to 60%.

Based on the simulation results, we can make the following conclusions. First, our proposed MAC protocol is shown to be more efficient than the other two protocols not only from throughput aspect, but also from unused channels' utilization aspect. Second, our proposed MAC protocol greatly improves the quality of service for secondary users and guarantees the fairness of channels allocation for groups.

6. Conclusions

In this paper, we have proposed a novel DSA-driven MAC protocol for cognitive radio networks, In which, secondary users are divided into several non-overlapping groups, and each group uses bonding/aggregation technique to transmit data. All leftover channels are allocated among groups taking the groups' bandwidth requirements into consideration. Moreover, the allocation of vacant channels can be adjusted dynamically when members join/leave groups or primary users return/leave the current network. The simulations indicate that our proposed MAC protocol greatly improves the quality of service for secondary users and maximizes the utilization ratio of spectrum resources.

7. References

[1] M. Mchenry, "Spectrum white space measurements," New America Foundation Broadband Forum, June 2003.

[2] FCC, "Et docket no. 03-237," November 2003. http://hraunfoss.fcc.gov/edocs public/attachmatch/FCC-03-289A1.pdf

[3] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," Computer Networks, 2006.

[4] K. G. Shin and B. Hamdaoui, "OS-MAC: An efficient MAC protocol for spectrum-agile wireless networks," Transactions on Mobile Computing, 2007.

[5] J. Jia, Q. Zhang, and X. Shen, "HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management," IEEE Journal on Selected Areas in Communications, Vol. 26, No. 1, pp. 106–117, January 2008.

[6] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks," IEEE Journal on Selected Areas in Communications, Vol. 26, No. 1, pp. 118–129, January 2008.

[7] T. Mansi, V. S., and P. Ravi, "CSMA-based MAC protocol for cognitive radio networks," World of Wireless, Mobile and Multimedia Networks, IEEE International Symposium, pp. 1–8, June 2007.

[8] A. Mishra, "A multi-channel MAC for opportunistic spectrum sharing in cognitive networks," Proceedings of MILCOM 2006, October 2006.

[9] L. Ma, X. Han, and C. C. Shen, "Dynamic open spectrum sharing MAC protocol for wireless ad hoc networks," in IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, pp. 203–213, 2005.

[10] J. So and N. Vaidya, "Multi-channel MAC for ad-hoc networks: Handling multi-channel hidden terminals using a single transceiver," in ACM Proceedings of MOBIHOC, 2004, pp. 222–233.

[11] J. Zhao, H. Zheng, and G. H. Yang, "Distributed coordination in dynamic spectrum allocation networks," in IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, pp. 259–268, 2005.

[12] N. Nie and C. Comaniciu, "Adaptive channel allocation spectrum etiquette for cognitive radio networks," in IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, pp. 269–278, 2005.

[13] C. Corderio, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: An introduction to the first wireless standard based on cognitive radios," Journal of Communications, Vol. 1, No. 1, pp. 38–47, April 2006.

[14] IEEE Standard 802.11 – 1999, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," November 1999.

[15] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part I - carrier sense multiple-access modes and their throughput-delay characteristics," IEEE Transactions on Communications, No. 23, pp. 1400–1416, December 1975.

[16] P. Apkarian and H. D. Tuan, "Robust control via concave minimization local and global algorithms," IEEE Transactions on Automatic Control, Vol. 45, No. 2, pp. 299–305, February 2000.

[17] W. Wang, T. Peng, and W. Wang, "Optimal power control under interference temperature constraints in cognitive radio network," Wireless Communications and Networking Conference, WCNC 2007 IEEE, pp. 116–120, March 11–15, 2007.

[18] The network simulator (ns-2). http://www.isi.edu/nnsnam/ns/.

