International Journal of Communications, Network and System Sciences

ISSN: 1913-3715

Volume 3, Number 2, February 2010

www.scirp.org/journal/ijcns/

Scientific Research

JOURNAL EDITORIAL BOARD

ISSN 1913-3715 (Print)   ISSN 1913-3723 (Online)
http://www.scirp.org/journal/ijcns/

Int. J. Communications, Network and System Sciences, **2010, 3, 107-211**
Published Online February 2010 in SciRes (http://www.SciRP.org/journal/ijcns/).

TABLE OF CONTENTS

Volume 3 Number 2 February 2010

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

Journal Information

SUBSCRIPTIONS

The *International Journal of Communications, Network and System Sciences* (Online at Scientific Research Publishing, www.SciRP.org) is published monthly by Scientific Research Publishing, Inc., USA.

E-mail: service@scirp.org

Subscription rates: Volume 3 2010
Print: $50 per copy.
Electronic: free, available on www.SciRP.org.
To subscribe, please contact Journals Subscriptions Department, E-mail: service@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: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: service@scirp.org

COPYRIGHT

Statements and opinions expressed in the articles and communications are those of the individual contributors and not the statements and opinion of Scientific Research Publishing, Inc. We assumes no responsibility or liability for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained herein. We expressly disclaim any implied warranties of merchantability or fitness for a particular purpose. If expert assistance is required, the services of a competent professional person should be sought.

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: ijcns@scirp.org

Int. J. Communications, Network and System Sciences, **2010, 3, 107-115**
doi:10.4236/ijcns.2010.32016 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

The Downlink Adjacent Interference for Low Earth Orbiting (LEO) Search and Rescue Satellites

Shkelzen CAKAJ[1,2], **Mickey FITZMAURICE**[3], **Jesse REICH**[3], **Eric FOSTER**[3]
[1]*Post and Telecommunication of Kosovo (PTK), Dardania, Prishtina, Kosovo*
[2]*Fulbright Scholar Researcher at NOAA, Maryland, USA*
[3]*National Oceanic and Atmospheric Administration (NOAA), NOAA Satellite Operational Facility (NSOF), Maryland, USA*
Email: *Shkelzen.cakaj@ptkonline.com, Shkelzen.cakaj@fulbrightmail.org, Mickey.Fitzmaurice@noaa.gov, Jesse.Reich@noaa.gov, Eric.Foster@noaa.gov*

Received November 19, 2009; *revised December* 20, 2009; *accepted January* 12, 2010

Abstract

NOAA (National Oceanic and Atmospheric Administration) - LEO environmental satellites provide continuous coverage of Earth, supplying high-resolution global meteorological, oceanic and space observation data. In addition, these satellites are part of the international COSPAS – SARSAT program, which aides search and rescue teams worldwide. The USA segment, referred to as SARSAT (Search and Rescue Satellite Aided Tracking) system, is designed to provide distress alert and location data to assist on search and rescue operations. SARSAT locates distress beacons (406MHz) activated at distress locations. The system calculates a location of the distress event using Doppler processing techniques. Processed data is continuously retransmitted through the SARSAT downlink to Local User Terminals (LUT) when satellites are in view. The downlink adjacent interference is expected when two satellites operate in close proximity and share the same frequency. The downlinks of all SARSAT LEO satellites use the same 1544.5 MHz frequency. In cases where the satellites are within the main lobe of the local user terminal antenna, transmissions from adjacent satellites act as interference to one-another, effectively decreasing the signal-to-noise ratio of the desired downlink. This can result in missed distress beacon bursts or no stored solutions received at the LUT, consequently no data is provided about a distress location. Analysis on interference prediction, impacts on system operation and recommendations for mitigating interference periods where the duration may be significant, are presented in this paper.

Keywords: LEO, LUT, Beacon, Doppler, Satellite, Adjacent, Interference

1. Introduction

COSPAS-SARSAT is an international, humanitarian satellite based search and rescue system which operates continuously, detecting and locating transmissions from emergency beacons carried by ships, aircrafts and individuals. This system was originally sponsored by Canada, France, the former Soviet Union and the USA [1,2]. The success of the rescue operation depends crucially on accurate rapid determination of the distress location. The accuracy of location determination and the time required to alert rescue authorities depends on the communication reliability between the LUTs and the satellites [3,4]. The communication link is established when the satellite flies within a LUT's visibility. This 'fly-over' is called a *satellite pass.* Communication reliability during a satellite pass may be degraded when satellites sharing the same downlink frequency are adjacent and interfere to each other, consequently degrading the received signal at the LUT [5,6].

A general overview of COSPAS-SARSAT search and rescue system is briefly presented. Since interference analysis relate to the SARSAT system, the space and ground segment are described in greater detail. Finally, consideration of the LUT antenna gain pattern, satellite path geometry and separation distance between adjacent satellites are described in the prediction of significant periods of interference. Interference mitigation of significant duration, with attached measurement results is also presented.

2. COSPAS-SARSAT System Concept

The basic COSPAS - SARSAT concept is illustrated in Figure 1 [1,2]. The operation of this system is further described.

1) In situations of distress anywhere in the world, when and where lives are at risk, the emergency beacons are activated manually or automatically.

2) Emergency alerts received by the satellites are retransmitted to 45 automated (unstaffed) ground stations worldwide, with several more becoming operational each year [7,8]. These satellite ground stations are called Local User Terminals.

3) Alerts are routed to a Mission Control Center (MCC) in the country that operates LUT. Routed alerts include beacon location computed at the LUT received by one of the system Low - Earth - Orbiting (LEO) satellites.

4) After validation processing (based on Doppler Effect) alerts are relayed depending on beacon location or country of registration to either another MCC or to appropriate Rescue Coordination Center (RCC) [1,2].

3. SARSAT System

The USA portion of COSPAS-SARSAT system is operated by NOAA. NOAA’s environmental satellites carry SARSAT packages. The Mission Control Center (USMCC) is located in Suitland, Maryland [1]. US RCCs are operated by the US Coast Guard and the US Air Force [1,2]. The SARSAT system uses two different types of satellites: polar-orbiting satellites in low Earth orbit (LEO) and satellites in geosynchronous orbit (GEO). (Note: The downlinks for LEO and GEO are orthogonally polarized, LEO-Left Hand Circularly Polarized, GEO-Right Hand Circularly Polarized, providing about 16-20 dB of isolation.) It is shown that the supplemental use of both reduces the time delay and increases coverage area. Only LEOs adjacent interference is discussed [6].

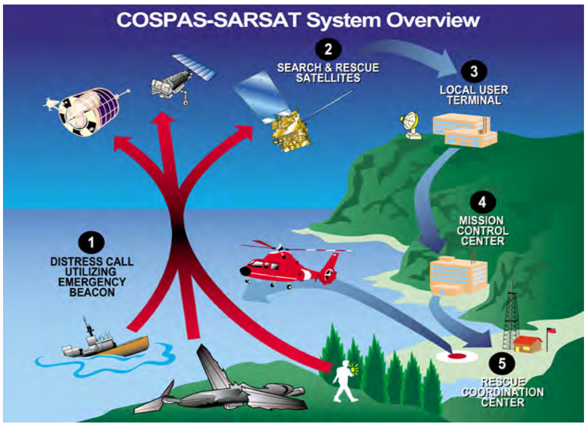

COSPAS-SARSAT System Overview
2
SEARCH & RESCUE SATELLITES
3
LOCAL USER TERMINAL
4
MISSION CONTROL CENTER
1
DISTRESS CALL UTILIZING EMERGENCY BEACON
5
RESCUE COORDINATION CENTER

Figure 1. COPSAS-SARSAT concept.

3.1. Space Segment

The SARSAT satellite constellation is presented in Figure 2 [1]. While GEO satellites continually view large areas of the Earth, and can provide immediate alerting and identification of 406 MHz beacons, there is no Doppler shift of the received beacon carrier since the geostationary satellites are by definition stationary with respect to the Earth, [9–11]. Another issue is that GEOs can not cover the Polar Regions adequately since the antenna footprint is limited to latitudes of about 75º- 80º [12–14].

LEO satellites in polar orbits cover these potential distress regions and allow for Doppler shift processing to be applied in distress alert location determination. The polar-orbiting approach with on-board memory provided by the SARP (Search and Rescue Processor) enables each satellite to provide complete coverage of any point on the Earth twice a day. Polar-orbiting LEO satellites are in Sun synchronized orbits [15].

For search and rescue missions, LEO satellites use two modes of operation:

Repeater mode (Local)–Search and Rescue Repeater (SARR) immediately retransmits received beacon signals to any LUT in the satellite’s footprint. This mode is possible when the spacecraft is visible to both the beacon

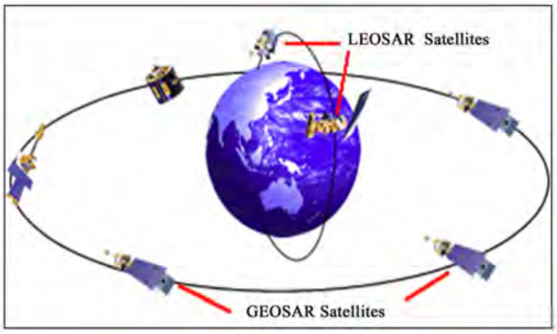


Figure 2. SARSAT system.

Table 1. SARSAT LEOSAR status.

Satellite	Orbital Parameters & Payload Instruments				
	Mean Motion (rev/day)	Altitude (km)	Orbit Period (hr:min:s)	406 MHz	Global
SARSAT - 7	14.2475	809.45	01 : 41 : 04.2	F	F
SARSAT - 8	14.1251	850.91	01 : 41 : 56.7	F	F
SARSAT - 9	14.2405	811.80	01 : 41 : 07.2	F	F
SARSAT-10	14.1125	855.21	01 : 42 : 02.2	F	F
SARSAT-11	14.2149	820.43	01 : 41 : 18.1	F	F
SARSAT-12	14.1095	856.25	01 : 42 : 03.5	F	F

and ground station simultaneously.

Store and forward mode (Global)–is applied when the spacecraft does not see the beacon and ground station simultaneously. The on board Search and Rescue Processor (SARP) receives and records search and rescue beacon transmissions and repeatedly retransmits them as part of a 3 minute continuously cyclical memory dump to LUTs when they are visible as the satellite orbits the Earth. This provides true global coverage [1,2].

Both modes, utilize a satellite downlink carrier frequency of 1544.5 MHz transmit LHCP (Left Hand Circular Polarization) to transmit to any LUT in view. Table 1 presents the orbital parameters and the status of SARSAT LEOSAR payload instruments as of October 2009. *Global* is related to global coverage, and F means Fully Operational [9–11].

3.2. Ground Segment

Receive-only ground stations, specifically designed to track the search and rescue satellites as they pass across the sky are called Local User Terminals. The LUTs are fully automated and completely unmanned at all times [8]. There are two-types of LUTs: Low Earth Orbiting LUTs (LEOLUT) and Geostationary LUTs (GEOLUT). These LUTs and their corresponding MCC (Mission Control Center) to whom these LUTs are interconnected, creates the US SARSAT ground segment. The distress signal is received on the satellite uplink and then it is transmitted to LEOLUTs by downlink. The beacon location is random and LUT locations are fixed and known. The main functions of a LEOLUT are:

- Track the SARSAT satellites
- Recover beacon signals
- Perform error checking
- Perform Doppler processing
- Send alert to Mission Control Center

The LEOLUT system consists of a satellite receive antenna, a digital processing system, an operator display and the software which implements all of the control, monitoring and processing functions. Since LEOLUTs track satellites in low orbits which move quickly relative to a fixed point on Earth, the antenna includes an Antenna Control Unit (ACU) and a tracking mount mechanism with azimuth range of 360º and elevation up to 90º. The appropriate antenna software controls the pointing of the antenna. This ensures that antenna tracks the satellite as it passes over the LEOLUT.

When a satellite receives a beacon signal from a distress location, the Search and Rescue Processor (SARP) on board the LEO satellite performs Doppler processing and generates an entry into the 2.4 kb/s Processed Data Stream (*pds*) that is continuously "dumped" on 3 minute intervals to any LUT in view of the satellite's downlink footprint. LEOLUT software accepts the satellite's down- link data stream, then decodes and extracts beacon data messages. From each satellite pass taken by the LEOLUT, software selects data from each detected beacon and validates time, frequency and message content. Data from each pass, and for each beacon identification number, is then passed to the solution processing software. The solution processing software determines an optimum location based on a Doppler frequency curve. The best curves are used to estimate the beacon location. If the curve cannot be determined, the solution is declared "*unlocated*". Once a signal is processed at the LUT, then the data stream which provides solution and status data is transmitted through a fully automatic communication link to the mission control center (MCC) that operates that particular LUT.

A mission control center (MCC) serves as the hub to collect, store, and sort alert data from other LUTs and other MCCs. The main function of an MCC is to distribute alert data to RCCs and other MCCs. The United States Mission Control Center (USMCC) in Suitland, MD serves as the focal point of the U.S. SARSAT program. NOAA operates 11 LEOLUTs in six locations, as presented in Table 2. These multiple LEOLUTs provide total system redundancy and allow for a maximization of satellite tracking within US Areas of Responsibility (AOR). There are two LEOLUTs in each of the following locations, except for Maryland. Two independent functionally and physically identical systems manufactured by "EMS Technologies" (a Canadian company), are implemented in:

- Miami, Florida (FL1&FL2)
- Vandenberg, California (CA1&CA2)
- Fairbanks, Alaska (AL1&AL2)
- Wahiawa, Hawaii (HI1&HI2)
- Andersen, Guam (GU1&GU2)
- Suitland, Maryland (LEO Support Equipment)

Since each LUT operates independently, they are denoted as 1 and 2. The LEOLUT in Maryland is used as support equipment for tests, software and hardware upgrades and analysis.

4. Downlink Interference

The adjacent satellite interference manifests when two

Table 2. LEOLUTs coordinates.

LEOLUT Locations	Latitude	Longitude
Maryland (MDLUT)	38.85	-76.94
Florida (FLLUT)	25.61	-80.38
California (CALUT)	34.66	-120.55
Alaska (ALLUT)	64.97	-147.51
Hawaii (HILUT)	21.52	-151.99
Guam (GULUT)	13.34	144.56

satellites sharing the same downlink frequency are located close to each other from the perspective of the receiving ground station antenna, as presented in Figure 3 for LUTs of U.S. SARSAT ground segment.

The downlink of all SARSAT LEO satellites uses the same 1544.5 MHz frequency. If the transmitted EIRP from each satellite is similar, for two satellites close to each other, the two signals will act as interference to each other, severely degrading the received signal [16].

Downlink interference between S11 and S9 was documented by France in March and April 2009, when S9 and S11 were close to each other. The 8 March 2009 occurrence of interference between these two satellites caused four passes, over a period of three orbits, which produced no *pds* solutions. The 16 April 2009 occurrence of interference caused three passes with no *pds* solutions over a period of three orbits, presented in Table 3. But, the number of no *pds* solutions alone cannot accurately gauge the amount of interference in the downlink. It is a significant variability in the number *pds* bursts received by the satellite during each orbit depending on the path.

The received carrier frequency provides a useful measure of the interference level. The carrier frequency of the transmitter is 1544.5 MHz, but the relative velocity between the satellite and LUT causes a Doppler shift in the received frequency, and a plot over time shows the characteristic Doppler curve of a LEO satellite. As the orbital positions of the two satellites converge, so do their relative velocities to the LUT and Doppler curves [11].

The implemented software enables prediction of Doppler curves, as presented in Figure 4 for satellites S9 and S11 having interference conflict with each other. Slight differences in relative velocity between the two interfering satellites cause two distinct curves of carrier frequency. When the difference in relative velocity and angular separation is minimal, the Doppler curves of the carrier frequency become almost identical. In Figure 4, the angular separation of satellites viewed from LUT's antenna is presented on the right axis. Figure 5 shows the real time received carrier frequency for Florida-1 and Florida-2 LEOLUTs. Florida-1 is tracking S9 and Florida-2 is tracking S11. When receiver locks on the interfering signal, a jump in the received carrier frequency is seen. These interruptions in carrier lock results in loss of downlink capability and can visually show when interference has occurred. Figure 5 shows that for the most of the pass, each LUT is successfully locked on its desired signal. Two vertical lines show the period of interference. This can result in missed bursts or no solutions received at all. The similarity of predicted and real time recorded curves is obvious.

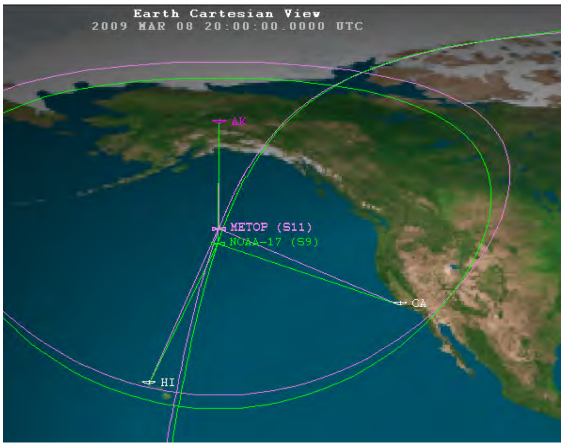


Figure 3. Adjacent satellites seen from the ground station.

Table 3. Passes affected by interference.

Date	DOY	AOS	LOS	LUT	SAT	Orbit	Reason
03.08.09	067	16:31	16:43	LSE	S9	34844	No *pds* solution
03.08.09	067	16:35	16:47	CA2	S9	34844	No *pds* solution
03.08.09	067	18:08	18:18	AK2	S9	34845	No *pds* solution
03.08.09	067	19:56	20:04	CA1	S9	34846	No *pds* solution
04.16.09	106	19:49	19:58	CA2	S9	35401	No *pds* solution
04.16.09	106	21:34	21:44	HI2	S9	35402	No *pds* solution
04.17.09	107	0:58	1:11	GU2	S11	12932	No *pds* solution

4.1. Adjacent Satellites

Visual inspection shows pairs of satellites with similar

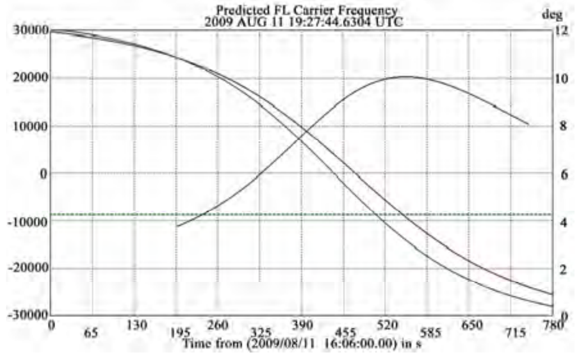


Figure 4. Predicted doppler curves.

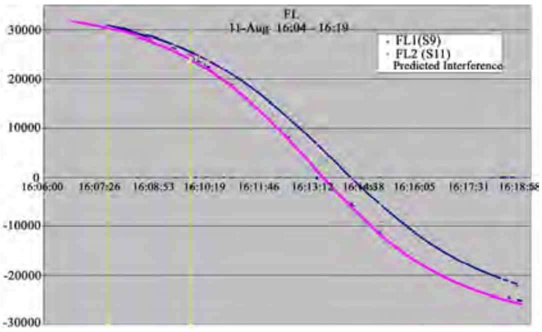


Figure 5. Real time doppler curves.

ground tracks. Kepler elements or the two line orbital elements can be used to analyze the in-track separation. The nearly identical orbital periods and ground tracks will result in long durations where the satellites are in close proximity of each other. In the case where the satellites are in close proximity within the main lobe of the receiving ground station antenna, long periods of interference between the two satellites can manifest itself. During these periods, the downlink of the adjacent satellites may be severely impaired [6]. Three pairs of operational SARSAT satellites are susceptible to this interference condition: S10/S12, S9/S11, and S7/S8 are identified and presented in Table 4 with their respective orbital periods and differences between them.

The small difference in orbital periods of the S10/S12 pair is particularly concerning. The *Orbit repeat cycle* indicates the number of orbits that satellite should pass through to achieve the same position relative to the adjacent satellite and to the fixed ground station. Mathematically, *Orbit repeat cycle* is the ratio of orbit period and orbital difference, as calculated and presented in Table 4. Further, for this cycle to be expressed in days, it should be divided by the mean motion from Table 1. The USA documented that the launch of S12 (NOAA-19) into an orbital plane similar to S10 (NOAA-18), and with nearly identical orbital periods, created long periods of adjacent interference. The first period of extended interference occurred from 15 September 2009 to 20 September 2009.

Table 4. SARSAT adjacent satellites.

Satellite	Orbit Period	Difference	Orbit Repeat Cycle	Repeat Cycle (days)
SARSAT12	01:42:03.53	00:00:01.30	4710	334
SARSAT10	01:42:02.23			
SARSAT11	01:41:18.10	00:00:10.92	556	39
SARSAT 9	01:41:07.18			
SARSAT 8	01:41:56.75	00:00:52.55	116	8
SARSAT 7	01:41:04.20			

Table 5. SARSAT adjacent satellites.

Satellite	Slant range (at 9.25°) (km)	Slant range (at 5°) (km)	Separation Distance (km)
SARSAT12	2544.5		416.4
SARSAT10		2903.8	
SARSAT11	2470.3		399.1
SARSAT 9		2812.9	
SARSAT 8	2534.4		341.4
SARSAT 7		2806.5	

Table 6. SARSAT adjacent satellites.

Satellite	Velocity (km/s)	Duratio n (s)	Interference Repeat Cycle (#Orbits)	Repeat Cycle (#days)
SARSAT12	7.423	56.3	43.3	3.10
SARSAT10	7.423			
SARSAT11	7.446	53.9	4.9	0.35
SARSAT 9	7.441			
SARSAT 8	7.447	46.1	0.9	0.06
SARSAT 7	7.426			

4.2. Duration of Interference

To determine the duration of the interference periods, one must find the minimal angular separation between satellites as seen from the ground station, when interference occurs. This is highly dependent on the gain pattern and pointing accuracy of the LUT antenna. For a typical LEO-LUT antenna gain pattern, the -3dB (half power) beamwidth is found to be ± 4.25°. This beamwidth represents the necessary angular separation to prevent undesired signals from being highly amplified. As the angular separation increases, the gain of the interfering source decreases. Since the distance between the two satellites is relatively constant during a singular pass, it can be seen that the apparent angular separation is greatest when the satellites are at their maximum elevation (closest approach) [17]. Thus, minimum angular separation occurs when the satellites are at minimum elevation. Thus, the cases with low elevation are of interest from the interference aspect.

Let us consider a LUT with antenna aperture of ± 4.25. This antenna is tracking a satellite, which is moving ahead relative to another satellite which is seen at minimum elevation above the horizon (5°), as shown in Figure 6.

These adjacent satellites, seen at low elevation, and with a very low separation angle, have great potential to interfere each other. The slant range is calculated for elevations of 9.25° and 5° (5° horizon with 4.25° separa-

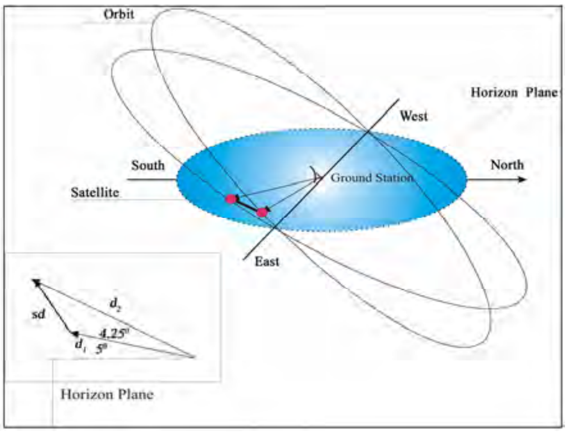


Figure 6. Adjacent satellites under beamwidth angle.

tion) from a ground station. Spatially the separation angle is the spherical angle from 0° to 4.25°. The 0° point is on the desired satellite, and 4.25° point is the -3dB interference point, consequently it is the largest possible distance for interference from another satellite. The general formula for the slant range (d) under elevation ε_0 is [12–14]:

$$d = R_e\left[\sqrt{\left(\frac{H+R_e}{R_e}\right)^2 - \cos^2\varepsilon_0} - \sin\varepsilon_0\right] \quad (1)$$

where, $R_e = 6378$ km is Earth radius and H is orbital altitude. The separation distance (sd) can then be determined using a small angle approximation and applying cosines theorem, as:

$$sd = \sqrt{d_1^2 + d_2^2 - 2d_1d_2\cos 4.25^\circ} \quad (2)$$

where d_2 is the slant range of pointed satellite from the ground station and d_1 is the slant range of the adjacent satellite potential to interfere. Altitude H of each satellite is in Table 1. The slant ranges and separation distances are presented in Table 5. For more exact calculations these separation distances should be multiplied by cosines of separation angle (projection of separation distance in its own orbit), which for too low angles can be considered as 1. This separation distance when interference may occur, and the difference in orbital periods can then be used to find the duration of possible interference. Considering that these satellites are always moving with a particular velocity v, the question is how long they can be together within a separation angle of 4.25°. This represents the ***Duration*** of possible interference. The frequency of these events and their duration relative to the fixed ground stations depend on the difference of orbital period times (Table 4). The ratio of interference time duration to time difference in orbital periods represents *Interference repeat cycle per orbit.* This cycle is expressed in days when divided by mean motion (Table 1). Considering separation distance, predictions for the interference repeat cycle of satellite pairs are listed in Table 6.

In general, the difference in orbital period between the two satellites will dictate the duration and repeatability of interference intervals. From Table 6 it is obvious that the S10/S12 pair experiences the highest interference repeat cycle, consequently the longest possible interference disturbance, because of too close orbital periods. The S7/S8 pair is the least experiences with interference.

5. Downlink Interference Mitigation

As another approach, and for results comparison, a satellite orbit analysis program using the known LUT antenna gain pattern is applied. Again, when the both satellites are within -3dB beamwidth (separation angle of 4.25°) from the point of view of the LUT, it was determined that interference is possible.

Table 7. Timeline of significant future interference events.

Satellite pair	Start of interference	End of interference	Duration (days-hh:mm:ss)
S9/S11	5.25.09, 18:06:59	5.26.09, 02:26:06	0 – 08:19:07
S9/S11	7.03.09, 13:11:29	7.03.09, 19:37:15	0 – 06:25:46
S9/S11	8.11.09, 01:29:55	8.11.09, 07:41:06	0 – 06:11:11
S9/S11	9.18.09, 09:20:34	9.18.09, 14:40:31	0 – 05:19:57
S10/S12	9.20.09, 10:49:37	9.23.09, 19:54:12	3 – 09:04:35
S9/S11	10.26.09, 10:28:57	10.26.09, 16:34:48	0 – 06:05:51
S9/S11	12.03.09, 07:23:15	12.03.09, 12:37:25	0 – 05:14:10

5.1. Timelines and Events

Considering events on March and April from Table 3, a period from May to December is analyzed. The beginning of the period of possible interference was designated as the first pass at a USA LUT where S10 and S12 would be within 4.25° of each other at any point during the pass. The predicted periods of interference were generated by a satellite orbital analysis program. Table 7 shows the timeline of these significant events.

The duration of the S10/S12 interference is of particular concern, and has been verified through this secondary method to be about 3 days (See Table 6 and Table 7, approximately the same results from mathematical analysis and simulation results). The duration of S9/S11 interference periods decreases as predictions are made farther in the future. Periods of S8/S7 interference are approximately one orbit in duration repeating every 8 days, and therefore, are not listed.

Thus, considering antenna pattern and satellite pass geometry, analytical models can be built to predict the time and duration of interference based on the angular separation between the two satellites [18]. The interference mitigation efforts should be performed if operational impacts become severe. Mitigation efforts must be relatively benign. Canada developed a procedure to interrupt RF transmission from the satellite with a minimal chance of irrecoverable failure. The USA executed this procedure when the operational impacts of interference became evident. The USA analyzed the downlink characteristics during the periods both before and after the mitigation actions were taken. This process is further described.

5.2. Interference Records

Satellite pair S10/S12, as the worst case of adjacent interference is further analyzed. The Canadian procedure to interrupt the downlink RF transmission from the satellite is considered to be applied as a method to mitigate adjacent satellite interference. The turnoff transmission was planned for S10. Further plots presented in Figure 7, Figure 8, Figure 9 and Figure 10 show various passes, in chronological order, before the turnoff of the downlink of S10. The received carrier frequency is shown on the left axis, and modulation index mean and RMS (Route Mean Square) on the right axis. Modulation index indicates the quantity by how much the modulated variable varies around its unmodulated level. Considering downlink phase modulation, this index relates to the variations in the phase of the carrier signal.

Figure 7 shows the case with relatively high maximal elevation of 76º and Figure 8, the case with medium maximal elevation of 36º. Figure 7 shows the interference during AOS (Acquisition of Satellite), as bit and frame sync is being established. Frequency jumps in the downlink carrier can be seen in the upper left corner of Figure 7, and are reflected by a high mean modulation index at the same time. Further, as the satellite moves toward higher elevation there is no interference (medium part of figure) and then again there is interference near LOS (Loss of Satellite). In Figure 8 it is very expressive modulation index and frequency jump during the loss of satellite. Figure 9 and Figure 10; show the cases with low maximal elevation, respectively of 12º and 9º respectively. The jump in frequency is present in both, particularly under 9º.

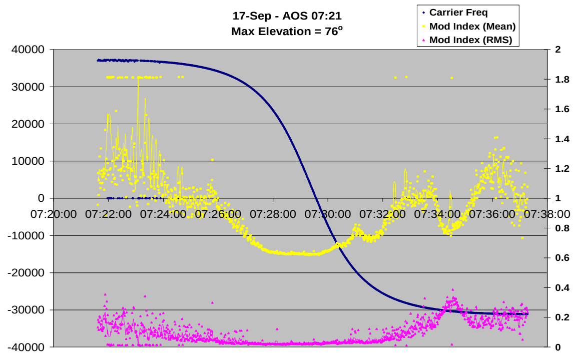


Figure 7. Doppler curve for maximal elevation of 76 º.

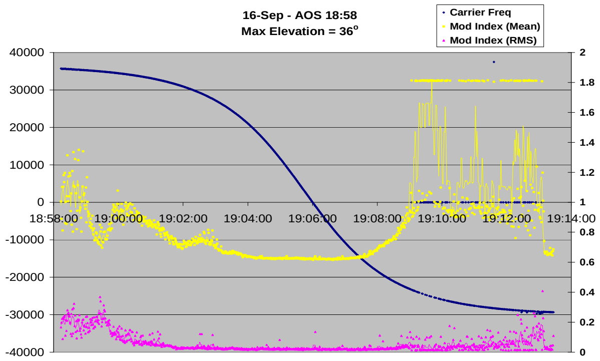


Figure 8. Doppler curve for maximal elevation of 36 º.

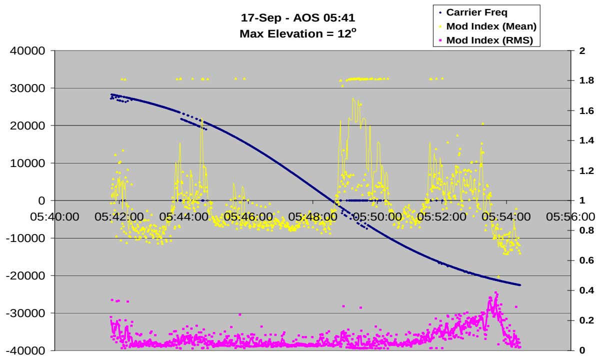


Figure 9. Doppler curve for maximal elevation of 12 º.

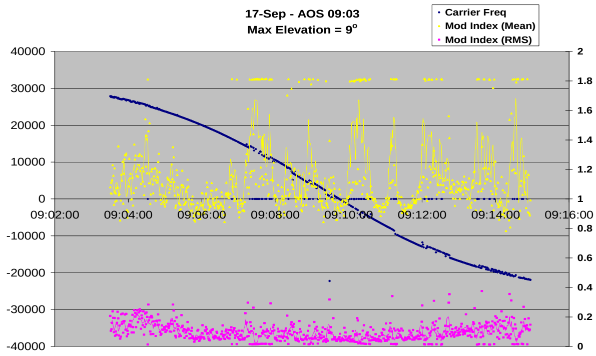


Figure 10. Doppler curve for maximal elevation of 9º.

In all figures before the turnoff of the S10 downlink, the received carrier frequency can be seen jumping from one satellite's downlink to the other one, causing the degradation of downlink capabilities. The modulation indices are higher during these times since the receiver cannot lock on only one carrier. The modulation indices are typically lower (less interference) during the middle of the pass when the apparent separation of the satellites is greatest. The total magnitude of interference is greater for low elevation passes, and it becomes even greater as the peak period of interference approaches (from one pass to the next).

Figure 11 and Figure 12 show the same plots after the downlink of S10 had been turned off. They show that the only increase of the modulation indices occurs near AOS and LOS, when the signal is the weakest. Figure 11 and Figure 12 are typical of what you would see during a nominal pass with no interference. The procedures developed by Canada and executed by the USA were successful in interference mitigation. The worst of the inter ference was completely mitigated through the coordinated efforts of the ground and space segment providers.

6. Conclusions

U.S. SARSAT is data communication system dedicated to search and rescue purposes oriented on determination of distress locations worldwide, thus the performance of

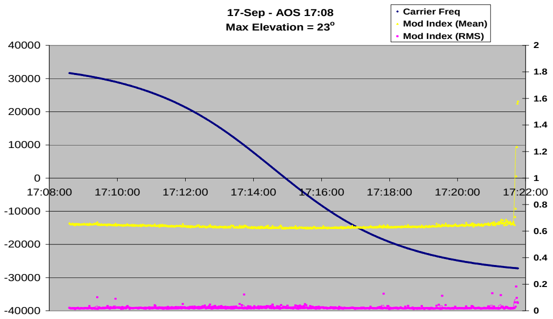


Figure 11. No interference doppler curve for Max. El. 23°.

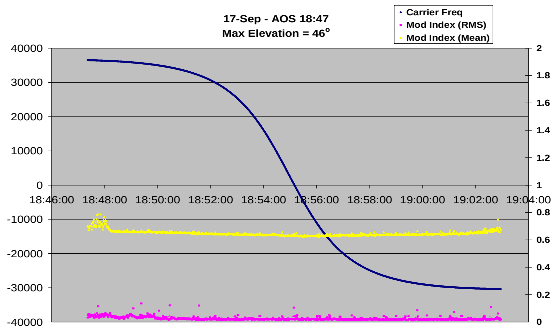


Figure 12. No interference doppler curve for Max. El. 48°.

the ground station is of high importance to this process. It is confirmed that adjacent SARSAT satellites with short differences in orbital period interfere with each other. During these interference periods, significant degradation of downlink occurs.

The procedure to interrupt the downlink RF transmission from the "undesired" satellite is applied as a method to mitigate adjacent satellite interference. It has been confirmed that the interference was mitigated using this method. For newly built terminals though, larger antennas with a narrower beamwidth may also reduce the adjacent interference issue and impacts.

The DASS (Distress Alert Satellite System) is a newly developed & future approach intended to enhance the international COSPAS-SARSAT program. In this effort the satellite-aided search and rescue (SAR) system will install 406 MHz SAR instruments on the Medium Earth Orbit (MEO) navigational satellites [GPS (US), Galileo (EU), and Glonass (Russian Federation)]. With an expected 80 satellites expected once fully operational, new processing algorithms and interference mitigation strategies should also be considered. Because of the much higher altitudes of MEO satellites, a larger separation distance exists, and the adjacent interference will be less pronounced. This is just one more significant factor in favour of DASS approach.

7. References

[1] http://www.sarsat.noaa.gov/

[2] http://www.cospas-sarsat.org/

[3] D. Ludwig, R. Wallace, and Y. Kaminsky, "Proposed new concept for an advanced search and rescue satellite system," IAF, International Astronautical Congress, 36th, Stockholm, Sweden, pp. 18, October 1985.

[4] I. W. Taylor and M. O. Vigneault, "A neural network application to search and rescue satellite aided tracking (SARSAT)," in Proceedings of the Symposium/ Workshop on Applications of Experts Systems in DND, pp. 189–201, Royal Military College of Canada, 1992.

[5] S. Cakaj and K. Malaric, "Rigorous analysis on performance of LEO satellite ground station in urban environment," International Journal of Satellite Communications and Networking, UK, Vol. 25, No. 6, pp. 619–643, November/December 2007.

[6] F. Vataralo, G. Emanuele, C. Caini and C. Ferrarelli, "Analysis of LEO, MEO and GEO global mobile satellite systems in the presence of interference and fading," IEEE Journal on Selected Areas in Communications, Vol. 13, No. 2, pp. 291–299, February 1995.

[7] J. S. Landis and J. E. Mulldolland, "Low cost satellite ground control facility design," IEEE, Aerospace & Electronic Systems, Vol. 2, No. 6, pp. 35–49, 1993.

[8] L. Losik, "Final report for a low–cost autonomous, unmanned ground station operations concept and network design for EUVE and other NASA Earth orbiting satellites," Technology Innovation Series, Publication 666, Center for EUVE Astrophysics, University of California, Berkeley, California, July 1995.

[9] "Specification for COSPAS" – SARSAT406MHz Distress Beacons, C/T T.001, No. 3 – Revision 9, October 2008.

[10] COSPAS –SARSAT System Monitoring and Reporting, C/S A.003, No. 1, Revision 15, October 2008.

[11] COSPAS – SARSAT 406MHz Frequency Management Plan, C/T T.012, No. 1 – Revision 5, Probability of Successful Doppler Processing and LEOSAR System Capacity, October 2008.

[12] G. Maral and M. Bousquet, "Satellite communication systems," John Willey & Sons, Ltd, Chichester, England, 2002.

[13] D. Roddy, "Satellite communications" McGraw Hill, New York, 2006.

[14] M. Richharia, "Satellite communication systems" McGraw Hill, New York, 1999.

[15] Sh. Cakaj, M. Fischer, and A. L. Schotlz, "Sun synchronization of Low Earth Orbits (LEO) through inclination angle," in Proceedings of 28th IASTED International Conference on Modelling, Identification and Control, MIC 2009, Innsbruck, Austria, pp. 155–161, Feruary16–18, 2009.

[16] J. D. Kanellopoulos, T. D. Kritikos, and A. D. Panagopoulos, "Adjacent satellite interference effects on the outage performance of a dual polarized triple site diversity scheme", IEEE Transaction on Antennas, Vol. 55, Issue 7, pp. 2043–2055, July 2007.

[17] S. Cakaj, "Practical horizon plane and communication

duration for Low Earth Orbiting (LEO) satellite ground stations", WSEAS Journal Transactions on Communications, Vol. 8, No. 4, pp. 373–383, April 2009.

[18] S. Cakaj, K. Malaric, and A. L. Schotlz, "Modelling of interference caused by uplink signal for Low Earth Orbiting satellite ground stations," in Proceedings of 17th IASTED International Conference on Applied Simulation and Modelling, ASM 2008, Corfu, Greece, pp. 187–191, June 23–25, 2008.

Int. J. Communications, Network and System Sciences, 2010, 3, 116-118
doi:10.4236/ijcns.2010.32017 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Design and Measurements of Ultra-Wideband Antenna

Giorgos TATSIS, Vasilis RAPTIS, Panos KOSTARAKIS
Physics Department, University of Ioannina, Ioannina, Greece
Email: *gtatsis@grads.uoi.gr, vraptis@grads.uoi.gr, kostarakis@uoi.gr*

Received October 17, 2009; *revised November* 21, 2009; *accepted November* 29, 2009

Abstract

This paper describes the design, realization and experimental measurements of an antenna element to operate at ultra-wideband (UWB) spectrum. The type of this antenna is a circular disk monopole (CDM), with two notches opposite to each other at two sides of the disk. The feed of the antenna is a coplanar waveguide (CPW). The effect of the presence of the notches is studied through simulations and tested experimentally.

Keywords: UWB, Antenna, CPW, CDM

1. Introduction

In every wireless telecommunication system, antennas are some of the most essential elements that characterize the system. Antennas are responsible for effective propagation of electromagnetic energy from transmitter to receiver through the wireless channel. Design and implementation of antennas varies from system to system, depending on the special characteristics of the antennas, such as spectrum occupancy, transmitting power level, directionality, etc. The optimum design of an antenna, to meet those characteristics, has become a great challenge, especially in new technologies like Ultra-Wideband (UWB) technology. In an UWB system a very large spectrum bandwidth must be used. By definition [1], a transmitting signal is considered as UWB if the absolute bandwidth spectrum, at -10dB, exceeds 500MHz or the fractio nal bandwidth is greater than 20%. The Federal Communications Commission (FCC), released a spectrum from 3.1GHz to 10.6GHz, to be used in such a system, with the maximum power density of -41.3 dBm/M-Hz. This is the area of interest that concerns the UWB antennas. Important efforts have been made [2–5], in the design of such antennas. In this paper a modified coplanar feed circular disk monopole (CPW-CDM), is studied and constructed. At the sides of the printed disk two symmetrical notches are placed. It is shown that this modification affects the performance of the antenna significantly. Measurements of the constructed antennas and simulations are presented.

2. Description

Figure 1, shows the geometry of the antenna. A single metallic layer is used upon a substrate of dielectric. This type of antennas is very easy to be constructed, low cost and suitable for portability due to its small size. It consists of a coplanar waveguide in which a sma connector is attached, a circular disk, truncated symmetrical with two notches, and a ground plane, all printed at the same layer. The coplanar waveguide is designed to have 50 Ohm impedance to match the coaxial cable impedance. This impedance is controlled by the two parameters W and G, the width of the feed line and the gap between line and ground plane respectively. The most critical parameters, affecting the performance of this antenna [2], are the disk radius R, the distance H between the disk and the ground plane and the width X of the ground plane. The aforementioned parameters of the constructed antenna are given in Table 1. The modification of the known CPW-CDM antenna that took place in our design is the truncation of the circular disk as shown in Figure 1 by two symmetrical notches. The depth of the two notches is controlled by the variable A.

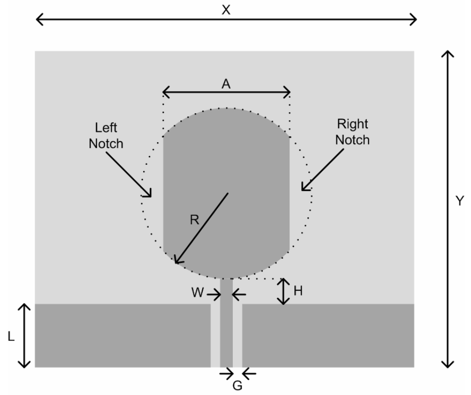


Figure 1. UWB antenna schemat.

3. Fabrication and Results

In order to test the performance of the antenna and to study the effect of the notches four antennas were fabricated using printed circuit techniques, one without notches and three with variable notch depth. The disk diameter was 36mm and the variable A took the values, 32mm, 28mm and 24mm. The dielectric substrate is an FR4 epoxy with relative permittivity of 4.6 and 1.5mm thickness. We used these values to calculate the width and gap of the coplanar feed to match the 50 Ohm impedance of the coaxial cable. The calculated impedance of the transmission line using the values of Table 1, is found to be 50.4 Ohm. For the measurement tests a network analyzer was used and we evaluate the performance of the antennas by measuring the reflection coefficient S11. The results are shown in Figure 2. One can see that the presence of the notches is affecting the reflection coefficient of the antenna. S11 increases at lower frequencies, between 2-5GHz but the peak around 5GHz decreases. One could use that effect to exploit a less frequency selective region, for example between 2-6GHz for the rhomb marked curve. For better distinction of this effect we depict the initial antenna's S11 and the one's with A=28mm in Figure 3. In some cases as for Ultra Wideband Impulse Radio less frequency selectivity introduce less pulse distortion which is very significant in this technology. Furthermore a series of simulation were done to have a comparison with the experimental results. The simulation software used, manages frequency domain calculations of the electromagnetic equations utilizing the method of moments. The simulated return loss is shown in Figure 4. We can notice that the results are fairly consistent with the measurements. In addition we have the opportunity to investigate the whole spectrum, released by the FCC (3.1GHz-10.6GHz). The spectrum from 9GHz to 11GHz was not measured experimentally since the maximum frequency of the network analyzer used was 9GHz. By checking both results it is reasonable to assume that the antenna in fact doesn't exceed the -10dB limit in the region between 9GHz-10.6GHz.

Table 1. Antenna parameters.

Parameter	Value (mm)
X	60
Y	50
R	18
W	2
G	0.35
H	0.4
L	10

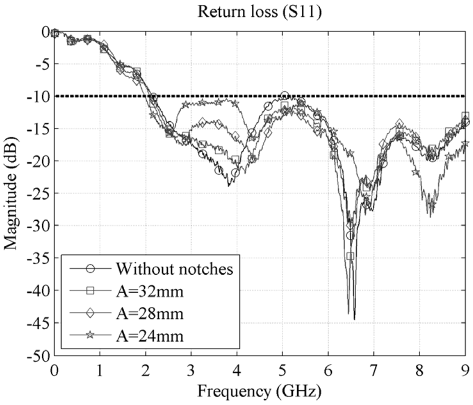


Figure 2. Measured return loss of antennas.

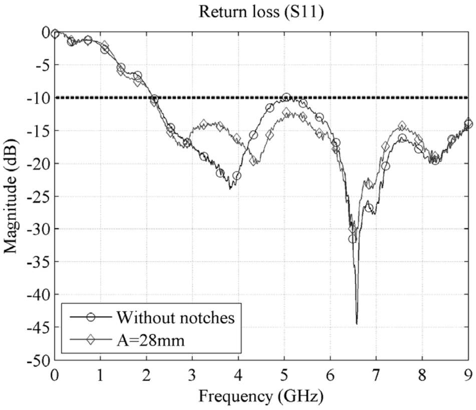


Figure 3. Return loss comparison of the antenna without notches and the one with A=28mm.

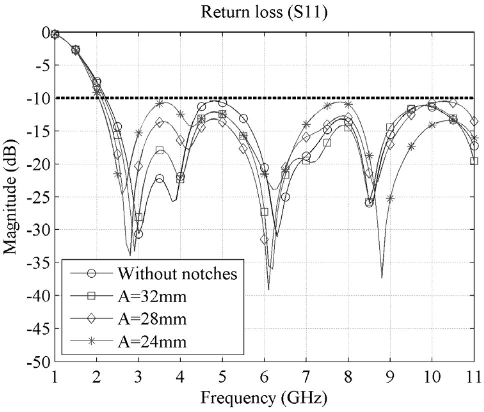


Figure 4. Simulated return loss of antennas.

4. Conclusions

A modified coplanar waveguide circular disk monopole antenna, for ultra wideband applications was described in this article. The effect of two symmetrical notches of the circular disk was under investigation. The impact of the truncated monopole was calculated with simulations using the method of moments and compared with the measured results. Measurements results show that the antenna has an ultra-wideband performance with reflection coefficient less than-10dB in the range 2.2-9Ghz. Simulations have shown a good agreement with measurements and an acceptable S11 performance below -10dB in range

2.2-11GHz. From this study we concluded that the modified antenna leads to less signal distortion of ultra wideband pulses.

5. Acknowledgment

This research project (PENED) is co-financed by E.U.-European Social Fund (80%) and the Greek Ministry of Development-GSRT (20%).

6. References

[1] L. Q. Yang and G. B. Giannakis, "Ultra-wideband communications," IEEE Signal Processing Magazine, November 2004.

[2] J. X. Liang, L. Guo, C. C. Chiau, and X. D. Chen, "CPW-fed circular disc monopole antenna for UWB applications," IEEE 2005, pp. 505–508.

[3] Z. N. Low, J. H. Cheong, and C. L. Law, "Low-cost PCB antenna for UWB applications," IEEE Antennas and Wireless Propagation Letters, Vol. 4, 2005.

[4] Y. Kim and D. H. Kwon, "CPW-fed planar ultra wideband antenna having a frequency band notch function," Electronics Letters, Vol. 40, April 2004.

[5] J. Jung, W. Y. Choi, and J. Choi, "A compact broadband antenna with an L-shaped notch", IEICE Transactions on Communications, Vol. E89-B, No. 6, June 2006.

Int. J. Communications, Network and System Sciences, **2010, 3, 119-125**
doi:10.4236/ijcns.2010.32018 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

A Topology-Aware Relay Lookup Scheme for P2P VoIP System

Xiuwu ZHANG[1], Ruifeng GUO[1,2], Weimin LEI[1,2], Wei ZHANG[1]
[1]*School of Computer Science and Technology, University of Science and Technology of China, Hefei, China*
[2]*Shenyang Institute of Computer Technology, Shenyang, China*
Email: *zhangxiuwu@sict.ac.cn*, grf@sict.ac.cn, *leiwm@sict.ac.cn, zhangwei@sict.ac.cn*
Received November 9, 2009; *revised December* 29, 2009; *accepted January* 6, 2010

Abstract

Because of the best-effort service in Internet, direct routing path of Internet may not always meet the VoIP quality requirements. Thus, many researches proposed Peer-to-Peer VoIP systems such as SIP+P2P system, which uses relay node to relay RTP stream from the source node to the destination node and uses application-layer routing scheme to lookup the best relay nodes. The key of those systems is how to lookup the appropriate relay nodes, which we call relay lookup problem. This paper presents a novel peer relay lookup scheme based on SIP+P2P system. The main ideas are to organize the P2P network using a Cluster overlay and to use topology-aware to optimize relay selection. We introduce the mechanism in detail, and then evaluate this mechanism in NS2 network simulation environment. The results show that our scheme is scalable and can get high relay hit ratio, which confirm the feasibility of a real system. We also make comparison with traditional schemes and the results show that our scheme has good path quality.

Keywords: Relay Lookup, Cluster, QoS, Peer-to-Peer VoIP

1. Introduction

Unfortunately, the current Internet provides a Best-effort service, which cannot guarantee VoIP service qualities on the IP-layer sometimes. That is, sometimes direct routing path in the Internet may not always provide good path quality. For the good quality of service, it is important to select alternate paths between the source node and the destination node in the application layer. Thus, it has been popular to build overlay routing for VoIP application based on Peer-to-Peer (P2P) networks in the Internet [1–3].

Many previous research works [2,4,5] show that overlay routing can effectively improve VoIP session's performance by avoiding IP-layer's path failure. In application-layer routing scheme for VoIP systems, we should first lookup appropriate relay nodes in the application layer to bring an intermediate routing scheme. We call it the relay lookup problem. The definition of relay node is one which can not only give relay services for local users, but also serving as a bridge for remote nodes in the topology. The role of relay node is not only to support two/-multiple node's communication (traverse a NAT or firewall), but also to avoid congestion or failure in IP-layer to improve the quality of communication capacity. Many researches [2,6] show that relaying RTP packet is an effective way to improve the quality of VoIP communication.

There are many application-layer routing projects which propose relay lookup scheme, such as MIT's resilient overlay network (RON) [4,5], Ohio state University's AS-Aware Peer-Relay Protocol (ASAP) [1], which all are to solve the Internet path failure through one-hop overlay routing. Besides these systems, there is one commercial software which is Skype [3,7,8]. As we know, Skype successfully use P2P technology and supports millions of online users, and the most important technique is the relay mechanism. Skype profits from the relay mechanism, but the fact is that, Skype encrypts its packets so that the exact routing methods are unknown. These defects hinder its further growth.

In our initial work, SIP+P2P system [2], we propose a P2P relay mechanism, which is to make those nodes with good network situation and machine performance become relay nodes, and to use a P2P overlay to organize those nodes. In the original system, we use a simple relay lookup scheme which we call one-side optimal selection

An early work of this paper will be presented in part at the 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2009, September, Beijing, China.

scheme. In one-side optimal selection, client nodes cache a list of relay nodes and keep heartbeat detecting. When need relays, the caller client node selects several nearest relay nodes in cache by means of RTT and negotiates with peer client to find proper relays. But this mechanism is a unilateral optimal method, not suitable for bilateral optimal requirement in VoIP communication.

Thus, how to design an effective scheme in SIP+P2P system to find the appropriate nodes is really a field worth researching. In this paper, we propose a novel peer relay lookup scheme based on cluster participation, which can be used in SIP+P2P system. The main idea is to use topology-aware to optimize the peer relay selection. We divide the whole P2P relay network into many clusters, and organize the whole P2P network using Cluster overlay. When one node joins the system, it will automatically upgrade its identity to be a landmark node or just be an ordinary node according to its own network situation and relative location. If it is upgraded to be a landmark node, it will be in charge of building clusters in the overlay network, and other ordinary nodes will join in the cluster. The landmark node is the responsible one of the cluster, and it will record the relay nodes in its cluster which have the ability to relay with other clusters. When the node in the cluster needs relay nodes during its communication, the landmark will give out a relay node candidate list, and the communication node chooses the right relay node from the list after testing and negotiating with peer node. In the aspect of cluster topology constructing, we use path quality as the measurement metric, which is quite different from ASAP. ASAP use AS information from BGP routing table to build cluster topology, which is unrealistic in large scale network and mobile network. Instead, our measurement metric of path quality also reflects the proximity relationship in topology distance, and is easy to realize and suitable for large scale network.

2. Related Work

The resilient overlay network (RON) [4,5] is a solution for general application-layer routing problems, which aims to provide optimized packet routing performance through routing packets in an overlay network. Using link state routing protocol, RON server detects all the links in regular cycle, thus it gets the update link state of the whole network, including the delay between any paired nodes. When there is a failure in the direct path, the nodes ask the server to find relay nodes. The server will check all the links around the source node and the destination node, and then return a proper node list which is near to both nodes. RON is only efficient for small network application in which the nodes number is below 50. Besides, RON is not specifically designed for VoIP system, thus, is not suitable for relay service in VoIP system.

Many researchers attempt to gain insights into the Skype relay selection algorithm in a black-box manner [3,7, 8]. They capture the packets to study the protocol in Skype and analyze the key techniques according to these traces. From the perspective of specific technologies, Skype's application layer routing has several advantages, including quick routing, lower operating costs and network overhead. Then the system can greatly reduce the burden of the central server. Nevertheless, on the one hand, the relay selections do not take topology into consideration, so it can not always find proper relays. On the other hand, due to its private signal and encrypting packet transmission, we still do not know the detailed design of the system up to now.

Our previous work, SIP+P2P system [2], proposed a P2P relay mechanism to improve the VoIP quality. We select nodes which have good network situation and machine performance to become relays and provide relay service for VoIP sessions. Then we use a P2P overlay to organize those nodes. In our original system, we used a simple peer relay lookup scheme: all client nodes cache a list of relay nodes and keep alive; when a node needs relays to communicate, it will select several nearest relays and negotiate with the peer node; after detecting and negotiating, the two nodes find appropriate relay nodes for their communication. In order to optimize the selection procedure, client node evaluates relay nodes in its cache by means of *ping*/*pong* RTT value, and sorts them according to the evaluation result. We also classify all relay nodes into different ISP sets to enhance path diversity and reduce the cache space. For all that, this mechanism can not always find the near relays for both sides yet, because it is a unilateral optimal selection. But VoIP systems often require that the selected relays are near to both the source node and the destination node. Thus this relay lookup mechanism needs to be further improved.

ASAP (AS-Aware Peer-Relay Protocol) [1] is specially designed for VoIP systems and as it experienced, they got a good result for VoIP communication nodes. The main idea is to divide the whole network into AS domains. According to IP prefix, the system classifies all nodes that have joined the system into different ASs, which are called clusters. There are Bootstrap servers in the system in charge of the IP prefix of clusters and the overall AS graph. There are responsible nodes, also called proxy nodes, in every cluster. These proxy nodes are chosen from the clusters with better network state. They get the location of AS and the IP prefix table from the Bootstrap server, and then make a record of nodes in the cluster. When nodes join the system, they can find their cluster nodes based on its own IP, and let the cluster record. Also proxy node will detect the surround AS and record the near ones. When two communicating nodes need relays, they will ask proxy nodes for the near clusters' information and the two clusters' information will make an intersection and the nodes in the set will be the relay candidates. The two nodes will then detect the relay candidates to find the appropriate relays. The key of this system is to find the IP prefix to give nodes basic judgment to join the right cluster. But getting the IP prefix from the BGP routing table is not easy to bring out for large scale network, and the real-time update one is even more difficult and needs lots of work, so the actual deployment of system in large scale network is unrealistic. But the system itself is great reference for relay lookup mechanism.

As introduced above, both Skype and SIP+P2P system have some limits: 1) their peer selection is suboptimal; 2) there are a large number of unnecessary probes. Those systems can be further improved by using topology-aware technology. ASAP uses AS-Aware to optimize relay selection, but it still has some limitations. Thus, we propose a novel peer relay lookup based on SIP+P2P system, which still uses a Peer-to-Peer relay mechanism. We organize the P2P relay network using Cluster overlay, and use topology-aware to optimize relay lookup. When build cluster overlay and select relay, we use path quality as its measurement metric, which is easier to realize in large scale network.

3. Design of Peer Relay Lookup Scheme

Our design still confirms to the framework of SIP+P2P system, in which the SIP signal is client/server based, but the media stream is transmitted in P2P relay network. We still adopt one-hop relay scheme, and use topology-aware technology to help selecting the proper relays. We divide the whole relay network into different clusters according to the path quality, which denotes the distance between two nodes. Thus, nodes which have good quality path (e.g. MOS>3.5) to each other are near in topology and will join in the same cluster. We use delay and packet loss to evaluate the path quality, because it is generally thought that the voice quality mainly relates to the delay and the packet loss among all network factors. When the one-way delay is lower than 200ms and the packet loss less than 2%, the voice quality is usually acceptable. In this cluster system, there are three kinds of nodes:

- Landmark node: Landmarks are chosen from ordinary nodes when they have good hardware condition and network surrounding. The landmark is the management center of the cluster.
- Relay node: Relay nodes provide relay service for other nodes. They usually have good network situation. When a node becomes a relay, it will record its address on landmark.
- Ordinary node: Ordinary nodes will join the cluster by being recorded on the landmark and maintain the states with periodic heartbeat method.

Relay nodes and ordinary nodes are all cluster nodes. Sometimes, a landmark node may also be a relay node. When communication nodes need relaying, they will ask the recording landmark for proper relay nodes candidates, and then detect the relay to see if it can be the relay node.

3.1. System Architecture

As shown in Figure 1 and Figure 2, the system is composed of some Bootstrap Servers, a cluster layer and a logical relay layer. The Bootstrap Server is only server-like component in the system which needs to be deployed previously. It is responsible for node login, landmark management, and maintains a bootstrap's data structure, including: a cluster graph, a cluster landmark IP table. Cluster layer is used to organize nodes and maintain the real topology. Cluster layer is formed with several clusters. In each cluster, nodes are tagged as landmarks or cluster nodes. Each cluster node is attached to only one landmark. In normal cases, the quality of path between a node and its landmark is very good. The landmark node maintains its near clusters to form the topology, as show in Figure 1. Relay logical layer is used to help nodes in cluster to find appropriate relay nodes. Relay layer is a virtual layer with all nodes mapped from cluster layer; the connection relationship is kept the same as the one in cluster layer, simply adding some relay related information on each node to help find relay.

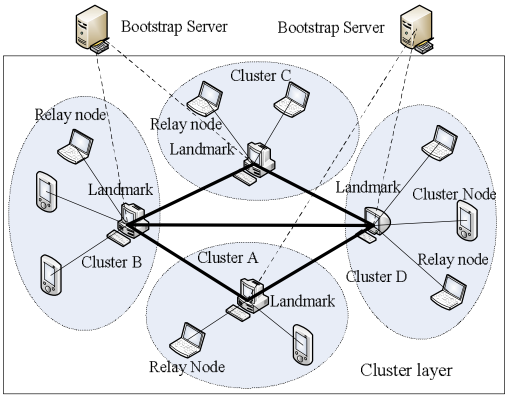


Figure 1. Deployment of system architecture.

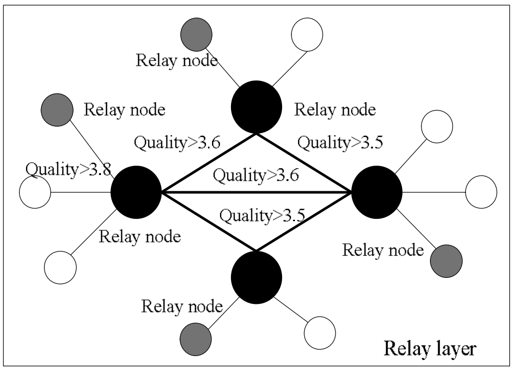


Figure 2. A relay logical layer mapped from the cluster-layer.

3.2. Module Design on the Peer

Figure 3 shows the module design on one peer. There are two main modules, including cluster management module and relay module. There are two parts in cluster management module: the landmark logic and the cluster node logic. And relay module also consists of the relay server logic and the relay client logic.

Cluster management module: Its main function is to maintain the cluster topology. The landmark logic maintains a neighbor landmark set, a cluster node set and a relay node set. Landmark logic is opened when this node becomes a landmark. The cluster node logic maintains its landmark set and a standby landmark set. There will be one candidate landmark which is decided by landmark competing. All nodes keep heartbeat communications with other nodes, and will help them to find the relay nodes candidates. The node will change the landmark set according to the dynamic network states.

Relay module: The main function is to provide relay service for upper SIP sessions. Relay server logic is an independent function. When a cluster node can become a relay, it will open this function and record its address on the landmark. Relay client logic mainly helps to find candidate relay nodes by detection. At the same time, it will cache a local relay list which has been found to be the proper relays. But the relay client will first ask the cluster management module to find proper relay nodes; if it is failed, it will then use the local relays.

3.3. Node Operations

Node joining: There is a fixed address list of the bootstrap servers stored in every client. The cluster node will send a *RTT* detection request to the bootstrap server and receive the response when it logins. Then according to the *RTT*, the normal node requests a list of landmarks from the nearest bootstrap server. After got a list of landmarks, this node detects them and finds out the nearest landmark according to the path quality. We set a threshold when judging whether a node belongs to a cluster. If there is no corresponding cluster landmark, this node will be automatically promoted to be a landmark to form a new cluster, and then will get a cluster ID from the bootstrap server which is used to represent the new cluster.

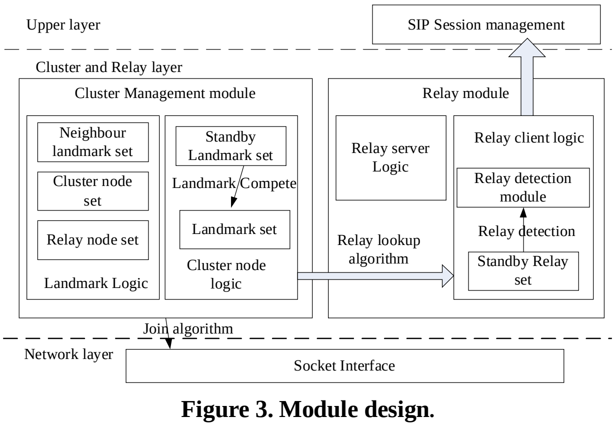


Figure 3. Module design.

The maintenance of topology: After a node joins a cluster, it will request near landmark list from its cluster landmark and detect them. If there is a landmark that is nearer than the current one, this node will join the new cluster by attaching to that landmark. This situation is common in mobile network because of node migration. In addition, each landmark stores a list of nearby ones. At the same time these landmarks will exchange their stored nearby landmark lists with each other so that the new joined landmark can be broadcasted in the network in time. Meanwhile, the landmark will also notify messages of new landmarks to its cluster nodes simultaneously. Then the nodes in the cluster will check their own condition and do some adjustments.

The maintenance of cluster node: the landmark maintains heartbeats with its cluster nodes and also with the relay nodes it stores.

Collecting information of relay: when a normal node joins the network, it will check out whether it could be a relay or not. If a relay node knows more than two landmarks whose path qualities (from itself to one landmark) are greater than a constant Q_0 value, it will notify the landmark. The landmark will also send back the list of previous landmarks addresses. The structure is called relay-cluster data structure in the system.

The initial communication: when two nodes communicate and need relay, they will first test the quality of the direct connection. And then, they will begin new inquiry for relays. Our SIP+P2P system supports multiple paths in one session, and uses optimizing routing [2,6] to improve QoS performance.

Inquiry for Relay: a normal node will ask landmark for candidate relay nodes by notifying the destination node's cluster, otherwise it will find candidate relay nodes locally. The candidate relay nodes are found in Relay cluster date structure. The landmark first searches the relay-cluster data structure for available relays of the corresponding destination cluster's relay nodes set and then sends the result to the source peer for detecting. If there are no proper relay nodes at last, the source node will ask for the near cluster set (which is got from the neighbor landmark set) and make an intersection with the destination node's near cluster set. Then it asks the cluster's landmark in the cluster set for relay nodes, makes them for the initial candidate relay nodes, and begins the detecting. If both of the two methods fail, the source node will randomly search nodes in its neighbors set for candidate relay nodes. Of course, this is the worst condition.

Building the communication: whether the source node gets the candidate relay list from local or from the landmark, these relays must be negotiated with the destination node, and then be detected. After the negotiation and detection, those suitable relay nodes will provide the application layer routing service.

3.4. Summary

Our system is a typical distributed system. As depicted above, the bootstrap server in our method is similar to a central server, and the landmark is similar to local server in CDN network. The cluster node should register and log off at the landmark and maintain a certain link with the landmark through heartbeat packages. All the landmarks make up a coverage network through the network topology, and maintain the cluster topology by heartbeat detecting. The landmark node and the cluster node compose a relation of tree. Landmarks exchange information with each other so that they could keep a close watch on the change of the network. And the landmark also maintains a list of relay of its autonomy cluster. The whole cluster overlay is similar to the Internet topology. The communication nodes request relays from its landmark and do the detection. If it fails, it will get the intersection of landmarks of the source node and the destination node. These landmarks and the relays in their clusters will be the candidate relays. This will reduce the load of the bootstrap server greatly and make the network very scalable. The node detection and relay detection are finished when building the coverage network topology, so the cost is shared with the process of building the coverage network topology. And this mechanism has a high hit ratio in the simulation system.

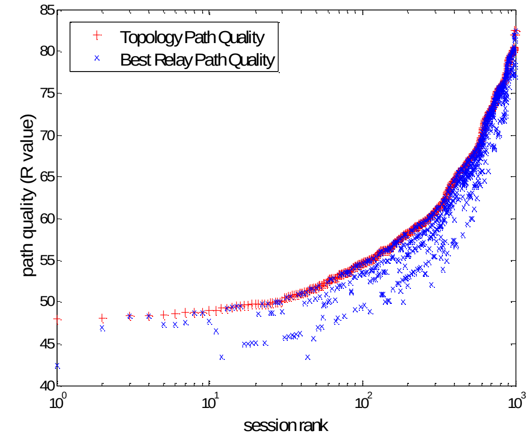


Figure 4. Topology path quality vs best relay path quality.

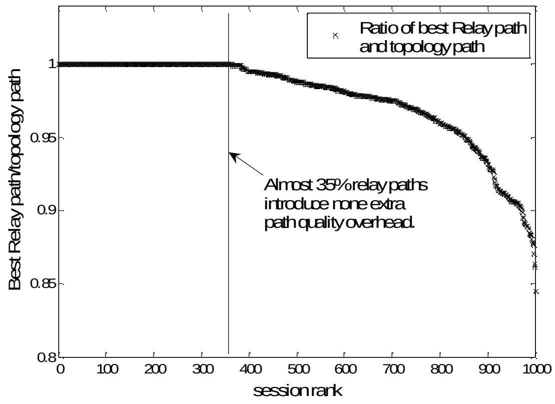


Figure 5. Distribution of relative quality overhead.

4. Experiments and Analyses

In this section, we will first introduce our experiment environments, including the simulation platform and the topology generator. Then we will introduce our evaluate methods. We mainly consider the performance of cluster overlay and the efficiency of relay lookup algorithm.

We evaluate our cluster overlay mainly from two aspects:

1) Firstly, we evaluate the similarity degree of topology between our cluster overlay and the Internet.

2) Then, we build cluster overlay in different network scales to evaluate its scalability.

The simulation results show that our cluster overlay topology is similar to Internet and has high scalability.

When evaluate the performance of relay lookup scheme, we mainly consider two evaluations below:

1) Hit ratio in different network scales.

2) Comparing path quality of our scheme with original SIP+P2P system and a simple random selection.

The simulation results show that our scheme has higher hit ratio and performs better than other schemes.

4.1. Experiment Environments

We simulate our design in the NS2 simulation platform. There are three main steps in our experiment:

1) Firstly, we use a topology generator tool to generate network topology which is as same as Internet.

2) Secondly, we build our cluster overlay based on this network topology.

3) Finally, we run relay lookup algorithms on our cluster overlay, and evaluate its performance.

In the NS2 environment, there is a tool to build network topology called GT-ITM [9]. We adopt the network model of random connecting of Transit AS with Stub AS (Transit-Stub model), which is most similar to the real network. With different sets of delay and packet loss between nodes (read from a static configure file), we guarantee that the topology is the same in different experiments for a given network scale. We build cluster overlay networks in different network scales to evaluate the scalability, and compare node pair's quality in cluster overlay with quality in original topology. After building cluster overlay, we run different relay lookup scheme to evaluate the performance of our topology aware algorithm.

4.2. Similarity Degree of Topology

We define relative quality ratio as the quality ratio of best relay path and topology path. When the relative quality ratio is higher, our cluster overlay topology is similar to Internet. We select 1000 node pairs and compute there topology path qualities and the best relay path qualities, which are plot in Figure 4.

Then, we compute the relative quality ratio, and plot its distribution below. As we can see in Figure 5, almost 35% relay paths introduce none extra quality overhead and 90% of relative quality ratio are above 0.9. These results prove our cluster overlay is very similar with real Internet topology.

4.3. Scalability of Cluster Topology

The scalability of this system requires that when the network scale increases greatly, the cluster scale does not increase greatly. The scalability is very important to our cluster overlay, because when the cluster overlay has high scalability, the maintenance cost of cluster topology does not increase greatly according to the network scale. As show in Figure 6, if the network scale increases, the number of landmarks and cluster nodes is increasing as well. But, the number of cluster nodes is increasing greatly and the number of landmark has no clear changes. This result proves that our scheme has good scalability.

4.4. Hit Ratio of the Best Relay Nodes

The object of this scheme is to find the best relays in P2P relay network which are near to both the source node and the destination node. This proportion of best relay found in this scheme is called hit ratio. In this experiment, we compare the relay nodes resulted from simulation with those resulted from calculation according to topology graph, and get a hit ratio. As show in Figure 7, we compare the hit ratio in different network scales for SIP+P2P and our new scheme. The results show that our scheme has better hit ratio than original SIP+P2P system. The reason is that our new scheme uses topology-aware to help look up the best relay.

4.5. Path Quality of Peer Relay Routing

In this section, we evaluate the QoS performance of our new scheme. We choose a network with 3000 AS nodes for simulation, and randomly select 10^4 pairs of nodes sessions need relay. We only select those node pairs which are in different ASs, and whose path qualities are bad. Then we use different scheme to select relays and compare their path quality (Random selection, One-side optimal selection, Topology Aware, and Optimal search calculated according to topology graph):

As we can see in Figure 8, our new scheme performs better than One-side optimal and random selection, and is very close to optimal search. However, when using topology-aware lookup algorithm, there are 8% sessions can not find relay. As mentioned above, when a client lookup relay failed, it will use a relay in local cache.

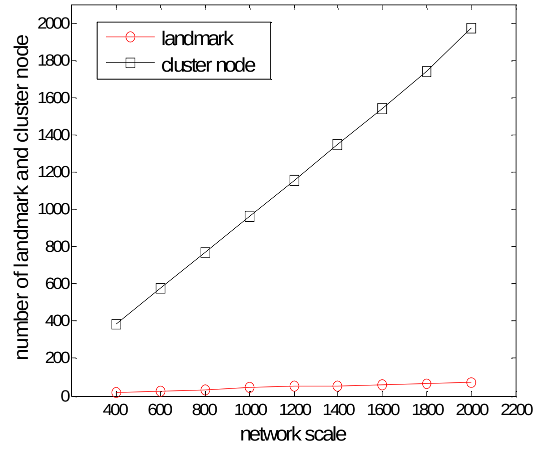


Figure 6. The number of cluster node and landmark.

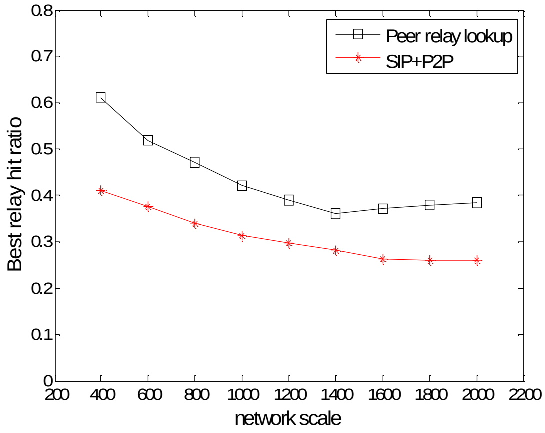


Figure 7. Hit ratio of the best relay.

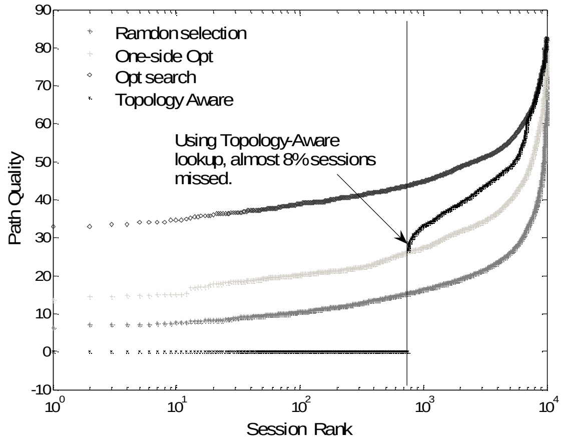


Figure 8. Path quality of peer relay routing

5. Conclusions

In this paper, we propose a novel peer relay lookup scheme based on SIP+P2P system to improve its relay selection. This approach uses topology-aware technology to optimize the relay selection. First, we divide the whole

P2P relay network into different clusters. When building the cluster topology, we use path quality between two nodes as the measurement metric. Then, when communicating node needs relay, it asks its landmark for proper relays and the cluster landmark will choose candidate relays according to the network topology. This new scheme can guarantee that the selected relay is near to both the source node and the destination node, which is better than the original SIP+P2P system.

We introduce the whole design in detail, including the system architecture, the peer design, and the node operation procedure. Then we evaluate our design in NS2 simulation, and the results show that our scheme is scalable. We also make comparison with traditional schemes and the results show that our scheme can get high best relay hit ratio and better path quality. Those experimental results confirm the feasibility of a real system.

Our future work is to realize this scheme in our SIP+P2P UA [2], and to evaluate it in a real network environment. Moreover, to modify this scheme to make it suitable for mobile network is another future work.

6. References

[1] S. Ren, L. Guo, and X. Zhang, "ASAP: An as-aware peer-relay protocol for high quality VoIP," Proceedings of 2006 IEEE ICDCS, pp. 70–79, 2006.

[2] X. W. Zhang, W. M. Lei, and W. Zhang, "Using P2P network to transmit media stream in SIP-based system," Proceedings of ICYCS 2008, pp. 362–367, November 2008.

[3] S. Guha and N. Daswani, "An experimental study of the skype peer-to-peer VoIP system," Proceedings of 2006 IPTPS, pp. 677–686, 2006.

[4] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris, "Resilient overlay networks," ACM SIGCOMM Computer Communication Review, Vol. 32, No. 1, pp. 75–83, 2001.

[5] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris, "The case for resilient overlay networks," Proceedings of 2001 Workshop on Hot Topics in Operating Systems, pp. 131–143, 2001.

[6] S. Tao, K. Xu, A. Estepa, T. Fei, *et al.*, "Improving VoIP quality through path switching," Proceedings of 2005 IEEE INFOCOM, pp. 515–523, 2005.

[7] K. Suh, F. Daniel, K. Jim, and D. Towsley, "Characterizing and detecting skype-relayed traffic," Proceedings of 2006 IEEE INFOCOM, pp. 292–301, 2006.

[8] H. Xie and Y. Yang, "A measurementbased study of the skype peer-to-peer VoIP performance," Proceedings of 2007 IPTPS, 2007.

[9] K. Calvert and E. W. Zegura, "GT-ITM: Georgia tech internetwork topology models," In http://www.cc.gatech.edu/projects/gtitm/, 1997.

Int. J. Communications, Network and System Sciences, **2010, 3, 126-132**
doi:10.4236/ijcns.2010.32019 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Selection of Design Parameters for Generalized Sphere Decoding Algorithms

Ping WANG, Tho LE-NGOC
Department of ECE, McGill University, 3480 *University, Montréal, Canada*
Email: *pingwang.mcgill@gmail.com, tho.le-ngoc@mcgill.ca*

Received November 11, 2009; *revised December* 15, 2009; *accepted January* 7, 2010

Abstract

Various efficient generalized sphere decoding (GSD) algorithms have been proposed to approach optimal ML performance for underdetermined linear systems, by transforming the original problem into the full-column-rank one so that standard SD can be fully applied. However, their design parameters are heuristically set based on observation or the possibility of an ill-conditioned transformed matrix can affect their searching efficiency. This paper presents a better transformation to alleviate the ill-conditioned structure and provides a systematic approach to select design parameters for various GSD algorithms in order to high efficiency. Simulation results on the searching performance confirm that the proposed techniques can provide significant improvement.

Keywords: Sphere Decoding (SD), Generalized Sphere Decoding (GSD), Maximum-likelihood (ML), Integer Least-Square (ILS), MIMO, λ-GSD, Multi-User Detection (MUD), CDMA, MC-CDMA

1. Introduction

Sphere decoding (SD) is an efficient searching method to obtain maximum-likelihood (ML) solution for NP-hard integer least-square (ILS) problems. SD can have polynomial time average complexity for many practical communication problems [1]. It offers large complexity reduction over the exhaustive search method in numerous applications, e.g., lattice decoding [2], multi-input, multi-output (MIMO) detection [3] and multi-user detection (MUD) [4]. However, when the problem is underdetermined, zero elements appear in the diagonal terms of the upper-triangular matrix generated by QR or Cholesky decomposition before searching, and the standard SD searching cannot apply. Such underdetermined ILS problems arise in many areas, e.g., MIMO detection with the number of transmit antennas larger than that of receiver antennas; MIMO detection with strongly correlated channel gains [5] or MUD for overloaded CDMA-related systems [8]. To solve such problems efficiently, generalized SD (GSD) algorithms, fully or partly based on SD, have been developed [6–12] for underdetermined or rank-deficient MIMO systems.

The GSD algorithms in [6,7] modify the underdetermined problem for constant-modulus constellation (e.g., QPSK) into a full-column-rank one by introducing a design parameter purely based on observations and then uses SD on the modified problem. For non-constant-modulus ones, e.g., 16/64QAM, they have to be transformed into multiple QPSKs, leading to larger dimension and hence increased complexity. To avoid this problem and obtain better efficiency, the λ-GSD algorithm proposed in [8], performs transformation without expanding the size of M-ary QAMs. Unlike the GSD algorithms in [6,7], the design parameter can be is upper-bounded [8,13] to guarantee near-ML performance for high QAMs and has little effect on the efficiency of λ-GSD [14].

In this paper, we first present an improved version of the λ-GSD algorithm to alleviate the ill-conditioned problem in the transformed matrix of [8]. Then we study the setting of the design parameter for the improved version of λ-GSD, as well as GSD in [6,7]. A more systematic approach is proposed to select design parameters for the above GSD algorithms to achieve high efficiency.

The remainder of the paper is as follows. Section 2 summarizes transformations employed in several popular GSD algorithms. Section 3 presents the improved version of λ-GSD over [8] and briefly discusses its performance characteristics; then we proceed to propose a systematic approach to choose the design parameter for it and also for that of [6,7] in Section 4. Section 5 provides illustrative results under flat-fading underdetermined MIMO scenarios. Section 6 concludes the paper.

2. Transformations by G SD Algorithms

The objective of ILS problem is to find the least-square (LS) solution for a linear system with integer unknowns. When the transmitted vectors (the unknowns) are from a finite set A in communications, i.e., $\mathbf{x}\in A\subset Z^n$, where Z^n denotes the n-dim vector space with integer entries, the objective function is

$$\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\|\mathbf{y}-\mathbf{Hx}\|_2^2,\ \text{for}\ \ \mathbf{y}=\mathbf{Hx}+\mathbf{v}, \tag{1}$$

where $\mathbf{y}\in\Re^m$ is a $m\times 1$ vector for a real-valued system[1], $\mathbf{H}\in\Re^{m\times n}$ the system matrix typically with full column rank (i.e., rank($\mathbf{H}$)=n) and $\mathbf{v}$ a $m\times 1$ zero-mean Gaussian noise vector, $\mathbf{v}\sim N(0,\sigma^2\mathbf{I})$. It's well known that SD algorithm and its variants are efficient approaches to solve such problems: it reduces the complexity of exhaustive-search greatly by conducting search within a hyper-sphere but still guarantees the ML optimality. One advantage of SD algorithms is that the search can be divided into layered enumeration of integer values in *one-dimensional* interval, so that it can efficiently decide which points are inside the hyper-sphere. However, SD algorithm was originally designed for full-column-rank system only. For underdetermined ILS problem when $\mathbf{H}\in\Re^{m\times n}$ has rank ($\mathbf{H}$) $=m, m<n$ in (1), the original SD algorithms can not proceed efficiently, as on certain searching layers, the candidate point set is no longer in a 1D-interval. [6–8] introduce transformations for underdetermined problem so that SD can be utilized on the transformed full-column-rank system. The transformations used are as follows.

The original underdetermined problem for constant-modulus constellation (e.g., QPSK) was modified into a full-column-rank system [6] and the equivalent new objective function is

$$\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\left(\|\mathbf{y}-\mathbf{Hx}\|_2^2+\lambda^2\|\mathbf{x}\|_2^2\right), \tag{2}$$

where $\lambda\neq 0$ is a design parameter. (2) is equivalent to

$$\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\left\|\begin{pmatrix}\mathbf{y}\\ \mathbf{0}\end{pmatrix}-\begin{pmatrix}\mathbf{H}\\ \lambda\mathbf{I}_{n\times n}\end{pmatrix}\mathbf{x}\right\|_2^2, \tag{3}$$

where $\begin{pmatrix}\mathbf{H}\\ \lambda\mathbf{I}_{\mathrm{n\times n}}\end{pmatrix}$ in (3) is full-column-ranked with $\lambda\neq 0$. If $\mathbf{x}$ is from constant-modulus constellations, $\|\mathbf{x}\|_2^2$ is constant and (3) is the same as (1); and when applying SD on the new problem (3), ML performance can still be achieved. For high QAMs, in order to guarantee ML performance, high QAMs are transformed into multiple QPSKs in [6] before using the above transformation, e.g., 16QAM vector $\mathbf{x}$ can be represented as $\left[\sqrt{2}\mathrm{x}_1\ \sqrt{2}\mathrm{x}_2/2\right]$ where $\mathbf{x}_1,\mathbf{x}_2$ are from QPSK. The 16QAM system is transformed into a QPSK one with

$$\mathbf{y}=\left[\sqrt{2}\mathbf{H}\ \ \frac{\sqrt{2}}{2}\mathbf{H}\right]\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix}+\mathbf{v}. \tag{4}$$

Hence the transformation in (3) can be used on (4) and SD can then be applied efficiently. Despite the fact that the complexity of the GSD algorithm in [6] depends on λ, the parameter is set arbitrarily in [6].

The transformation of high QAMs into multiple QPSK ones in [6] unavoidably lead to larger problem size for the ILS problem and increased search complexity. To avoid this problem, another transformation is utilized in [8] for both the constant-modulus and non-constant modulus constellations, i.e.,

$$\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\left(\|\mathbf{y}-\mathbf{Hx}\|_2^2+\lambda^2\|\bar{\mathbf{x}}_2\|_2^2\right)=\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\left\|\begin{pmatrix}\mathbf{y}\\ \mathbf{0}\end{pmatrix}-\begin{pmatrix}\bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2\\ \mathbf{0} & \lambda\mathbf{I}_{(n\text{-}m)\times(n\text{-}m)}\end{pmatrix}\begin{bmatrix}\bar{\mathbf{x}}_1\\ \bar{\mathbf{x}}_2\end{bmatrix}\right\|_2^2 \tag{5}$$

where $\bar{\mathbf{x}}_2$ is subset of $\mathbf{x}$ =[$\bar{\mathbf{x}}_1$ $\bar{\mathbf{x}}_2$] with length (n-m). SD searching is applied on (5) directly for both QPSK and high QAMs. Consequently, exact ML performance can be achieved for QPSK constellation with arbitrary value of λ whist close-to-ML performance can be achieved with sufficient small λ [8]. Moreover, an upper bound is provided in [8] for λ in high QAMs to approach ML performance with negligible loss and it's also shown that the efficiency is quite insensitive to λ [14] for λ-GSD, unlike GSD in [6]. Therefore, the choice of λ can be randomly set for QPSK and judiciously set to be within the upper bound for high QAMs.

Later in [7], by noticing the structure of the transformation in (5) is less efficient[2] and also targeting to alleviate the size expansion problem in [6,7] considers to utilize a transformation of adding the $\bar{\mathbf{x}}_2$ part of $\mathbf{x}$ =[$\bar{\mathbf{x}}_1$ $\bar{\mathbf{x}}_2$] with length (n-m+1) instead of (n-m), and then apply SD to achieve ML performance for constant-modulus constellations. The resulted new ILS problem is

$$\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\left(\|\mathbf{y}-\mathbf{Hx}\|_2^2+\lambda^2\|\bar{\mathbf{x}}_2\|_2^2\right)=$$
$$\min_{\mathbf{x}\in A\subset \mathbf{Z}^n}\left\|\begin{pmatrix}\mathbf{y}\\ \mathbf{0}\end{pmatrix}-\begin{pmatrix}\bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2\\ \mathbf{0} & \lambda\mathbf{I}_{(n-m+1)\times(n-m+1)}\end{pmatrix}\begin{bmatrix}\bar{\mathbf{x}}_1\\ \bar{\mathbf{x}}_2\end{bmatrix}\right\|_2^2 \tag{6}$$

For non-constant-modulus constellations, [7] adopts the same strategy as [6], i.e., to transform high QAMs into multiple QPSKs in order to achieve exact-ML performance in a similar way as [6,7] also suggests a value

[1] For complex-valued communication systems, linear transformations are generally used to obtain real-valued ones and apply SD.

[2] to be discussed in Section 3.

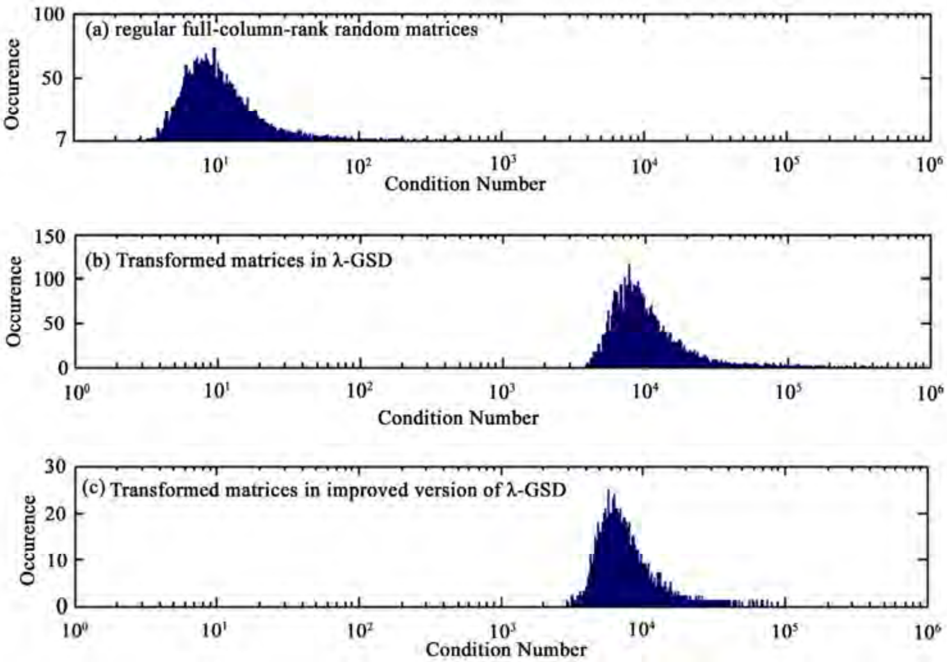


Figure 1. Histogram of condition number distribution of MIMO system matrices over 10^4 runs @12dB, QPSK. (a) 12x12 real matrices of regular i.i.d. 6x6 complex-valued matrices; (b) transformed 12x12 real matrices of λ-GSD at λ=10^{-3}; (c)transformed 14x12 real matrices of improved λ-GSD at λ=10^{-3}.

of λ purely based on experimental observations though apparently the efficiency depends on λ.

3. Improved Transformation Structure to Speed up λ-GSD Algorithm

For a negligible performance loss, λ-GSD keeps the problem size intact for high QAMs as mentioned, and hence it is more efficient than [6,7] for high QAMs. However, despite of the high efficiency of λ-GSD compared to its counterparts, the efficiency of λ-GSD is still significantly higher than SD on regular full-column-rank system. Ill-conditioned structure of the transformed matrix of λ-GSD partly accounts for the inefficiency. It is known that well-conditioned matrix generally leads to more efficient searching than ill-conditioned ones [16]. Figure 1 illustrates histograms for the occurrence of the condition number of the three categories of matrices over 10^4 runs for a QPSK MIMO system at 12dB, where (a) is for the 12x12 real-valued matrices of the i.i.d. regular 6x6 complex-valued system with full-column-rank; (b) is for the transformed 12x12 real matrices of λ-GSD in a 4x6 complex-valued system. (c) is for an improved version of λ-GSD to be proposed in this section. Comparing Figure 1(a) and (b), we can see that over 10^4 runs, the condition number of transformed matrices of λ-GSD is centered at around 8x10^3 whilst that of the regular full-column-rank ones is only centered at around 10. Next we will propose an improved version of λ-GSD, which combines the transformation in [7] and the idea of keeping problem size intact of λ-GSD. The modified version of λ-GSD to alleviate the ill-conditioned problem can be presented as follows.