Call for Papers

International Conference on Computational Intelligence and Software Engineering (CiSE)

December 11~13, 2009 Wuhan, China

http:// www.ciseng.org

The International Conference on Computational Intelligence and Software Engineering (CiSE) will be held on December 11~13, 2009 in Wuhan, China. This conference will cover issues in **computational intelligence, information security, multimedia and graphics technologies, and software engineering.** The conference proceedings will be published by IEEE, and all papers accepted will be included in IEEE Xplore and indexed by EI Compendex.

TOPICS

- Artificial Intelligence
- Automated Reasoning
- Autonomous Systems
- Cloud Computing
- Cluster Computing
- Cognitive Science
- Communication Networks and Protocols
- Computational Biology
- Computational Chemistry
- Computational Neuroscience
- Computational Physics
- Computer Graphics
- Data Modeling
- Database Mining
- Database Technology
- Evolvable Hardware
- Expert Systems
- Fuzzy Systems
- Geographical Information System
- Grid Computing
- Hardware Implementation
- Human-computer Interaction
- Hybrid Systems
- Image Processing
- Image Understanding
- Machine Learning
- Multi-Agent Systems
- Natural Neural Systems
- Neural Genetic Systems
- Neural-Fuzzy Systems
- Numerical Algorithms
- Operating Systems
- Pattern Recognition
- Programming Methodology
- Project Management
- Real Time Control
- Requirement Analysis
- Simulation and Modeling
- Software Engineering
- Software Maintenance
- Software Standards and Design
- Software Testing and Quality Control
- Symbolic Mathematics
- Technological Forecasting
- Visualization
- Web-based Services

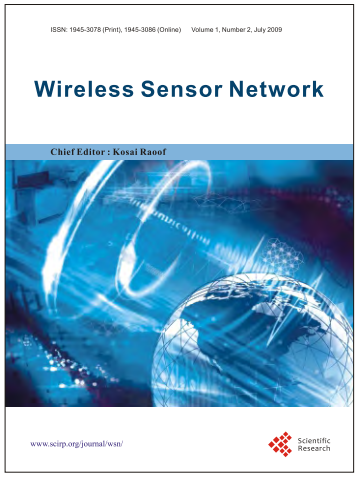
IMPORTANT DATES

- Paper submission due: Jun. 20, 2009
- Acceptance notification: Aug. 20, 2009
- Conference: Dec. 11–13, 2009

CONTACT INFORMATION

Website: http://www.ciseng.org

E-mail: info@ciseng.org



Wireless Sensor Network (WSN)

Call For Papers

http://www.scirp.org/journal/wsn

ISSN 1945-3078 (Print) ISSN 1945-3086 (Online)

WSN is an international refereed journal dedicated to the latest advancement of wireless sensor network and applications. The goal of this journal is to keep a record of the state-of-the-art research and promote the research work in these areas.

Editor-in-Chief

Dr. Kosai Raoof , GIPSA LAB, University of Joseph Fourier, Grenoble, France

Subject Coverage

This journal invites original research and review papers that address the following issues in wireless sensor networks. Topics of interest are (but not limited to):

- Network Architecture and Protocols
- Self-Organization and Synchronization
- Quality of Service
- Data Processing, Storage and Management
- Network Planning, Provisioning and Deployment
- Integration with Other System
- Software Platforms and Development Tools
- Routing and Data Dissemination
- Energy Conservation and Management
- Security and Privacy
- Developments and Applications
- Network Simulation and Platforms

We are also interested in short papers (letters) that clearly address a specific problem, and short survey or position papers that sketch the results or problems on a specific topic. Authors of selected short papers would be invited to write a regular paper on the same topic for future issues of the WSN.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. Authors are responsible for having their papers checked for style and grammar prior to submission to WSN. Papers may be rejected if the language is not satisfactory. For more details about the submissions, please access the website.

Website and E-Mail

http://www.scirp.org/journal/wsn

Email: wsn@scirp.org

Wireless Sensor Network, 2009, 2, 61-121
Published Online July 2009 in SciRP (http://www.SciRP.org/journal/wsn/).

TABLE OF CONTENTS

Volume 1 Number 2 **July 2009**

9771945307008 02