Consider a real-valued underdetermined system $\mathbf{y}=\mathbf{H}\mathbf{x}+\mathbf{v}$ in (1) where 2<*rank*($\mathbf{H}$)=*m*<*n*. Partition $\mathbf{H}=[\bar{\mathbf{H}}_1\,\bar{\mathbf{H}}_2]$ where $\bar{\mathbf{H}}_1,\bar{\mathbf{H}}_2$ are $m\times(m-2)$ and $m\times(n-m+2)$ matrices respectively, instead of $m\times m$ and $m\times(n-m)$ in the original λ-GSD (5). Then partition $\mathbf{x}=[\bar{\mathbf{x}}_1{}^{\mathrm{T}}\ \ \bar{\mathbf{x}}_2{}^{T}]^{T}$ so that $\mathbf{y}=[\bar{\mathbf{H}}_1\ \bar{\mathbf{H}}_2][\bar{\mathbf{x}}_1{}^{\mathrm{T}}\ \ \bar{\mathbf{x}}_2{}^{T}]^{T}$ $+\mathbf{v}$. Adding an equation $\mathbf{0}=\lambda\bar{\mathbf{x}}_2-\lambda\bar{\mathbf{x}}_2$ where $\mathbf{0}$ is a zero-vector with length (*n-m+2*) and λ >0 is a weighting factor, the transformed new ILS problem becomes,

$$\min_{x\in A\subset Z^n}\left\|\begin{pmatrix}\mathbf{y}\\ \mathbf{0}\end{pmatrix}-\begin{bmatrix}\bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2\\ 0 & \lambda\mathbf{I}_{(n-m+2)\times(n-m+2)}\end{bmatrix}\begin{bmatrix}\bar{\mathbf{x}}_1\\ \bar{\mathbf{x}}_2\end{bmatrix}\right\|_2^2 = \min_{x\in A\subset Z^n}\left(\|\mathbf{y}-\mathbf{H}\mathbf{x}\|_2^2+\lambda^2\|\bar{\mathbf{x}}_2\|_2^2\right) \tag{7}$$

Again, SD algorithms then can be used on (7) to solve the transformed full-column-rank ILS problem. Note that the transformation in (7) appears very analogous[3] to that

[3] The only difference in transformation is the length of (n-m+2) in $\bar{\mathbf{x}}_2$ is used in improved version of λ-GSD, instead of (n-m+1) in [7]. Using (n-m+2) in real-valued ILS problem is actually equivalent to using ((n-m)/2+1) in corresponding complex-valued problem; i.e,, "+1" is assumed to operate in complex-valued domain, whilst [7] applies "+1" directly on real-valued domain.

used in [2], i.e., (6), yet they are distinct for the case of high QAMs: the improved version of λ-GSD keeps the problem size intact for high QAMs and hence SD algorithm is directly applied on (7) without further transformation. Performance of the modified approach can be analyzed following the analysis for λ-GSD in [13], and the conclusions are similar. Specifically, for constant-modulus modulations, the modified GSD can achieve exact ML performance and for high QAMs, the choice of λ has to be sufficiently small to keep the performance close to optimal. Here we omit the detailed upper bound analysis for simplicity and only provide the conclusions for a $R_x \times T_x\ (R_x < T_x)$ complex linear system using uncoded square *M*-QAM with average symbol energy E_s: Approximately, assuming ε denotes the difference in the upper bound error probability between λ-GSD and ML performance, the upper bound for λ is $\lambda \le (32\varepsilon)^{1/4}\left[3(T_x - R_x + 1)E_s / \sigma^2\right]^{-1/2}$.

Compared to original λ-GSD, we expect the improved version is more efficient as it has better transformed matrix structure, which can be seen from the illustrative histogram of condition number in Figure 1(c) for the transformed 14x12 real-valued matrices in the improved version of λ-GSD over 10^4 runs for a QPSK MIMO system. We can see that the condition number in this example is centered at around $6\text{x}10^3$. Compared to $8\text{x}10^3$ in Figure 1(b), the ill-conditioned situation gets alleviated. The complexity efficiency of the improved version of λ-GSD will be illustrated in Section 5 via simulations.

4. Selection of Design Parameters for the GSD Algorithms

It was mentioned in [6,7] that there may exist an optimal value for the design parameter in their GSD algorithms. It is very hard, if not impossible, to derive a closed-form expression for the optimal value in a strict sense. In this section, alternatively, we propose one criterion to choose design parameters for these GSD algorithms, which enables the underlying algorithms with high efficiency. The proposed criterion also works for the improved version of λ-GSD algorithm in QPSK systems, since the underlying transformation shares similarity with that for GSD in [7]. Therefore, we mainly derive the design parameter for the improved version of λ-GSD in QPSK systems for simplicity, which can be applied to GSD in [7] as well. Note that for high QAMs, the selection of the design parameter for the improved version of λ-GSD is limited by the upper bound as mentioned in previous session.

For SD searching, if the first point found is closer to ML point, the search tends to be faster [15]. For the improved version of λ-GSD and GSD in [6,7], the first point obtained in search is different for various values of λ. Let us consider the transformation (5) of GSD [6] first. Since the first point using Shnorr-Euchner enumeration[4] is the midpoint, hence, the elements of the first point can be written as [15]

$$x_{init\_k} = \left\lfloor (\bar{y}_k - \textstyle\sum_{i=k+1}^{n} R_{k,i} x_{init\_i}) / R_{k,k} \right\rceil,\ \mathit{for}\ k = n, \cdots, 1, \tag{8}$$

where x_{init_k} is the ZF-DFE point [15] for the *transformed* ILS problem; $\bar{y}_k = (\mathbf{Q}^T \tilde{\mathbf{y}})_k$ where$(\mathbf{Q}, \mathbf{R})$ are the resulted matrices after QR decomposition of the transformed matrix; $R_{i,j}$ is the (*i*,*j*)-th element of $\mathbf{R}$ and $\tilde{\mathbf{y}}$ is the equivalent received vector for the transformed ILS problem. Then at *k*=*n*, based on (5), we have

$$\begin{aligned} x_{init\_n} &= \left\lfloor \left(\begin{pmatrix} \bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2 \\ & \lambda \mathbf{I}_{n\times n} \end{pmatrix}^T \begin{pmatrix} \bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2 \\ & \lambda \mathbf{I}_{n\times n} \end{pmatrix} \right)^{-1} \begin{pmatrix} \bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2 \\ & \lambda \mathbf{I}_{n\times n} \end{pmatrix}^T \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix} \right\rceil_n \\ &= \left\lfloor \left(\mathbf{H}^T \mathbf{H} + \lambda^2 \mathbf{I}_{n\times n} \right)^{-1} \mathbf{H}^T \mathbf{y} \right\rceil_n \end{aligned} \tag{9}$$

Clearly, x_{init_n} depends on λ. From (8) and (9), we

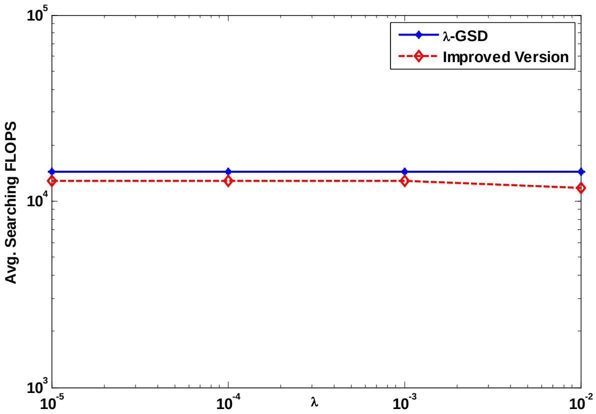


Figure 2. Avg. searching FLOPS vs. λ @ 2x4 MIMO, 16QAM, 28dB, 10^4 runs.

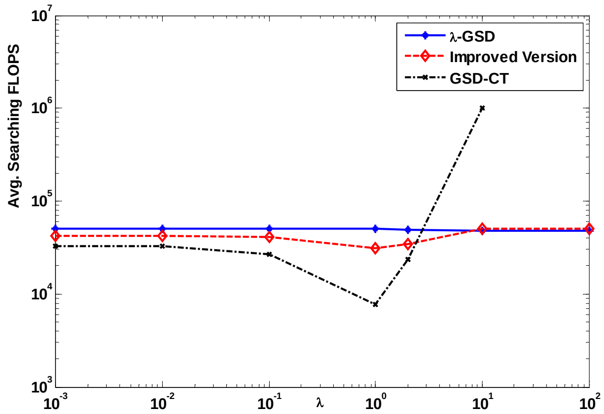


Figure 3. Avg. searching FLOPS vs. λ @ 4x8 MIMO, QPSK, 12dB, 10^4 runs.

know that when λ is 0, the first element of (9), x_{init_n}, is the first element of ZF-DFE point of the original ILS, i.e, $x_{init\_n} = \left[x_{zf-dfe}\right]_n$; Thus, the first point, $\mathbf{x}_{init}$, is its ZF-DFE point. On the other hand, when λ is equal to $1/\sqrt{\gamma_s}$, where γ_s represents the SNR ratio, (9) becomes $x_{init\_n} = \left[\left(\mathbf{H}^T\mathbf{H} + \mathbf{I}_{n\times n}/\gamma_s\right)^{-1}\mathbf{H}^T\mathbf{y}\right]_n$ which is the first element of MMSE-DFE point for the original ILS, i.e, $x_{init\_n} = \left[x_{mmse-dfe}\right]_n$ and $\mathbf{x}_{init}$ is the MMSE-DFE point. It's known that MMSE detector is the best *linear* estimator to the ILS problem based on mean-square error criterion and MMSE-DFE is generally closer to the optimal ML point than ZF-DFE; Using MMSE-DFE point as the first point in the searching procedure could lead to a fast convergence [15]. Therefore, to achieve high efficiency, we suggest using $\lambda = 1/\sqrt{\gamma_s}$ in GSD [6], so that the first valid point is the MMSE-DFE point of the original ILS.

Next, we continue to derive the suggested value of λ for the improved version of λ-GSD for QPSK system as follows. From (8), we obtain x_{init_n} in the improved version of λ-GSD:

$$x_{init\_n} =$$

$$\left[\left(\begin{pmatrix}\bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2\\ 0 & \lambda\mathbf{I}_{(n\text{-}m+2)\times(n\text{-}m+2)}\end{pmatrix}^T\begin{pmatrix}\bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2\\ 0 & \lambda\mathbf{I}_{(n\text{-}m+2)\times(n\text{-}m+2)}\end{pmatrix}\right)^{-1}\begin{pmatrix}\bar{\mathbf{H}}_1 & \bar{\mathbf{H}}_2\\ 0 & \lambda\mathbf{I}_{(\text{n-m+2})\times(\text{n-m+2})}\end{pmatrix}^T\begin{pmatrix}\mathbf{y}\\ \mathbf{0}\end{pmatrix}\right]_n$$

$$= \left[\begin{pmatrix}\mathbf{H}_1{}^T\mathbf{H}_1 & \mathbf{H}_1{}^T\mathbf{H}_2\\ \mathbf{H}_2{}^T\mathbf{H}_1 & \mathbf{H}_2{}^T\mathbf{H}_2 + \lambda^2\mathbf{I}_{(n\text{-}m+2)\times(n\text{-}m+2)}\end{pmatrix}^{-1}\mathbf{H}^T\mathbf{y}\right]_n \tag{10}$$

Comparing with the first element of MMSE-DFE point,

$$x_{init\_n} = \left[\left(\mathbf{H}^T\mathbf{H} + \frac{\mathbf{I}_{n\times n}}{\gamma_s}\right)^{-1}\mathbf{H}^T\mathbf{y}\right]_n =$$

$$\left[\begin{pmatrix}\mathbf{H}_1{}^T\mathbf{H}_1 + \dfrac{\mathbf{I}_{(m-2)\times(m\text{-}2)}}{\gamma_s} & \mathbf{H}_1{}^T\mathbf{H}_2\\ \mathbf{H}_2{}^T\mathbf{H}_1 & \mathbf{H}_2{}^T\mathbf{H}_2 + \dfrac{\mathbf{I}_{(n-m+2)\times(n\text{-}m+2)}}{\gamma_s}\end{pmatrix}^{-1}\mathbf{H}^T\mathbf{y}\right]_n \tag{11}$$

We would like to choose λ such that (10) and (11) are as similar as possible. Thus, we propose to choose λ such that the two following matrices, $\mathbf{A}_1 \triangleq \begin{pmatrix}\mathbf{H}_1{}^T\mathbf{H}_1 & \mathbf{H}_1{}^T\mathbf{H}_2\\ \mathbf{H}_2{}^T\mathbf{H}_1 & \mathbf{H}_2{}^T\mathbf{H}_2 + \lambda^2\mathbf{I}_{(n\text{-}m+1)\times(n\text{-}m+1)}\end{pmatrix}$ in (10) and $\mathbf{A}_2 \triangleq \begin{pmatrix}\mathbf{H}_1{}^T\mathbf{H}_1 + \mathbf{I}_{(m-2)\times(m\text{-}2)}/\gamma_s & \mathbf{H}_1{}^T\mathbf{H}_2\\ \mathbf{H}_2{}^T\mathbf{H}_1 & \mathbf{H}_2{}^T\mathbf{H}_2 + \mathbf{I}_{(n-m+2)\times(n\text{-}m+2)}/\gamma_s\end{pmatrix}$ in (11) are as similar as possible. One way to measure the similarity between the two matrices is the norm of the difference. Here we use Frobenius norm to measure the difference and choose λ to minimize $\left\|\mathbf{A}_1 - \mathbf{A}_2\right\|_F$, i.e.,

$$\min_{\lambda>0}\left\|\mathbf{A}_1 - \mathbf{A}_2\right\|_F =$$

$$\left\|\begin{pmatrix}\mathbf{I}_{(m-2)\times(m\text{-}2)}/\gamma_s & \mathbf{0}\\ \mathbf{0} & (1/\gamma_s - \lambda^2)\mathbf{I}_{(n-m+2)\times(n\text{-}m+2)}\end{pmatrix}\right\|_F$$

$$= \frac{n}{\gamma_s} - \lambda^2(n-m+2) \tag{12}$$

Clearly, when $\lambda = \sqrt{n/[(n-m+2)\gamma_s]}$, $\left\|\mathbf{A}_1 - \mathbf{A}_2\right\|_F$ achieves the minimum, zero.

5. Illustrative Results

This section illustrates the merits of the improved ver-

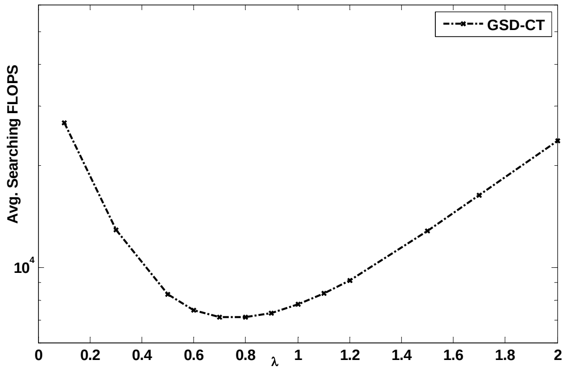


Figure 4. Avg. searching FLOPS vs. λ @ 4x8 MIMO, QPSK, 12dB, 10^4 runs.

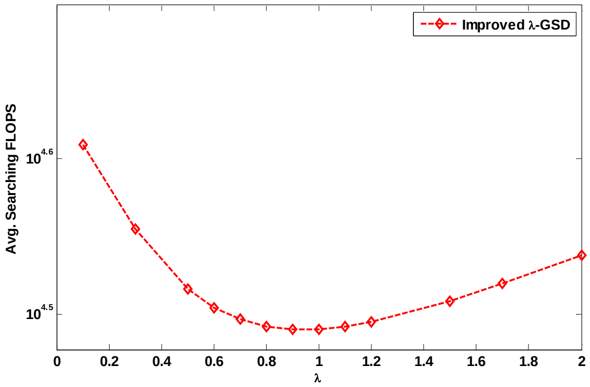


Figure 5. Avg. searching FLOPS vs. λ @ 4x8 MIMO, QPSK, 12dB, 10^4 runs.

[4] Shnorr-Euchner enumeration refers to the ordering of the integer points in each layer in a zig-zag manner, which is more efficient than natural ordering and recommended to be used for SD searching recently.

sion of λ-GSD and verify the efficiency of the selection of design parameter for the improved version of λ-GSD and GSD [6] via simulation in a flat fading underdetermined MIMO channel with *n* transmit and *m* receive antennas (*m*<*n*). In simulations, we use the "accelerated" standard SD algorithm in [17] as the sub-algorithm for GSDs. The received signal can be represented as $\mathbf{y} = \mathbf{Hx} + \mathbf{v}$, where $\mathbf{x}$ is the transmitted vector from M-QAM and $\mathbf{v}$ is an $m \times 1$ AWGN noise vector with variance $E(\mathbf{v}^H\mathbf{v}) = 2\sigma^2\mathbf{I}_m$. $\mathbf{H} = [h_{ij}]$ is an underdetermined $m \times n$ matrix where h_{ij} follows Rayleigh fading model $\sim CN(0,1)$. We define $SNR = 10\log_{10}(nE_s / (2q\sigma^2))$ dB, where E_s is the average M-ary QAM symbol energy. Note that the complex model would be transformed into an equivalent real model with *2n* unknowns (from $L = \sqrt{M}$-PAM equivalently) before using $\lambda - GSD$. The average *searching* FLOPS are counted as complexity measurement for simulation comparison for the GSDs.

Figure 2 provides the average searching FLOPS comparison between the original λ-GSD and the improved version for various λ values under a 2x4 16QAM flat-fading MIMO scenario at 28dB. Note that the upper bound for λ in these two GSD algorithms with ε=1% is 0.0167 and 0.0149 respectively. We can see that the complexity is reduced by about 5% with the better structure. Besides, Figure 2 indicates that in the improved version of λ-GSD for 16QAM, the complexity is quite insensitive to the choice of the design parameter λ as long as the value of λ is within the upper bound, similar to the case of the original λ-GSD.

Figure 3 plots the complexity comparison of three GSD algorithms for 4x8 QPSK flat-fading MIMO scenarios at SNR=12dB, including the original λ-GSD, the improved version and GSD in [6] (i.e., GSD-CT). Figure 3 shows the average searching FLOPS w.r.t. various values for the design parameter in all the three GSD algorithms for QPSK system. We can see that the complexity of the original λ-GSD is insensitive to the choice of the design parameter; yet the other two GSD algorithms are more sensitive to the design parameter and apparently there is an optimal value for the lowest complexity as expected.

Thus, we calculate the $\lambda = 1/\sqrt{\gamma_s}$ value for a 4x8 MIMO QPSK system at 12dB and the resulted value for λ is λ_0=0.710468. Then we simulate the complexity for different λ values in the range (0.1, 2) in Figure 4, which indicates that the lowest complexity is achieved approximately at λ_0=0.710468, i.e., choosing λ equal to λ_0 can lead to very good efficiency. Correspondingly, we calculate the $\lambda = \sqrt{n / [(n - m + 2)\gamma_s]}$ value of the improved version of λ-GSD for the same 4x8 MIMO QPSK system at 12dB and the resulted value for λ is λ_0= 0.8987. Then we simulate its corresponding complexity for different λ value in the range of (0.1, 2). Figure 5 indicates that the lowest complexity is achieved for λ around λ_0, i.e., choosing λ equal to the suggested value of λ_0= 0.8987 also can lead to very good efficiency for the improved version of λ-GSD.

6. Conclusions

An improved version of λ-GSD algorithm was first proposed for underdetermined linear systems by transforming the original problem into a full-column-rank one with better structure, so that standard SD can be efficiently applied. As the introduced transformation maintains the original problem dimension and has better system structure than the original one, lower complexity can be obtained compared to its counterparts, especially for high QAMs. Selection of design parameter λ was then studied for the improved version of λ-GSD. The resulted systematic approach to select the design parameter can also be used for GSD in [6], [7] in which the design parameters were only selected randomly or based on observations. Simulation results confirmed the high efficiency achieved by such choices of design parameters for these GSDs.

7. References

[1] E. Agrell, T. Eriksson, A. Vardy, and K.Zeger, "Closest point search in lattices," IEEE Transactions Information Theory, pp. 2201–2214, August 2002.

[2] O. Damen, A. Chkeif. and J. C. Belfiore, "Lattice code decoder for space-time codes," IEEE Communication Letters, pp. 161–163, May 2000.

[3] H. Vikalo, B. Hassibi, and T. Kailath, "Iterative decoding for MIMO channels via modified sphere decoding," IEEE Transactions on Wireless Communication, Vol. 3, No. 6, pp. 2299–2311, November 2004.

[4] L. Brunel, "Multiuser detection techniques using maximum likelihood sphere decoding in multicarrier CDMA systems," IEEE Tranactions on Wireless Communication, Vol. 3, No. 3, pp. 949–957, May 2004.

[5] T. Cui and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," IEEE VTC. Fall, 2004.

[6] T. Cui and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," IEEE Communication Letters, Vol. 9, No. 5, pp. 423–425, May 2005.

[7] X. Chang and X. Yang, "An efficient regularization approach for underdetermined MIMO system decoding", IWCMC, Hawaii, USA, August, 2007.

[8] P. Wang and T. Le-Ngoc, "An efficient multi-user detection scheme for overloaded group-orthogonal MC-CDMA systems," *IET Communications*, Vol. 2, No. 2, pp.346–352, February 2008.

[9] M. Damen, K. A. Meraim, and J. C. Belfiore, "Generalised sphere decoder for asymmetrical space-time communication architecture," Electronics Letters, Vol. 36, No. 2, pp. 166–167, January 2000.

[10] P. Dayal and M. K. Varanasi, "A fast generalized sphere decoder for optimum decoding of under-determined MIMO systems," 41st Annual Allerton Conference on Communication Control and Computer, October 2003.

[11] Z. Yang, C. Liu, and J. He, "A new approach for fast generalized sphere decoding in MIMO systems," IEEE Signal Processing Letters, Vol. 12, No. 1, pp. 41–44, January 2005.

[12] X. W. Chang and X. H. Yang, "A new fast generalized sphere decoding algorithm for underdetermined sphere decoding algorithm for underdetermined MIMO systems," 23rd Biennial symposium on Communications, Kingston, ON, Canada, pp. 18–21, June 2006.

[13] P. Wang and T. Le-Ngoc, "A low complexity generalized sphere decoding approach for underdetermined linear communication systems: performance and complexity evaluation," IEEE Transactions on Communications, Vol. 57, No. 11, pp. 3376–3388, November 2009.

[14] P. Wang and T. Le-Ngoc, "On the expected complexity analysis of a generalized sphere decoding algorithm for underdetermined linear communication systems," IEEE International Conference Communication (ICC), Glasgow, June 2007.

[15] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," IEEE Transactions on Information Theory, Vol. 49, No. 10, pp. 2389–2402, October 2003.

[16] H. Artes, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: a geometrical approach to approximate ML detection," IEEE Transactions on Signal Processing. Vol. 51, No. 11, pp. 2808–2820, November 2003.

[17] J. Boutros, N. Gresset, L. Brunel, and M. Fossorier, "Soft-input soft-output lattice sphere decoder for linear channels," IEEE Globecom 2003, pp. 1583–1587, 2003.

Int. J. Communications, Network and System Sciences, **2010, 3, 133-139**
doi:10.4236/ijcns.2010.32020 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Community Analysis of Social Network in MMOG

Sheng PANG, Changjia CHEN
School of Electric & Information Engineering, Beijing Jiaotong University, Beijing, China
Email: *pangsheng_bjtu@yahoo.com.cn, changjiachen@sina.com*

Received November 6, 2009; *revised November* 29, 2009; *accepted December* 27, 2009

Abstract

Massive Multiplayer Online Games (MMOG) have attracted millions of players in recent years. In MMOG, players organize themselves voluntarily and fulfill collective tasks together. Because each player can join different activities, one player may show different social relationship with others in different activities. In the paper we proposed the incremental label propagation algorithm to search the cliques accurately and quickly. Then we analyzed community structure characteristics on multi-activities. It's shown that the existing guild organization cannot satisfy the requirements of multi-activities in MMOG, which motivates us to devise new community communication channels and platforms in future.

Keywords: MMOG, Behavior, Community, Social Network

1. Introduction

About ten years ago, Internet grows explosively in China. Making friends through online chatting is fabulous and very popular at that time. Though a “net friend”, who is sitting in front of a certain PC at the other side of the wire, is mysterious and dangerous maybe, this platform indeed creates a new way to make friends. Nowadays, Internet has become a widely used tool for us to communicate with others. Similarly, the massive multiplayer online games (MMOG) also provide us a platform for making friends as well as entertaining. In MMOG, the virtual relationship network of the players resulted from their behaviors, which we call “relationship network” in briefly in this paper, has also become part of the real social relationship network of people, and the virtual world is becoming a field for the study of social behavior [1,2]. On the other hand, the players' behaviors in terms of their interaction can affect the performance of the game’s network system [3,4].

We name players of the game as nodes, and the players who participate in the same activities at the same time as neighbors. Thus the relationship network structure is formed. According to the pattern of multiple activities in the WoW, we divide these activities into two categories: Player vs. Environment (PVE) and Player vs. Player (PVP). PVE is composed of raid activities and party ones, while PVP contains battle ground activities. To show the characteristics of the players’ behaviors and reveal the unreasonable point of the existing policies and organizational style of the game, we go deep into the relationship network in the pattern of multiple activities in the World of Warcraft (WoW).

In the paper we proposed the time-based incremental label propagation algorithm to reveal the community structure of relationship network formed in one activity. And then we analysed the characteristic of overlap between these structures of different activities. Our major contributions are as follows:

1) Most networks will evolve with time instead of keeping unchanged in reality. Therefore, we proposed the time-based incremental label propagation algorithm, which is based on label propagation algorithm [5]. Our time-based algorithm will only take the local changed vertexes into consideration. The computation time is greatly decreased, while vertex (edge) is added, deleted or modified. The algorithm will definitely converge since the original version of label propagation algorithm converges.

2) Just as in real life, the clique of friends is different from the clique of workmates. In MMOG, there is little correlation among network structures of different activities. However, it is the overlap of these structures of different activities that extend the whole relationship network which turns the path reachable from unreachable. This indicates that the present guild organization cannot meet the requirements of multi-activities, which motivates us to devise new community communication channels and platforms in future.

The paper is organized as follows. In Section 2 we will present some related works on MMOG. We will go on to introduce the WoW and trace collection in Section 3. In Section 4, the time-based incremental abel propagation algorithm will be proposed, and we will present the characteristics of multi-layer network structures which orrespond to multiple activities in Section 5. We will give some oncluding remarks and future work in Section 6.

2. Related Work

This work analyzes the player relationship network resulted from the players' behavior in terms of their interaction. We adopt the measurement traces of Shaolong Li's work [6], and put our emphasis on the characteristics of multi-activities in the paper.

2.1. Players' Behaviors Analysis

As the virtual relationship network has also become part of the real social relationship network, many studies focus on psychological factors. N. Ducheneaut [1] observed player-to-player interaction in two locations in the game Star Wars Galaxies. They analyzed user interaction patterns, mainly gestures and utterances, and discussed how they were affected by the structure of the game. In [7], the authors believed it was important to study social interactions within the virtual communities. They examined the WoW as an online community, and investigated the degree to which it exhibits characteristics of a new tribalism.

However, it is equally important to understand users' behavior for network researchers, because how users act determines how well network systems perform, such as online games. In [3], authors analyzed player behavior in term of their social interaction, and revealed several featured patterns of player interaction. And then they proposed some suggestions to improve game system based on their studies. Tobias Fritsch [2] defined "hardcore" player and classified four game types (FPS, RTS, RPG and SG). They analyzed "hardcore" players' distribution of age and education level respectively for the four game types. The results of their statistical analysis can be used to improve the current game design and retail strategies. Lia C. Rodrigues [8] employed self-organizing maps to find clusters from MMOG, and designed a fuzzy system to improve the performance of the algorithm. But the algorithm must be trained for every MMOG before using it.

2.2. Clique Finding

The GN algorithm [9], proposed by Newman and Girvan in 2004 is used to find cliques in the graph, which is substantially classic and widely-used. GN algorithm has two definitive features: firstly, it involves iterative removal of edges from the network to split it into communities, the edges removed being identified using any one of possible "edge betweenness" measures; secondly, these measures are, crucially, recalculated after each removal. They also propose a measure to estimate the strength of the community structure found by our algorithms, which gives us an objective metric for choosing the number of communities into which a network should be divided. They also demonstrate that GN algorithm is highly effective at discovering community structure in both computer-generated and real-world network data [9,10]. But the disadvantage is the high time complexity of calculation, thus it is only applicable in small-scale network.

Raghavan proposed a localized community detection algorithm [5] based on label propagation. Each node is initialized with a unique label and at every iteration step of the algorithm each node adopts a label that a maximum number of its neighbors have, with ties broken uniformly randomly. As the labels propagate through the network in this manner, densely connected groups of nodes form a consensus on their labels. At the end of the algorithm, nodes having the same labels are grouped together as communities, as shown in Figure 1. The advantage of this algorithm over the others is its simplicity and time efficiency. Its time complexity can reach nearly linear time. In the paper we modified the algorithm to meet the demands of real time analysis and demonstrations.

Nearly all the community detecting methods cant tell us when the communities found are good ones or which of the divisions are the best ones for a given network. To answer these questions, Newman and Girvan define a measure of the quality of a particular division of a network modularity [11], denoted by *Q*, which is a numerical index to assess how good a particular division is. For a division with g groups, a g £ g matrix *e* is defined whose element *eij* is the fraction of edges in the original network that connect vertices in group *i* to those in group *j*. ai which is the sum of any row (or column) of e corresponds to the fraction of links connected to group *i*. Then the modularity is defined as follows:

$$Q = \sum_i (e_{ii} - a_i^2) \tag{1}$$

Physically, *Q* is the fraction of all edges that lie within communities minus the expected value of the same quantity in a graph in which the vertices have the same degrees but edges are placed at random without regard for the communities. A value of *Q* = 0 indicates that the community structure is no stronger than would be expected by random chance and values other than zero represent deviations from randomness. Local peaks in the modularity during the progress of the community struc-

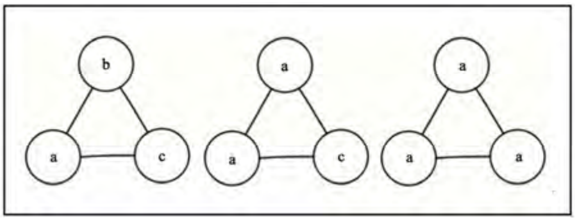


Figure 1. An example of label propagation algorithm.

ture algorithm indicate particularly good divisions of the network. The definition and application of the modularity is independent of the particular community structure algorithm used, and it can therefore also be applied to any other algorithm [3]. In practice, the modularity Q typically falls in the range from about 0.3 to 0.7 [9].

3. Game Virtual World and Datasets

World of Warcraft (WoW) is one of the most popular MMOGs, which has attracted 11.5 millions players online till Dec. 24, 2008. The virtual game world consists of many kinds of regions, and those game world regions can be classified into two categories: some regions which can be reached by all players are called world regions, while the others which only accept group players who must join in a team are called dungeons. Thus, we can identify which dungeons and activities the players are participating by the locations of them. At the same time, one dungeon can accept multiple groups simultaneously, but each group is in its independent environment. Therefore, players of the same group can not be in another group dungeon environment, while players in the same dungeon may belong to different groups. By the way, almost all players in one dungeon leave with their teammates concurrently.

To get application-level traces, we set up a monitor as a client. The monitor is a PC which is equipped with a 2.4 GHz Pentium 4 CPU and 1GB RAM. The monitor is attached to game program as a plugin. The modified client disguises itself as a normal player, who sends requests to game server periodically. A little like database query, the client can initiate many kinds of requests according to different query conditions, such as level, race and profess sion. Every response includes the information about all players who satisfy the query standard. The information contains the profile of each player: level, race, profession, location, guild, name, and online time. A round of complete scan for players lasts for 10 minutes which contain about 1200 requests.

The datasets record the information from the server-named ‘AoLaJiEr’ in the 2nd (Beijing) game district between Apr. 9, 2006 and Apr. 16, 2006. There are 221053 items in total and each of them records state information of one online player at that moment. There are 91 record hours and 7137 players involved in our datasets. Table 1 and Table 2 display data statistics of our collected traces.

Table 1. Datasets statistics (WoW).

date	time duration	number of records	players
Apr, 9 - Apr, 16	91hrs	22105	7137

Table 2. Region statistics (WoW).

	dungeon	world region
number	53	126
players	5143	11892
records	67068	153985

4. Time-Based Incremental Label Propagation Algorithm

In reality most networks will evolve with time instead of keeping unchanged. For example, in P2P networks, every client will go online and off-line from time to time, thus the topology of the overlay networks will definitely change with the dynamics of these peers. Upon every change of the networks, traditionally the algorithm has to run again to get the results. The disadvantage of this kind of algorithms is its high computational cost and long time consumption. To deal with the above problems, we propose the incremental label propagation algorithm.

4.1. Method

The label propagation algorithm was proposed in [5]. It’s a fast algorithm which tries to find out the cliques with nearly linear time complexity. The idea is rather simple. Each vertex will be allocated the label (group) number randomly at first, and then the label will be changed based on the neighbors’ labels. The vertex will be given the label which the majority of its neighbors have. The algorithm iterates until no change can be made. The label propagation algorithm is proposed in the following steps:

1) bInitialize the labels at all nodes in the network. For a given node, its label is n.

2) Arrange the nodes in the network in a random sequence X.

3) Each node changes its label to maximum number of the same label among its neighbors in the order of sequence X.

4) Iterate the above two steps 2 and 3 until no labels can be changed.

It has been proved that it works well for static networks. However, in reality most networks will evolve with time instead of keeping unchanged. For example, in P2P networks, every client will go online and off-line from time to time, thus the topology of the overlay networks will definitely change with the dynamics of these peers. Upon every change of the networks, traditionally the algorithm has to run again to get the results. The disadvantage of this kind of algorithms is its high computational cost and long time consumption.

To deal with the above problems, we propose the time-based incremental label propagation algorithm in the paper. The algorithm tries to deal with the network changes incrementally. That is, when the new vertex (edge) joins or the old vertex (edge) leaves the network, the algorithm will be executed locally instead of globally. Our idea is simple, but it works effectively. Time domain is discretized into time intervals of the same length. For each round, the algorithm only considers the vertexes or edges changed in previous interval. The algorithm will be run locally and iteratively until no labels can be changed.

In our algorithm, we add the time interval sequence as well as the label number in the label. Thus, a vertex can be labeled as (*t*, *g*) in which t denotes the time sequence while g represents the group this vertex belongs to. When the new edge is added, the labels of the two vertexes which are adjacent to the edge will both be updated. It should be noted that, when there are two different labels which same number of neighbors have, the vertex will follow the label which has bigger *t*. Then these updated vertexes will calculate the new labels until no changes can be made. One example which adopts our time-based incremental label propagation algorithm is drawn in Figure 2.

The formal time-based incremental label propagation algorithm can be defined as follows:

1) For each new edge which is added during time interval t, the two vertexes which are incident to the edge will be labeled as (*t*, *m*) and (*t*, *n*). These two vertexes are recorded as the new labeled nodes.

2) All new labeled nodes and their neighbors are added to the vertex calculation sequence *X* in terms of random order.

3) Each element in sequence *X* is fetched one by one and the new label is determined according to the original label propagation algorithm. Similarly, vertexes whose labels are changed will be recorded as new labeled nodes. If there are two or more different labels with same number of neighbors, we will follow the label of the neighbors with bigger *t*.

Iterate the above two steps 2 and 3 until no labels can be changed.

4.2. Validity of the Algorithm

Newman [11] proposed that the divisions of the network can be evaluated using a measure they call modularity *Q*. We use a real network and the modularity to prove our algorithm's validity.

The network of the well-known "karate club study of Zachary includes 34 vertices, as shown in Figure 3(a). This network is divided into two communities correctly by our algorithm, as shown in Figure 4. The value of the modularity *Q* for this partition is 0.488, which is better than the value 0.381 reported by A. Clauset [12]. The number 3 vertex is grouped correctly with our method, whereas it is grouped wrongly by GN algorithm [11].

However, the algorithm does not converge at the unique result, when the random calculate sequence *X* is different. In Figure 3(a), the partition of the network is exactly same with the real community structure. But the result of Figure 3(b) is different from the real structure.

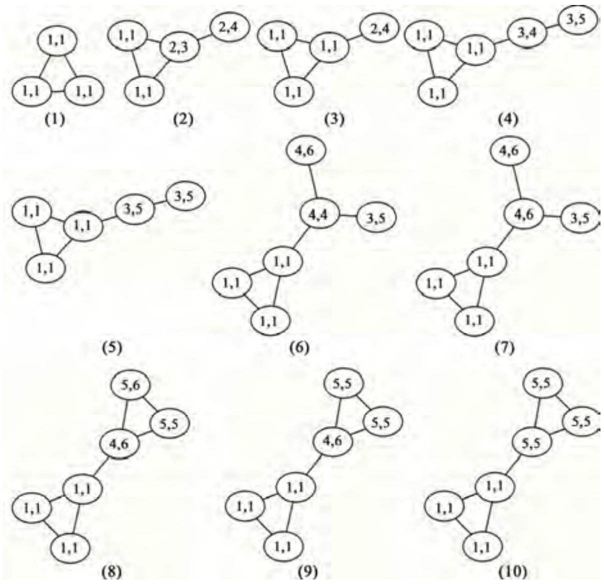
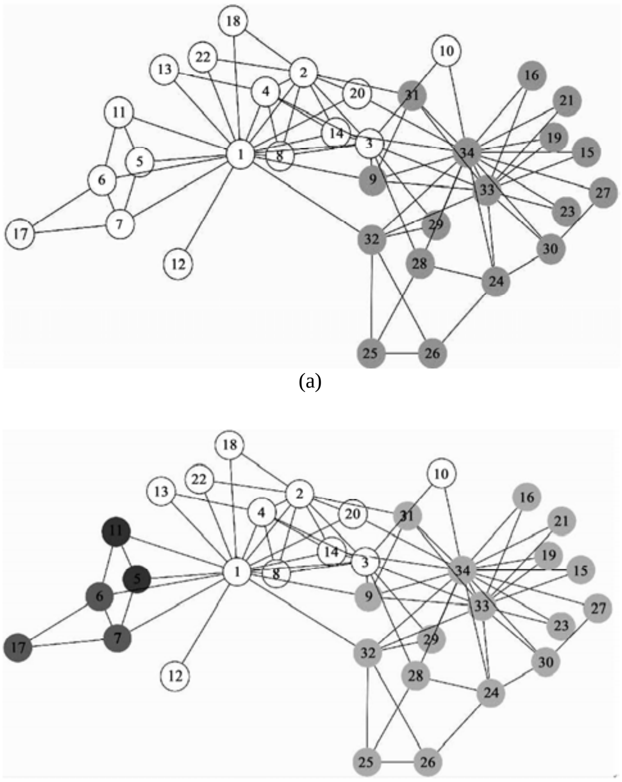


Figure 2. An example of time-based incremental label propagation algorithm.



(a)

(b)

Figure 3. The network of friendships between individuals in the karate club study of Zachary [9]. The administrator and the instructor are represented by nodes 1 and 33, respectively.The communities can be identified by their shades of grey colors. (a) The result is same with real community structure; (b) The result is different from the real community structure.

4.3. Time Complexity

Label propagation algorithm takes near-linear time run its completion [5]. Each iteration of the label propagation algorithm takes linear tim in the number of edges *O*(*m*)).They also found 95% of the nodes or more are classified correctly by the end of iteration 5. When the algorithm terminates, it is possible that two or more disconnected groups of nodes have the same label (the groups are connected in the network via other nodes of different labels). In such case, after the algorithm terminates one can run a simple breadth-first search on the sub-networks of each individual groups to separate the disconnected communities. This requires an overall time of *O*(*m* + *n*).

When the new edge is added, the labels of the two vertexes which are adjacent to the edge will both be updated. The vertexes of the new-added edges and their neighbor vertexes are added into local random calculation sequence X_l It just takes time in the number of local edges connected to the local vertexes ($O(m_l)$). Similarly to the propagation algorithm, the overall time is $O(m_l + n_l)$ in the worst case.

5. Structure Analysis on Multi-Activities

Players can participate in different activities at any time, therefore, for each activity we can get the network of players. For example, in 2006, WoW can provide players with raid activity, party activity and battle ground activity. Among them, only well-organized guilds can play the raid activity, while the party and battle ground activities are free for every player. However, it will be revealed that the players in battle ground are strongly organized too. Thus, different activities which WoW provides combined with the behaviors of the players will result in the network structure of multiple layers, and each layer corresponds to one activity.

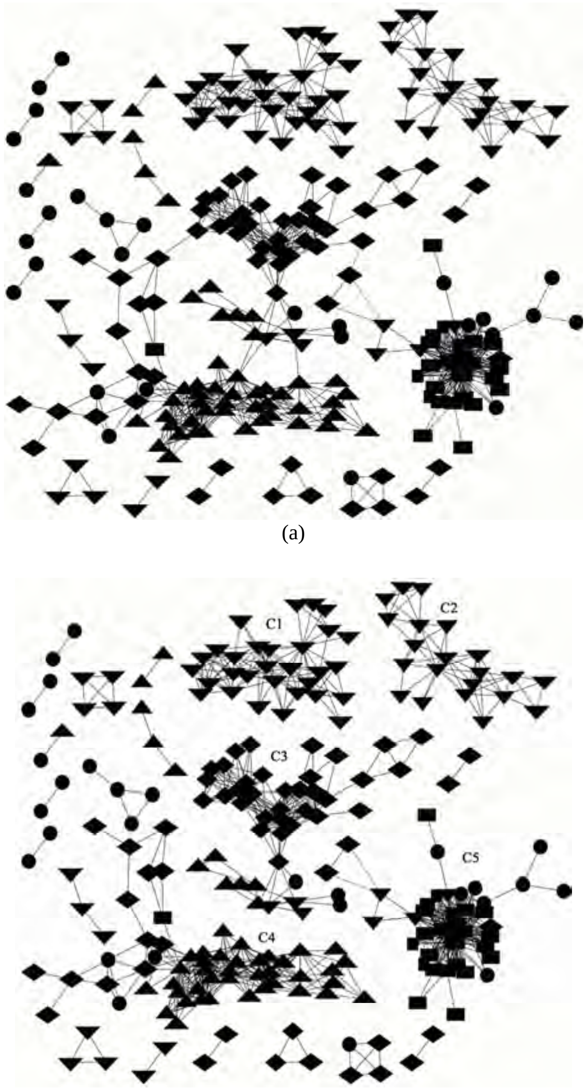
Based on our traces, we preprocess our traces for the convenience of better analysis. Because players with short online time and low level will not possible to participate in all activities, the core network is composed of players with long online time. We only take the 60-level players who kept online for 240 minutes a week in five big guilds as the objects for network structure analysis.

In general, the average shortest length of WoW player network is 3.097 and the clustering coefficient is 0.553, which show that the network is rather dense for players above 30 online minutes a week. Table 3 shows the player number and online time statistics for multiple activities. It's revealed that the numbers of players who take only one or two (and above) activities are 879 and 848 respectively. The number is similar while the total online time is greatly different (the latter is 1.85 times of the former)! It indicates that attracting players to take multiple activities will increase their online time effectively, and the profits of the game operators will be greatly improved for the time-based charging strategies in most MMOG games.

In the section we choose datasets of 60-level players who come from 5 large guilds and play 240 minutes per week as our research targets. The goal of the section is to analyze the community structures of different activities and the relationship among these structures.

Table 3. Table type styles.

number of activities	number of players	total online time (min)
1	879	223370
2	510	209570
3 or above	338	202950

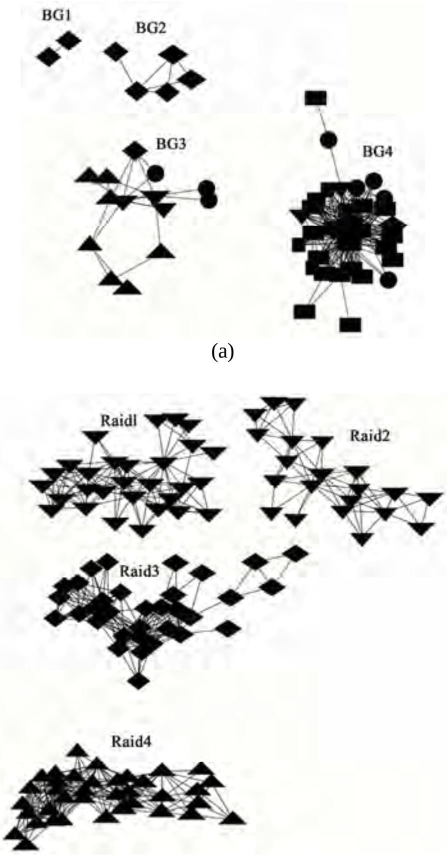

(a)

(b)

Figure 4. Guild vs. activities. (a) Original graph for all activities. (b) Structure graph for all activities.

5.1. Multi-Layer Player Networks

In WoW a player can only take one kind of activity at one time slot, however, a player can take different activities at different time slots. Therefore, supposing we draw a relationship network for each activity, some players who take several kinds of activities can exist in multiple relationship networks. If we project all these layers of networks into one plane, we will get the whole player network. In Figures 4(a) and (b), it's plotted the whole player network. The independent relationship networks of battle ground, raid and party activities are drawn in Figure 5(a), (b) and (c) respectively. Some clusters of players only attend one activity, such as BG4 in (a), Raid1 and Raid2 in (b). However, it's the players who take multiple activities that make the whole player network connectable. For example, the cluster BG3 in (a) connects cluster Raid3 and cluster.



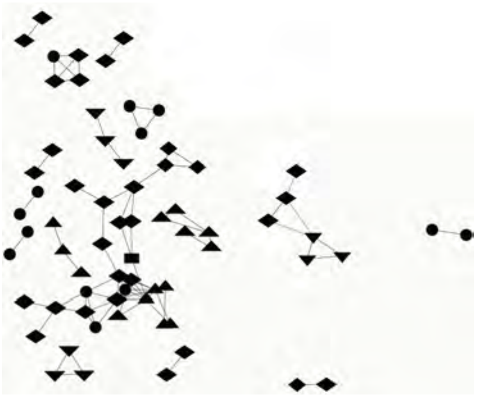


Figure 5. Guild vs. activities. (a) Battle ground activity subgraph; (b) Raid activity subgraph; (c) Party activity subgraph.

Furthermore, considering the extent of overlap between two clusters, we define the overlap coefficient correspondingly. Suppose that there are V_i players in i cluster and V_j ones in j cluster, the number of players who are in both clusters is V_{ij}, thus the overlap coefficient O_{ij} is defined as $O_{ij} = V_{ij} / (V_i + V_j - V_{ij})$. We choose two relationship networks of raid and battle ground activities, and calculate the overlap coefficient between any two clusters from different activities. The maximum and average coefficient for these clusters are 0.182 and 0.022, respectively, which indicates that there are overlaps among the clusters from different activities.

5.2. Guild Organization vs. Activities

There are strong cluster characteristics in battle ground and raid activities, which can be observed in Figure 5(a) and (b) respectively. From the design philosophy of WoW, raid activity will definitely show strong cluster effect because the task must be well organized in advance. To our surprise, from the random organized battle ground activity, we can also observe the effect, and the clustering coefficient of BG4 is rather high (0.581). The cluster is gradually formed instead of planned beforehand.

In addition, in Figure 4 and Figure 5 different node shapes are adopted and they represent different guilds in WoW. Guild in an in-game organization is for real-time communication and chatting. In WoW a player can only join in one guild. This mechanism does works well in some activities, such as raid activities in Figure 5(b). However, in some activities, one cluster is often composed of players who come from several guilds, such as BG3 and BG4 Figure 5(a) in battle ground activity, though there is strong connectivity in the cluster.

To measure the relationship between guild structures and activities, we propose the concept of guild organization coefficient C_g. Firstly, the clusters are found by the time-based incremental label propagation algorithm. In a cluster, if there are fewer guilds in the cluster, or there is higher ratio of nodes in the biggest guild, the cluster shows stronger organization. The definitions of the parameters are shown in Table 4. We define the guild organization coefficient as the average of $C_i = \max$

$(G_k / V_i) / g$ throughout all clusters in certain activity.

In Table 5 we list the guild organization coefficient on each activity. The guild organization coefficient for well-organized activity, such as raid activity, is 1. Though there is no need for the organization of party activity, the coefficient reaches to 0.536, too. For battle ground activity, we have shown in previous subsection that stable organization is required since there are strong clusters in the activity. However, the relative low coefficient (0.376) shows that the present guild organization cannot meet this requirement. In all, the guild regulations should be developed according to the activities for player may play different activities with different swarm of players. The limit that one player can only join in one guild should be eliminated.

6. Conclusions

In the paper, we measured one of the most popular MMOG called WoW and traced the behaviors of many players. It's revealed the structure characteristics how the players are organized. The overlap among different activities can make the whole player network connectable. Some key players who are active in multiple activities take responsibility to connect players of these activities. Furthermore, there are some activities which show great organization while some don't. In addition, the guild organization works well for some activities and doesn't for some others. New policies and organization style should be promoted.

In the future, we will collect datasets for longer time and try to analyze these characteristics theoretically. We'll also try to propose a network structure model to satisfy the development of multi-activities.

7. Acknowledgment

This work is supported by China 973 2007CB307101 and the National Natural Science Foundation of China (60772043, 60672069).

8. References

[1] N. Ducheneaut and R. J. Moore, "The social side of gaming: A study of interaction patterns in a massively multiplayer online game," In CSCW'04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 360–369, 2004.

[2] T. Fritsch, B. Voigt, and J. Schiller, "Distribution of online hardcore player behavior (How hardcore are you?) ," Netgames'06, 2006.

[3] K. T. Chen and C. L. Lei, "Network game design: Hints and implications of player interaction," Netgames'06, 2006.

[4] K. Luyten, K. Thys, S. Huypens, and K. Coninx, "Telebuddies on the move: Social stitching to enhance the networked gaming experience," Netgames'06, 2006.

[5] U. N. Raghavan, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," Physical Review E, Vol. 76, 036106, 2007.

[6] S. L. Li, C. J. Chen, and L. Li, "Using group interaction of players to prevent in-game cheat in network games," First International Symposium on Data, Privacy and E-Commerce, pp. 47–49, 2007.

[7] T. W. Brignall, III, and T. L. Van Valey, "An online community as a new tribalism: The world of warcraft," Proceedings of the 40th Annual Hawaii International Conference on System Sciences, 2007.

[8] L. C. Rodrigues, C. A. M. Lima, P. P. B. de Oliveira, and P. N. Mustaro, "Clusterization of an online game community through self-organizing maps and an evolved fuzzy system," Fourth International Conference on Natural Computation, 2008.

[9] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," Physics Review E, Vol. 69, 026113, 2004.

[10] M. E. J. Newman, "Detecting community structure in network," The European Physical Journal B - Condensed Matter, Vol. 38, No. 2, 2004.

[11] M. E. J. Newman, "Modularity and community structure in networks," Physics 0602124, Vo1, 17, February 2006.

[12] Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," Physics Review E, Vol. 70, 06111, 2004.

Int. J. Communications, Network and System Sciences, **2010, 3, 140-151**
doi:10.4236/ijcns.2010.32021 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Class of Service Support Layer for Wireless Mesh Networks

Jukka SUHONEN, Timo D. HÄMÄLÄINEN, Marko HÄNNIKÄINEN
Department of Computer Systems, Tampere University of Technology, Tampere, Finland
Email: *jukka.suhonen@tut.fi, timo.d.hamalainen@tut.fi, marko.hannikainen@tut.fi*
Received November 12, 2009; *revised December* 23, 2009; *accepted January* 9, 2010

Abstract

This paper presents an add-on Class of Service (CoS) layer for wireless mesh networks. The proposed protocol is applicable for contention-based MACs and is therefore compatible with most of the Wireless Local Area Network (WLAN) and Wireless Sensor Network (WSN) protocols. The protocol has a locally centralized control for managing data flows, which either reserve a fixed bandwidth or are weighted by fair scheduling. The protocol reduces transmission collisions, thus improving the overall throughput. IEEE 802.11 adhoc WLAN has been taken as a platform for simulations and prototyping for evaluating the protocol performance. Network Simulator Version 2 (NS2) simulations show that the CoS protocol efficiently differentiates bandwidth, supports bandwidth reservations, and reaches less than 10 ms transfer delay on IEEE 802.11b WLAN. Testing with a full prototype implementation verified the performance of the protocol.

Keywords: Wireless LAN, Wireless Sensor Network, Scheduling, Traffic Control (Communication), Quality of Service

1. Introduction

A wireless mesh network consist even thousands of devices communicating with multihop routing. Examples of wireless mesh technologies are Wireless LANs (WLANs) and Wireless Sensor Networks (WSNs). Compared to WSNs, WLAN has typically higher capacity and comprises more powerful devices such as laptops and cell phones. A WSN consists of small, low cost, and autonomic devices combining environmental sensing, data processing, and wireless communication. Unlike the traditional computer networks, a WSN is application-oriented and deployed for a specific task e.g. monitoring a physical phenomenon, object tracking, or control purposes. The WSN application space ranges from home and industrial automation to military uses. WSN nodes are typically deployed without planning in large quantities, necessitating self-configurability and distributed operation.

In general, the wireless networking offers easy deployment and mobility but has limited capacity due to constraints in radio frequency bands. Still, bandwidth is needed e.g. for file transfer and video conferencing in WLANs or the retrieval of surveillance images in multimedia WSNs [1]. As several devices share the common wireless medium, a network may congest causing rapidly increasing and varying transfer delays. The network congestion is prone to occur, when multiple users or devices are utilizing the wireless medium such as in open access WLANs or dense WSNs. Thus, Quality of Service (QoS) guarantees are required to ensure sufficient bandwidth and delays for each application.

Varying QoS requirements necessitates traffic differentiation to ensure that the important traffic is preferred when a network congests. Lower priority traffic should receive least bandwidth while higher priority traffic should get low channel access delays. Also, to prevent service degradation, a certain bandwidth should be guaranteed for constant bit rate applications such as voice and video.

The wireless medium can be shared with contention-based or contention-free approaches. The contention based approach, such as Carrier Sense Multiple Access (CSMA), has low overhead and divides bandwidth on-demand basis. However, as the offered load increases, additional methods are required to coordinate transmissions to avoid capacity loss due to collisions. Contention-free channel access eliminates collisions by assigning dedicated (reserved) communication times for a device but requires coordination between devices thus introducing additional overhead. As adjusting reservations to varying traffic requirements is complex, the majority of the popular WLANs and WSNs use CSMA based approaches or reservations are optional extensions for CSMA.

This paper presents the design, implementation, and performance evaluation of a Class of Service (CoS) support layer for wireless mesh networks referred to as Class of Service Protocol (CoSPro). CoS is an approach for delivering QoS by dividing traffic into several classes and providing differentiated service for each class. The key principle is to elect a controller device to manage traffic and grant other devices within its vicinity permissions to access medium.

Bandwidth reservation and traffic differentiation algorithms are individually well researched in the literature. However, in the existing protocols and research proposals, the coexistence of these methods requires a specific Medium Access Control (MAC) protocol. This paper presents a novel per flow QoS model that provides both bandwidth guarantees with reservations and priority based traffic differentiation. CoSPro assumes only contention-based channel access and is compatible with wide range of existing networks such as IEEE 802.11 WLAN and IEEE 802.15.4 Low Rate Wireless Personal Area Network (LR-WPAN).

The overhead of the protocol is reduced by avoiding signaling on lightly loaded network when contention-based channel access is sufficient to guarantee acceptable performance. Thus, CoSPro achieves substantially lower delays than traditional polling protocols. Unlike traditional resource reservation protocols, unused reservations are not wasted in CoSPro as the bandwidth control algorithm assigns unused capacity to the priority based traffic flows. The feasibility of the protocol is verified with simulations and an implementation on IEEE 802.11 WLAN. IEEE 802.11 is used as an example of throughput oriented ad-hoc network technology. The results are generalizable to other contention-based MAC protocols and multimedia WSNs utilizing bandwidth critical applications.

The paper is organized as follows. Section 2 presents the related research proposals. Section 3 describes the design of CoSPro. Section 4 provides the protocol performance evaluation and simulation results. Section 5 presents the implemented prototype and gives the measurement results and Section 6 concludes the paper.

2. Related Work

This section discusses the relation and applicability of CoSPro to the popular wireless WLAN and WSN standards and the related wireless QoS proposals. As CoSPro is a MAC layer enhancement, only the standards that define channel access are considered. In the related QoS proposals, proprietary QoS proposals that define completely new MAC or routing layers are not listed due to incompatibility with the existing standards.

2.1. Wireless MAC Standards

IEEE 802.11 is the most widely utilized WLAN standard that provides the nominal data rates of 11 Mbit/s, 54 Mbit/s, and 150 Mbit/s with 802.11b, 802.11a/802.11g, and 802.11n extensions. The 802.11n standard can achieve even 600 Mbit/s data rate by multiplexing spatially up to four data streams. The standard defines two topologies, infrastructure and ad-hoc. In the ad-hoc topology, stations communicate directly with each other under the Distributed Coordination Function (DCF). In the infrastructure mode, stations communicate through an Access Point (AP) under DCF or with an optional Point Coordination Function (PCF) that provides contention-free communication via reservations. A common MAC protocol based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) is utilized on top of the heterogeneous radio layer. The probability of collisions is reduced by carrier sensing and waiting a predetermined Inter-Frame Space (IFS) and a randomized backoff period before the transmission of a frame. Also, an optional Request to Send/Clear to Send (RTS/CTS) procedure can be used to avoid collisions due to hidden node problem [2]. However, the RTS/CTS handshake before each transmission causes additional sending delays and introduces a relatively large overhead when the size of a data payload is small.

IEEE 802.11e extension defines QoS support for 802.11 MAC by introducing two new communication modes: Enhanced Distributed Channel Access (EDCA) and Hybrid Coordination Function (HCF). EDCA is an extension to DCF, while HCF is used in the infrastructure mode. Traffic differentiation is realized by defining up to 8 traffic categories (TC) per station. A traffic category is associated with certain CW and IFS settings, in which lower values increase the likelihood for channel access and thus better throughput and lower delays. Similarly to the PCF, HCF allows stations to request resources and thus gain a reserved throughput.

The performance of IEEE 802.11e has been extensively studied. Although studies have shown that 802.11e provides relatively good service differentiation [3,4], the standard has also shortcomings. While the QoS provided by EDCA have been found notably better than best-effort, it does not have real QoS guarantees [5,6]. Also, the adjustment of CW and IFS sizes is problematic [7]. With CW, high-priority traffic susceptible to degradation when heavy low-priority exists. IFS size adjustment has been found to provide more efficient service differentiation, but it tends to starve lower priority traffic.

IEEE 802.11s is draft amendment for mesh networking. It specifies a mesh routing protocol and a new coordination function that allows reservations without a centrally coordination device. The reservation method is based on scheduled channel usage. A source device negotiates either a periodic or one-time dedicated channel access time with the destination device. During the negotiation, the destination exposes the reservation schedules of its neighbors, which allows finding a non-conflicting communication time [8]. The priority in which reservations are granted is based on the IEEE 802.11e and shares its benefits and problems.

Unlike WLANs, WSNs do not have a dominating standard as resource constrained hardware necessitates a trade-off between energy, complexity, and performance thus denying a fit-for-all protocol. IEEE 802.15.4 LR-WPAN, WirelessHART, and ISA 100.11a are the three most prominent standards defining WSN MAC. In addition to these standards, the use of proprietary, deployment specific techniques is popular.

IEEE 802.15.4 uses CSMA/CA approach that is similar to the 802.11. A high-band physical layer operating in the 2.4 GHz band uses 250 kbit/s nominal data rate, while the nominal data-rates in 868 MHz and 915 MHz bands are 20 kbit/s or 40 kbit/s, respectively. IEEE 802.15.4 network supports three types of network devices: a Personal Area Network (PAN) coordinator, coordinators, and devices. The PAN coordinator initiates the network and operates often as a gateway to other networks. Coordinators collaborate with each other for data routing and network self-organization. Devices do not have data routing capability and can communicate only with coordinators. An optional low duty cycle operation allows nodes to save energy as the transceiver is active only part of the time. To reduce complexity, RTS/CTS procedure is not supported and backoff period is not increased upon collisions, which increases the probability of collisions and therefore decreases performance. IEEE 802.15.4 allows contention-free slots but does not otherwise support QoS.

WirelessHART and ISA 100.11a define a whole protocol stack including application profiles, routing and transport, and MAC. WirelessHART builds on top of the IEEE 802.15.4 physical layer but defines own MAC protocol. WirelessHART guarantees certain capacity and latency by globally scheduling transmission times for flow on each device. While the approach is acceptable for industrial applications in which reliability is the main concern and network is static and predictable, the approach is too inflexible for generic networks as it requires global knowledge of the traffic patterns. ISA 100.11a reuses IEEE 802.15.4 MAC and physical layers, but adds channel hopping and blacklisting. The network and transport layers in ISA 100.11a are IPv6 based and support various transport services, such as best-effort and real-time.

In WSNs, the scope of the CoSPro differs from Wireless HART and ISA 100.11a that are targeted at industrial, low-traffic networks. CoSPro is targeted at higher rate multimedia WSNs where network congestion is a problem.

CoSPro is compatible with IEEE 802.11 and 802.15.4 standards and solves several of their shortcomings. Differentiated traffic flows can utilize unused reservations, unlike in 802.11/802.15.4 standards where the unused reservations waste capacity. In addition, the CoSPro traffic control limits competition on air interface, and thus prevents performance drop caused by collisions typically seen in these protocols [9–11].

2.2. Wireless QoS Proposals

Several protocols and methods have been introduced for providing QoS over the Internet Protocol (IP). Popular choices are Resource Reservation Protocol (RSVP) and Differentiated Services (DiffServ) defined by the Internet Engineering Task Force (IETF). RSVP uses resource reservation, while DiffServ offers CoS by utilizing the IP header Type of Service (TOS) field and defining a basic set of rules for differentiating packet forwarding in the routers. DiffServ and RSVP can be used in conjunction of other protocols. For example, the IETF defined Multi-Protocol Label Switching (MPLS) has a support for DiffServ. MPLS makes network management easier for QoS by creating a specific path for a sequence of packets. Also, Subnet Bandwidth Manager (SBM) is a RSVP based admission control protocol targeted for IEEE 802-style LANs. The IP-based methods are complementary to the CoSPro, as they ensure end-to-end QoS and can utilize CoSPro to manage per-hop service.

Several modifications for providing CoS for WLAN have been proposed in research papers. Most of these concentrate on improving the WLAN MAC protocol. The proposed modifications often utilize the changing of the size of the contention window or backoff interval of the CSMA-algorithm [12,13]. Another method is to change the length of the IFS. Since a shorter IFS gives a higher priority access, some proposals present an enhanced service support by changing the type or size of IFS [14,15]. While the concepts of these methods can be used with both DCF and PCF, few modifications concentrate solely on PCF [16,17]. As the IEEE 802.15.4 channel access is conceptually similar to the 802.11 WLAN, similar backoff and IFS based QoS have been proposed [10,18,19]. Other related methods for achieving CoS on wireless networks include reservation based protocols [20], enhanced support and adaptation of a particular higher layer protocol (mainly UDP and TCP) [21,22], and admission control [23].

The protocol presented in [23] estimates available bandwidth and prevents congestion with admission control for real-time flows and rate control for non-real time flows. In the protocol, the bandwidth usage estimation is a critical issue and the protocol needs to be integrated with MAC to get the extensive channel usage information.

A number of research articles address the performance problems encountered in the IEEE 802.11 and 802.15.4

channel access. Performance can be enhanced by adapting to the channel in order to compensate channel errors [24,25], by tuning the CSMA algorithm to improve throughput [26], or by avoiding hidden node collisions in multihop networks [27]. However, these MAC layer proposals do not usually provide differentiated QoS.

A problem with the MAC layer research proposals for QoS is the guaranteeing of a certain bandwidth for a single flow. The bandwidth reservation can be provided by a network layer solution, but all traffic belonging to the bandwidth reserved with such method must still contend with other traffic on the link layer. The limited capacity and varying radio environment in wireless networks demand constraining offered load [28,29].

Unlike most of the related proposals, CoSPro provides both bandwidth control according to the application requirements and solves the performance degradation seen in IEEE 802.11 and 802.15.4 with a high number of devices. CoSPro can be implemented without modifying the underlying MAC layer. Also, the bandwidth control supports both bandwidth reservation and traffic classification.

3. CoSPro Design

The lack of QoS is generally not a problem if network is lightly loaded, as all devices get desired bandwidth and buffering delays remain small. The design of CoSPro relies on this fact by controlling bandwidth usage to prevent network congestion and thus performance problems. Minimizing control overhead is another key factor in the design; control messaging is relaxed when the network is not congested.

CoSPro uses locally centralized approach to control traffic in one-hop radius. The design extends to mesh networks, where separate parts of the network are controlled independently. In mesh networks, nodes make forwarding decisions independently resulting into a self-configuring and self-healing multi-hop topology. With CoSPro, each mesh node belongs to a certain local traffic control area managed by a controller device. The controlled devices are referred to as end devices. Data transfer can take place between the controller and end devices or directly between end devices. In mesh networks, a controller manages traffic in a virtual cluster of devices as shown in Figure 1. The clustering algorithm should minimize the number of clusters by selecting one controller within communication range. Suboptimal clustering (several or no controllers for an end device) lowers performance but does not prevent device operation, as CoSPro is built on top of a contention-based channel access that will work regardless of the existence of a CoSPro controller. Defining the clustering algorithm is outside the scope of this paper, but examples of suitable algorithms are presented in [30] and [31]. CoSPro can be applied directly to clustered networks, such as in IEEE 802.15.4 operating in cluster-tree mode. In these networks, cluster heads forward traffic from their child nodes. This way, a child node saves energy as its transceiver can be turned off most of the time.

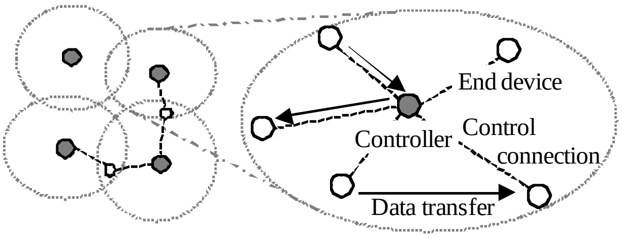


Figure 1. The CoSPro mesh network topology.

CoSPro controls traffic per hop basis and can thus be realized either on the mesh routing or MAC layer as shown in Figure 2. The implementation on MAC layer decreases overhead as messages and status information can be piggybacked to existing Protocol Data Units (PDUs), but may not be feasible e.g. when MAC protocol is hardware implemented. An application can utilize CoSPro directly through an Application Programming Interface (API), in which case application packets are fitted into a CoSPro PDU, or indirectly by differentiating traffic based on its contents (e.g. TOS field). In the latter case, traffic is controlled by predefined rule sets and CoSPro is transparent to applications, allowing e.g. the use of IP protocols with the IPv6 over Low power WPANs (6LoWPAN) technique. With 6LoWPAN, CoSPro is implemented on mesh routing layer (mesh under) instead of IP layer (route over). The protocol analysis and implementation in this paper assume the routing layer and the API approaches.

The functional model of the CoSPro end device is presented in Figure 3. The CoSPro classifier prepares application data for sending and puts data to transmission queue according to its traffic class. The local scheduler retrieves packets from queue according to the flow settings. The phases in sending of a PDU are managed by the CoSPro control, which communicates with the controller device. The CoSPro control also allows an application to set flow priorities and ask for bandwidth reservations. The received data is passed transparently to the application.

The model of the controller device is presented in Figure 4. The network scheduler is used for handling differentiated traffic and the reservation table holds reserved flow information. CoSPro uses the medium adaptation to provide information about the usage of the medium and its capabilities, such as data rate. The usage of the medium affects the decision to enter or leave the congestion mode, whereas reservations utilize knowledge of the available data rates. A configuration for limiting allocations and the value of the congestion threshold is determined by the policy control.

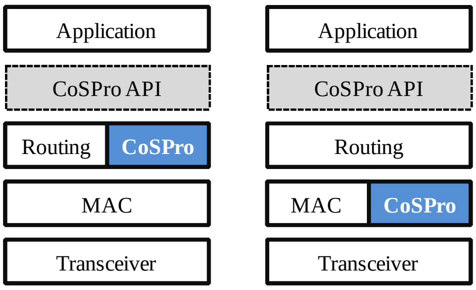


Figure 2. CoSPro protocol architecture alternatives.

3.1. Traffic Management

The goal of the CoSPro traffic management is a versatile traffic differentiation while assuming only simplest of contention based channel access protocols. Both prioritized channel access and bandwidth guarantees are supported, and applications using different types of traffic can coexist seamlessly in a network.

CoSPro manages traffic in data flows. Each end device can originate or terminate one or more flows. A flow uses one of the two modes provided by CoSPro, which are the reserved bandwidth mode and differentiated bandwidth mode. The modes define how the available bandwidth is divided between active devices. Both reserved and differentiated mode flows coexist in the network. The reserved bandwidth mode allocates a fixed bandwidth for a flow and is thus suitable for satisfying flows with strict bandwidth requirements. CoSPro uses the remaining capacity for the differentiated bandwidth mode traffic, which provides controllable bandwidth division according to traffic classes. A controller manages flows only in one-hop radius. For multi-hop reservations, CoSPro can be used with conjunction of end-to-end bandwidth reservations protocols such as RSVP.

The CoSPro controller keeps record of average bandwidth usage in the network by receiving periodic channel usage updates from active devices. A device that has not transmitted traffic does not need to send updates. An end device can send freely with the differentiated mode as long as the used bandwidth remains under a network specific congestion threshold. When the threshold is exceeded, CoSPro considers the network to be congested and the devices must request a permission to send differentiated mode flows. The use of congestion threshold significantly reduces signaling overhead on lightly loaded network, since there is no need to exchange request and grant signals.

The QoS states used in a CoSPro end device is depicted in Figure 5. An end device changes its congestion state upon receiving an indication from the controller. All flows in a device share the congestion state. A flow is initially in the differentiated bandwidth mode, but enters the reserved bandwidth mode when the end device containing the flow requests and receives a reservation update from the controller device. The flow is set back to differentiated mode, if the reservation is dropped.

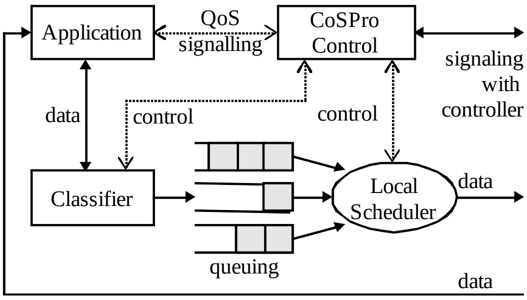


Figure 3. Functional architecture of the CoSPro end device.

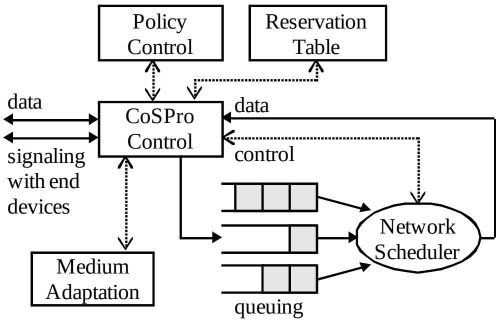


Figure 4. Functional architecture of the CoSPro controller.

3.2. Traffic Classes

Traffic class is a mean to define and provide support for different traffic types. For ease of configurability, a traffic class must be defined in an understandable manner. For example, the level of service in 802.11e defined traffic classes (background, best-effort, video, and voice) related to each other is not obvious. Adjusting the traffic classes for the exact level of service requires experimenting with different values. The definition of CoSPro traffic class allows predictable and easily configurable level of service: bandwidth is divided fairly based on defined weights (priorities).

A CoSPro traffic class is consists of priority and aging time parameters. Each flow is assigned with a traffic class that can be negotiated between controller and an end device. For example, in the prototype implementation, an end device defines the used traffic class by sending the parameters during an initial handshake procedure with the controller. The aging time defines the maximum transfer time for a PDU. After the aging time has elapsed, a queued packet is discarded as obsolete. The interpretation of the priority parameter differs slightly between the reserved and differentiated bandwidth modes. In the reserved traffic mode, priority indicates whether the controller device should accept or reject a new allocation request. In the differentiated mode, the priority parameter defines how the bandwidth is di-

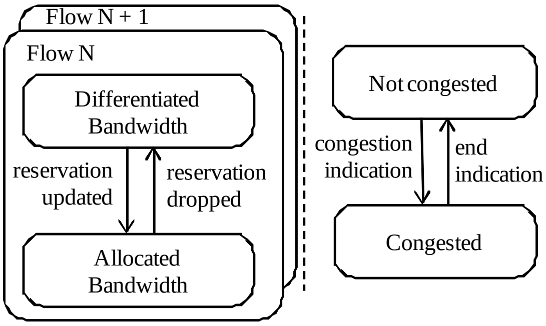


Figure 5. QoS states in a CoSPro end device.

vided among traffic flows. Also, low priority reservations can be dropped if differentiated traffic contains a lot of high priority flows.

3.3. Reserved Bandwidth Algorithm

In the reserved bandwidth mode, controller grants a portion of total capacity for a flow. The reserved bandwidth mode relies on the fact that contention based MAC protocols support strict QoS requirements, when offered traffic load is controlled and not near the maximum capacity [32]. Therefore, by ensuring that used bandwidth does not exceed the maximum capacity, collisions are rare and each flow gets statistical throughput and delay guarantees.

A device reserves bandwidth by sending a request containing minimum required and preferred bandwidth to the controller. Controller assigns each reserved flow at least the minimum requested bandwidth. If the requested bandwidth exceeds available capacity, lowest priority flows are dropped. Surplus bandwidth is assigned first to the higher priority flows. Controller may redefine or cancel the reservation to enable admission for a higher priority flow. The local scheduler of an end device limits traffic to the negotiated bandwidth by delaying the retrieval of next packet for a flow, thus limiting the number of reserved flows accessing air interface simultaneously.

3.4. Differentiated Bandwidth Algorithm

In the differentiated bandwidth mode, an end device requests the controller for permission to send data. This is a common approach in polling mechanisms that has the main drawback of increased overhead and latency [33]. CoSPro introduces several methods to improve the polling efficiency. First, polling is used only when the network is congested. Second, a node may send several packets when it gets the permission, which reduces the per packet overhead. Third, instead of assigning each device a transmission opportunity periodically [34], CoSPro grants transmission periods dynamically based on the knowledge of flow activity and queued frames. Therefore, CoSPro is able to control latency and bandwidth.

The differentiated bandwidth mode uses a weighted fairness algorithm, in which the flows are given bandwidth in proportion to their weight. In CoSPro, a priority parameter p_i is used as the weight of a flow f_i. A flow f_i gets an average bandwidth of b_i as

$$b_i = \frac{p_i}{\sum_{j=0}^{N-1} p_j} B_{diff} , \tag{1}$$

where N is the amount of active flows and B_{diff} is the bandwidth available to differentiated mode flows. The adjustment of priority is intuitive, for example a class with priority of four receives twice as much bandwidth as class with priority of two. The value for B_{diff} is calculated as

$$B_{diff} = B_{total} - \sum_{i=0}^{N-1} r_i , \tag{2}$$

where B_{total} is the total available bandwidth and r_i is the actual bandwidth utilized by reserved mode flow f_i.

A device signals the status of its local queue within a transmission request containing information about each active flow. The controller queues the requests, uses the network scheduler to determine the next request, and sends an allowed transmit indication to an appropriate device. The indication contains the allowed traffic class, the length of the transmission period, and a limit to the sending bandwidth. The bandwidth limit is determined by (2). Thus, the limit ensures that enough capacity is available for reserved flows, but also utilizes bandwidth that is unused by reserved flows. Consequently, a differentiated mode flow can send data without bandwidth limitations when reservations do not exist.

A device stops sending, when the transmission period length is reached. Next, the device sends end transmission indication to the controller, which contains the same information than traffic request. The message denotes that the device has completed sending during this period and eliminates the need to update the flow status with a new traffic request message. The controller may interrupt an active device with deny transmit indication, if the controller has a request that must be scheduled immediately. If the controller does not receive end indication, it schedules next request after the transmission period has ended.

3.5. Differentiated Mode Scheduling Algorithm

This section presents one possible algorithm for providing bandwidth differentiation as shown in (1). The algorithm realizes bandwidth allocation by scheduling transmission opportunities based on the flow priority values, requests sent by the devices, and the amount of successfully transmitted traffic. The use of transmitted traffic ensures that bandwidth differentiation is fair between flows using different packet sizes and devices having differrent packet error rates. For example, a device experiencing high packet error rate can still get more bandwidth than lower priority flows from devices having a better link. The use of an algorithm that considers aging time as a scheduling deadline could improve real-time scheduling for traffic with same priorities [35,36]. However, it would not provide differences in bandwidths, which is the approach taken in CoSPro.

The scheduler in the CoSPro controller device has a list of active flows *F*. A flow *f* is marked active after a traffic request concerning it is received. Each request contains the number of PDUs waiting for sending at the end device, denoted as c, and the average size of waiting PDUs, denoted as a. The scheduler selects one of the active flows and gives it a permission to send PDUs for the duration of *d*. A flow is marked inactive after it is scheduled.

For the scheduling decision, the scheduler keeps a record of received traffic from each flow, denoted as t_i. The scheduled flow is selected by weighting the amount of received traffic by the priority of the flow. The weighted traffic value w_i is derived as

$$w_i = t_i / p_i. \tag{3}$$

The flow f_i that has the smallest weighted traffic value w_i is selected. If there are several flows having the same weight, the one with the highest priority is selected. The purpose of the selection is to keep the traffic weights equal. As the weights are equal, the transferred traffic and therefore the obtained bandwidth are proportional to the priority values as (1) requires. For ensuring that the w_i calculation is not distorted by different device activities, the scheduler reduces the recorded received traffic counters (t_i) periodically. Otherwise, the devices that have not sent for a long time would get more bandwidth than defined.

After the flow has been selected, the scheduler determines the time the device is allowed to send PDUs belonging to the flow. This is done by calculating the length of the transmission period *d*. First, the amount of PDUs is determined, denoted as *n*. Then, *n* is converted to a particular transmission period length *d*.

If *F* contains only one flow, the *n* is set directly to the value c_i contained in the request. Otherwise, the *n* is calculated as

$$n = \min\left(c_i, \frac{(w_k - w_i)p_i}{a_i}\right), \tag{4}$$

where w_k corresponds to flow f_k, which is the flow with the second smallest weighted traffic value in *F*. By comparing the two smallest weighted traffic values, the scheduler can allocate the length of the transmission period for the selected flow before the other flows (from which the traffic requests has been received at the moment) need to be scheduled for transmission. In (4), the actual number of bytes to be sent is derived by multiplying the subtracted weighted traffic values with priority pi. The amount of bytes is converted to the number of PDUs by dividing it with the average PDU size a_i. Taking the minimum prevents any larger n value than the device has requested.

Finally, the length of the transmission period is calculated from the PDU count with the average PDU size and the known network bandwidth *b* as

$$d = na_i \;/\; b. \tag{5}$$

The scheduler has two configurable parameters for controlling the average sending durations: D_{min} and D_{max}. The *d* is assigned within these limits. The limits affect to the overall CoSPro trade-off between transfer delay and throughput. The purpose of the D_{min} parameter is to improve throughput by ensuring that several PDUs can be sent during a transmission period. The D_{max} is used to control the delay that a flow experiences while waiting for a transmission period to be reserved.

4. Performance Analysis

The performance is analyzed when CoSPro is implemented on top of the IEEE 802.11 MAC operating under DCF. The DCF mode was chosen because the proposed centralized control architecture of CoSPro would be interfered by PCF.

4.1. Throughput Analysis

The throughput of CoSPro using differentiated mode is compared to the standard 802.11 with and without the RTS/CTS procedure. CoSPro reserved mode does not include additional messaging and has throughput similar to the standard WLAN. The values are calculated by assuming IEEE 802.11b having 11 Mb/s basic data rate with direct sequence spread spectrum. The overhead caused by IFS sizes, backoff times, and acknowledgements for WLAN MAC PDUs have been accounted in the calculations. The calculated bandwidth presents the best-case throughput for a single device, since the effect of collisions, bit-errors, and retransmissions is ignored. CoSPro overhead includes also transmission requests, allow transmit, and the end of transmission indications.

The calculated throughputs are shown in Figure 6. The frame overhead and acknowledgements cause considerable throughput penalty with frames having a small payload. The throughput penalty of CoSPro differentiated flows depends on the length of the transmission period. Differentiated mode flows shows better results than the standard WLAN with RTS/CTS, when the transmission period length exceeds 10 ms. Compared to the standard WLAN with 1500 B payload, CoSPro has 4.5% lower throughput with 50 ms transmission period when

RTS/CTS is not used and 17% higher throughput when RTS/CTS is used. In addition, CoSPro has a good throughput with a small frame payload, because it allows more frames to be sent within a single transmission period. As the transmission period length is increased, the throughput of CoSPro approaches the standard WLAN values. Because the length of the transmission period is controlled with D_{min} and D_{max} scheduler parameters, the overhead penalty can be adjusted to an acceptable range. In general, longer transmission periods reduce the overhead.

4.2. Delay Analysis

The delivery delay for a PDU contains the time required to transfer it between MAC layer entities and the time it waits for scheduling. The delay can be evaluated by calculating the time an arbitrary flow k must wait for scheduling. Let t_k denote the amount of data belonging to the kth flow. Before sending the kth flow again, the device must wait until equal amount of data has been transferred via other flows. The amount of data is determined by the priority values and can be obtained from (1) and (3) as

$$t_i = (p_i / p_k)t_k , \quad (6)$$

where t_i is the traffic that must be transferred via any other ith flow before the device can sent to the kth flow again. The value of t_i can be converted to time d_i as

$$d_i = (p_i / p_k)d_k , \quad (7)$$

where d_k denotes the time taken to send traffic t_k. It is easily seen that a long transmission period for the kth flow increases the transmission periods of the other flows as well. Since a flow must wait for other flows to finish, the use of long transmission periods increases delays. Also, it should be noted that (7) does not include the time caused by the MAC layer overhead or the CoSPro signaling. For equal sized PDUs, the delay will be divided in the fraction of the flow priorities similarly to the bandwidth.

4.3. Simulated Performance

CoSPro performance was verified with the Network Simulator (NS) version 2 [37] and compared to IEEE 802.11 EDCA [38]. CoSPro was implemented in NS as a new protocol layer, which was located on top of the simulated 802.11 MAC layer. The simulation settings used 11 Mb/s physical link rate and did not include the generation of packet losses, RTS/CTS signaling, or fragmentation.

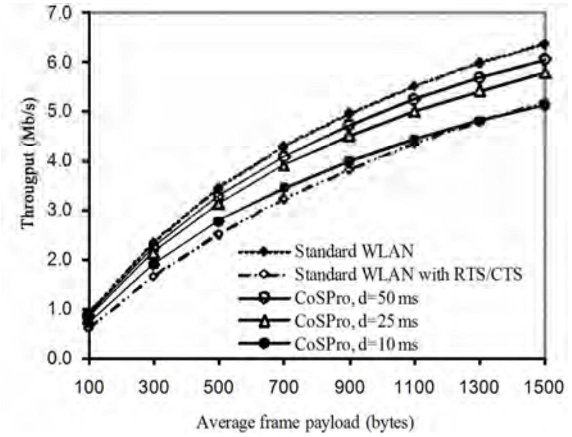


Figure 6. The effect of payload size and transmission period d in differentiated mode CoSPro.

Table 1. Simulation settings with three different traffic flows.

Flow	Packet size (B)	Offered load (kbit/s)	CoSPro Traffic mode	CoSPro Priority	802.11 EDCA AC
Background	1400	400	Differentiated	1	Background
Best-effort	800	200	Differentiated	2	Best-effort
Real-time	204	64	Reserved	-	Video

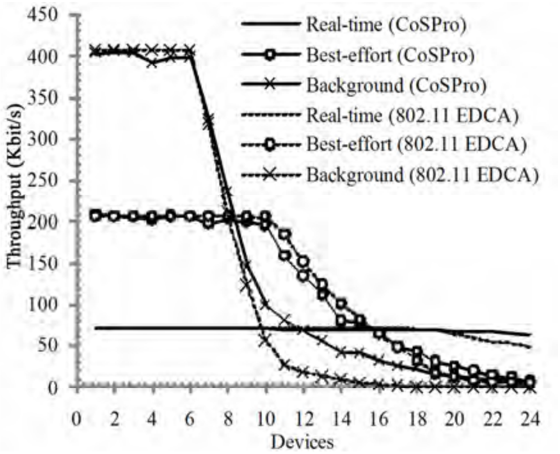


Figure 7. Average throughput of a traffic flow with CoSPro and IEEE 802.11 EDCA.

The simulation measured traffic differentiation and the effect of competition in air interface, by increasing the number of active devices. Each device utilized three flows with real-time, background, and best-effort priorities. Real-time flow was presented typical voice traffic, while background and best-effort flows presented typical data traffic. The offered throughput and simulation settings with CoSPro and IEEE 802.11 EDCA are presented in Table 1. The 802.11 EDCA access categories determine IFS and backoff values and are the same as defined in IEEE 802.11e standard. Video access category was used for real-time traffic, because voice access category performed poorly in EDCA when the number of devices was increased.

Figure 7 presents the average throughput of traffic flows. After 6 active devices, network capacity is exceeded and the throughputs of low priority flows decrease. CoSPro provides better throughput for real-time traffic when the number of active devices increases. The performance drop with 802.11 EDCA is caused by the increased number of collisions, whereas CoSPro limits the number of sending devices thus reducing the competition. In 802.11 EDCA, higher priority traffic causes starvation to the lower priority traffic. CoSPro does not have this problem. Instead, the aggregate throughputs of best-effort and background traffic in CoSPro converge with 24 active devices because of used minimum scheduler time D_{min}, which prevents low priority traffic from starving. As the competition over air interface increases, 802.11 EDCA real-time traffic drops, while CoSPro is able to provide real-time traffic flow sufficient throughput without notable performance drop.

End-to-end packet delays with CoSPro and 802.11 EDCA with 6 active devices are shown in Figure 8. 802.11 EDCA real-time traffic has the smallest delays due to its small IFS and backoff values. As the CoSPro uses the legacy 802.11 IFS and backoff values that are slightly higher than EDCA real-time values, the CoSPro real-time traffic has slightly higher delays. However, the delays are still low enough for the flows to be used for multimedia traffic.

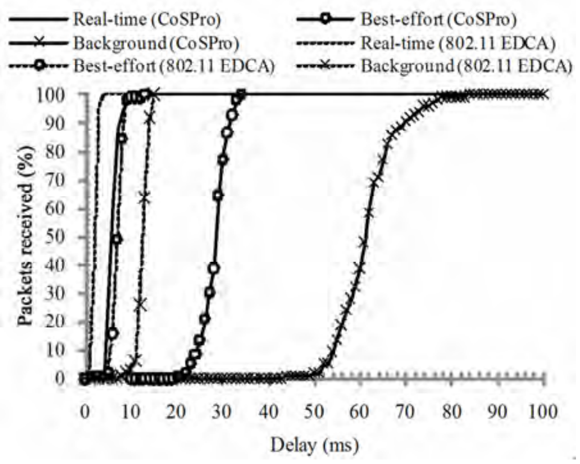


Figure 8. End-to-end packet delay in CoSPro and IEEE 802.11 EDCA with 6 active devices.

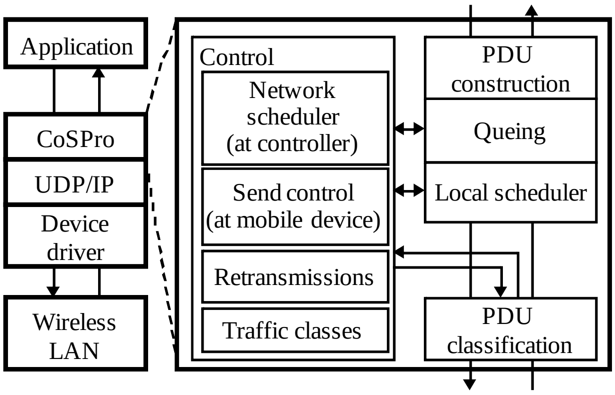


Figure 9. Prototype implementation architecture of CoSPro.

5. Prototype Implementation

The prototype implementation places CoSPro between the application and WLAN link protocol layers, as presented in Figure 9. The prototype consists of the controller and end device software modules that implement CoSPro functionality. In the implemented prototype, CoSPro also operates on top of the User Datagram Protocol/Internet Protocol (UDP/IP) layers. The prototype was built on that layer in order to enable testing with different existing network cards and drivers. Although UDP/IP increases overhead, the layers are otherwise transparent and do not affect the operation of the prototype, since the layers do not employ flow control or acknowledgements.

The prototype architecture is the same for both the end device and the controller device except for the control part. An application data packet is fitted inside CoSPro PDU in the PDU construction function. A CoSPro PDU contains the identification of the used traffic class and the information needed to separate CoSPro control and application data in the PDU classification part of the protocol. The PDU is queued until the local scheduler selects it and sends it by using the lower protocol layers. The local scheduler is controlled by the send control part of CoSPro. The send control communicates in the congestion mode with the network scheduler, which is located at the controller device. The protocol control provides retransmissions for the CoSPro control messages.

The CoSPro modules were implemented as Microsoft Windows 32-bit executables. The executable programs utilized Windows Sockets 2 application programming interface (WinSock2) for networking. The prototype used common PC hardware with Avaya Wireless PC Card, Nokia C110/C111 and Nokia D211 WLAN cards meeting the IEEE 802.11b standard.

6. Prototype Measurements

The measurements were carried out using the implemented prototype. The end devices were located around a central device, which was also the CoSPro controller device. The distances between the devices were less than ten meters, as the purpose was to provide similar link conditions to all devices and to minimize the near-far effect. The central device was used to receive all data sent by the other devices. The results obtained with CoSPro were compared to the standard WLAN. In the standard WLAN tests, devices sent data also to the central device, but the CoSPro software modules were not used.

The measurements utilized the 11 Mbit/s nominal bandwidth of IEEE 802.11b. Each measurement used the packet size of 1400 B, which was selected to minimize the header overhead and to avoid IP datagram fragmentation. In the used WLAN cards, the RTS/CTS procedure and fragmentation were disabled, and the number of transmission retry limit was set to four. Also, the default values were used for other MAC parameters.

6.1. Performance of Priority Scheduling

The priority scheduling test evaluated the performance of the CoSPro scheduling algorithm. The test used four devices utilizing priority values of 2, 4, 6, and 8. Each device started sending at 700 kbit/s and increased the offered throughput to 2.7 Mbit/s during the test.

Figure 10 presents the measured throughput with CoSPro. The CoSPro congestion threshold is exceeded when a device sends more than 1.0 Mbit/s, which totals to more than 4 Mbit/s network load. As a result of the protocol, the devices with higher priorities get better throughput with the expense of lower priorities. The device having the lowest priority traffic carries the penalty of additional overhead, while the other devices are more or less unaffected by the overhead. Throughput of the device using the highest priority increases, until the sending bandwidth reaches the point where the priority settings do not allow the device to increase its share of bandwidth.

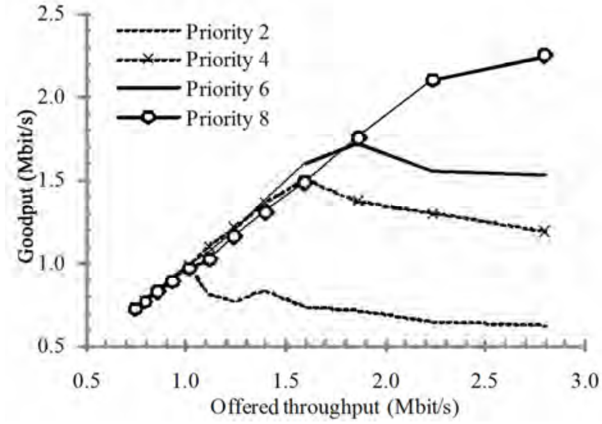


Figure 10. Device throughput with CoSPro.

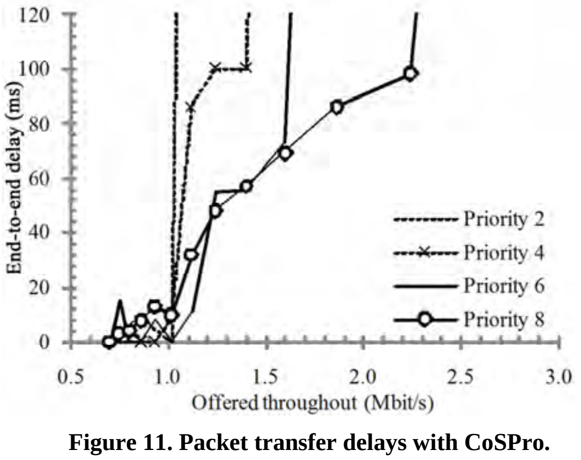


Figure 11. Packet transfer delays with CoSPro.

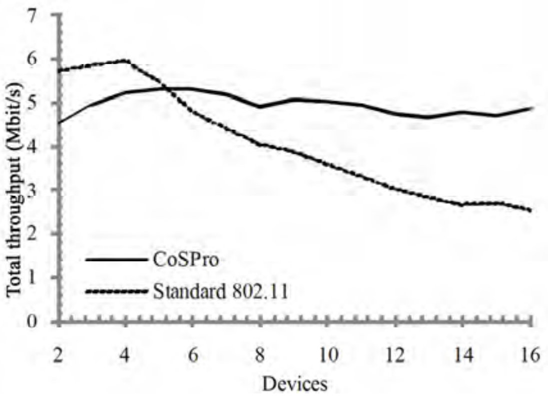


Figure 12. Measured network throughput with and without CoSPro.

The test setup used a scheduler configuration that preferred higher throughput and considered 50-100 ms delays acceptable (D_{min} and D_{max}). Therefore, while the test results show that CoSPro did not cause a notable penalty on throughput, it adds a small delay. The packet queuing and transfer delays from an end device to the controller are presented in Figure 11. End-to-end delays rise when the network congests and packets have to wait for sending. The traffic with a high priority endures well against network congestion.

6.2. Throughput in a Congested Network

The effect of an increased traffic load and congestion on the total network throughput was tested by increasing the number of active devices, while each device offered 5.0 Mbit/s throughputs. Thus, the offered throughput exceeds the IEEE 802.11b practical bandwidth limit already with two active devices.

The measured total network throughput is presented in Figure 12. The network throughput in the standard WLAN achieves its peak with four active devices. After the peak, throughput quickly drops as the number of collisions on the link increases. The signaling overhead of CoSPro is evident with small number of active devices. However, when the number of devices increases, CoSPro limits competition in air interface and provides a higher throughput than standard WLAN. In addition to outperforming the standard WLAN, the throughput stays steady as the number of devices increases.

7. Conclusions

This paper presented a CoS support layer for wireless mesh networks referred to as CoSPro. The proposed protocol is applicable for contention-based MACs and is therefore compatible with most of the popular WLAN and WSN protocols. The proposed protocol is simple but efficient, which makes it easy to implement in existing networks. CoS is realized by managing traffic flow activity with a controller device. Both the simulations and the measurements with the implemented prototype confirm that the protocol differentiates bandwidth and transfer delays for each traffic class. The protocol adds 5-20% overhead depending on configured latency/throughput trade-off, but increases overall throughput on congested network by avoiding collisions. In a non-congested network, CoSPro does not cause any overhead, but the traffic control is activated only when offered load exceeds a defined threshold. The total throughput of CoSPro exceeds the standard 802.11b WLAN with more than 6 competing devices.

The performance of CoSPro can be increased by integrating it more tightly with wireless MACs. The traffic request and response signaling can be implemented in the RTS/CTS procedure. Also, in synchronized MACs, the beacon frame can be utilized to signal the congestion state indication thus reducing CoSPro signaling overhead. On the other hand, when the protocol is implemented on higher layer, it is compatible and complementary to the wireless standards, such as IEEE 802.11 and IEEE 802.15.4.

8. References

[1] I. F. Akyildiz, T. Melodia, K. R. Chowdury, “Wireless multimedia sensor networks: a survey,” IEEE Wireless Communications, pp. 32–39, December 2007.

[2] G. Bianchi, “Performance Analysis of the IEEE 802.11 Distributed Coordination Function,” IEEE Journal on Selected Areas in Communications, vol. 18, pp. 535–547, March 2000.

[3] Y. Xiao, “Enhanced DCF of IEEE 802.11e to support QoS,” in Proceedings of Wireless Communications and Networking Conference (WCNC), 2003, Vol. 2, pp. 1291–1296, March 2003.

[4] P. Garg *et al*., “Using IEEE 802.11e MAC for QoS over Wireless,” in Proc. Performance, Computing, and Communications Conference, pp. 537–542, 2003.

[5] H. Zhu *et al.* “A survey of quality of service in IEEE 802.11 networks,” IEEE Wireless Communication, August 2004.

[6] D. He and C. Q. Shen, “Simulation study of IEEE 802.11e EDCA,” in Proceedings of VTC, 2003.

[7] J. W. Robinson and T. S. Randhawa, “Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function,” IEEE JSAC, Vol. 22, No. 5, June 2004.

[8] G. Hiertz, S. Max, and T. Junge, “IEEE 802.11s – mesh deterministic access,” in Proceedings of the 14th European Wireless Conference, pp. 1–8, 2008.

[9] M. Hännikäinen, M. Niemi, and T. Hämäläinen, “Performance of the Ad-hoc IEEE 802.11b wireless LAN,” in Proceedings of International Conference on Telecommunications, pp. 938–945, June 2002.

[10] S. Pollin *et al.* “Performance analysis of slotted carrier sense IEEE 802.15.4 medium access layer,” IEEE Transactions on Wireless Communications, Vol. 7, No. 9, pp. 3359–3371, September 2008.

[11] T-J Lee, H. R. Lee, and M. Y. Chung, “MAC throughput limits of slotted CSMA/CA in IEEE 802.15.4 WPAN,” IEEE Communication Letters, Vol. 10, No. 7, pp. 561–563, July 2006.

[12] A. Dugar, N. Vaidya, and P. Bahl, “Priority and fair scheduling in a wireless LAN,” in Proceedings Military Communications Conference, Vol. 2, pp. 993–997, 2001.

[13] D. Qiao and K. G. Shin, “Achieving efficient channel utilization and weighted fairness for data communications in IEEE 802.11 WLAN under the DCF,” in Proceedings of Tenth IEEE International Workshop on Quality of Service, pp. 227–236, May 2002.

[14] S. Sheu and T. Sheu, “A bandwidth allocation/sharing/extension protocol for multimedia over IEEE 802.11 ad hoc wireless LANs,” IEEE Journal on Selected Areas in Communication, Vol. 19, October 2001.

[15] I. Aad and C. Castelluccia, “Differentation mechanisms for IEEE 802.11,” in Proceedings of IEEE Infocom, 2001.

[16] T. Suzuki and S. Tasaka, “Performance evaluation of priority-based multimedia transmission with the PCF in an IEEE 802.11 standard wireless LAN,” in Proceedinds of International Symposium on Personal, Indoor and Mobile Radio Communications, Vol. 2, No. 30.

[17] L. Jacob, Q. Qiang, R. Radhakrishna Pillai, and B. Pradhakaran, “MAC protocol enhancements and a distributed scheduler for QoS guarantees over the IEEE 802.11 wireless LANs,” in Proceedings of Vehicular Technology Conference (VTC), Vol. 4, pp. 2410–2413, September 2002.

[18] C. K. Singh, A. Kumar, and P. M. Ameer, “Performance evaluation of an IEEE 802.15.4 sensor network with a star topology,” Wireless Network, Vol. 14, pp. 543–568, 2008.

[19] F. Shu, “Performance evaluation of the IEEE 802.15.4 CSMA-CA protocol with QoS differentiation,” in Proceedings of International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp. 475–480, December 2008.

[20] S. T. Sheu, T. F. Sheu, C. Wu, and J. Y. Luo, “Design and implementation of a reservation-based MAC protocol for voice/data over IEEE 802.11 ad-hoc wireless networks,” in Proceedings of IEEE International Conference on Communications (ICC), Vol. 6, pp. 1935–1939, June 2001.

[21] S. Sharma, K. Gopalan, and N. Zhu, “Quality of service guarantee on 802.11 networks,” in Proceedings Hot Interconnects 9, pp. 99–103, 2001.

[22] S. Garg, M. Kappes, and M. Mani, “Wireless access server for quality of service and location based access control in 802.11 networks,” in Proceedings of Interna-

tional Symposium on Computers and Communications (ISCC'02), pp. 819–824, 2002.

[23] Q. Shen, X. Fang, Pan Li, and X. Fang, "Admission control based on available bandwidth estimation for wireless mesh networks," IEEE Transactions on Vehicular Technology, Vol. 58, No. 5, July 2009.

[24] A. Giorgetti, G. Pasolini, and R. Verdone, "Performance evaluation of a channel adaptive wlan polling protocol," in Proceedings of Vehicular Technology Conference, 56th IEEE, Vol. 3, pp. 1379–1383, 2002.

[25] M. Gidlund, "An approach for using adaptive error control schemes in wireless LAN with CSMA/CA MAC protocol," in Proceedings of Vehicular Techology Conference (VTC), Vol. 1, pp. 224–228, May 2002.

[26] F. Calì, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," IEEE/ACM Transactions on Networking, Vol. 8, No. 6, December 2000.

[27] A. Koubaa, R. Severino, M. Alves, and E. Tovar, "Improving quality-of-service in wireless sensor networks by mitigating 'hidden-node collisions'," IEEE Transactions on Industrial Informatics, pp. 299–313, Vol. 5, No. 3, August 2009.

[28] B. Luca, F. D. Priscoli, T. Inzerilli, P. Mähönen, and L. Muñoz, "Enhancing IP service provision over heterogeneous wireless networks: A path toward 4G," IEEE Communications Magazine, pp. 74–81, August 2001.

[29] G. Xylomenos and G. C. Polyzos, "Link layer support for quality of service on wireless internet links," IEEE Personal Communications, pp. 52–60, October1999.

[30] S. Jardosh and P. Ranjan, "A survey: Topology control for wireless sensor networks," in Proceedings of IEEE International Conference on Signal processing, Communications, and Networking, pp. 422–427, January 2008.

[31] B. Cărbunar, A. Grama, and J. Vitek, "Redundancy and coverage detection in sensor networks," ACT Transactions on Sensor Networks, Vol. 2, pp. 94–128, February 2006.

[32] H. Zhai, X. Chen, and Y. Fang, "How well can the IEEE 802.11 wireless LAN support quality of service," IEEE Transactions on Wireless Communications, Vol. 4, pp. 3084–3094, November 2005.

[33] H. Levy and M. Sidi, "Polling systems: Applications, modeling, and optimization," IEEE Transactions on Communications, Vol. 38, pp. 1750–1760, October 1990.

[34] O. Sharon and E. Altman, "An efficient polling MAC for wireless LANs," IEEE/ACM Transactions on Networking, Vol. 9, pp. 439–451, August 2001.

[35] M. Adamou, S. Khanna, I. Lee, I. Shin, and S. Zhou, "Fair real-time traffic scheduling over a wireless LAN," in Proceedings Real-Time Systems Symposium, 22nd IEEE, pp. 279–288, 2001.

[36] H. Aydin, R. Melhem, D. Mossé, and P. Mejía-Alvarez, "Optimal reward-based scheduling for periodic real-time tasks," IEEE Transactions on Computers, Vol. 50, February 2001.

[37] K. Fall and K. Varadhan, "The ns manual (formerly ns notes and documentation)," [Online]. Available: http://www.isi.edu/nsnam/ns/doc/index.html.

[38] S. Wiethölter, C. Hoene, and A. Wolisz, "Perceptual quality of internet telephony over IEEE 802.11e supporting enhanced DCF and contention free bursting," Technical Report TKN-04-11, Telecommunication Networks Group, Technische Universität Berlin, September 2004.

Int. J. Communications, Network and System Sciences, **2010, 3, 152-166**
doi:10.4236/ijcns.2010.32022 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Adaptive Power Saving Receiver for DVB-H Exploiting Adaptive Modulation and Coding

Tallal Osama ELSHABRAWY, Sherif Hassan Abdel WAHED
Faculty of Information Engineering & Technology, German University, Cairo, Egypt
Email: *Tallal.El-Shabrawy@guc.edu.eg, Sherif.Mohamed@student.guc.edu.eg*
Received November 12, 2009; *revised December* 11, 2009; *accepted January* 18, 2010

Abstract

Broadcasting live digital TV to a small battery-powered handheld device is very challenging. One of the most promising technologies to provide such services is DVB-H (Digital Video Broadcasting over Handheld). Power consumption has always been one of the most crucial challenges for handheld devices. In this paper, a novel Adaptive Modulation and Coding (AMC) framework is proposed for DVB-H systems to address the challenging problem of power consumption. The proposed power saving AMC framework operates by rearranging the transmitted frames in a pre-defined pattern. The adaptive receiver selects the appropriate modulation technique and/or code rate, one that achieves a target Bit Error Rate (BER), and then could be switched off and/or powered down resulting in significant potential for saving of reception and processing powers. Simulation of the DVB-H system under the proposed framework proved that the proposed power saving AMC framework is capable of achieving power saving up to 71.875% in COST207 Typical Urban 6-paths (TU6) channel. Furthermore, numerical analysis for the power saving potential and BER performance of the proposed framework is performed for both flat Rayleigh channel and multipath TU6 channel.

Keywords: Broadcast, DVB-H, Power Consumption, Power Saving, Adaptive Modulation and Coding, AMC, RCPC Codes.

1. Introduction

Multimedia services demand has grown rapidly in the previous decade. As a result, the Digital Video Broadcasting (DVB) project was founded in 1993 by the European Telecommunications Standards Institute (ETSI) with the goal of standardizing digital television services. In order to meet market demands, the DVB project released several standards regulating digital video broadcasting via satellite (DVB-S), cable (DVB-C), and terrestrial television (DVB-T) [1].

Recently along with the wireless era, there emerged a growing demand for multimedia services over handheld devices (defined as small and lightweight battery-powered devices). Accordingly, the ETSI ratified the standard for digital video broadcasting for handheld devices (DVB-H) in November 2004 [2], which is an amendment of the DVB-T standard [3] for handheld devices. DVB-H specifications were defined to achieve IP (Internet Protocol) data broadcasting to handheld devices. One of the most promising features of DVB-H is its synergy with interactive cellular platforms such as GPRS/UMTS [4,6], allowing for IP-based interactive broadcasting "on the move". Since its release, DVB-H technology has been employed worldwide by a number of mobile operators in order to add new multimedia services such as Mobile TV [7]. In March 2008, DVB-H was officially endorsed by European Union as the preferred technology for terrestrial mobile broadcasting [8].

The DVB-H is a superset of DVB-T with additional three main features to meet the specific requirements of *battery-powered* handheld devices:

•MPE-FEC: Multi-Protocol Encapsulation - Forward Error Correction (MPE-FEC) employs a powerful channel coding on top of the DVB-T channel coding. Intensive testing of DVB-H showed that MPE-FEC reduces Signal to Noise Ratio (SNR) requirements up to 8 dB [9].

•4K OFDM Mode: A new 4K Orthogonal Frequency Division Multiplexing (OFDM) mode is introduced in order to facilitate network planning. The new 4K mode offers a compromise between the good Doppler performance of the 2K mode and the good suitability for large Single Frequency Networks of the 8K mode.

•Time-Slicing: Time-Slicing operates by transmitting data in *bursts* corresponding to a single service rather than *continuously* multiplexed with other services. No data is transmitted for a service between two consecutive bursts. Time-slicing is capable of reducing the receiver average power consumption up to 90% by switching off the receiver between consecutive bursts [9,10].

Handheld devices are usually small and lightweight to facilitate ease of mobility. Hence, batteries are restricted to be small and lightweight as well resulting in a low power battery. Therefore, battery life time is a crucial limitation for handheld devices and should be used efficiently. DVB-H addressed this problem by introducing time slicing which is able to reduce the receiver average power consumption up to 90%. Since then, a new field of research emerged looking for techniques to reduce the power consumption furthermore.

Power saving techniques for DVB-H could be categorized into cooperative techniques and compression techniques. The power saving potential in DVB-H networks exploiting cooperation among handheld devices is investigated in [11]. It is shown using numerical results that power saving of over 50% can be achieved by cooperative networks of three handheld devices in fully cooperating mode. The power saving potential due to employing progressive video codecs is studied in [12]. It is proposed to use progressive video codecs to provide a trade-off between video quality and receiver power consumption. JPEG2000 proved to be the best compression algorithm as it allows the receiver to control power consumption depending on quality level.

Adaptive Modulation and Coding (AMC), first introduced in the late 1970s [13], is one of the promising techniques to reduce power consumption. In [14], it is shown that AMC could result in power savings (at the transmitter) up to 40% over fading channels. However, AMC is impractical for broadcast systems as it is unfeasible to transmit at different code rates and modulation schemes to each user. To adopt the AMC concept in broadcast systems, the transmitter must adapt to the user (i.e. receiver) experiencing the worst channel. This constitutes a penalty to other users experiencing good channels.

In this paper, a novel power saving AMC framework is proposed for DVB-H systems. The proposed framework features three power saving schemes: Adaptive Coding (AC), Adaptive Modulation (AM), or Adaptive Modulation and Coding (AMC), in order to reduce the receiver power consumption significantly. The novelty arises from the fact that the proposed framework is initiated from receiver in contrast to traditional AMC schemes that rely on the existence of a feedback channel (from receiver to transmitter). Such feedback channels are unfeasible in broadcast systems. It is shown that the proposed framework is capable of achieving power savings up to 71.875%. The proposed power saving AMC framework operates by rearranging the transmitted frames in a pre-defined pattern. The adaptive receiver selects the appropriate modulation technique and/or code rate, one that achieves a target Bit Error Rate (BER), and then could be switched off and/or powered down resulting in significant potential for saving of reception and processing powers.

A power saving approach similar to the proposed power saving AC scheme is introduced in [15]. This approach saves power by leaving out some FEC columns in the MPE-FEC frame once the receiver has received all the error-free data packets instead of always receiving the full frame. The maximum power saving potential for this approach is limited to 25%. The power saving AC scheme proposed in this paper operates rather over the physical layer and could reap the benefits from switching off and/or powering down circuitry components from the receiver frontend in saving both reception and processing powers where receiving a bit stream usually consumes much more power than processing it.

The organization of this paper is as follows. In Section 2, the DVB-H system is briefly illustrated. In Section 3, the novel power saving AMC framework is proposed. In Section 4, theoretical performance of the proposed framework is analyzed numerically. In Section 5, simulation results are given to illustrate the power saving potential of the proposed framework. Finally, a conclusion is given in Section 6.

2. DVB-H System

2.1. System Overview

DVB-H transmits IP-based services. The IP datagrams are typically produced via MPEG-4/AVC coding of video/audio signals. The IP datagrams are encapsulated by the IP encapsulator. The output is then modulated by a Coded OFDM (COFDM) modulator with 2K, 4K, or 8K carriers. The modulation scheme used could be QPSK, 16-QAM, or 64-QAM. The transmitted signal is organized into frames. Each OFDM frame consists of 68 OFDM symbols. Four frames constitute one super frame.

2.2. Channel Coding

The DVB-H system employs powerful concatenated channel coding consisting of a variable rate punctured Convolutional code as the inner code and a Reed-Solomon (RS) code as the outer code. This combination provides great performance due to the optimum performance

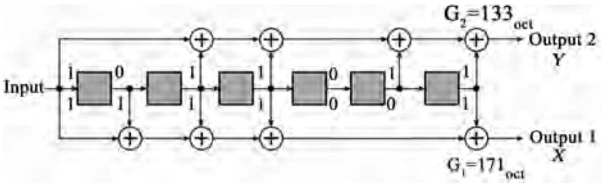


Figure 1. Schematic diagram of DVB-H conv. encoder.

of Viterbi decoder and ability of RS code to correct burst errors produced by Viterbi decoder.

The DVB-H standard employs a (204, 188) shortened RS code that corrects up to 8 erroneous bytes in a received word. In addition, it employs a variable rate punctured convolutional code based on a powerful 1/2 mother code compiled by Odenwalder [16]. A schematic diagram of the convolutional encoder is shown in Figure 1. The code rates supported by DVB-H are 1/2, 2/3, 3/4, 5/6, and 7/8. The puncturing patterns for the different code rates are given in Table 1.

Table 1. Puncturing patterns for different code rates [17].

Code Rate	Puncturing Pattern
1/2	X: 1 Y: 1
2/3	X: 1 0 Y: 1 1
3/4	X: 1 0 1 Y: 1 1 0
5/6	X: 1 0 1 0 1 Y: 1 1 0 1 0
7/8	X: 1 0 0 0 1 0 1 Y: 1 1 1 1 0 1 0

1 and 0 denote transmitting and puncturing of coded bits, respectively.

Table 2. Power saving of DVB-H’S code vs RCPC code.

Code Rate	2/3	3/4	5/6	7/8
DVB-H Power Saving (%)	1.9	5.71	17.14	42.86
RCPC Power Saving (%)	25	33.33	40	42.86

Power saving of different code rates was calculated compared to worst-case of receiving the whole transmit frame (i.e. code rate 1/2).

Table 3. Puncturing patterns for Hagenauer RCPC codes [18] (M = 6 and P = 8).

Code Rate	Puncturing Pattern	
8/9	X: **1111** 0**111** Y: **1**000 **1**000	Transmission ↓
4/5	X: 1111 **1**111 Y: 1000 1000	
2/3	X: 1111 1111 Y: 10**1**0 10**1**0	
4/7	X: 1111 1111 Y: 1**1**10 1**1**10	
½	X: 1111 1111 Y: 111**1** 111**1**	

Table 4. Puncturing patterms for proposed AC scheme employing Hagenauer RCPC codes (M = 6 and P = 8).

Code Rate	Puncturing Pattern	Puncturing Pattern
8/9	Q_1	X: 1111 0111 Y: 1000 1000
4/5	Q_2	X: 0000 1000 Y: 0000 0000
2/3	Q_3	X: 0000 0000 Y: 0010 0010
4/7	Q_4	X: 0000 0000 Y: 0100 0100
1/2	Q_5	X: 0000 0000 Y: 0001 0001

3. Proposed Power Saving AMC Framework

In this section, a novel power saving AMC framework is proposed for DVB-H systems. The three power saving schemes: adaptive coding, adaptive modulation, or adaptive modulation and coding, supported by the AMC framework are proposed in Subsections 3.1, 3.2, and 3.3, respectively. Hardware aspects of the proposed framework are discussed in Subsection 3.4.

3.1. Novel Power Saving AC Scheme

3.1.1. Novel Power Saving AC Scheme Framework

Since for broadcast systems the transmit power is much less crucial than handheld receiver power, therefore transmission using the strongest code (i.e. lowest code rate) is proposed in the power saving AC scheme introduced in this paper. The transmitted frame is arranged in such a way that higher code rates (i.e. weaker codes) are transmitted first and then incremental bits (constituting stronger codes) are transmitted afterwards. The incremental bits are chosen according to pre-defined puncturing patterns. This gives the receiver the flexibility to receive few more bits in order to upgrade the code to a stronger one, in case higher rate codes are not sufficiently powerful to meet the target BER. The transmitted frame size defines how often the code rate is allowed to change (i.e. adaptation rate). The adaptation rate is a very important design parameter that depends on time variation of the channel (i.e. Doppler frequency).

On the other end, the adaptive power saving receiver selects the appropriate code rate, one that achieves a target BER, according to certain thresholds and then could be switched off and/or powered down resulting in significant potential for saving of receiving and processing power. The adaptive receiver is incremental in nature. It works by initially receiving the weakest coded bits, and then using the received SNR a decision is made whether the receiver should receive the bits of the next stronger code.

3.1.2. DVB-H Convolutional Code

Employing DVB-H standard punctured convolutional code in the proposed scheme would face the following crucial limitation. Close inspection of the puncturing patterns given in Table 1 reveals that not all bits of high rate codes are used by lower rate codes. Therefore, by arranging bits of the transmit frame in accordance to the proposed scheme, the adaptive receiver would be occasionally forced to receive more bits than required for the decoding of stronger codes. This introduces losses in the power saving potential compared to the case of using ideal RCPC codes as shown in Table 2.

The power saving of the proposed AC scheme when RCPC code of rate R_i is employed at the receiver could be simply calculated as follows,

$$P_i = \left(1 - \frac{R_{WC}}{R_i}\right) \times 100 \qquad (1)$$

where R_{WC} denotes the worst case scenario code rate. The power saving of the proposed AC scheme when DVB-H convolutional code is employed was calculated numerically using puncturing patterns summarized in Table 1.

3.1.3. RCPC Codes

In order to get over the crucial limitation of DVB-H's convolutional code, employment of RCPC codes is proposed. RCPC codes are variable rate punctured convolutional codes with *rate-compatibility restriction* on puncturing patterns. The rate-compatibility restriction ensures that *all* code bits of high rates code are used by the lower rate codes. RCPC codes were first introduced by Hagenauer in [18]. Hagenauer produced families of RCPC codes with rates between 8/9 and 1/4.

Hagenauer RCPC code with memory $M = 6$ and puncturing period $P = 8$, that is based on the same DVB-H standard 1/2 mother code, is employed in the proposed AC scheme. The puncturing patterns of employed code rates are summarized in Table 3 where the bits in bold are the ones required to be transmitted by the proposed scheme for each code rate.

For example, the proposed power saving AC scheme employing Hagenauer RCPC code is illustrated in Figure 2. This scheme assumes portable reception (i.e. v= 3km/hr, thus channel is quasi-stationary over the period of one super frame). Hence, an adaptation rate of super frame is adopted. Each transmitted super frame is arranged such that higher code rates are transmitted first followed by incremental bits in order to upgrade the code. The incremental bits are chosen according to the pre-defined puncturing patterns summarized in Table 4. The adaptive receiver must rearrange the received bits according to the puncturing patterns given in Table 3 before the Viterbi decoder. Assume in Figure 2 that the receiver decides that the 4/5 code is sufficient to meet the target BER. Then, the receiver will receive 170 instead of 272 OFDM symbols. This translates into a power saving of 37.5% compared to the worst case of receiving the whole super frame. The power saving of the proposed AC scheme employing Hagenauer RCPC code is summarized in Table 5. It is worth noting that generating RCPC code with the same code rates of DVB-H standard convolutional code would require a puncturing period of 210 which would make finding the best puncturing pattern for each code rate very tedious.

3.2. Novel Power Saving AM Scheme

3.2.1. Novel Power Saving AM Scheme Framework

In order to adopt the AM concept in broadcast systems, users experiencing good SNR conditions should be allocated a high modulation technique such as 16-QAM, while users experiencing bad SNR conditions should be allocated a more robust modulation technique such as QPSK. However, transmitting two different modulation techniques is unfeasible. Instead, it is proposed to combine both constellations in a hierarchical fashion where lower modulation symbols are embedded inside higher modulation ones. For instance, the constellation diagram of an AM scheme supporting QPSK and 16-QAM is depicted in Figure 3 where the two left most bits of each 16-QAM symbol could be treated as a QPSK symbol embedded inside the 16-QAM symbol. In this manner, only the two right most bits would not be retrieved by receivers employing QPSK. Hence, retransmission using QPSK modulation of those non-embedded bits is required.

Table 5. Power saving of proposed AC scheme employing Hagenauer RCPC code.

Code Rate	2/3	3/4	5/6	7/8
Power Saving (%)	12.5	25	37.5	43.75

Power saving of different codes was calculated using (1).

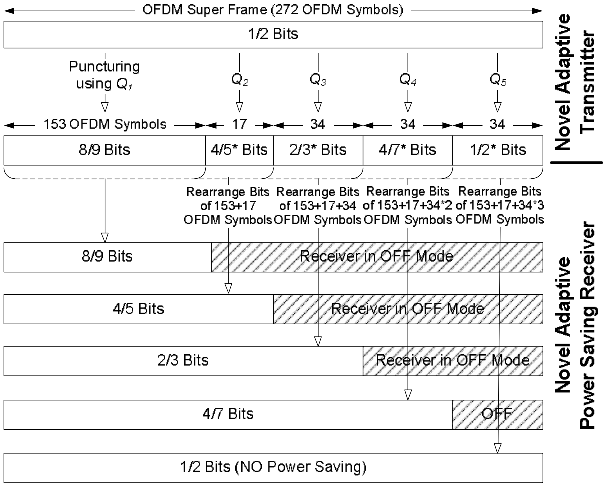

OFDM Super Frame (272 OFDM Symbols)
1/2 Bits
Puncturing using Q_1
Q_2
Q_3
Q_4
Q_5
153 OFDM Symbols
17
34
34
34
8/9 Bits
4/5* Bits
2/3* Bits
4/7* Bits
1/2* Bits
Rearrange Bits of 153+17 OFDM Symbols
Rearrange Bits of 153+17+34 OFDM Symbols
Rearrange Bits of 153+17+34*2 OFDM Symbols
Rearrange Bits of 153+17+34*3 OFDM Symbols
8/9 Bits
Receiver in OFF Mode
4/5 Bits
Receiver in OFF Mode
2/3 Bits
Receiver in OFF Mode
4/7 Bits
OFF
1/2 Bits (NO Power Saving)
Novel Adaptive Transmitter
Novel Adaptive Power Saving Receiver

Figure 2. Schematic of the proposed AC scheme (adaptation rate: super frame).

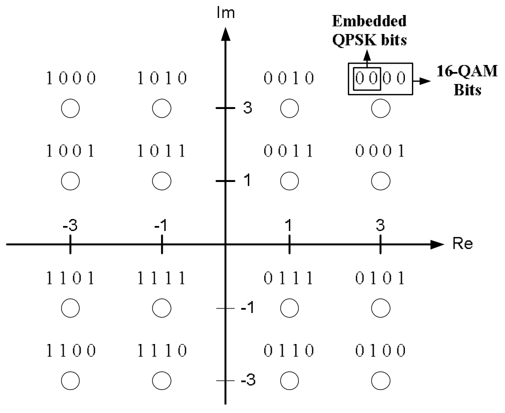


Figure 3. QPSK symbols embedded inside 16-QAM symbols.

Accordingly in the proposed power saving AM scheme, the transmitted frame constituting a block of convolutionally coded bits is subdivided into L clusters. In order to provide users (i.e. receivers) having good SNR conditions the highest power saving potential, the first cluster consists of all coded bits which are then modulated using the highest modulation technique (i.e. modulation technique having the largest constellation size). As a result, users having good SNR conditions are capable of reliably receiving the highest modulation symbols and then could switch off and/or power down their receiver frontend resulting in significant power saving potential. For users experiencing bad SNR conditions (where the highest modulation technique is insufficient to meet the target BER), the second cluster of the transmit frame consists of the non-embedded coded bits (i.e. coded bits that will be lost when employing lower modulation techniques such as QPSK) which are then modulated using a lower modulation technique. As a result, users having bad SNR conditions are capable of receiving the first two clusters and then could switch off and/or power down their receiver frontend resulting in significant power saving potential. The remaining clusters are arranged in a similar fashion to accommodate coded bits that are not embedded in previous clusters (i.e. are not embedded in higher modulation symbols).

The transmitted frame size defines how often the modulation technique is allowed to change (i.e. adaptation rate). The adaptation rate is a very important design parameter that depends on time variation of the channel (i.e. Doppler frequency).

On the other end, the adaptive power saving receiver selects the appropriate modulation technique, one that achieves a target BER, according to certain thresholds and then could be switched off and/or powered down resulting in significant potential for saving of receiving and processing powers. The adaptive receiver is incremental in nature. It works by initially receiving the highest modulation symbols, and then using the received SNR a decision is made whether the receiver should receive the symbols of the next lower modulation technique.

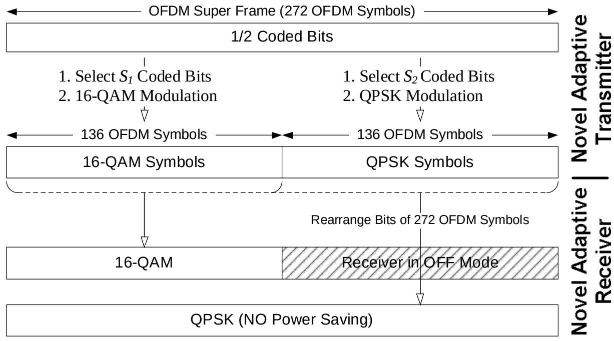


Figure 4. Schematic of proposed AM scheme (adaptation rate: super frame).

Table 6. Selection patterns employed by the proposed AM scheme.

Modulation	Selection Pattern	Selected Bits
16-QAM	S_1	1, 2, 3, 4, 5, 6, …, *K*
QPSK	S_2	1, 2, 5, 6, 9, 10…, *K*-3, *K*-2

3.2.2. Employment of AM Scheme in DVB-H

For example, the proposed power saving AM scheme employing two modulation techniques (QPSK and 16-QAM) is depicted in Figure 4. This scheme assumes portable reception (i.e. v=3km/hr, thus channel is quasi-stationary over the period of one super frame). Hence, an adaptation rate of super frame is adopted. Furthermore, 1/2 rate convolutional code is employed in order to maintain reliable performance in bad SNR conditions. Each transmitted super frame is arranged such that higher modulation symbols (in this example 16-QAM symbols) are transmitted first followed by lower modulation symbols (in this example QPSK symbols). The selected bits are chosen according to the pre-defined selection patterns summarized in Table 6. Assume the receiver decides that the 16-QAM is sufficient to meet the target BER. Then, the receiver will receive 136 instead of 272 OFDM symbols. This translates into a power saving of 50% compared to the worst case of receiving the whole super frame.

3.2.3. Power Saving Potential

The power saving of the proposed AM scheme when modulation technique M_i is employed at the receiver could be simply calculated as follows,

$$P_i = \left(1 - \frac{\log_2(M_{WC})}{\log_2(M_i)}\right) \times 100 \qquad (2)$$

where M_{WC} denotes the worst case scenario modulation technique. The power saving for the three modulation techniques supported by DVB-H is given in Table 7.

3.3. Novel Power Saving AMC Scheme

3.3.1. Novel Power Saving AMC Scheme Framework

In the proposed AMC scheme, the transmitted frame is arranged in such a way that higher modulation symbols are transmitted first followed by lower modulation symbols comprising symbols that were not embedded in higher modulation. Those non-embedded symbols must be retransmitted as receivers employing lower modulation technique will not be able to retrieve them (similar to proposed AM scheme, see Figure 3). Each M_i-QAM (i = 1, 2, …, L) transmitted symbols are arranged such that higher code rates (i.e. weaker codes) are transmitted first followed by incremental bits (constituting stronger codes). The incremental bits are chosen according to predefined puncturing patterns. The transmitted frame size defines how often the modulation technique and code rate are allowed to change (i.e. adaptation rate). The adaptation rate is a very important design parameter that depends on time variation of the channel (i.e. Doppler frequency).

3.3.2. Employment of AMC Scheme in DVB-H

For example, the proposed AMC scheme employing two modulation techniques (QPSK and 16-QAM) and five code rates is depicted in Figure 5. This scheme assumes portable reception (i.e. v=3km/hr, thus channel is quasi-stationary over the period of one super frame). Hence, an adaptation rate of super frame is adopted. The transmitted bits are first rearranged using the five puncturing patterns summarized in Table 4. Then, the bits are further divided into two groups according to the selection patterns summarized in Table 6. Finally, the first and second group of bits (S_1 and S_2) are modulated using 16-QAM and QPSK modulation techniques, respectively.

The adaptive receiver selects the appropriate modulation and coding configuration, one that achieves a target BER, according to pre-defined SNR thresholds and then could switch off and/or reduce power from any possible circuitry in the frontend resulting in significant potential for saving of reception and processing powers. The receiver is incremental in nature. It works by initially receiving bits of the highest modulation technique and weakest code (i.e. 16-QAM, 8/9), and then using the received SNR a decision is made whether the receiver should receive symbols of the next stronger code. Assume the receiver decides that 16-QAM and 2/3 code are sufficient to meet the target BER. Then, the receiver will receive and process 102 instead of 272 OFDM symbols. This translates into a power saving of 62.5%.

3.3.3. Power Saving Potential

The power saving of the proposed AMC scheme when modulation technique M_i and code rate R_j are employed at the receiver could be simply calculated as follows,

$$P_{i,j} = \frac{\log_2(M_{WC})}{\log_2(M_{BC})}\left(\frac{\log_2(M_i)}{\log_2(M_{WC})} - \frac{R_{WC}}{R_j}\right)\times 100 \qquad (3)$$

Table 7. Power saving of proposed AM scheme.

Modulation	QPSK	16-QAM	64-QAM
Power Saving (%)	–	50	66.67

Power saving of different modulation techniques calculated compared to worst-case scenario of receiving the whole transmit frame (i.e. QPSK).

Table 8. Power saving (%) of proposed AMC scheme.

Modulation	Code Rate				
	1/2	4/7	2/3	4/5	8/9
16-QAM	50	56.25	62.5	68.75	71.875
QPSK	-	6.25	12.50	18.75	21.875

Power saving of different AMC configurations was calculated compared to worst-case scenario of receiving whole transmit frame (i.e. QPSK, 1/2).

where M_{WC} and M_{BC} denote the worst and best case modulation technique, respectively. R_{WC} depicts the worst case code rate. Equation (3) is valid only for a scheme supporting two modulation techniques. Power saving of the proposed AMC scheme is given in Table 8.

3.4. Hardware Aspects

It is emphasized that the proposed AMC framework operates at the physical layer leading to power saving potential in the Radio Frequency (RF) frontend (a main source of battery power drainage in DVB-H devices). By switching off the receiver, this do not necessarily mean the receiver is going to sleep mode. As in conventional TDMA systems, the receiver during inactivity periods has the ability to switch off and/or reduce power from any possible circuitry in the frontend [19]. In the proposed AMC framework, power reduction could be achieved by using a DC-to-DC converter [20] to reduce supply voltage during receiver inactivity periods. Typical supported switching rate in DC-to-DC converters is 6 MHz (i.e. 0.16 μs), for example; Texas Instruments TPS62601. This high switching rate achieves the adaptation rates employed by the proposed framework.

4. Numerical Analysis

In this section, the theoretical performance of the proposed power saving AMC framework is analyzed numerically. For simplicity and without loss of generality,

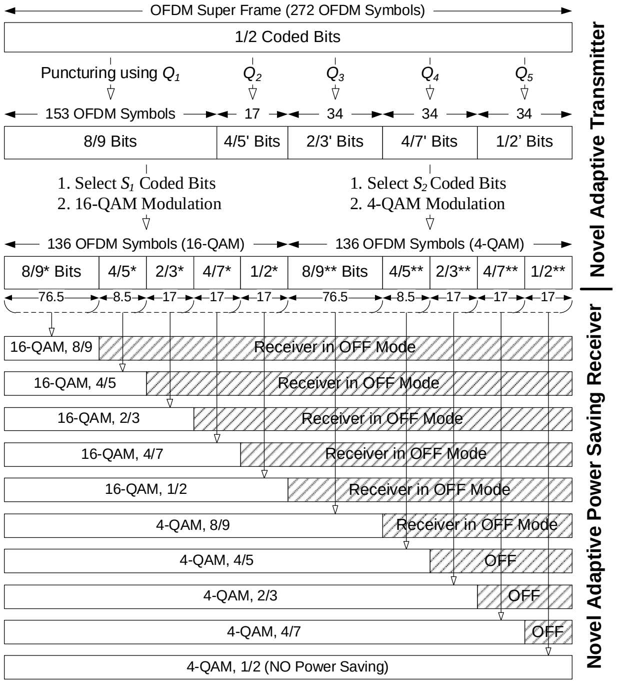


Figure 5. Schematic of proposed AMC scheme (adaptation rate: super frame).

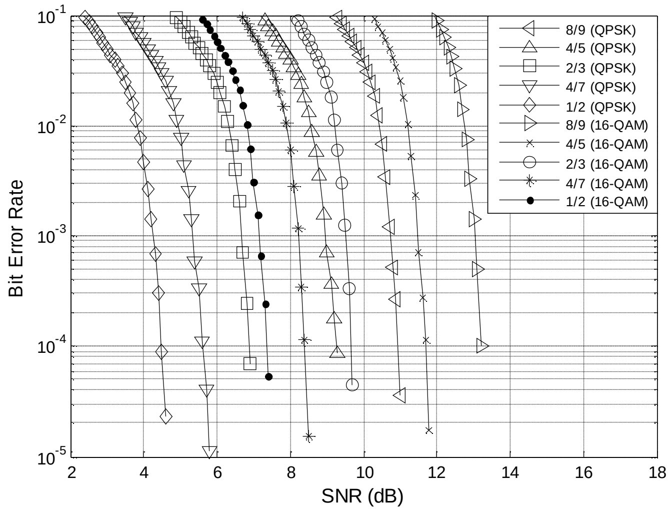


Figure 6. BER performance of fixed-rate AMC scheme in AWGN channel.

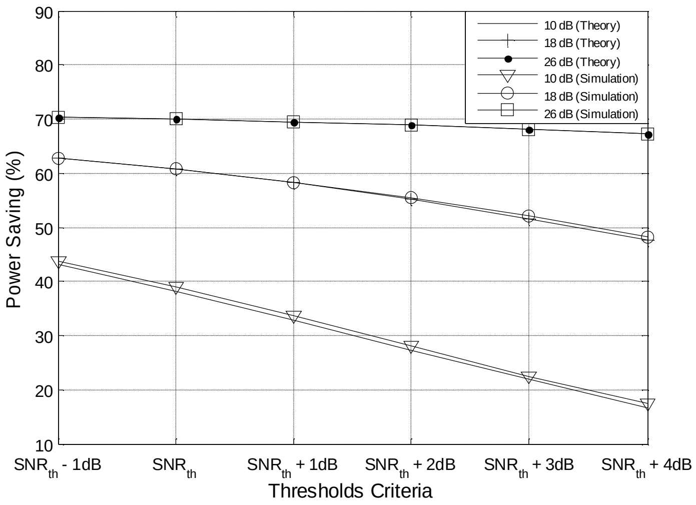


Figure 7. Power saving of proposed AMC scheme in flat Rayleigh channel.

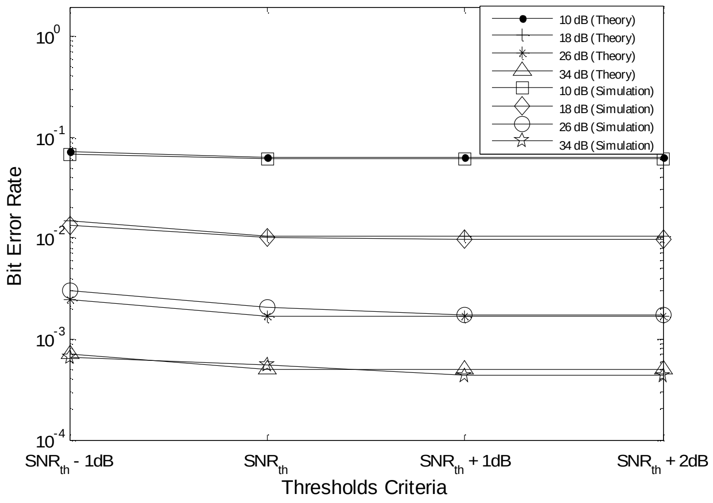


Figure 8. BER performance of proposed AMC scheme in flat Rayleigh channel.

perfect Channel State Information (CSI) is assumed at the receiver. The numerical analysis is performed for the generic case of an AMC scheme. The numerical analysis of AC and AM schemes could be regarded as special cases. The numerical analysis is performed in two environments: flat Rayleigh channel and COST207 Typical Urban 6-paths (TU6) channel [21]. The power delay profile of the TU6 channel is given in Section 5.1.

4.1. Flat Fading Channel

The power saving is given by,

$$P_{Saving} = \sum_{i=1}^{L} \sum_{j=1}^{N} P_{i,j}\, p_{i,j} \qquad (4)$$

where $P_{i,j}$ denotes the power saving due to employing modulation technique M_i and code R_j at the adaptive receiver and is given by (3). L and N are the number of modulation techniques and code rates supported by the proposed AMC scheme. $p_{i,j}$ depicts the probability of selecting modulation technique M_i and code R_j, and is given as,

$$p_{i,j} = \int_{x_{i,j}^{(L)}}^{x_{i,j}^{(U)}} f_X(x)\, dx \qquad (5)$$

where $x_{i,j}^{(L)}$ and $x_{i,j}^{(U)}$ denote the lower and upper SNR thresholds of using modulation technique M_i and code R_j, respectively. $f_X(x)$ is the Probability Density Function (PDF) of the received SNR. Since a Rayleigh channel is considered, the received SNR follows an exponential distribution. Hence, $f_X(x)$ is defined as,

$$f_X(x) = \frac{1}{l} e^{-\frac{x}{l}} \quad x \ge 0 \qquad (6)$$

where λ depicts the average channel SNR.

The BER in flat Rayleigh channel is given by,

$$BER = \sum_{i=1}^{L} \sum_{j=1}^{N} \int_{x_{i,j}^{(L)}}^{x_{i,j}^{(U)}} BER_{i,j}(x)\, f_X(x)\, dx \qquad (7)$$

where $BER_{i,j}(x)$ denotes the BER due to employing modulation technique M_i and code R_j at the adaptive receiver while having SNR x. It is obtained from the Monte Carlo simulations of DVB-H system in AWGN given in Figure 6.

The appropriate modulation and code are selected by comparing the SNR to a set of predefined thresholds that guarantee a target BER. A target BER of 10^{-3} is chosen. The 10^{-3} SNR thresholds (depicted SNR_{th} in Figs.) could be obtained easily from Figure 6. Such thresholds are incrementally increased and decreased by steps of 1 dB to provide a tradeoff between power saving and BER. The simulation results depicted in Figure 6 show that (16-QAM, 1/2) mode performs better than (QPSK, 4/5) and (QPSK, 8/9) modes while providing better power saving. As a result, those two modes were eliminated from the proposed AMC scheme.

Figures 7 and 8 illustrate the power saving and BER of the proposed scheme, respectively, where the modulation and code rate are allowed to change every convolutional block through-out the simulations (The DVB-H standard defines the convolutional 1/2 coded block to be 3264 bits). Simulation results are in excellent agreement with theoretical results. The BER performance is limited by the deep fades of the Rayleigh channel. This explains the BER floor experienced in the low SNR region where the AMC scheme cannot guarantee target BER anymore (even with the aid of 1/2 code). At SNR of 26 dB, the AMC scheme is capable of achieving almost its full power saving potential of 71%.

4.2. Multipath TU6 Channel

For simplicity and without loss of generality, COST207 TU6 channel having Doppler frequency of 40 Hz and an AMC scheme employing an adaptation rate of quarter frame (i.e. modulation and coding configuration is allowed to change every 17 OFDM symbols) are considered. Numerical analysis of the BER performance in TU6 multipath channel is very tedious due to the frequency-selective nature of the channel. To the best of the authors' knowledge, there is no upper bound on the BER performance of OFDM employing concatenated RS-convolutional codes in frequency-selective fading channels due to the complicated nature of this problem. However, an approximate method is adopted in order to perform numerical analysis of the BER performance in the frequency-selective TU6 channel. The approximate BER estimation method works by averaging the SNR in time (i.e., over the quarter frame interval) and frequency (i.e. across sub-carriers) referred to as mean quarter frame SNR. Then, the BER could be obtained from BER performance of DVB-H system employing proposed AMC scheme in 40 Hz TU6 channel (depicted in Figure 9). Hence, Equation (7) could be used to deduce the BER performance where $BER_{i,j}(x)$ denotes the BER due to employing modulation technique M_j and code rate R_i at the adaptive receiver while having x as the mean quarter frame SNR. $BER_i(x)$ is obtained from Figure 9. For the distribution of the mean quarter frame SNR, $f_X(x)$ employed in Equations (5), (7) is approximated as a log-normal distribution where the PDF of the received mean quarter frame SNR is defined as,

$$f_X(x) = \frac{1}{xs\sqrt{2p}} e^{-\frac{(\ln x - m)^2}{2s^2}} \quad x \ge 0 \qquad (8)$$

where μ and σ depict the average and standard deviation of the nature logarithm of x, the mean quarter frame SNR.

Based on simulation experiments carried on a 40 Hz

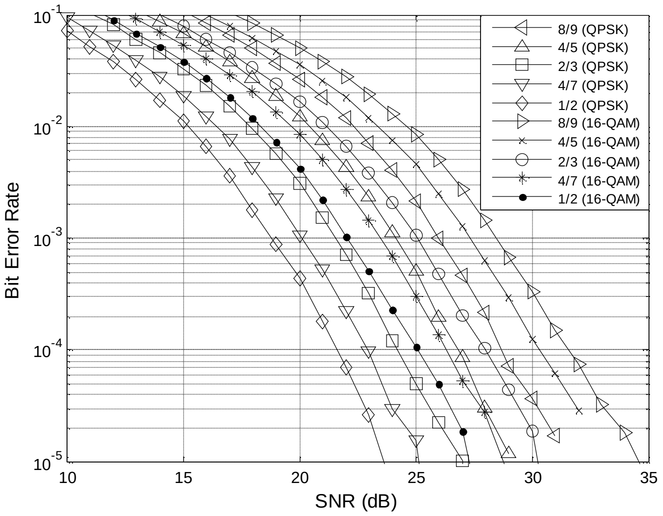


Figure 9. BER performance of fixed-rate AMC scheme in 40 Hz TU6 channel.

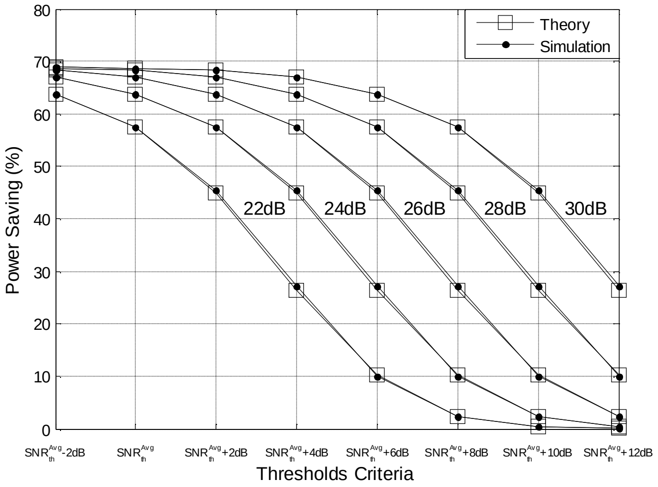


Figure 10. Power saving of proposed AMC scheme in 40 Hz TU6 channel.

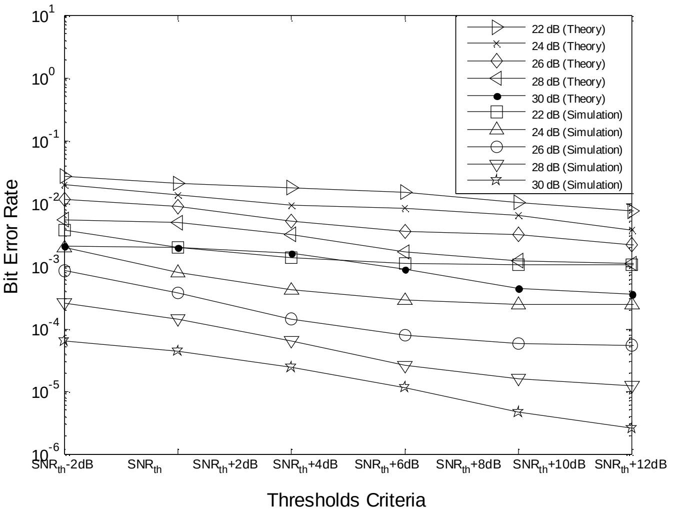


Figure 11. BER performance of proposed AMC scheme in 40 Hz TU6 channel.

TU6 channel, the PDF of the received mean quarter frame SNR could fit a log-normal distribution with parameters:

- $\mu = 4.046256349081641$
- $\sigma = 0.529152417564057$

The appropriate modulation and coding configuration is selected by comparing the mean quarter frame SNR to a set of predefined SNR thresholds that guarantee a target BER. A target BER of 10^{-3} is chosen. The assignment of the 10^{-3} SNR thresholds is defined in the next section. These thresholds are incrementally increased and decreased by steps of 2 dB to provide a tradeoff between power saving and BER. Figures 10 and 11 illustrate the power saving and BER of the proposed AMC scheme, respectively. The simulation results of the power saving potential are in excellent agreement with theoretical ones. However, the simulation results of BER performance are better than theoretical ones. As a result, the theoretical BER results (obtained via proposed approximate method) could be regarded as an upper bound on the BER performance.

5. Simulation Results

In this section, the simulation results of DVB-H system employing the proposed power saving AMC framework are given and discussed comprehensively. Section 5.1 defines the channel model and adaptation rates used throughout the simulations. Section 5.2 summarizes the simulation parameters of the DVB-H system employing the proposed AMC framework. Simulation results illustrating power saving potential and BER performance of DVB-H system employing AC, AM, and AMC schemes are given in Sections 5.3, 5.4, and 5.5, respectively. Finally, the three power saving schemes are compared in Section 5.6.

Table 9. TU6 channel power delay profile.

Tap Number	1	2	3	4	5	6
Delay (μs)	0.0	0.2	0.5	1.6	2.3	5.0
Power (dB)	-3	0	-2	-6	-8	-10
Doppler Spectrum	Rayleigh					

Table 10. Reception model and velocity range for different Doppler Frequencies.

Doppler Frequency	Receiver	Velocity Range
10 Hz	Moderately Mobile	13.5 – 54 km/h
40 Hz	Severely Mobile	54 – 216 km/h

Table 11. SNR thresholds (dB) (adaptation rate: quarter frame, 40HZ TU6 channel).

Code Rate	8/9	4/5	2/3	4/7
SNR Threshold	[18, ∞[	]16, 18]	]13, 16]	]11, 13]

5.1. Channel Model and Adaptation Rates

The COST207 TU6 channel model [21] has proven to be a very good representative for typical mobile reception. The power delay profile of the TU6 channel is given in Table 9. In order to model various types of receivers, two Doppler frequencies were considered. The Doppler frequencies along with their corresponding velocity ranges are summarized in Table 10. The velocity range is calculated starting from carrier frequency 800 MHz (upper part of Band V) to carrier frequency 200 MHz (lower part of Band III).

Accordingly, the proposed power saving schemes were implemented employing three adaptation rates (to accommodate different Doppler frequencies and to provide a trade-off between performance and complexity). The three supported adaptation rates are: super frame, frame and quarter frame (where as the name implies, the modulation and/or coding configuration is allowed to change every super frame, frame, and quarter frame corresponding to 272, 68, and 17 OFDM symbols, respectively).

5.2. Simulation Model

For simulation experiments, a DVB-H system is employed assuming 2K OFDM mode with 1/4 guard interval. The DVB-H standard convolutional code is replaced by the Hagenauer RCPC code in order to exploit full power saving potential. 8 MHz channel is considered. Perfect CSI is assumed at the receiver. Simulations are run for 500 OFDM super frames. MATLAB/SIMULINK [22] is the simulation tool used. For simulations of DVB-H employing the proposed AC scheme, an AC scheme that supports five code rates (1/2, 4/7, 2/3, 4/5, and 8/9) is considered. Furthermore, QPSK is chosen as the modulation technique. For simulations of DVB-H employing the proposed AM scheme, an AM scheme that supports two modulation techniques (QPSK and 16-QAM) is considered. In addition, the convolutional code rate is fixed to 1/2. For simulations of DVB-H employing the proposed AMC scheme, an AMC scheme that supports two modulation techniques (QPSK and 16-QAM) and five code rates (1/2, 4/7, 2/3, 4/5, and 8/9) is considered.

5.3. Power Saving AC Scheme

For TU6 multipath channel, defining the SNR thresholds is challenging due to the frequency-selective nature of the channel. Hence, the mean SNR (averaged over time and frequency, i.e. across OFDM subcarriers) is adopted as the decision criteria. A target BER of 10^{-3} is defined. The 10^{-3} average thresholds (depicted $\mathrm{SNR}_{\mathrm{th}}^{\mathrm{Avg}}$ in Figures.) were obtained by gathering statistics of appropriate code rates that guarantee the target BER averaged over

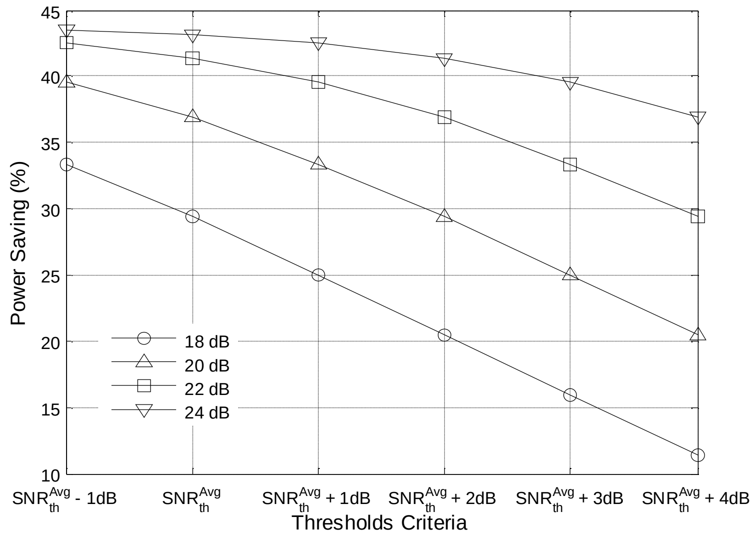


Figure 12. Power savings of proposed AC scheme in 40 Hz TU6 channel.

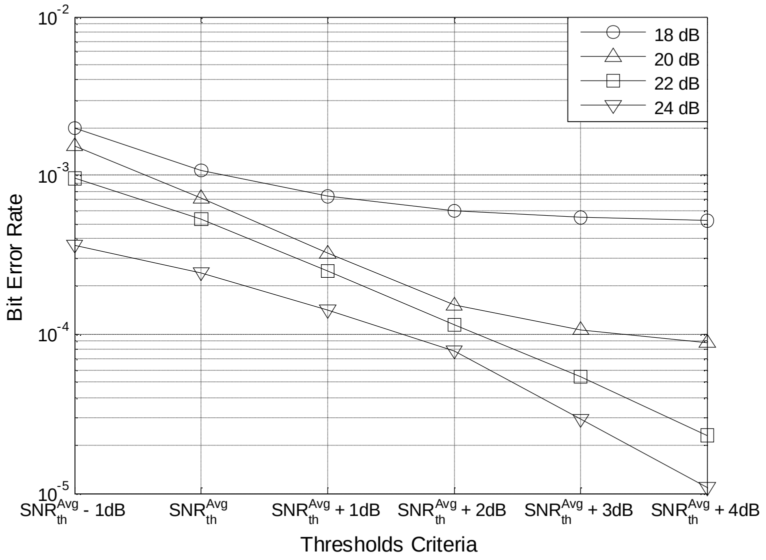


Figure 13. BER performance of proposed AC scheme in 40 Hz TU6 channel.

5000 super frames. For example, consider an AC scheme with an adaptation rate of quarter frame in a 40 Hz TU6 channel. The 10^{-3} thresholds are given in Table 11. Power saving and BER of the AC scheme at different values of average channel SNR are depicted in Figures 12 and 13, respectively.

Figures 12 and 13 demonstrate the tradeoff of increasing/decreasing the decision thresholds. Increasing the thresholds, improves the BER performance on the expense of power saving. Figures 12 and 13 could be combined as depicted in Figure 14. For a target BER, Figure 14 could be used to determine the corresponding power saving given the average channel SNR. Moreover, the appropriate thresholds to be employed by the receiver could be deduced. For example at average channel SNR of 18 dB power saving of approximately 29% could be achieved for a target BER of 10^{-3}. At high average channel SNR, maximum power saving potential of 43.75% could be achieved.

The power savings versus average channel SNR that guarantee a target BER of 10^{-3} for different adaptation rates (super frame, frame, and quarter frame) and Doppler frequencies (10 Hz and 40 Hz) are summarized in Figure 15. As expected, increasing the adaptation rate helps the adaptive receiver increase the power saving potential particularly for fast fading channels. This is obvious for the 40 Hz channel where the AC scheme employing an adaptation rate of quarter frame saves more power than the one employing an adaptation rate of super frame. For instance at an average channel SNR of 18 dB, the faster AC scheme saves 6% more power. In the 10 Hz channel, increasing the adaptation rate saves

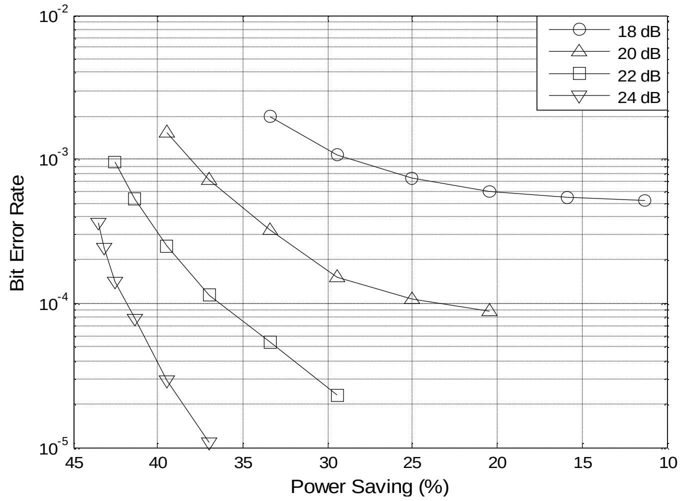


Figure 14. BER vs. power saving of proposed AC scheme in 40 Hz TU6 channel.

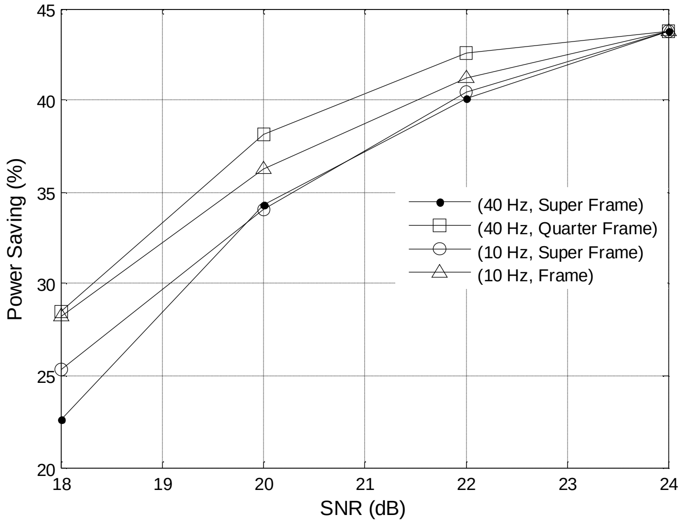


Figure 15. Power saving vs. SNR of proposed AC scheme for different adaptation rates and Doppler frequencies in TU6 channel (Target BER = 10^{-3}).

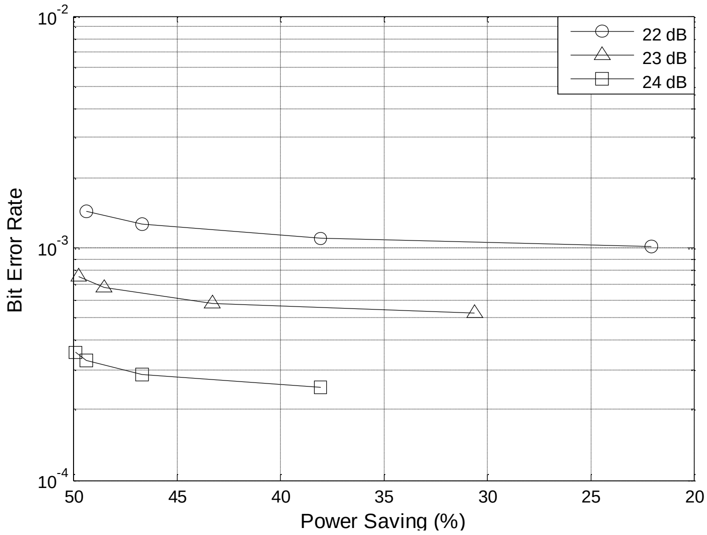


Figure 16. BER vs. power saving of proposed AM scheme in 40 Hz TU6 channel.

3% more power at 18 dB. At high average channel SNR, all schemes achieve maximum power saving potential of 43.75%.

5.4. Power Saving AM Scheme

Consider an AM scheme with an adaptation rate of quarter frame in a 40 Hz TU6 channel. The 10^{-3} SNR thresholds could be defined in a similar manner to AC scheme. The power saving versus the BER performance is shown in Figure 16. For a target BER, Figure 16 could be used to determine the corresponding power saving given the average channel SNR. Moreover, the appropriate thresholds to be employed by the receiver could be deduced. For instance, at average channel SNR of 22 dB power saving of 22% could be achieved. At high average channel SNR, maximum power saving potential of 50% could be achieved.

5.5. Power Saving AMC Scheme

Consider an AMC scheme with an adaptation rate of quarter frame in a 40 Hz TU6 channel. The 10^{-3} SNR thresholds could be obtained in a similar manner to AC scheme. The power saving versus the BER performance is shown in Figure 17. For a target BER, Figure 17 could be used to determine the corresponding power saving given average channel SNR. Moreover, the appropriate thresholds to be employed by the adaptive receiver could be deduced. For instance, at average channel SNR of 22 dB power saving of 30% could be achieved. At high average channel SNR, maximum power saving potential of 71.875% could be achieved.

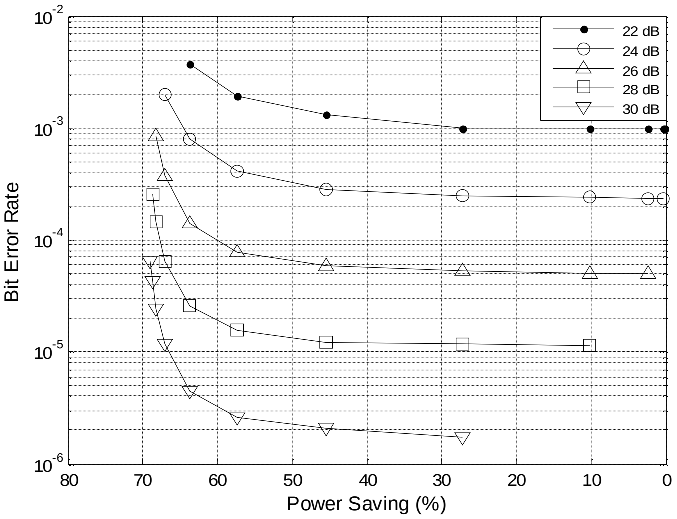


Figure 17. BER vs. power saving of proposed AMC in 40 Hz TU6 channel.

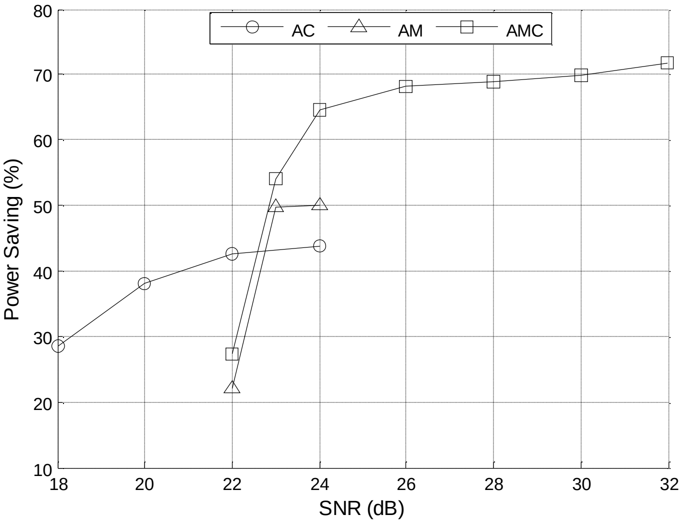


Figure 18. Power saving vs. average channel SNR of the proposed three power saving schemes in 40 Hz TU6 channel (target BER = 10^{-3}).

5.6. Comparison of Three Power Saving Schemes

The proposed three power saving schemes are compared in terms of power saving potential in Figure 18 The proposed AC scheme provides significant power saving (28.5–43.75%) for a wide range of SNRs (18–24 dB). On the other hand, the proposed AM scheme provides more power saving (50%) compared to the AC scheme but for a narrower range of SNRs (23–24 dB). This narrow range of SNRs is due to the limited power saving modes (only two modulation techniques, opposed to the five code rates supported by the AC scheme). Note that in the low SNR region (18–22 dB) the AC scheme provide higher power saving potential than AM scheme where at very low SNRs (18–20 dB) the AM scheme fails to provide any power saving for the target BER of 10^{-3}.

For high average channel SNR (above approximately 22.5 dB), the proposed AMC scheme is clearly the better choice. It provides significant power saving (42–71.87-5%) for a very wide range of SNRs (22.5–32 dB). The proposed AMC scheme is unable to provide any power saving (for the target BER of 10^{-3}) in the low SNR region (18–20 dB) due to the fact that the QPSK modulation technique is actually embedded inside the 16-QAM constellation (see Figure 3) and hence making it much less robust than the normal non-embedded QPSK modulation technique employed by the AC scheme. This also justifies why the AC scheme outperforms the AM scheme in the very low SNR region.

6. Conclusions

In this paper, a novel power saving AMC framework was proposed for DVB-H systems to address the challenging problem of power consumption. The proposed framework features three schemes: adaptive coding, adaptive modulation, and adaptive modulation/coding in order to reduce the receiver power consumption. The proposed AMC framework is initiated from receiver in contrast to traditional AMC schemes that rely on the existence of a feedback channel. It operates by rearranging transmitted frames in a pre-defined pattern. The adaptive receiver selects the appropriate modulation and/or coding configuration, one that achieves a target BER, and then could be switched off and/or powered down resulting in significant potential for saving of reception and processing powers. It was shown that the proposed power saving AMC framework is capable of achieving remarkable power saving (up to 71.875%) in TU6 channel environment.

Furthermore, numerical analysis for the power saving potential and BER performance of the proposed AMC framework is performed where theoretical ones were shown to be in excellent agreement with simulation results for the case of Rayleigh channel. For TU6 channel, the numerical results could serve as an upper bound on the BER performance of the proposed scheme.

In addition, the effect of adaptation rate on power saving was studied for different Doppler frequencies. As expected, increasing the adaptation rate helps the adaptive receiver increase the power saving potential particularly for fast fading channels. This is obvious for the 40 Hz TU6 channel where the proposed AC scheme employing an adaptation rate of quarter frame saves 6% more power than the one employing an adaptation rate of super frame.

Finally, the three schemes supported by the proposed framework (AC, AM, and AMC) were compared in terms of power saving potential. It was shown that for environments characterized by high average channel SNR (greater than 22.5 dB), the AMC scheme would be the best scheme where 42–71.875% power saving was achieved for the target BER of 10^{-3}. On the other hand, for environments characterized by low SNR (less than 22.5 dB), the AC scheme would yield the highest power saving potential (28.5–42%).

7. References

[1] U. Reimers, “DVB—The family of international standards for digital video broadcasting,” in Proceedings of IEEE, Vol. 94, No. 1, January 2006.

[2] ETSI EN 302 304: “Digital Video Broadcasting (DVB); transmission system for handheld terminals (DVB-H),” European Telecommunication Standard, November 2004.

[3] ETSI EN 300 744: “Digital Video Broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television (DVB-T)”, European Telecommunication Standard, January 2004.

[4] C. Rauch, W. Kellerer, and P. Sties, “Hybrid mobile interactive services combining DVB-T and GPRS,” in Proceedings of European Personal Mobile Communication Conference, Vienna, Austria, February 2001.

[5] E. Stare and S. Lindgren, “Hybrid broadcast-telecom systems for spectrum efficient mobile broadband internet access,” in Nordic Radio Symposium, Sweden, 2001.

[6] G. Gardikis, G. Kormentzas, G. Xilouris, H. Koumaras, and A. Kourtis, “Broadband data access over hybrid DVB-T networks,” in Proceedings of 3rd Conference on Heterogeneous Network, Ilkley, UK, July 2005.

[7] A. Kumar, “Mobile TV: DVB-H, DMB, 3G systems and rich media applications,” Focal Press, 2007.

[8] J. Gozalvez, “The european union backs the DVB-H standard [mobile radio],” IEEE Vehicular Technology Magazine, Vol. 3, pp. 3–12, June 2008.

[9] ETSI TR 102 401: “Digital Video Broadcasting (DVB); transmission system for handheld terminals (DVB-H),” Validation Task Force Report, May 2005.

[10] G. Gardikis, H. Kokkinis, and G. Kormentzas, “Evaluation of the DVB-H data link layer,” in Proceedings of European Wireless, Paris, France, April 2007.

[11] Q. Zhang, F. H. P. Fitzek, and M. Katz, “Cooperative

power saving strategies for IP-services supported over DVB-H networks," in Proceedings of IEEE Wireless Communication and Network Conference, March 2007.

[12] E. Belyaev, T. Koski, J. Paavola, A. Turlikov, and A. Ukhanva, "Adaptive power saving on the receiver side in digital video broadcasting systems based on progressive video codec," The 11th International Symposium on Wireless Personal Multimedia Communication, 2008.

[13] J. F. Hayes, "Adaptive feedback communications," IEEE Transactions on Communication Technology, Vol. COM-16, pp. 29–34, February 1978.

[14] A. Goldsmith and S. G. Chua, "Adaptive coded modulation for fading channels," IEEE Transactions on Communication, Vol. 46, pp. 595–602, May 1998.

[15] E. D. Balaguer, F. H. P. Fitzek, and O. Olsen, "Performance evaluation of power saving strategies for DVB-H services using adaptive MPE-FEC decoding," The 16th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communication, Berlin, Germany, September 2005.

[16] J. Odenwalder, "Optimal decoding of convolutional codes, " Ph.D. dissertation, Department of Systems Sciences, School of Engineering and Applied Sciences, University of California, Los Angeles, 1970.

[17] Y. Yasuda, K. Kashiki, and Y. Hirata, "High-rate punctured convolutional codes for soft decision viterbi decoding," IEEE Transactions on Communication, Vol. COM-32, pp. 315–319, March 1984.

[18] J. Hagenauer, "Rate-compatible punctured convolutional Codes (RCPC Codes) and their applications," IEEE Transactions on Communication, Vol. 36, pp. 389–400, April 1988.

[19] T. S. Rappaport. Wireless communications: Principles & practice," Prentice Hall, New Jersey, 1996.

[20] F. L. Luo, and H. Ye, Essential DC/DC Converters. CRC Press, 2006.

[21] COST207, "Digital land mobile communications," Commission of the European Communities, Directorate General Communications, Information Industries and Innovation, 1989, pp. 135–147.

[22] A. Gilat, "MATLAB: An introduction with applications," 3rd Edition, John Wiley and Sons, 2008.

Int. J. Communications, Network and System Sciences, **2010, 3, 167-172**
doi:10.4236/ijcns.2010.32023 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

An Approach to Dynamic Asymptotic Estimation for Hurst Index of Network Traffic

Xiaoyan MA, Hongguang LI
School of Information Science and Technology, Beijing University of Chemical Technology, Beijing, China
Email: *maxy@mail.buct.edu.cn, lihg@mail.buct.edu.cn*

Received November 14, 2009; *revised December* 27, 2009; *accepted January* 27, 2010

Abstract

As an important parameter to describe the sudden nature of network traffic, Hurst index typically conducts behaviors of both self-similarity and long-range dependence. With the evolution of network traffic over time, more and more data are generated. Hurst index estimation value changes with it, which is strictly consistent with the asymptotic property of long-range dependence. This paper presents an approach towards dynamic asymptotic estimation for Hurst index. Based on the calculations in terms of the incremental part of time series, the algorithm enjoys a considerable reduction in computational complexity. Moreover, the local sudden nature of network traffic can be readily captured by a series of real-time Hurst index estimation values dynamically. The effectiveness and tractability of the proposed approach are demonstrated through the traffic data from OPNET simulations as well as real network, respectively.

Keywords: Network Traffic, Hurst Index, Dynamic Asymptotic Estimation, Long-Range Dependence

1. Introduction

A large number of studies have shown that the real network traffic has self-similarity and long-range dependence, the characteristic exists in the traffic streaming and video streaming of LAN, MAN, WAN, ISDN, CDPN, CDMA, GPRS, wireless networks and Adhoc networks [1–6]. Hurst index is a primary parameter to describe the sudden nature of network traffic. Hurst index estimation methods are mainly two types [7]: one is time-domain methods, including the absolute value estimation, variance, R/S and the IDC method. The other is the frequency domain or wavelet domain methods, including Whittle’s maximum likelihood estimation, periodgram method and wavelet domain estimation method. Time-domain methods calculate the law of power function between data statistics value and the aggregated order. Similarly, the frequency domain or the wavelet domain methods find the law of power function between frequency domain spectrum or energy and time scale.

Contrary to the methods using limited data series, this paper presents an approach to dynamic asymptotic method for Hurst index estimation using infinite time series, which is strictly consistent with the asymptotic property of long-range dependence. With the evolution of network traffic over time, more and more data are generated. Based on the calculations in terms of the incremental part of time series, the algorithm enjoys a considerable reduction in computational complexity. The algorithm can also capture the local sudden information of network traffic at the same time by a series of Hurst index values. Wei Jinwu [8] proposed a long-range dependence sliding window time-varying estimation algorithm to capture local sudden information. But its Hurst index estimation is still based on the limited time series. Hurst index estimation without previous network traffic information is not accurate.

The second part of this paper introduces the network traffic self-similarity and long-range dependence theory, uses ON/OFF model in OPNET simulation software to generate the self-similar traffic, and applies the traditional R/S algorithm to estimate Hurst index. The third part presents an approach to dynamic asymptotic estimation for Hurst index of Network Traffic. The fourth part shows the effectiveness and tractability of algorithm using simulated data and real network traffic. The last part concludes the paper.

2. Self-Similarity, Long-Range Dependence and Traditional Hurst Index Estimation

$X=(X_1,X_2,...)$ is a broad stationary stochastic process, with a constant mean *μ=E[Xi]*, finite variance $\sigma^2 = E[(X_i-\mu)^2]$. Its auto-correlation function $r(k)=E[(X_i-\mu)(X_{i+k}-\mu)]/\sigma^2$ is only with the k-related, *(k = 0,1,2, ...).* $X^{(m)}=(X_1{}^{(m)},X_2{}^{(m)},...)$, $X_i{}^{(m)}=(X_{im-m+1}+...+X_{im})/m, i\ge 1$ express an m-order aggregation of broad stationary random process. The autocorrelation function of $X^{(m)}$ is $r^{(m)}(k)$.

Definition 1 [9]: For stochastic process X, if $\mathrm{var}[X^{(m)}]\sim m^{-\beta}\mathrm{var}(x)$ and $r^{(m)}(k)=r(k), k\ge 0$, *m= 1,2,3,…,*then *X* is called second-order accurate self-similar process. Its self-similarity parameter (Hurst parameter) is *H = 1-β/2,* in which *0 <β <1.*

Definition 2[9]: For stochastic process *X,* if k is big enough, $\mathrm{var}[X^{(m)}]\sim m^{-\beta}\mathrm{var}(x)$ and $r^{(m)}(k)=r(k)$, $m\to\infty$, then *X* is called a asymptotic second-order self-similar process. Its self-similarity parameter is *H = 1-β/2,* in which *0 <β <1.*

Theory 1[9]: For a stationary process *X,* if $\sum_{k=0}^{\infty} r(k)=\infty$, $r(k)\sim c_1 k^{r-1}, r\in(0,1)$, *r=2H-1, 1/2<H<1,* the process *X* has long-range dependence. The spectral density is attenuated according to the hyperbolic form, $\Gamma_x(v)\sim c_2|v|^{-r}, v\to 0, r\in(0,1)$, $c_2=2(2\pi)^{-r}\sin((1-r)\pi/2)c_1$.

Definition 2[9]: For random process *X,* if the tail distribution function approximates in power law, $P[X>x]\sim cx^{-\alpha}, x\to\infty$, *0<$\alpha$ <2, c>0,* then *X* is called heavy-tailed distribution.

A significant feature of heavy-tailed distribution is that it has infinite variance. One of the most commonly used heavy-tailed distributions is the Pareto distribution. The distribution function is $F(x)=P[X\le x]=1-\frac{b^a}{x}, x\ge b$, *0<α<2, α* for the shape parameter determining the severity of trailing of tail distribution function and *b* for the location parameter.

The transmission of network business includes the application layer, transport layer, network layer and data link layer. The application layer is the data source of network transmission, presents the self-similarity in wide time range, for example the heavy-tailed distribution of the file size and packet arrival time interval. The heavy-tailed distribution in application-layer is considered the main physical characteristic of network traffic self-similarity. The self-similarity in application layer is thus mapped and introduces the self-similarity to the underlying network layer.

Theory 2[9]: For a given time series $X=\{X_i, i=1,2,...\}$, partial summation is $Y(n)=\sum_1^n X_i$, the sample variance is $s^2(n)=\frac{1}{n}\sum_{i=1}^n X_i^2-\frac{1}{n^2}Y^2(n)$, and then the *R/S* statistic is as follows:

$$R(n)/S(n)=\frac{1}{S(n)}[\max_{0\le t\le n}(Y(t)-\frac{t}{n}Y(n))-\min_{0\le t\le n}(Y(t)-\frac{t}{n}Y(n))]$$

If $E(R(n)/S(n))\sim cn^H, n\to\infty$, 1/2<*H*<1, *c* is the normal number which is independent with n, then the time series has long-dependence.

The *R/S* estimation method for Hurst index is as follows:

- Divide the time series $X=(X_1,X_2,...X_N)$ into K groups. The length of each group is *n=N/K*, $X_k(i)=\{X_{(k-1)n+i},...,X_{(k-1)n+i}, i=1,...,n; k=1,...,K\}$;
- Calculate the mean and variance of each group, $k=1,...,K$,

$$\bar{X}_k(n)=(X_{(k-1)n+1},...,X_{(k-1)n+n})/n$$

$$S_k^2(n)=\frac{1}{n}\sum_{i=1}^n X_k{}^2(i)-(\bar{X}_k(n))^2;$$

- Calculate $R_k(n)$ of each group, $k=1,...,K$,

$$R_k(n)=\max_{0\le t\le n}(Y_k(t)-t\bar{X}_k(n))-\min_{0\le t\le n}(Y_k(t)-t\bar{X}_k(n))$$

$$Y_k(0)=0, Y_k(t)=\sum_{j=1}^t X_{(k-1)n+j};$$

- Calculate the mean of $R_k(n)/S_n(n)$,

$$E\{R_k(n)/S_k(n)\}=\tfrac{1}{k}\sum_1^k R_k(n)/S_k(n);$$

- Repeat the above steps to get several $E\{R_k(n)/S_k(n)\}$ for different *n* and *K*;
- Draw all the points $(\log n, \log E\{R_k(n)/S_k(n)\})$ in the coordinate diagram; fit a straight line through these points according to the least mean square criteria, and then the slope of this line is the Hurst index.

Superposition of a large number of independent ON/OFF sources can generate self-similar volume of business. In the ON period, the packets enter the network, in the OFF period, no packet generated. ON/OFF duration is Pareto distribution. When 1<*α*<2, the infinite number of such ON/OFF sources will generate self-similar volume of business, of which Hurst parameter is *H = (3-α)/2.* When a sufficient not infinite number of superposition of independent ON/OFF sources, we will get a very high degree of self-similar volume of business. In this section, simulation software OPNET is used. The packet arrival time interval is the 0.2s in each ON cycle, so the send rate is 5packet/s. The superposetion of 50 such ON/OFF source will generate network similar traffic, of which the average packet arrival rate is $\lambda=R\times N/2=125 packets/s$, R for the sending rate of each ON/OFF sources and *N* for the superposition number of ON/OFF sources. In OPNET, each ON/OFF source packet inter-arrival interval is the Pareto distribution. When the value of α is set to 1.8, 1.6, 1.4 and 1.2 respectively, the corresponding Hurst index of the self-similar traffic is 0.6, 0.7, 0.8 and 0.9 accordingly. In the above modeling, simulation time was 24576 seconds and the data time series length is 8192, as shown in Figure 1.

Figure 2 shows the Hurst index estimation results of ON/OFF simulated flows using the R/S estimation method. Not any Hurst index estimation algorithm is generally applicable to any situation; there is always the estimate error in the different circumstances. Each estimation algorithm uses the different statistics; different factors have an impact on the corresponding statistics, and therefore cause the algorithms the different degrees of estimation error. The main factors are non-stationary nature and periodic component, and the white noise when sampling the data series. In addition, the various types of algorithms are based on the global domain summation and average, so the variability of data series will be smoothed out. The stronger variability the data series have, the bigger estimation error the algorithm will cause.

In addition, Hurst index estimation based on limited time series will also cause some degree of estimation error. According to the definition of long-range dependence, we need to estimate the Hurst index with infinite time series. We know that with the evolution of network traffic over time, more and more data are generated. Based on the gradually increased data set, we can get a series of Hurst index estimation values which asymptotically tend to theory value. Moreover, a single Hurst index estimation based on limited time series is difficult to reflect the sudden nature of network traffic, but a series of Hurst index estimation values can capture the sudden information in local network traffic dynamically.

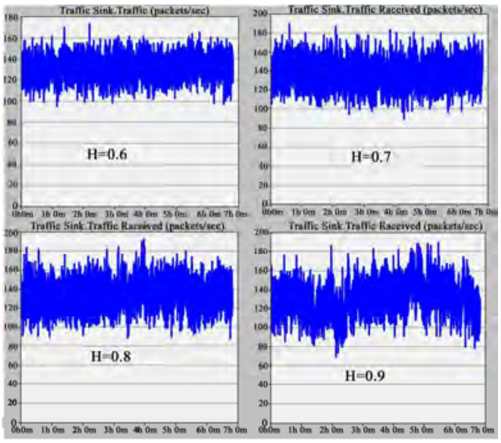


Figure 1. The simulation of time series.

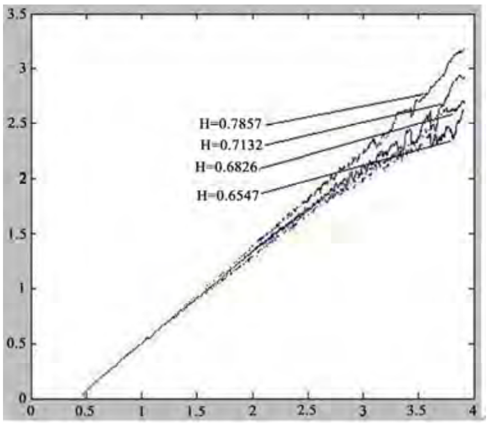


Figure 2. The hurst index estimation using the R/S method.

3. The Dynamic and Asymptotic Algorithm of Hurst Index Estimation

The length of network traffic time series will increase limitless in real-time sampling process. Strictly speaking, Hurst index estimation should not use the time series with limited length, because the mathematical definitions of self-similarity and long-range dependence are asymptotic. According to the inference method of mathematical statistics [10–11], this paper presents the dynamic and asymptotic algorithm using infinite time series.

The length of time series increase gradually, though the algorithm only computes the incremental part of time series to improve the execution speed and reduce the computational complexity. The algorithm estimates the Hurst index with the current computation values and the previous results in order not to lose previous network traffic information. The algorithm is introduced below using R/S method as an example.

The original data series is X, the initial data series is X^0 with the length of n_0, the following data series are $X^1, X^2, \ldots, X^m$ with the growing length $n_0 < n_1 < \cdots < n_m$, the value of n_m is equal to or close to the length of the data series X.

- STEP1: For time series $X^0 = \{X_i, i = 1,2,\ldots,m1\}$, divide it into *K1* groups, the length of each group is *n*, $X_k(i) = \{X_{(k-1)n+i},\ldots,X_{(k-1)n+i}, i = 1,\ldots,n; k = 1,\ldots,K1\}$, *K1=m1/n*;
- STEP2: Calculate the mean and variance of $X_k(i), k = 1,\ldots,K1$, $\bar{X}_k(n) = (X_{(k-1)n+1},\ldots,X_{(k-1)n+n})/n$, $S_k^2(n) = \frac{1}{n}\sum_{i=1}^{n}(X_k(i) - \bar{X}_k(n))^2$;
- STEP3: Calculate $R_k(n), k = 1,\ldots,K1$ of each group, $R_k(n) = \max r_k(n) - \min r_k(n)$, $\max r_k(n) = \max_{0\le t\le n}(Y_k(t) - t\bar{X}_k(n))$, $\min r_k(n) = \min_{0\le t\le n}(Y_k(t) - t\bar{X}_k(n))$, $Y_k(0) = 0, Y_k(t) = X_{(k-1)n+1},\ldots,X_{(k-1)n+t}, t = 1,\ldots,n$;
- STEP4: Calculate the mean of $R_k(n)/S_k(n)$, *k=1*

,…,*K1*, $avg\,(n) = \frac{1}{K}\sum_{\mathrm{k}=1}^{\mathrm{K1}} \mathrm{R}_{\mathrm{k}}(\mathrm{n}) / S_k(n)$,
$E\{R_k(n)/S_k(n)\} = avg(n)$;

- STEP5: Repeat above steps to get more value of $E\{R_k(n)/S_k(n)\}$ for different n;
- STEP6: Draw all the points ($\log n, \log E\{R_k(n)/S_k(n)\}$) in the coordinate diagram, fit a straight line through these points according to the least mean square criteria, and then the slope of this line is the Hurst index of time series which length is *m1*;
- STEP7: When the length of time series is gradually increased to *m2*, $X^1 = \{X_i, i = 1,2,...m1,...,m2\}$. The partial data, $\{X_i, i = n * K1,...,m2\}$, hasn't been used for Hurst index estimation. Divide these data into *(K2-K1)* groups. The length of each group is n, $X_k(i) = \{X_{(k-1)n+i},...,X_{(k-1)n+i}, i = 1,...,n; k = K1+1,...,K2\}$ *K2=m2*;
- STEP8: Calculate the mean and variance of $X_k(i), k = K1+1,...,K2$, ,
  $\bar{X}_k(n) = (X_{(k-1)n+1},...,X_{(k-1)n+n})/n$
  $S_k^2(n) = \frac{1}{n}\sum_{i=1}^{n}(X_k(i) - \bar{X}_k(n))^2$;
- STEP9: Calculate $R_k(n), k = K1+1,...,K2$ of each group, $R_k(n) = \max r_k(n) - \min \mathrm{r}_k(n)$,
  $\max r_k(n) = \max_{0 \le t \le n}(Y_k(t) - t\bar{X}_k(n))$,
  $\min r_k(n) = \min_{0 \le t \le n}(Y_k(t) - t\bar{X}_k(n))$,
  $Y_k(0) = 0$, $Y_k(t) = X_{(k-1)n+1},..., X_{(k-1)n+t}, t = 1,..., n$;
- STEP10: Calculate the mean of $R_k(n)/S_k(n)$,
  $avg\,(n) = \frac{1}{K_1+K_2}(K1 * avg\,(n) + \sum_{k=K1+1}^{\mathrm{k2}} \mathrm{R}_{\mathrm{k}}(\mathrm{n})/\mathrm{S}_{\mathrm{k}}(\mathrm{n}))$
  , $E\{R_k(n)/S_k(n)\} = avg(n)$,
  so the old value is revised in this step based on the incremental data.
- STEP11: For different n, repeat above steps to get all the new revised value of $E\{R_k(n)/S_k(n)\}$.
- STEP12: Draw all the points ($\log n, \log E\{R_k(n)/$
- $S_k(n)\}$) in the coordinate diagram, fit a straight line through these points according to the least mean square criteria, and then the slope of this line is the Hurst index value of time series which length is *m2*.
- STEP13: *m1=m2,K1=K2;*
- STEP14: If the length of time series increases continually, repeat the steps from 7 to 13 to revise all the values of $E\{R_k(n)/S_k(n)\}$, $k = K1+1,...,K2$ based on the incremental data accordingly, or the loop is finished.

From above steps we can see that, by calculating the incremental part of data series, we get a series of Hurst index values which asymptotically tend to theory value. These Hurst index are continually revised based on a new period of data series, so the new revised Hurst index value can reflect the degree of sudden nature of current local network traffic dynamically. At the same time, when the length of time series increases to infinite, the obtained Hurst index value will reflect the long-range dependence of the overall network traffic.

4. Algorithm Validations

4.1. The Simulation Data Validation

In OPNET, the ON/OFF source packet inter-arrival interval is the Pareto distribution. The value of α is set to 1.6. According to *H=(3-α)/2*, the corresponding Hurst index of the self-similar traffic is 0.7. The simulation time is extended to 904.8 hours, and the length of time series D is extended to 524288.

We apply the proposed algorithm to estimate the Hurst index of the data series D. The value of variable i is increased to 9, and the length of the data series is increased as follows, $n_i = 2^{10+i} =$ 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, 131072, 262144 and 524288 respectively. As shown in Figure 3, the time-scales are from 9 to 19. A series of Hurst index fluctuate in form of the asymptotic trend around the theory value of 0.7, proving that data series D has long-range dependence. The Hurst index is not static and will change with the evolution of network traffic. Any estimation of Hurst index based on the limited data series will draw the wrong conclusion.

Secondly, we use the data series E with the Hurst index of 0.5 to verify the algorithm presented in this paper. Same as the data series D, the simulation time is 904.8 hours, and the length is 524288. The ON/OFF source packet inter-arrival interval is the exponential distribution.

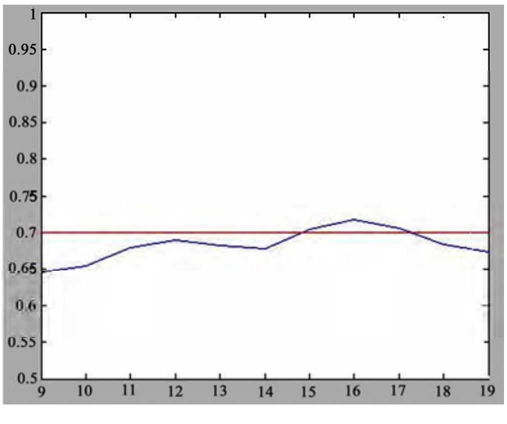


Figure 3. Hurst index estimation for data series D.

As can be seen in Figure 4, a series of Hurst index values also fluctuate slowly. The difference from Figure 3 is that Figure 4 is in form of the asymptotic trend around 0.5. The reason is due to the exponential distribution in the ON/OFF source packet inter-arrival interval. So the time series E has not characteristics of self-similar and long-range dependence. But if the Hurst index estimation is based on the limited length of time series of time scale 9, 10, or 11, then Hurst index will be greater than 0.5 and the wrong conclusions may be drawn. The algorithm presented in this paper considers the asymptotic trend of a series of Hurst index, and comes to the conclusion that the busty traffic of time series E becomes weaker and weaker to zero, and do not have the characteristics of long-range dependence.

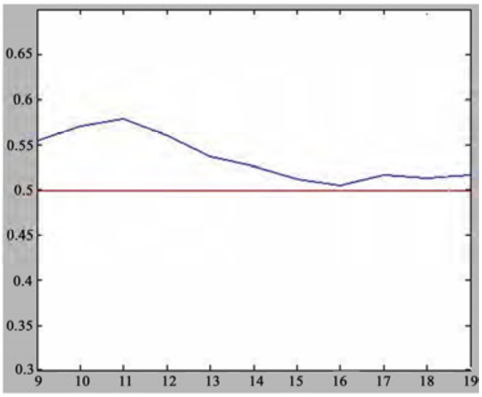
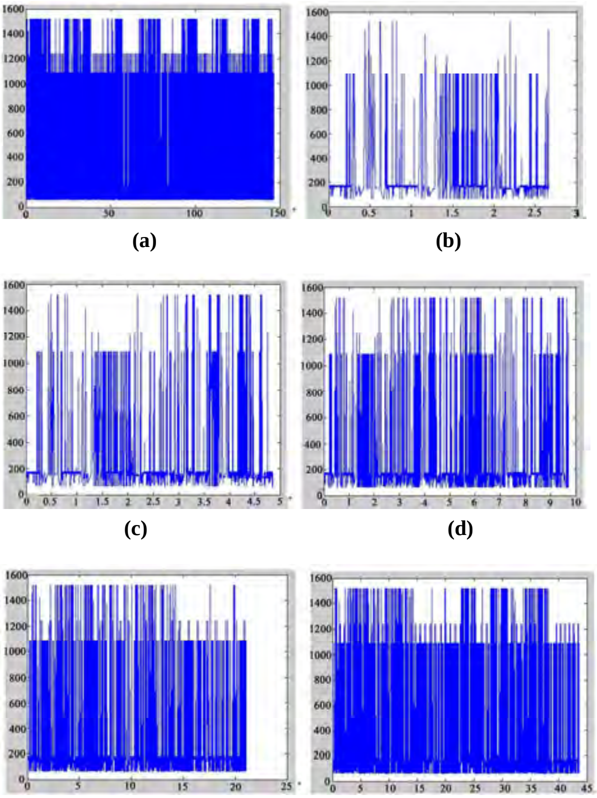


Figure 4. Hurst index estimation for data series E.



(a)

(b)

(c)

(d)

(e)

(f)

Figure 5. BC-pAu989 data series. (a)BC-pAu989 data series; (b)Length of 1024; (c) Length of 2048; (d) Length of 4096; (e) Length of 8192; (f) Length of 16384.

4.2. The Actual Network Traffic Data Validation

Finally, the proposed algorithm is applied to the BC-pAu989 data series. BC-pAu989 data series is a real network traffic data series collected in Bellcore [12].

The real data series has a clear evidence of self-similarity, shown in Figure 5(a). The sudden nature of data series with length of 1024 shown in Figure 5(b) is week. The sudden nature is enhanced significantly in Figure 5(c) decreased slightly in Figure 5(d), decreased significantly in Figure 5(e), and enhanced significantly again in Figure5(f).

The algorithm is applied to all these data series, a series of Hurst index estimation values are shown in Table 1. We can see that the Hurst index values change with the degree of sudden network traffic accordingly. The time-varying Hurst index estimation values dynamically track the local sudden degree of BC-pAu989 network traffic analyzed above.

Certainly, Figure 6 shows that this series of Hurst index fluctuate in form of the asymptotic trend around the theory value of 0.72, proving that the BC-pAu989 time

Table 1. Hurst index estimation on BC-pAu989 data series.

Data Series	Hurst Index
1024	0.7047
2048	0.7133
4096	0.7026
8192	0.6762
16384	0.7232
32768	0.7227
65536	0.7010
131072	0.7124
262144	0.7362

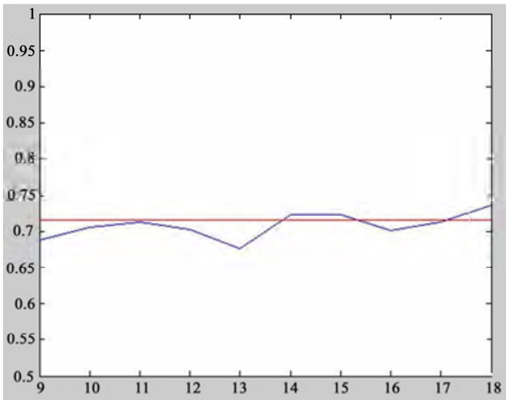


Figure 6. Tracking the local sudden traffic dynamically.

series has long-range dependence, which is consistent with the results of the literature [13].

5. Conclusions

Hurst index is an important parameter to describe the sudden nature of network traffic. To avoid the estimation error, the dynamic and asymptotic algorithm of Hurst index estimation is proposed in this paper. As the length of data series is gradually increased, the algorithm only calculates the incremental part of data series to reduce the computational complexity. A series of Hurst index values will be getting quickly in real-time. This series of Hurst index will change asymptotically from near to far infinite time scales, which is strictly in line with the mathematical definition of long-range dependence. At the same, the time-varying Hurst index values also track the local sudden information of network traffic dynamically. The effectiveness and tractability of the algorithm are validated by the simulated data series generated in OPENNET software and the real network traffic respectively. The algorithm can truly reflect the local sudden nature and the long-range dependence of network traffic.

6. References

[1] K. Park and W. Willinger, "Self-similar network traffic and performance evaluation," Wiley-Interscience, New York, 2000.

[2] S. Y. Yin and X. K. Lin, "Traffic self-similarity in mobile ad hoc networks," in Proceedings of Second IFIP International Conference Wireless and Optical Communications Networks, pp. 285–289, 2005.

[3] A. Athanasopoulos, E. ToPalis, C. D. Antonopoulos *et al.*, "Evaluation analysis of the permance of IEEE802.1lb and IEEE802.llg standards," in Proceedings of International Conference Networking, International Conference on Systems and International Conference on Mobile Communications and Leaning Technologies, available at: http://doi.ieeecomputersociety.org/l0.ll09/ICNICONSMCL, 2006.

[4] M. Jiang, M. Nikolic, S. Hardy *et al.* "Impact of self-similarity on wireless data network Performance," ICC 2001, IEEE International Conference Communications, Vol. 2, No. 11–14, pp. 477–481, 2001.

[5] J. S. Zh., H. Ming, and N. B. Shroff, "Sudden data over CDMA: MAI self-similarity, rate control and admission control and admission control," in proceedings of IEEE INFOCOM 2002, Vol. 1, No. 23–27, pp. 391–399, 2002.

[6] R. Kalden and S. Ibrahim, "Searching for self-similarity in GPRS[C]," The 5th Annual Passive & Active Measurement Workshop, PAM 2004, France, April 2004.

[7] Muradtaqqu. Methods. http://math.bu.edu/people /murad/ methods/index.html, September 2005.

[8] J. W. Wei, J. Zhang, and J. X. Wu, "A long-range dependence sliding window time-varying estimation algorithm for network traffic," Journal of Computer Research and Development, Vol. 45, No. 3, pp. 436–442, 2008.

[9] O. Cappe, E. Moulines, A. PetroPulu *et al.*, "Long-range dependence and heavy-tail modeling for teletraffic data," IEEE Signal Processing Magazine, Special Issue on Analysis and Modeling of High-Speed Data Network Traffic," Vol. 19, No. 5, pp. 14–27, May 2002.

[10] D. R. Figueiredo, B. Liu, V. Misra, and D. Towsley, "On the autocorrelation structure of TCP traffic, Computer Networks, Vol. 40, No. 3, pp. 339–361, 2002.

[11] D. R. Figueiredo, B. Y. Liu, A. Feldmann, V. Misra, D. Towsley, and W. Willinger, "On TCP and self-similar traffic," Performance Evaluation, 2005.

[12] P. Danzig, J. Mogul, and V. Paxaon, "Traces available in the interact traffic archive," http://ira.ee.1b1.gov/html /tmces.html. September 2005.

[13] W. E. Leland, M. S. Taqqu, and W. Willinger, *et a1.* "On the self-similar nature of ethernet traffic," IEEE/ACM Transactions on Networking, Vol. 2, No. 1, pp. l–15. 1994.

Int. J. Communications, Network and System Sciences, **2010, 3, 173-176**
doi:10.4236/ijcns.2010.32024 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Efficient Adaptive Algorithms for DOA Estimation in Wireless Communications

J. G. ARCEO-OLAGUE[1], D. H. COVARRUBIAS-ROSALES[2], J. M. LUNA-RIVERA[3], A. ÁNGELES-VALENCIA[4]
[1]*Electrical Engineering, Communications and Electronics, UAZ, Zacatecas, México*
[2]*Electronics and Telecommunication Department, CICESE Research Center, Ensenada, México*
[3]*Electronics Department, College of Sciences, UASLP, San Luis Potosi, México*
[4]*CITEDI Research Center, Tijuana, México*
Email: *arceojg@ymail.com, dacoro@cicese.mx, mlr@fciencias.uaslp.mx, alfang@ieee.org*

Received October 29, 2009; *revised December* 21, 2009; *accepted December* 28, 2009

Abstract

The problem of direction-of-arrival (DOA) estimation in mobile communication systems requires efficient algorithms with a high spatial resolution and a low computational complexity when moving sources are considered. MUSIC is known as a high resolution algorithm for DOA estimation but, this method demands high computational complexity to compute the signal subspace of the time-varying data of the correlation matrix. This paper focuses on MUSIC using subspace tracking methods, such as Bi-SVD and PAST, to carry out iterative DOA estimation. Accuracy and the processing time of both methods are evaluated and compared with the results of MUSIC. These results show the potential of Bi-SVD and PAST to reduce the processing time and to improve the accuracy when the number of snapshots and the source angular variation increases, assessing the source location for a dynamic mobile cellular environment.

Keywords: Direction-of-Arrival, Accuracy, Computational Complexity, Subspace Tracking, Array Antennas

1. Introduction

In mobile communication systems source localization with distributed sensor arrays can be performed by estimating the direction of arrival (DOA) of the signals coming from mobile terminals (sources). DOA is one of the most challenging problems which one has to solve for localizing and tracking multiple moving source as in radar, mobile communications and in other areas.

Within the class of high-resolution source localization methods, the MUSIC (MUltiple SIgnal Classification) method [1] has received the most attention and has been widely studied [2–4]. However, when the problem of moving sources is addressed, this method demands heavy computational load due to the decomposition of the subspace of the vectors of the correlation matrix, which is estimated based on an N-size sample [4–7]. For fast DOA tracking, the subspace tracking methods iteratively estimate the signal subspace which in this case is involved in the MUSIC spectrum. To address this problem, Bi-SVD (Bi-Iteration Singular-Value Decomposition) [4,5] and PAST (Projection Approximation Subspace Tracking) [6] have been proposed and investigated due to their capabilities of successively updating (tracking) eigenvectors in the signal subspace of a correlation matrix [7–9]. Compared with the rectangular window in MUSIC conventional, in this signal subspace algorithms an exponential window is used for the signal processing. This window considers a forgetting factor to reduce the effect of the past observations which are down-weighted.

In this work, we present and evaluate a scheme that incorporates those subspace tracking methods into MUSIC, and shows a comparison of DOA successive estimation performance. For a small sample size in [4] is found that all the methods show the same root mean square error (RMSE) after convergence, and that all the methods can satisfactorily estimate DOAs. In contrast, we use a larger sample size to estimate the DOA considering the source's mobility. Compared with MUSIC conventional, our experimental analysis on simulated data demonstrates the accuracy improvement of both Bi-SVD and PAST algorithms to assess the sources DOA for a dynamic environment in terms of the root mean square error (RMSE).

2. System Model

Let ***x(t)*** be the *Kx1* data vector at the output of an uniform linear array of K elements spaced a distance d between two consecutive elements. Assuming that the array of K sensors receive D narrowband signal waves from far-field sources with the same known center frequency, then the output vector $\boldsymbol{x(t)}$ can be described as:

$$\boldsymbol{x}(t)=\sum_{l=0}^{D-1} s_l(t)\boldsymbol{a}_l(\theta_l)+\boldsymbol{n}(t)=\boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{n}(t) \qquad (1)$$

where the $Dx1$ vector $\boldsymbol{s(t)}=[s_0(t), \ldots, s_{D-1}(t)]^T$ and the $Kx1$ vector $\boldsymbol{n(t)}=[n_1(t), \ldots, n_K(t)]^T$ denote the complex amplitudes of the signals and the measurement noise at time t. The KxD steering matrix $\boldsymbol{A}$ is composed by D steering vectors $(Kx1)$, with $\boldsymbol{a}_l(\theta_l)$=[1, exp($-j(2\pi/\lambda)d\ sen\theta_l$), …, exp($-j(2\pi/\lambda)d(K-1)sen\theta_l$)], θ_l is the direction of arrival and λ is the carrier wavelength. The noise samples are assumed to be zero mean with variance σ^2.

From the eigen decomposition of the correlation matrix $\boldsymbol{R}=E[\boldsymbol{x(t)x}^H(t)]$ are obtained $\boldsymbol{E}_s(t)$ and $\boldsymbol{E}_n(t)$ corresponding to the signal and noise subspaces respectively, [1]. The signal subspace is formed by the D eigenvectors corresponding to the D largest eigenvalues. Usually, the MUSIC spectrum is computed with the noise subspace to improve the spectral resolution. Besides that, the subspace tracking algorithms approaches the signal subspace $\boldsymbol{E}_s(t)$, then, the identity $\boldsymbol{E}_n(t)\boldsymbol{E}_n(t)^H=\boldsymbol{I}-\boldsymbol{E}_s(t)\boldsymbol{E}_s(t)^H$ must be employed to track the DOA from the MUSIC spectrum as follows [4]:

$$P_{MU}(\theta,t)=\frac{1}{\boldsymbol{a}^H(\theta)(\boldsymbol{I}-\boldsymbol{E}_s(t)\boldsymbol{E}_s(t)^H)\boldsymbol{a}(\theta)} \qquad (2)$$

In what follows, are explained Bi-SVD and PAST algorithms for the $\boldsymbol{E}_s(t)$ signal subspace tracking.

2.1. Bi-SVD Algorithm

Let $\boldsymbol{x(t)}$ be a random complex process with the correlation matrix $\boldsymbol{R}$=E[$\boldsymbol{x(t)x(t)}^H$] and the signal subspace $\boldsymbol{E}_s(t)$ contains the D dominant eigenvectors. For subspace tracking, the data matrix $\boldsymbol{X(t)}$ can be updated in the time according [5]:

$$\boldsymbol{X}(t)=\begin{bmatrix}(1-\beta)^{1/2}\,\boldsymbol{x}^T(t)\\ \beta^{1/2}\boldsymbol{X}(t-1)\end{bmatrix} \qquad (3)$$

where $\boldsymbol{x(t)}$ is the current data vector, $0<\beta<1$ is the forgetting factor and $\boldsymbol{X(t\text{-}1)}$ is the data matrix at time instant $t\text{-}1$. Consider the following bi-iteration applied on the (NxK) data matrix $\boldsymbol{X(t)}$ as given in [5]:

$$\boldsymbol{Q}_A(0)=\begin{bmatrix}\boldsymbol{I}_D\\ \boldsymbol{0}\end{bmatrix};\ \left[\begin{array}{c}\boldsymbol{B}(t)=\boldsymbol{X}(t)\boldsymbol{Q}_A(t-1)\\ \boldsymbol{B}(t)=\boldsymbol{Q}_B(t)\boldsymbol{R}_B(t)\\ \hline \boldsymbol{A}(t)=\boldsymbol{X}^H(t)\boldsymbol{Q}_B(t)\\ \boldsymbol{A}(t)=\boldsymbol{Q}_A(t)\boldsymbol{R}_A(t)\end{array}\right], \qquad (4)$$

where $\boldsymbol{Q}_A(0)$ is the initial value for $\boldsymbol{Q}_A(t)$. Here, $\boldsymbol{Q}_A(t)$ denotes an estimate of the signal subspace $\boldsymbol{E}_s(t)$, $\boldsymbol{R}_A(t)$ and $\boldsymbol{R}_B(t)$ are D-dimensional upper-triangular matrices obtained from QR decomposition. Notice that the data matrix $\boldsymbol{X(t)}$ in (3) is updated recursively then it results in a growing matrix. A solution of this problem is to approximate $\boldsymbol{X(t)}$ as [5]:

$$\hat{\boldsymbol{X}}(t)=\boldsymbol{Q}_B(t)\boldsymbol{R}_B(t)\boldsymbol{Q}_A^H(t-1). \qquad (5)$$

With the suboptimal approximation in (4), a fast exponential window based Bi-SVD subspace tracking algorithm was developed in [5]. Now, projecting $\boldsymbol{x(t)}$ into the space spanned by $\boldsymbol{Q}_A(t\text{-}1)$ the complement of its orthogonal projection is

$$\boldsymbol{x}_\perp(t)=\boldsymbol{x}(t)-\boldsymbol{Q}_A(t-1)\boldsymbol{h}(t), \qquad (6)$$

where $\boldsymbol{h}(t)=\boldsymbol{Q}_A^H(t-1)\boldsymbol{x}(t)$. The normalization of $\boldsymbol{x}_\perp(t)$ results in

$$\bar{\boldsymbol{x}}_\perp(t)=e_x^{-1/2}(t)\boldsymbol{x}_\perp(t), \qquad (7)$$

with $e_x(t)=\boldsymbol{x}_\perp^H(t)\boldsymbol{x}_\perp(t)$. From (5), the decomposition for $\boldsymbol{x(t)}$ becomes:

$$\boldsymbol{x}(t)=e_x^{1/2}(t)\bar{\boldsymbol{x}}_\perp(t)+\boldsymbol{Q}_A(t-1)\boldsymbol{h}(t). \qquad (8)$$

Using (5) and (8) into (3) and then substituting $\boldsymbol{X(t)}$ in (4), we obtain $\boldsymbol{Q}_A(t)$ after several computations as follows:

$$\boldsymbol{Q}_A(t)=\boldsymbol{Q}_A(t-1)\boldsymbol{\Theta}_A(t)+\bar{\boldsymbol{x}}_\perp(t)\boldsymbol{f}^H(t), \qquad (9)$$

where

$$\boldsymbol{f}(t)=\boldsymbol{Q}_A^H(t)\bar{\boldsymbol{x}}_\perp(t) \qquad (10)$$

$$\boldsymbol{\Theta}_A(t)=\boldsymbol{Q}_A^H(t-1)\boldsymbol{Q}_A(t) \qquad (11)$$

In Equation 9, the subspaces $\boldsymbol{R}_A(t)$, $\boldsymbol{R}_B(t)$ and $\boldsymbol{\Theta}_A(t)$ are updated successively with initial values $\boldsymbol{R}_A(0)=\boldsymbol{0}$, $\boldsymbol{R}_B(0)=\boldsymbol{I}_D$ and $\boldsymbol{\Theta}_A(0)=\boldsymbol{I}_D$.

2.2. PAST Algorithm

The PAST algorithm is based on the idea that $\boldsymbol{Q}_A(t)$ is the signal subspace when is minimized (12).

$$J(\boldsymbol{Q}_A(t))=\sum_{i=1}^{t}\beta^{t-i}\left\|\boldsymbol{x}(i)-\boldsymbol{Q}_A(t)\boldsymbol{Q}_A{}^H(t)\boldsymbol{x}(i)\right\|^2 \qquad (12)$$

An exponentially weighted sum is considered into $J(\boldsymbol{Q}_A(t))$, where β is the forgetting factor. At time t, all sample vectors are involved in the estimation of the signal subspace. The algorithm PAST approximates $\boldsymbol{Q}_A{}^H(t)\boldsymbol{x}(i)$ in (12) with $\boldsymbol{y}(i)=\boldsymbol{Q}_A{}^H(i\text{-}1)\boldsymbol{x}(i)$ by using the unknown projection of $\boldsymbol{x}(i)$ into the columns of $\boldsymbol{Q}_A(t)$ [6]. The approximated cost function is then given by:

$$J'(\boldsymbol{Q}_A(t))=\sum_{i=1}^{t}\beta^{t-i}\left\|\boldsymbol{x}(i)-\boldsymbol{Q}_A(t)\boldsymbol{y}(i)\right\|^2. \qquad (13)$$

The choice of $\boldsymbol{Q}_A(t)$ that minimize $J'(\boldsymbol{Q}_A(t))$ is [6]

$$\boldsymbol{Q}_A(t) = \boldsymbol{C}_{xy}(t)\boldsymbol{C}_{yy}^{-1}(t) \tag{14}$$

$$\boldsymbol{C}_{xy}(t) = \beta\boldsymbol{C}_{xy}(t-1) + \boldsymbol{x}(t)\boldsymbol{y}^H(t) \tag{15}$$

$$\boldsymbol{C}_{yy}(t) = \beta\boldsymbol{C}_{yy}(t-1) + \boldsymbol{y}(t)\boldsymbol{y}^H(t) \tag{16}$$

Therefore, the minimization process yields that $J(\boldsymbol{Q}_A(t))$ on only internal noise, resulting in $\boldsymbol{Q}_A(t)$ equal to the signal subspace matrix $\boldsymbol{E}_s(t)$, obtained recursively from (14).

3. Simulation Results

Successive update algorithms using Bi-SVD and PAST into MUSIC are compared in estimation accuracy with the conventional MUSIC. The estimation accuracy is evaluated by RMSE (Root Mean Square Error), defined as follows:

$$RMSE(t) = \sqrt{\frac{1}{D}\sum_{l=1}^{D}\left\{\frac{1}{P}\sum_{p=1}^{P}\left(\hat{\theta}_{lp}(t) - \theta_l\right)^2\right\}} \tag{17}$$

where $\hat{\theta}_{lp}$ defines the estimated DOAs, D denotes the number of sources and P the number of independent trials.

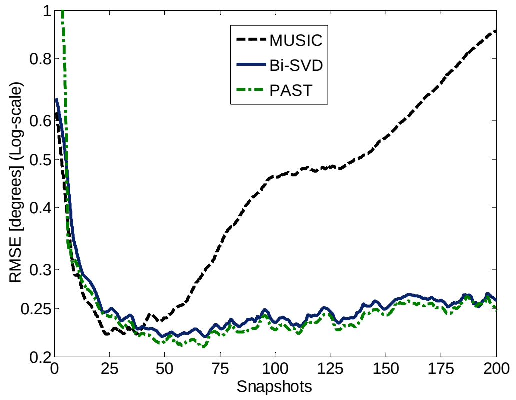


Figure 1. RMSE evaluation in DOA estimation (dynamic environment). SNR=20 dB spatial variation δ=0.01°.

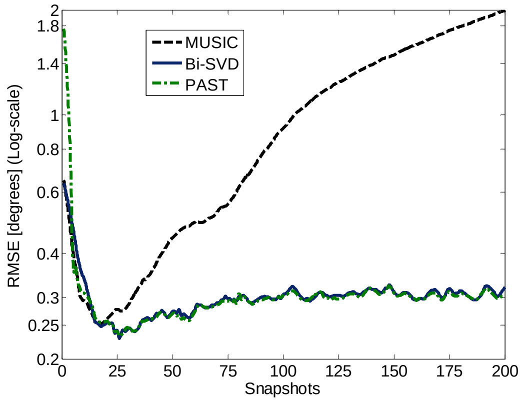


Figure 2. RMSE evaluation in DOA estimation (dynamic environment). SNR=20 dB spatial variation δ=0.02°.

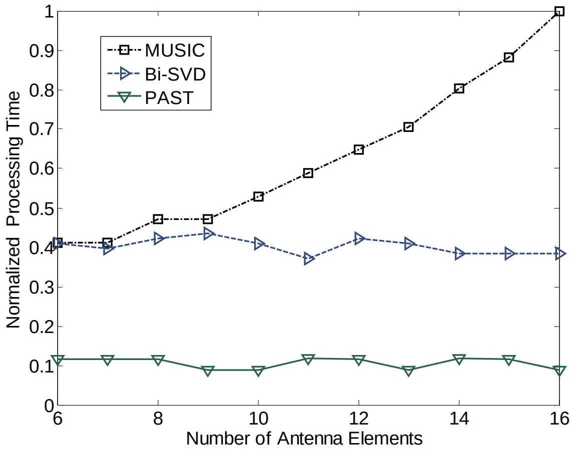


Figure 3. Normalized processing time. Signal subspace estimation considering variation of the number of array antenna elements.

For the simulations, we consider a uniform linear array (ULA) of K=10 sensors separated a half wavelength of the actual source signals. We assume the case of two narrowband signals in the far-field, located at [-30°, 50°] and separated from the antenna array by 930m and 425m respectively, transmitting with the same power and the same carrier frequency of *f*=1900 MHz. For this case, the number of independent trials is set to *P*=1000, assuming a forgetting factor of β=0.9.

Simulation results are obtained by incorporating over time the sources' spatial angular variation δ as the dynamic component. Figure 1 compares the RMSE obtained by the PAST, Bi-SVD and MUSIC methods in a dynamic environment. For each snapshot, the angular variation δ is increased a factor of 0.01 degrees, that is, the sources location at time *t* are given by [-30°+$t\delta$, 50°+ $t\delta$]. The results in Figure 1 show a small RMSE for PAST and Bi-SVD methods. After a 40 snapshots the accuracy of the MUSIC method is degraded with an accumulated RMSE value of 0.9° approximately. On the other hand, for Bi-SVD and PAST, their RMSE error increases slowly until a value close to 0.26°. In this case, comparing the subspace tracking algorithms into MUSIC with the conventional, they improve the accuracy as the number of snapshots increases.

In a second case, is considered an increase of angular variation (δ=0.02°). From Figure 2, we found that after 20 snapshots MUSIC conventional rapidly increases its RMSE error from 0.25° to 2° on the maximum number of snapshots considered. In this case, the accuracy of the MUSIC method is remarkably worse compared with the maximum RMSE error of 0.3° obtained by the subspace tracking algorithms. The forgetting factor used by subspace tracking algorithms improved the accuracy into MUSIC to estimate the DOA, while MUSIC conventional fails as the number of snapshots and angular variation increases.

Finally, the processing time of the methods is evaluated and normalized taking into account the maximum processing time. In Figure 3 the processing time is normalized with the maximum time obtained during the simulations. We observe that the processing time of the PAST algorithm is smaller than both Bi-SVD and MUSIC. More interestingly, the results show that PAST and Bi-SVD processing times are independent of the number of antenna array elements, situation that does not hold by MUSIC which grows linearly with the number of antenna elements.

4. Conclusions

The PAST and Bi-SVD subspace-tracking algorithms were applied to MUSIC in order to face the accuracy and processing time problems in DOA estimation. The simulations show good results as far as the application of the subspace-tracking algorithms. Our experimental analysis on simulated data demonstrates the potential reduction in processing time and accuracy improvement to assess the sources DOA when the number of snapshots and angular variation increases. It is shown that as source mobility increases, it produces a higher error for MUSIC conventional. However, the PAST and Bi-SVD algorithms show a lower estimation error, varying quietly with the mobility of sources. On the other hand, the processing time of PAST algorithm is smaller than both Bi-SVD and MUSIC conventional. The processing time required by these subspace tracking algorithms is independent of the number of antenna array elements.

5. Acknowledgments

This work was supported by PROMEP and the Mexican National Science and Technology Council, CONACyT, under Grant 52374.

6. References

[1] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Transactions Antennas and Propagation, Vol. 34, pp. 276–280, 1986

[2] D. H. Covarrubias and J. G. Arceo, “Improving resolution on DOA estimation in a multipath macrocell environment,” Proc. IEEE SympoTIC, Bratislava Slovak Republic, pp. 39–42, 2004.

[3] J. G. Arceo, D. H. Covarrubias, and J. M. Luna, “Efficiency evaluation of the unconditional maximum likelihood estimator for near-field source DOA estimation,” ETRI Journal, Vol. 28, pp. 761–769, 2006.

[4] N. Kikuma, “Iterative DOA estimation using subspace tracking methods and adaptative beamforming,” IEICE Transactions on Communication, Vol. E88-B, pp. 1818–1828, 2005.

[5] P. Strobach, “Bi-Iteration SVD subspace tracking algorithms,” IEEE Trans. Signal Processing, Vol. 45, pp. 1222–1240, 1997.

[6] B. Yang, “Projection approximation subspace tracking,” IEEE Transactions Signal Proceedings, Vol. 43, pp. 95–107, 1995.

[7] P. Strobach, “Bi-Iteration multiple invariance subspace tracking and adaptive ESPRIT,” IEEE Transactions on Signal Processing, Vol. 48, pp. 442–456, 2000.

[8] A. Kuchar, M. Tangemann, and E. Bonek, “A real-time DOA-based smart antenna processor,” IEEE Transactions on Vehicular Technology, Vol. 51, pp. 1279–1293, 2002.

[9] S. Ouyang and Y. Hua, “Bi-Iterative least-square method for subspace tracking,” IEEE Transactions on Signal Processing, Vol. 53, pp. 2984–2996, 2005.

Int. J. Communications, Network and System Sciences, **2010, 3, 177-182**
doi:10.4236/ijcns.2010.32025 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Online Detection of Network Traffic Anomalies Using Degree Distributions

Wuzuo WANG, Weidong WU
Department of Computer Science, Wuhan University of Science & Technology, Wuhan, China
Email: *wzwang*888@163.*com*, *wwdtylwt*@163.*com*

Received October 19, 2009; *revised December* 24, 2009; *accepted January* 10, 2010

Abstract

Diagnosing traffic anomalies rapidly and accurately is critical to the efficient operation of large computer networks. However, it is still a challenge for network administrators. One problem is that the amount of traffic data does not allow real-time analysis of details. Another problem is that some generic detection metrics possess lower capabilities on diagnosing anomalies. To overcome these problems, we propose a system model with an explicit algorithm to perform on-line traffic analysis. In this scheme, we first make use of degree distributions to effectively profile traffic features, and then use the entropy to determine and report changes of degree distributions, which changes of entropy values can accurately differentiate a massive network event, normal or anomalous by adaptive threshold. Evaluations of this scheme demonstrate that it is feasible and efficient for on-line anomaly detection in practice via simulations, using traffic trace collected at high-speed link.

Keywords: Anomaly Detection, Degree Distributions, Entropy

1. Introduction

Anomalies are unusual and significant changes in a network's traffic levels, which can create congestion in the network and stress resource utilization in a router. Network operators need to accurately detect traffic anomalies in a timely fashion. Without this kind of capability, networks are not able to operate efficiently or reliably. Researchers have approached traffic anomaly detection using various techniques from simple volume-based analysis [1–3] to network flow distribution-based analysis [4]. While recent studies demonstrate that entropy-based anomaly detection obviously has some advantages [5]. This approach is to capture fine-grained patterns in traffic distributions that simple volume based metrics cannot identify. What's more, the use of entropy for tracking changes in traffic distributions provides two significant benefits. First, the use of entropy can increase the sensitivity of detection to uncover anomalies incidents that may not manifest as volume anomalies. Second, using such traffic features provides additional diagnostic information into the nature of the anomalous incidents (e.g., making distinction among worms, DDoS attack, and scans) that is not available from just volume-based anomaly detection.

In general, most researchers consider flow-header features (e.g., IP addresses, ports, and flow-sizes) as candidates for entropy based anomaly detection. However, Port and address distributions with pair-wise correlation scores greater than 0.95, which arises due to the nature of the underlying traffic patterns [6]. Intrinsically, the anomalies detected by the port and address distributions overlap significantly. Furthermore, anomalous scan, DoS, and P2P activity are not subtly detected by port and address distributions, or only high-magnitude events can be detect that would have appeared as traffic volume anomalies. Considering the limited utility of port and address distributions, we should select traffic distributions as candidates for entropy based anomaly detection with care, and in particular we should look beyond simple port and address based distributions.

In this work we propose an anomaly detection mechanism using degree distributions to improve the detect abilities of port and address. We use in- and out-degree distributions to measure the number of distinct destination/source IP addresses that each host communicates with. For each value of in-degree (out-degree), we calculate the entropy to diagnose anomaly. Note that we chose source/destination IP addresses as unique candidate metric, not both address and port. There is no need to use different distributions of possessing same underlying properties to increase overheads of computation. To keep up with on-line traffic analysis, the essence to capture dynamic network traffic, we introduce a sliding windows mechanism with fixed time width.

The rest of this paper is organized as follows. Section 2 surveys related work. Section 3 briefly describes the basic theory of our detection scheme, including computation on the entropy values of degree distributions. Section 4 presents an overview of our scheme and describes the anomaly detection methodology. Section 5 evaluates the effectiveness of the proposed scheme. Section 6 concludes the paper.

2. Related Work

Anomaly detection has been studied widely, and has received considerable attention recently. Most of the work in the recent research and commercial literature (for e.g., [7–9]) has treated anomalies as deviations in the overall traffic volume (number of bytes or packets). Volume based detection schemes have been successful in isolating large traffic changes (such as bandwidth flooding attacks), but a large class of anomalies do not cause detectable disruptions in traffic volume. In contrast, we demonstrate the utility of a more sophisticated treatment of anomalies, as events that alter the distribution of traffic features.

Nowadays, a number of works have focused on using traffic distributions to diagnose anomalies. Feinstein *et al.* [10] used the distribution of source addresses to detect DDoS attack. Similarly, Karamcheti *et al.* [11] used inverse distributions of packet contents to detect malicious network traffic and Thottan *et al.* [12] used statistical distribution of the individual MIB variables to detect abrupt changes of network traffic. We use degree distributions to effectively profile traffic features, which can capture abnormal changes of traffic in a sensitive manner.

A variety of statistical anomaly detection techniques have been proposed to detect network-wide anomalies. Particularly, the entropy-based approaches have been demonstrated the accuracy and efficiency in detecting anomalies in the traffic matrix time series. Lakhina *et al.* [8] used entropy and subspace methods to mine traffic anomalies from network wide traffic data repositories. Gu *et al.* [13] used maximum and relative entropy to develop a behavior-based anomaly detection method. In [13], the maximum entropy-based baseline distribution is constructed from pre-labeled training data, but how this baseline is adapting itself to the dynamics of network traffic remains unclear. We propose a mechanism to construct adaptive baseline according to the dynamic network traffic during the measurement period, and adjust the baseline in a particular time span.

Online detection of anomalies suffers to compute real-time statistic from the large of traffic data. Xu *et al.* [14] used 5-tuple flow distribution (i.e., srcaddr, dstaddr, srcport, dstport, protocol) to do traffic analysis leads to intensive memory and high overhead on processing capacity. Some online intrusion detection systems, such as FlowMatrix [15] and Snort [16] match packets to a pre-defined set of rules, making them unable to detect unknown anomalies. In contrast, we consider the high correlation of address and port, and use address as unique metric, instead of 5-tuple, to compute entropy values of degree distributions for detecting anomalies. Our scheme not only alleviate overhead of computing during online analysis stage, but outperform rule-based approaches to uncover new anomaly types.

3. The Basic Theory

As we know, most traffic anomalies share a common characteristic [17]: they induce abnormal change in flow-header features distribution, such as source and destination addresses and ports, which show dispersed or concentrated distribution.

For example, Figure 1 displays flow-header features distribution of three types of attacks (graphs (a) (b) (c)). Let us highlight some interesting cases of graphs. Figure 1(a) displays a typical distributed denial-of-service (DDoS) attack. In such cases, a lot of hosts send traffic towards a particular (single) host. Similarly, many Internet worms spread by sending random probes (i.e., towards randomly generated a great number of destination IP addresses) from an infected computer to infect other vulnerable computers (Figure 1(b)). In some scan events, a single host scanning the random destination host or a set of random source host scanning a single destination host (Figure 1(c)).

We can conclude some information from the above analysis: In each type of attack, the source or destination

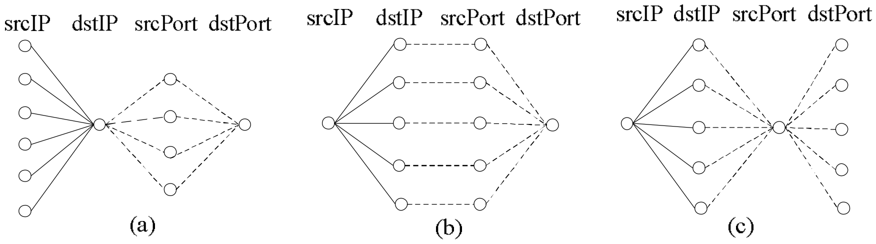


Figure 1. Source/destination IPs and ports distribution patterns of Anomalous events.

Table 1. IP distributions of anomalous events.

Anomaly Type	srcIP	dstIP
DDoS	Random	Fixed
Worms	Fixed	Random
Single Scanner	Fixed	Random

addresses present random or fixed state (shown in Table 1). One may naturally wonder 1) what metrics can accurately profile these anomaly traffic features, and obviously indicate the occurrence of such attacks mentioned above; 2) how to effectively quantify the magnitude of anomaly, and expose unusual traffic behavior.

From the Figure 1(a), we know six source hosts connect a specific destination host. Figure 1(b) illustrates a single source host connects five destination hosts. These inspired us to introduce in- and out-degree distributions to state relationship of source and destination address. For an end-host *X*, the out-degree is the number of distinct IP addresses that *X* contacts, and the in-degree is the number of distinct IP addresses that contact *X*. For example, in-degree of the destination host is six in Figure 1(a), out-degree of the source host is five in Figure 1(b).

Intuitively, in- and out-degree can effectively encapsulate and capture features of the underlying traffic distribution. In addition, entropy is an appropriate metric to manifest dispersed or concentrated state of degree distribution. The more concentrated it is, the less entropy values it is, or vice versa. Naturally, we use the entropy to determine and report changes of degree distributions, which changes of entropy values can sensitively represent variation of traffic feature distribution and designate unusual changes as an anomaly.

The natural definition of entropy in the context of this paper is the expression as follows:

$$H(x) \equiv -\sum_{i=1}^{n} p(x_i)\log(p(x_i)) \qquad (1)$$

where x_1, ..., x_N is the range of values for random variable *X*, and $p(x_i)$ represents the probability that *X* takes the value x_i. For each value of in-degree (out-degree) $x_{i,}$ we calculate the probability

$$p(x_i) = \frac{\text{Number of hosts with in-degree } x_i}{\text{Total number of hosts}} \qquad (2)$$

Useless otherwise specified, all log function in this paper are to the base 2 and we define 0log0 = 0.

Often it is useful to normalize the value to expediently compare entropy across different measurement periods. For this purpose, we define the standardized entropy (between zero and one) to be H/log *t*, where *t* is the number of distinct in-degree (similarly out-degree) values observed during the measurement interval.

4. Diagnosis Methodology

In this section, we first give an overview of the system model and the design notion of our scheme. Second, we describe our strategy of adaptive detection threshold setup. Then we present a proper algorithm for computing the entropy and self-adjusting the threshold to raise an alert when attacks happed. Finally, we show how our scheme works in detail.

4.1. System Model

The overall architecture of our scheme consists of three main functional parts: the processing engine (backend), database and WebGUI (front end). The processing engine carries out an explicit algorithm for communicating between WebGUI and database. The engine implement several aspects of task as follows: 1) it received NetFlow [18] records from capable source, such as routers, switches, firewalls, etc. in a particular manner, and store the data across a buffer into the database, 2) it obtained associated parameters are available to compute entropy values of degree distributions from the raw traffic statistics by using a single SQL query. This is a major benefit of keeping the raw traffic statistics in a database, 3) it can automatically adjust detection threshold according to the network state during the measurement period. The database provides structured storage for the traffic statistics and simplifies the computation about entropy values of degree distributions. The WebGUI frontend provides the flexibility of detection result graphically display.

4.2. Adaptive Detection Threshold Setup

To diagnosis anomalies, we must find a way to clearly differentiate network anomalies from normal behavior. Therefore we introduce a baseline method, which first define baseline values to represent steady "normal" behavior, and non-steady behavior which deviate from the baseline are then flagged at those points in time. But how far the deviation may be identified as anomalies we should take a further analysis.

During the measurement period, we first compute entropy values of degree distributions in each time interval, and then compute mean entropy as baseline in a particular time span. In addition, we use variance to reflect deviation between normal and abnormal behavior.

Let us assume, the measured entropy *Y* be a random variable with mean $E(y) = m$ and $\mathrm{var}(y) = s^2$. Then, the Chebyshev inequality states that:

$$p(|\, y - m \,| \geq e) \leq \frac{s^2}{e^2}, \text{ for any } e > 0 \qquad （3）$$

Therefore, we can define a band of $m \pm 2 * s$ as a normal region, where the proportion of observed entropy values falling in the region is at least 75%. Namely, the threshold is $m \pm 2 * s$. Beyond this normal region, the entropy represents traffic events is anomalous and assigned a severity level depending upon its deviation from the normal region.

Network traffic may change in different time or date. So the baseline will be changeable. One problem is how to automatically adjust the baseline to fit the normal behavior. From our experience, we determine a fixed time span (i.e., 30minutes) to self-adjust baseline according to network environment.

4.3. Algorithm

To keep up with on-line traffic analysis, our algorithm must be lightweight in terms of both store and retrieve data. Firstly, we design a buffer between data source and database to leverage store and retrieve. Secondly, considering many attacks today are only several minutes in duration, such as DDoS attacks generally last for only two minutes, we set a short time window (i.e.,30s) with limited (srcaddr, dstaddr) records to achieve high detection resolution and relatively low constraints on speed of store data and query database.

Conceptually, the algorithm can be divided into three stages. In the first stage we configure Netflow to page out traffic statistic in specific time span, and pre-define a threshold according to the training data to rule out anomalous entropy values, so as to accurately calibrate baseline during the measurement period. Note that adaptive threshold takes into effect in the detection process. In the second stage, processing stage, we repeatedly compute entropy values in fixed time interval with a sliding window. In the post-processing stage, we setup threshold by calculating mean entropy and variance for the next detection process. The pseudo code for this algorithm is shown in Algorithm 1.

```
Algorithm 1: Online anomaly detection algorithm
1.  Pre-processing stage
2.  Configure Netflow: paged out data to buffer every
    five minutes
3.  Initialize: pre-define threshold
4.  Processing stage
5.  Sliding window with fixed time, T(T=30s) and Load
    data using 2-tuple (srcaddr,dstaddr) into database
6.  SELECT the total number of host → sumhost
7.  SELECT the number of hosts with degree $x_i$
    →numhost[i]
8.  Count rows of different degree → numdiffdegree
9.  for i:= 1 to numdiffdegree do
10.         numhost[i]/sumhost →$p(x_i)$
11.         Compute and normalize the $H(x)$ →$y_i$
12. Repeat   5-11
13. Post-processing stage
14. Rule out $y_i$ which beyond the threshold
15. $avg(y_1,y_2,\ldots,y_{60}) \rightarrow$ baseline: $m$
16. $avg( ( Y- m )^2 ) \rightarrow s^2$          $Y=y_k$, $k=1,2,\ldots$
17. setup threshold:  $m \pm 2 * s$
```

4.4. Implementation Details

There are two working procedure in our anomaly detection scheme: deployment and measurement. First, our scheme must be deployed properly, such that it receives NetFlow records on available measurement network. We should configure internal NetFlow sources that handle traffic from corporate hosts to Internet and vice versa such as routers, switches and firewalls to export NetFlows to the processing engine server. For best result and more visibility make sure those sources deal with clear, not NATed traffic. Second, we assume that training traffic is devoid of any attack and the characterization of traffic features acts as a normal profile. The normal profile is used to calculate the pre-define thresholds. And then our scheme enters fully operational mode. In this mode the threshold is constantly compared with the current entropy value of degree distributions derived from incoming NetFlows. Alarms are generated if the entropy values differ beyond allowed tolerances. Note that associated thresholds are self-adjustable as they're calculated by the processed data itself (NetFlows) in particular time span and periodically update thresholds without requiring dedicated periodic training interval.

5. Performance Evaluation

To evaluate the effectiveness and performance of our scheme, implemented a software prototype that measures the entropy values of in- and out-degree and have tested it with real world traffic traces.

The traces we used were drawn from cisco 7609 router at our university's modern education information center, which handle three campus traffics exchanging with the commodity ISPs (Internet Service Provider). The time of-capture of analyzed traces was selected so that our methodology could be tested against a variety of network conditions.

Throughout our experimentation, both degree distributions show remarkable similarity except for few peaks. These exceptional entropy values represent magnitude of traffic feature's distributional variations during the measurement period. We picked sample snapshots of time where peaks are observed, and show work mechanism of our scheme in the link measurements.

From the Figure 2, it can be observed that the normal traffic region between lower and upper bound are determined by the threshold in the detection process. Intrinsically, a threshold is directly determined by its baseline. Figure 2 shows the baseline of in-degree and out-degree respectively adjust at points A, B, C or D, E, F according to consecutive network state. Note that baselines are adapting themselves to the dynamics of network traffic by implementing algorithm 1 when our scheme enters fully operational mode. Once entropy values of some event changes in an arbitrary manner, the event was designated in time as an anomaly. In addition, by measuring

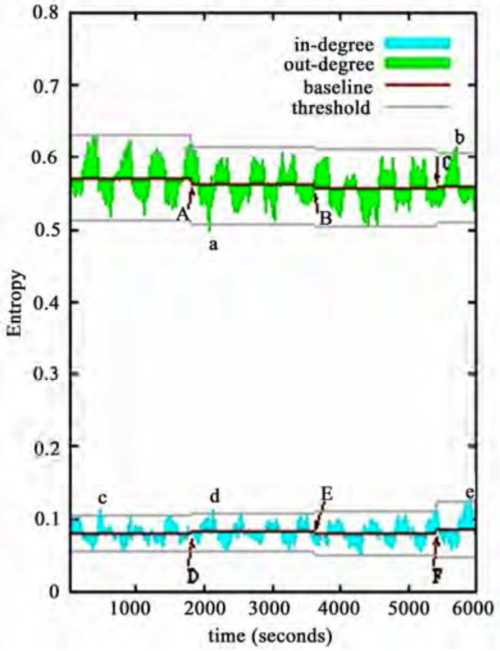


Figure 2. Sample snapshot of anomaly detection system.

the peak height and peak width of the entropy values in time series, one is able to begin to identify anomalous duration and relative intensity. Interestingly, such case happened at points a, b, c, d and e. At points c, d, e, there are peaks in entropy values of indegree but entropy values of out-degree do not show any corresponding peak. This observation has two implications. Firstly, the in-degree and out-degree are weakly correlated with each other. Secondly, entropy values of degree distributions are sensitive to these abnormal changes, even though subtle changes happened.

From the Figure 2, we also can conclude that while a network is not under attack, the entropy values for various degree distributions each fall in a narrow range. While the network is under attack, these entropy values exceed these ranges in a detectable manner.

In the following example, we choose two typical attacks which arouse traffic anomalies to validate efficiency of our approach in detail. Then we further discuss the reason for variation in entropy values of degree distributions.

Figure 3(a) shows the different changes before and during the worm outbreak. Before the outbreak time it can be seen that entropy values of in- and out-degree vary in a permitted scale, since source addresses and destination addresses do not obviously appear dispersed or concentrated state. However, during the outbreak of the worm the degree distributions change massively. The most obvious is that in- and out-degree plots change their values in different directions. Regarding the individual plots, it can be seen clearly that an obvious increase in entropy of out-degree at point f, while entropy values of

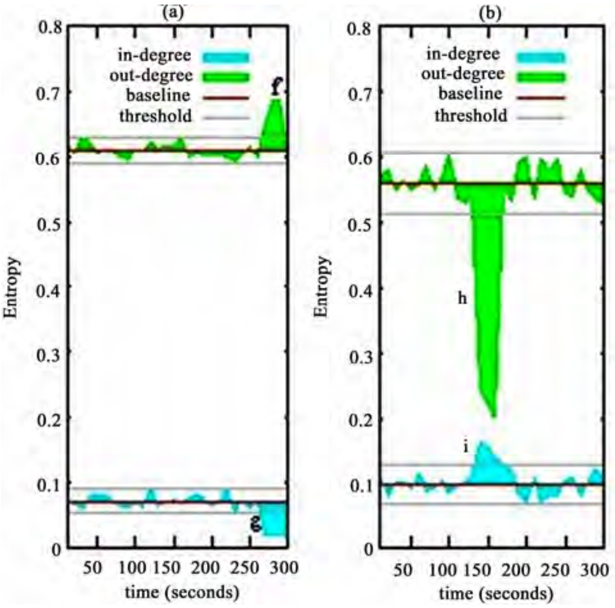


Figure 3. Examples of anomalies which we diagnose from link traffic.

in-degree leading to the opposite effect at point g. This divergent effect can be used to indicate worm anomalies or similar to worms. There is a persuasive reason behind this abnormal behavior. The change in IP address characteristics seen on a flow level (i.e. when packets belonging to a TCP connection or UDP data stream with same source/destination IP address is aggregated into one "flow") is relatively intuitive: a smaller number of infected hosts scan and connect to other vulnerable hosts in a random fashion. As a result, these flows grow to be a significant part of the set of flows seen in total, which give cause for variation to the whole traffic features distribution. On one hand, the source IP addresses of the infected hosts can be seen in many flows and since they are relatively few hosts. It means that some source IP addresses seen in flows become more fixed than in normal traffic, but the other source IP addresses show more dispersed distribution in the mass, which leading to out-degree distribution more dispersed, and hence the entropy value of out-degree significantly increase. On the other hand the destination IP addresses seen in flows will be much more random than in normal traffic, which causes a lot of hosts with in-degree 1, and hence indegree distribution appears to be more concentrated, entropy value of in-degree tends to decrease obviously.

Figure 3(b) plots an outbound DDoS attacks last less one minute (from 140s to 170s). The presence of these anomalies presents an interesting view in the structure of flow level traffic. These attacks were floods of 40-byte TCP SYN packets destined for the same host or server. The flood was reported as many "degenerate" flows, having only one packet per flow. And the flood packets had a lot of random source addresses and a fixed destination address. As a result, traffic will demonstrate a dispersed distribution for source IP addresses, namely the majority of the hosts with out-degree 1 connect to the same external destination IP addresses. It means that out-degree distribution tends to be more concentrated than in normal traffic, which sharply decreases the entropy value of out-degree at point h. From low to high rate DDoS attacks, the destination IP addresses show a small variation. But the in-degree distribution in the mass still show more concentrated state than in normal traffic, and hence the increases of entropy in in-degree at point i. This can clearly explain the onslaught of DDoS events.

6. Conclusions

In this paper we present degree distributions for detecting network traffic anomalies in IP flow data collected at our University's border router. We evaluate the scheme on network-wide traffic anomalies, which resulting from unusual changes in the real-time traffic features. We showed how to use our scheme to diagnose anomalies from simple and readily available link measurements. Rigorous experiments on real-world traffic validate our scheme obviously possess the following advantages: 1) it is accurate and efficient enough to use a little flow header features for capturing fine-grained patterns in traffic distributions. These not only reduce the on-line processing time but increase the detection abilities. 2) The use of entropy can increase the sensitivity of detection to uncover well-known or unknown anomalies and quantify traffic anomalies. 3) An adaptive threshold is available to lower false alarm rate.

Our ongoing work is further analysis traffic anomalous features, and extending the methodology proposed here to diagnose additional network-wide anomalies. In addition, lower result report latency is one of problems we consider.

7. Acknowledgment

We thank the modern education information center at Wuhan University of Science and Technology for supplying the real world traffic to experiment, and providing some technology support by their network engineers. We are also grateful to Zhi Wang in network laboratory at college of computer science and technology for helpful discussion on the subject matter.

8. References

[1] A. Lakhina, M. Crovella, and C. Diot, "Characterization of network-wide anomalies in traffic flows (short paper)," In IMC, 2004.

[2] D. Brauckhoff, B. Tellenbach, A. Wagner, A. Lakhina, and M. May, "Impact of traffic sampling on anomaly detection metrics," In Proceeding of ACM/USENIX IMC, 2006.

[3] P. Barford, J. Kline, D. Plonka, and A. Ron, "A signal analysis of network traffic anomalies," In Proceeding of IMW, 2002.

[4] A. Wagner, and B. Platter, "Entropy based worm and anomaly detection in fast IP networks," In Proceeding IEEE WETICE, 2005.

[5] A. Lakhina, M. Crovella, and C. Diot, "Mining anomalies using traffic feature distributions," In Proceeding of ACM SIGCOMM, 2005.

[6] G. Nychis, V. Sekar, D. G. Andersen, H. Kim, and H. Zhang, "An empirical evaluation of entropy-based traffic anomaly detection," In IMC, 2008.

[7] "Arbor networks," At http://www.arbornetworks.com/.

[8] A. Lakhina, M. Crovella, and C. Diot, "Characterization of network-wide anomalies in traffic flows (Short Paper)," In IMC, 2004.

[9] "Riverhead networks," At http://www.riverhead.com

[10] L. Feinstein, D. Schnackenberg, R. Balupari, and D. Kindred, "Statistical approaches to DDoS attack detection and response," In Proceedings of the DARPA Information Survivability Conference and Exposition, 2003.

[11] V. Karamcheti, D. Geiger, Z. Kedem, and S. Muthukri-Shnan, "Detecting malicious network traffic using inverse distributions of packet contents," In Proceeding of ACM SIGCOMM MineNet, 2005

[12] M. Thottan, and C. Ji, "Anomaly detection in IP networks," In IEEE TRANSACTIONS ON SIGNAL PROCESSING, August 2003.

[13] Y. Gu, A. McCallum, and D. Towsley, "Detecting anomalies in network traffic using maximum entropy estimation," In IMC, 2005.

[14] K. Xu, F. Wang, S. Bhattacharyya, and Z.-L. Zhang, "A real-time network traffic profiling system," In DSN, 2007.

[15] "FlowMatrix," At http://www.akmalabs.com/flowmatrix.php.

[16] M. Roesch, "Snort: Lightweight intrusion detection for networks," In USENIX LISA, 1999.

[17] T Karagiannis, K Papagiannaki, and M Faloutsos, "BLINC: Multilevel traffic classification in the dark," In Proceeding of ACM SIGCOMM, 2005.

[18] "CiscoNetflow," At http://www.cisco.com/en/US/tech/tk812/tsd_technology_support_technical_references_list.html.

Int. J. Communications, Network and System Sciences, **2010, 3, 183-189**
doi:10.4236/ijcns.2010.32026 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

A Modified T/2 Fractionally Spaced Coordinate Transformation Blind Equalization Algorithm

Yecai GUO[1,2], Xueqing ZHAO[1], Zhenxin LIU[1], Min GAO[1]
[1]*School of Electrical Engineering and Information, Anhui University of Science and Technology, Huainan, China*
[2]*College of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, China*
Email: *guo-yecai@126.com*

Received November 3, 2009; *revised December* 13, 2009; *accepted January* 19, 2010

Abstract

When T/2 Fractionally Spaced blind Equalization Algorithm based Constant Modulus Algorithm (T/2-FSE-CMA) is employed for equalizing higher order Quadrature Amplitude Modulation signals (QAM), it has disadvantages of low convergence speed and large Mean Square Error (MSE). For overcoming these disadvantages, a Modified T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA (T/2-FSE-MCTCMA) was proposed by analyzing the character of 16QAM signal constellations. In the proposed algorithm, real and imaginary parts of input signal of T/2 fractionally spaced blind equalizer are equalized, respectively, and output signals of equalizer are transformed to the same unit circle by coordinate transformation method, a new error function is defined after making coordinate transformation and used to adjust weight vector of T/2 fractionally spaced blind equalizer. The proposed algorithm can overcome large misjudgments of T/2 fractionally spaced blind equalization algorithm for equalizing multi-modulus higher order QAM. Simulation results with underwater acoustic channel models demonstrate that the proposed T/2-FSE-MCTCMA algorithm outperforms T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA (T/2-FSE-CTCMA) and the T/2-FSE-CMA in convergence rate and MSE.

Keywords: Blind Equalization; Underwater Acoustic Channel; Coordinate-Transformation; Fractionally Spaced Equalizer.

1. Introduction

In underwater acoustic communication system, blind equalization technique without training sequence is an important means to eliminate intersymbol interference (ISI) [1–5]. Among them, baud-spaced equalizer based on Constant Modulus Algorithm (BSE-CMA) has simple structure, but its convergence rate is slow and its steady error is large [6,7]. Whereas, fractionally-spaced equalizer based on constant modulus algorithm (FSE-CMA) is employed for equalizing constant modulus signal, its convergence rate is fast and its steady error is low [8–10]. When the FSE-CMA is used to equalize high-order QAM signal, it can produce large misjudgments and lead to large mean square error, because higher order QAM signal constellations distribute in the several known circles and its module value is not constant [11,12]. Therefore, intersymbol interference is not sufficiently eliminated.

In the paper, on the basis of analyzing the character of 16QAM signal constellations [13–15], T/2 fractionally-spaced equalizer, and the thought of coordinate transformation [16] real and imaginary parts of output signal of each sub-channel are equalized, respectively, a new constant modulus error function is defined after making coordinate transformation to output of each equalizer. A cost function based on this error function is given. Iterative formula of weight vector of T/2 fractionally-spaced equalizer is got by making the cost function minimization. Finally, a Modified T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA (T/2-FSE- MCTCMA) is established.

This paper is organized as follows. In Section 2, fractionally spaced blind equalization algorithm is described. In Section 3, the T/2-FSE-MCTCMA is proposed. The performance of the proposed T/2-FSE-MCTCMA is analyzed in Sections 4 and 5. Finally, some conclusions are obtained.

2. Fractionally Spaced Blind Equalization Algorithm

Relevant researches show that fractionally spaced equalizer is equivalent to multi-channel system model [17–19]. Its structure is shown in Figure 1 Input and output signals of this system have the same sampling rate.

In Figure 1, $a(k)$ is the transmitted signal sequence and its sampling period is T; $\boldsymbol{c}^{(i)}(k)\quad(i=0,1\cdots P-1)$ is an impulse response vector of the *i*th sub-channel and $\boldsymbol{c}^{(i)}(k)=c[(k+1)P-i-1]$; P is fractionally spaced sampling factor; $\boldsymbol{n}^{(i)}(k)$ is an additive noise vector of the *i*th sub-channel; $\boldsymbol{y}^{(i)}(k)$ is an input signal vector of the *i*th blind equalizer and written as

$$\boldsymbol{y}^{(i)}(k)=\sum_{j=0}^{N_c-1}a(j)\cdot c^{(i)}(j)+\boldsymbol{n}^{(i)}(k) \tag{1}$$

where N_c is the length of impulse response vector of channel in baud spaced equalizer.

$\boldsymbol{f}^{(i)}(k)$ is the weight vector of the *i*th sub-equalizer and written as

$$\boldsymbol{f}^{(i)}(k+1)=\boldsymbol{f}^{(i)}(k)+\mu z^{(i)}(k)e(k)\boldsymbol{y}^{(i)*}(k)\ (i=0,\cdots P-1) \tag{2}$$

where μ is defined as step size; $e(k)$ is an error function and given by $e(k)=R_2-|z(k)|^2$; R_2 denotes module value of signals and given by $R_2=\mathrm{E}\{|a(k)|^4\}/\ \mathrm{E}\{|a(k)|^2\}$.

The output signal of the whole system is given by

$$\begin{aligned}z(k)&=\sum_{i=0}^{P-1}\boldsymbol{f}^{(i)*}(k)\boldsymbol{y}^{(P-i-1)}(k)\\&=\sum_{i=0}^{P-1}\boldsymbol{f}^{(i)*}(k)[\boldsymbol{a}^{*}(k)\boldsymbol{c}^{(P-i-1)}(k)+\boldsymbol{n}^{(P-i-1)}(k)]\end{aligned} \tag{3}$$

where “*” denotes conjugate operator.

Fractionally Spaced blind Equalization algorithm based CMA(T/2-FSE-CMA）is only suitable to equalize constant modulus signals. When it is employed for equalizing multi-modulus QAM signal, it can easily produce large mean square error.

3. Modified T/2 Fractionally Spaced Coordinate Transformation Blind Equalization Algorithms

When the transmitted signal is higher order QAM signal, we make Figure 1 change in two aspects in order to obtain good equalization performance. At first, the real and imaginary parts of input signal $a(k)$ are equalized, respectively. It is equivalent to process real signals in the whole equalization process. Moreover, its computational complexity is decreased obviously comparison with that

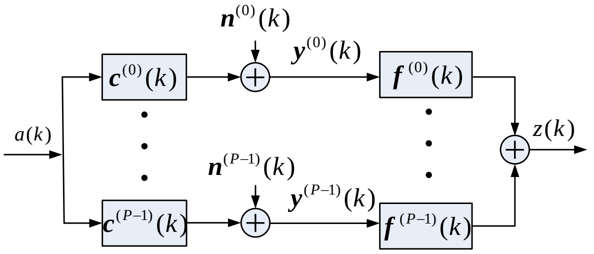


Figure 1. The structure of fractionally spaced blind equalizer.

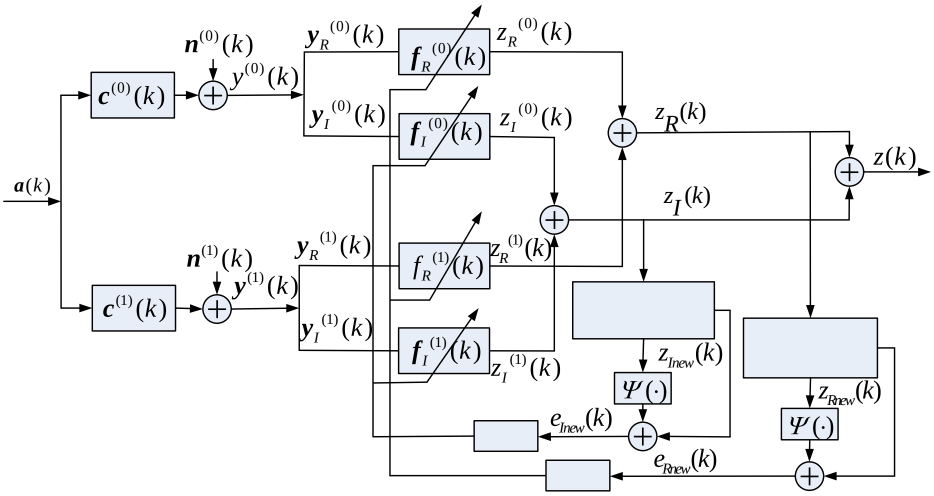


Figure 2. The structure of a modified T/2 fractionally spaced coordinate transformation blind equalizer.

of the complex signals. The second, two error functions are defined after making coordinate transformation to the real and imaginary parts of output signal $z(k)$ and the cost functions based on these two error functions are obtained. Accordingly, the updating formula of weight vector of modified equalizer is given by making the cost function minimization. The structure diagram of the modified equalizer is shown in Figure 2. We call the modified equalizer as T/2-FSE-MCTCMA(Modified T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA. In the proposed T/2-FSE-MCTCMA, according to Figure 2, the channel is divided into odd sub-channel $\boldsymbol{c}^{(0)}(k)$ and even sub-channel $\boldsymbol{c}^{(1)}(k)$, and $\boldsymbol{y}^{(0)}(k)$ and $\boldsymbol{y}^{(1)}(k)$ are input signals of each equalizer in the T/2-FSE-MCTCMA and written as

$$\boldsymbol{y}^{(0)}(k) = \boldsymbol{y}_R{}^{(0)}(k) + j\boldsymbol{y}_I{}^{(0)}(k) \tag{4}$$

$$\boldsymbol{y}^{(1)}(k) = \boldsymbol{y}_R{}^{(1)}(k) + j\boldsymbol{y}_I{}^{(1)}(k) \tag{5}$$

For sub-channel $\boldsymbol{c}^{(0)}(k)$, its has real and imaginary equalizer. The weight vectors of the real and imaginary equalizer are expressed as $\boldsymbol{f}_R{}^{(0)}(k)$ and $\boldsymbol{f}_I{}^{(0)}(k)$, respectively, and $z_R{}^{(0)}(k)$, $z_I{}^{(0)}(k)$ are output signals of the real and imaginary equalizer. As for sub-channel $\boldsymbol{c}^{(1)}(k)$, the weight vectors of real and imaginary equalizer are expressed as $\boldsymbol{f}_R{}^{(1)}(k)$ and $\boldsymbol{f}_I{}^{(1)}(k)$, the output signals of the real and imaginary equalizer are expressed $z_R{}^{(1)}(k)$ and $z_I{}^{(1)}(k)$, respectively. The real part of the final output is written as

$$z_R(k) = z_R{}^{(0)}(k) + z_R{}^{(1)}(k) \tag{6}$$

The imaginary part of the final output is written as

$$z_I(k) = z_I{}^{(0)}(k) + z_I{}^{(1)}(k) \tag{7}$$

The final output signal of the equalizer is written as

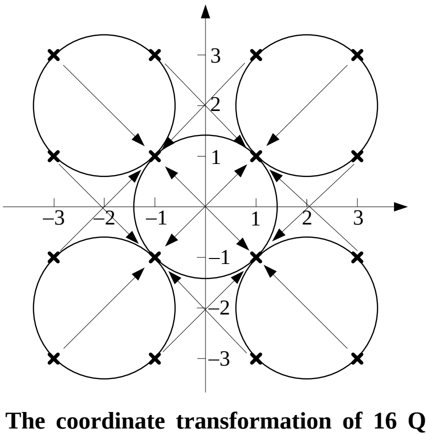


Figure 3. The coordinate transformation of 16 QAM signals.

$$z(k) = z_R(k) + jz_I(k) \tag{8}$$

In the process of equalization, coordinate transformation method is introduced into blind equalization algorithm, the principle diagram of the coordinate transformation is shown in Figure 3.

In Figure 3, "×" denotes ideal 16 QAM signal points after equalization, these points distribute in four known circles. We can make ideal 16 QAM signal points become four points A, B, C, and D via making coordinate transformation to output signals of the real and imaginary equalizer. A, B, C, and D must distribute in a circle.

When the CMA is used to equalize 16QAM signal, the error function is written as $e(k) = R_2 - |z(k)|^2$ (R_2 is a specific module value). Even if the channels are equalized completely, the error $e(k)$ is not zero. This affects equalization results. So, coordinate transformation method is used to make 16QAM signal points in different circles turn into A, B, C, D four points in the same circle. In other words, after multi-modulus 16QAM signals is become constant modulus 4QAM signals, the error is zero under the condition that the channels are equalized completely. The performance of the algorithm based on coordinate transformation method (T/2-FSE-MCTCMA) is optimal.

In Figure 2, $e_{Rnew}(k)$ and $e_{Inew}(k)$ are the error function of the real and imaginary part after coordinate transformation, respectively, and defined as

$$e_{Rnew}(k) = R_{Rnew}^2 - |z_{Rnew}|^2 \tag{9}$$

$$e_{Inew}(k) = R_{Inew}^2 - |z_{Inew}|^2 \tag{10}$$

where

$$z_{Rnew}(k) = z_R(k) - 2\text{sgn}[z_R(k)] \tag{11}$$

$$z_{Inew}(k) = z_I(k) - 2\text{sgn}[z_I(k)] \tag{12}$$

$$R_{Rnew}^2 = \frac{E\{|[a_R(k) - 2\text{sgn}[a_R(k)]]}{E\{|[a_R(k) - 2\text{sgn}[a_R(k)]]} \tag{13}$$

$$R_{Inew}^2 = \frac{E\{|[a_I(k) - 2\text{sgn}[a_I(k)]]}{E\{|[a_I(k) - 2\text{sgn}[a_I(k)]]} \tag{14}$$

The updating formula of weight vector of the real and imaginary equalizer are written as

$$\boldsymbol{f}_R{}^{(i)}(k+1) = \boldsymbol{f}_R{}^{(i)}(k) + \mu z_R{}^{(i)}(k) e_{Rnew}(k) \boldsymbol{y}_R{}^{(i)*}(k)\ (i = 0,1) \tag{15}$$

$$\boldsymbol{f}_I{}^{(i)}(k+1) = \boldsymbol{f}_I{}^{(i)}(k) + \mu z_I{}^{(i)}(k) e_{Inew}(k) \boldsymbol{y}_I{}^{(i)*}(k)\ (i = 0,1) \tag{16}$$

The final output signal of equalizer is written as

$$\begin{aligned} z(k) &= z_R(k) + jz_I(k) \\ &= \sum_{i=0}^{P-1} f_R{}^{(i)}(k) \cdot y_R{}^{(i)}(k) + j\sum_{i=0}^{P-1} f_I{}^{(i)}(k) \cdot y_P{}^{(i)}(k) \end{aligned} \tag{17}$$

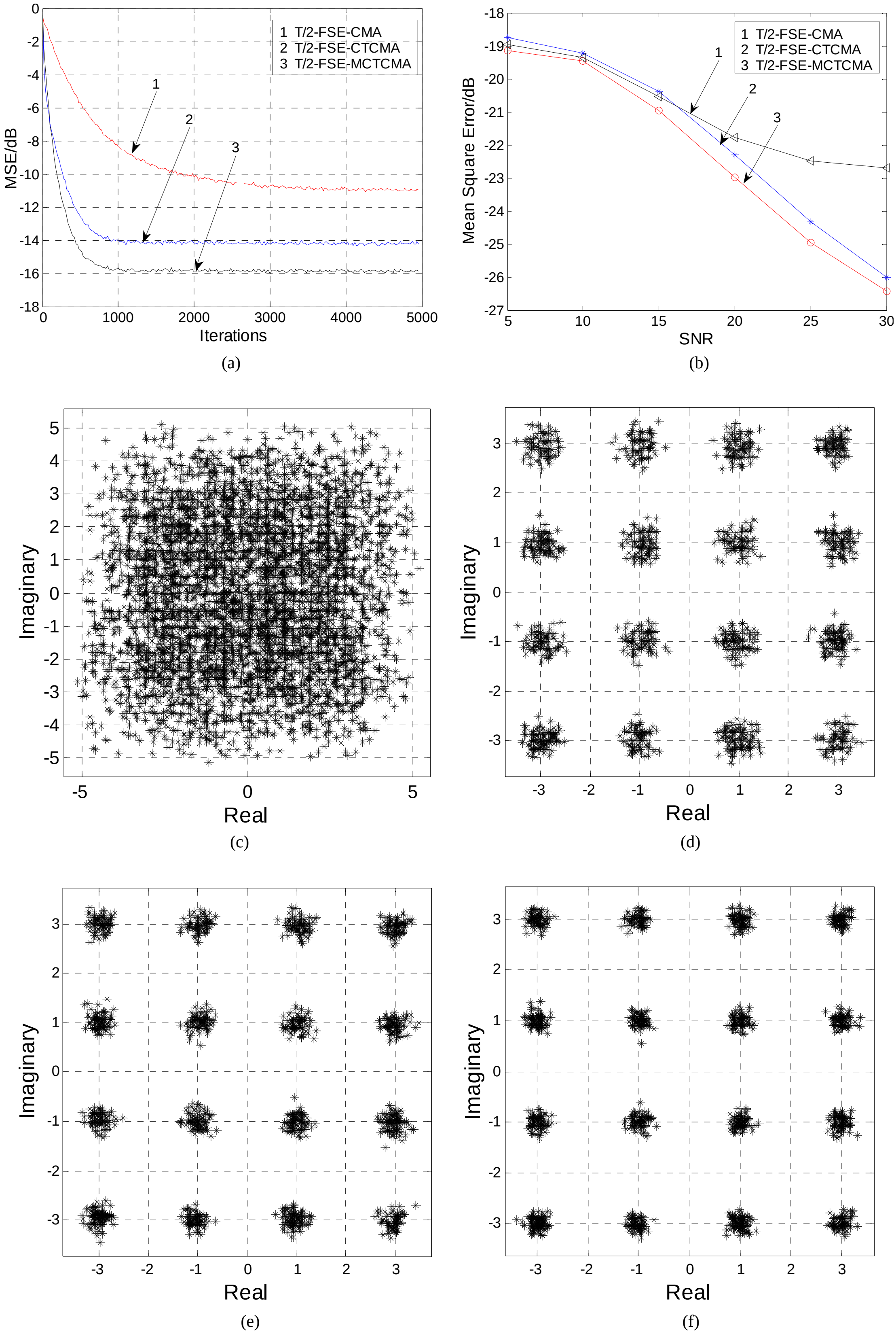


Figure 4. Simulation results. (a) Error curve; (b) Error curve of mean square root; (c) Input signals of equalizer; (d) Output signals of T/2-FSE-CMA; (e) Output signals of T/2-FSE-CTCMA; (f) Output signals of T/2-FSE-MCTCMA.

To 16 QAM signals, when the channel is equalized completely, Formula (9) is equal to zero. Until this, the T/2-FSE-MCTCMA is established. In this paper, we call the algorithm that the real and imaginary parts of input signal are not equalized, respectively, and only output signal is transformed as T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA (T/2-FSE-CTCMA).

4. Performance Analysis

4.1. Convergence Performance Analysis

Input signal is sampled by the rate of $T/2$ in T/2 fractionally spaced equalizer. It avoids spectrum aliasing by sub-sampling and compensates distortion of channel [19,20]. The real and imaginary parts of input signals in T/2 fractionally spaced equalizer are equalized, respectively, so it is equivalent to process real signal in the whole equalization process, and its computational complexity is decreased obviously. After the coordinate transformations of the real and imaginary parts of output signal are carried out, the mutli-modulus 16QAM signal is become constant modulus 4QAM signal. This treatment accelerates the updating speed of weight vector and when the channel is perfectly equalized, the error function tends to zero. So, the residual mean square error is decreased and the convergence rate is improved at end equalization.

4.2. Analysis of Computational Complexity

In the T/2-FSE-CMA, each weight vector iteration needs $4(N_f/2)$ multiplications and $3(N_f/2)+[(N_f/2)-1]$ additions (N_f is the length of equalizer). However, in the T/2-FSE-MCTCMA, the computation load of real part of each weight vector iteration is $N_f/2$ multiplications and $(N_f/2)-1$ additions. So, the total computation load of each weight vector iteration is N_f multiplications and N_f-2 additions. Based on above analysis, the computation load of the T/2-FSE-MCTCMA has a drop of about a half comparison with that of the T/2-FSE-CMA.

5. Simulation Results

In order to test the validity of the T/2-FSE-MCTCMA, we carried out simulation tests and compared the T/2-FSE-MCTCMA with the T/2-FSE-CTCMA and the T/2-FSE-CMA.

Simulation Test 1: 16QAM signals were transmitted to mixed-phase water acoustic channel, the impulse response vector of this channel was given by $\boldsymbol{c}$=[0.3132 -0.1040 0.8908 0.3134] [21]; The SNR was set to 25 dB; the weight length of equalizer was set to 32; the weight length of each sub-channel equalizer was set to 16; the center tap of the weight vectors of all equalizer were initialized to one; the step sizes $\mu_{\text{T/2-FSE}}$, $\mu_{\text{T/2-FSE-CTCMA}}$, $\mu_{\text{T/2-FSE-MCTCMA}}$ were set to 0.000006, 0.00003, 0.0009, respectively. Simulation results of 5000 Monte-Carlo times were shown in Figure 4.

Figure 4(a) shows that the MSE of the T/2-FSE-MCTCMA has a drop of about 2dB or 5dB comparison with that of the T/2-FSE-CTCMA or the T/2-FSE-CMA, respectively; the convergence rate of the T/2-FSE-MCTCMA is the fastest in all algorithms and performs an improvement of about 2000 steps comparison with the T/2-FSE-CMA. Root mean square error of the T/2-FSE-MCTCMA is minimum under the condition of the different SNR (see Figure 4(b)). The constellations of output signals in the T/2-FSE-MCTCMA is the clearest (see Figure 4(d), (e) and (f)). So, the T/2-FSE MCTCMA has great ability to suppress intersymbol interference.

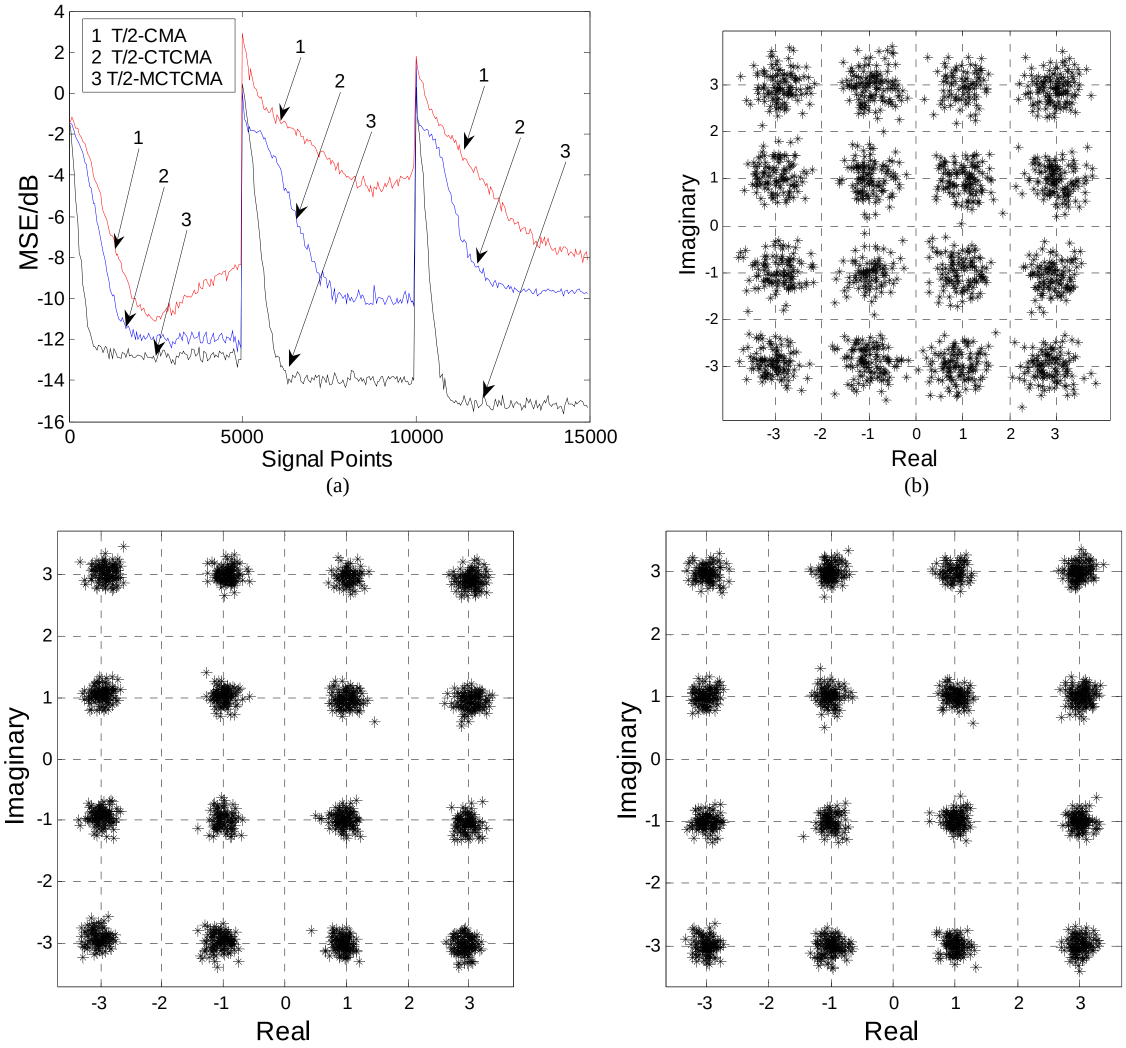
Simulation Test 2: transfer function of the channel $\boldsymbol{c}_1$ was given by $\boldsymbol{c}_1$=[0.9656 -0.0906 0.0578 0.2368] [21]. After 5000 signal points were transmitted, the channel $\boldsymbol{c}_1$ was changed into the channel $\boldsymbol{c}_2$, its transfer function was given by $\boldsymbol{c}_2$=[-0.35 0 0 1] [21]. After 10000 signal points were transmitted, the channel $\boldsymbol{c}_2$ was changed into the channel $\boldsymbol{c}_3$, its transfer function was given by $\boldsymbol{c}_3$=[0.3132 -0.1040 0.8908 0.3134] [21]. This established channel was called as time-varying channel.

The transmitted signals were 16QAM signal; the SNR was set to 25 dB; the weight length of equalizer was set to 32; the weight length of each sub-channel equalizer was set to 16; the center tap of the weight vectors of all equalizer were initialized to one. In the time-varying channel, the step sizes of the three algorithms were shown in Table 1. Simulation results of 500 times Monte-Carlo were shown in Figure 5.

Figure 5(a) illustrates that the T/2-FSE-MCTCMA outperforms the T/2-FSE-CTCMA and T/2-FSE-CMA in equalizing the time-varying channel and has strong restarted ability and can rapidly track time-varying channel

Table 1. The step size of three algorithms.

Channel	$\mu_{\text{T/2-FSE-CMA}}$	$\mu_{\text{T/2-FSE-CTCMA}}$	$\mu_{\text{T/2-FSE-MCTCMA}}$
Channel 1	0.0005	0.0001	0.000009
Channel 2	0.0006	0.0002	0.000006
Channel 3	0.001	0.0002	0.000009



(a)

(b)

(c)

(d)

Figure 5. Simulation Results. (a) Error curve; (b) Output signals of T/2-FSE-CMA; (c) Output signals of T/2-FSE-CTCMA; (d) Output signals of T/2-FSE-MCTCMA.

and that the output constellations of the T/2-FSE-MCTCMA are also the clearest.

6. Conclusions

In this paper, a Modified T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA(T/2-FSE-MCTCMA) is proposed, its computation load has a drop of about a half comparison with that of the T/2-FSE-CMA(T/2 Fractionally Spaced blind Equalization algorithm based on CMA). For 16QAM signals, the equalization performance of the T/2-FSE-MCTCMA is optimal. Simulation results with the different underwater acoustic channels indicate that the T/2-FSE-MCTCMA has faster convergence speed, the lower MSE, and the clearest constellations comparison with the T/2-FSE-CTCMA(T/2 Fractionally Spaced blind Equalization algorithm based on Coordinate Transformation and CMA) and the T/2-FSE-CMA. So, the T/2-FSE-MCTCMA can effectively eliminate Intersymbol Interference(ISI) and recovery signals real-timely.

7. Acknowledgment

This paper is supported by Specialized Fund for the Author of National Excellent Doctoral Dissertation of China (200753), Natural Science Foundation of Higher Education Institution of Jiangsu Province (08KJB510010) and “the peak of six major talent” cultivate projects of Jiangsu Province(2008026), Natural Science Foundation of Jiangsu Province(BK2009410).

8. References

[1] A. Naveed, I. M. Qureshi and A Hussain, *et al*. “Blind equalization of communication channels for equal energy

sources: Energy matching approach," Electronics Letters, Vol. 42, No. 4, pp. 247–248, 2006.

[2] Z. J. Liu, H. S. Xu, J. L. Wang, and K. C. Yi, "A novel hybrid blind equalization algorithm," Journal of Electronic and Information Technology, Vol. 81, No. 7, pp. 1606–1609, 2009.

[3] O. Dabeer and E. Masry, "Convergence analysis of the constant modulus algorithm," IEEE Transactions on Information Theory, Vol. 49, No.6, pp. 1447–1464, 2003.

[4] G. C. Li, C. B. Luo, X. G. Yang and Y. H. Gong, "New convex combination strategy for the MMSE blind equalization algorithms," Journal of Electronic Measurement and Instrument, Vol. 23, No. 1, pp. 37–41, 2009.

[5] Y. Wang and W. Guo, "Blind equalization of constant modulus based on Support Vector Regression," Journal of Electronic Measurement and Instrument, Vol. 22, No. 3, pp. 15–19, 2008.

[6] J. Liu and L. Y. Dai, "New Blind Equalization Algorithm Based on Multi-mode Error Switch," Journal of Data Acquisition and Processing, Vol. 19, No. 2, pp. 167–170, 2004.

[7] X. L. Li and X. D. Zhang, "A family of generalized constant modulus algorithms for blind equalization," IEEE Trans on communications, Vol. 54, No. 11, pp. 1913–1917, 2006.

[8] Y. P. Zhang and J. W. Zhao. "Blind equalization algorithms based on fractionally spaced underwater acoustic channels," Acoustics and Electronics Engineering, Vol. 78, No. 2, pp. 21–23, 2005.

[9] L. Zhou, J. D. Li and G. H. Zhang. "Novel DWPW system based on fractionally spaced equalizers and the maximum likelihood algorithm," Journal of Xidian University (Natural Science), Vol. 33, No. 4, pp. 509–513, 2006.

[10] Y. C. Guo and R. G. Lin, "Blind equalization algorithm based on T/4 fractionally spaced decision feedback equalizer," Journal of Data Acquisition and Processing, Vol. 23, No. 3, pp. 284–287, 2008.

[11] K. S. Chen and C. Y. Chu, "A propagation study of the 28GHz LMDs system performance with M-QAM modulations under rain fading," Progress In Electromagnetics Research, No. PIER 68, pp. 35–51, 2007.

[12] Y. P. Zhang, Y. C. Guo and J. Z. Liu, "Blind equalization algorithm suitable for 16QAM signals for carrier recovery of underwater acoustic channel," Journal of System Simulation, Vol. 20, No.1, pp. 156–158, 2008.

[13] Y. Q. Zhang, P. Li and Z. R. Zhang, "Dual-mode blind equalization algorithm for multi-lever QAM modulation based on Sign-CMA," Journal of China Institute of Communication, Vol. 25, No. 5, pp. 155–159, 2004.

[14] G. Q. Dou, J. Gao and P Wang, "A concurrent constant modulus algorithm and soft decision-directed algorithm for blind equalization," Signal Processing, Vol. 23, No. 6, pp. 833–835, 2007.

[15] W Rao and Y. C. Guo, "A new constant modulus algorithm based on dual-step size," International Symposium on Test Automation and Instrument, 2006.

[16] W. Rao, K. M. Yuan, Y. C. Guo and C. Yang, "A simple constant modulus algorithm for blind equalizer suitable for 16QAM signal," International Conference on Signal Processing Proceedings, pp. 1963–1966, 2008.

[17] Y. C. Wang, Z. L. Chen and Z. T. Liu, "Direct blind fractional spaced equalization algorithm based on channel output decorrelation," Journal of Data Acquisition and Processing, Vol. 20, No. 3, pp. 323–327, 2005.

[18] L. Zhou, J. d. Li and G. H. Zhang, "Novel DWPW system based on fractional spaced equalizers and the maximum likelihood algorithm". Journal of XiDian University, Vol. 33, No. 4, pp. 509–513, 2006.

[19] Y. C. Guo, "Adaptive blind equalization techniques," Hefei Industrial University Press, 2007.

[20] B. J. Kim and D. C. Cox, "Blind equalization for short burst wireless communications," IEEE Trans. On Vehicular Technology, Vol. 49, No. 4, pp. 1235–1247, 2000.

[21] F. Wang, "Higher-order statistics based on the acoustic channel blind equalization theory and algorithm," PhD thesis, Northwestern Industrial University, 2003.

Int. J. Communications, Network and System Sciences, **2010, 3, 190-195**
doi:10.4236/ijcns.2010.32027 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

CPW Fed Double T-Shaped Array Antenna with Suppressed Mutual Coupling

A. DANIDEH[1], A. A. Lotfi NEYESTANAK[2*]
[1]*Department of Electrical Engineering, Islamic Azad University (IAU),*
Science and Research Branch, Tehran, Iran
[2]*Department of Electrical Engineering, Islamic Azad University,*
Shahr_e_Rey Branch, Tehran, Iran
E-mail: *alexdanideh@gmail.com, alotfi@iust.ac.ir*

Received November 6, 2009; *revised December* 11, 2009; *accepted January* 16, 2010

Abstract

A compact CPW-fed double T-Shaped antenna is proposed for dual-band wireless local area network (WLAN) operations. For the proposed antenna, the -10 dB return loss bandwidth could reach about 25.5% for the 2.4 GHz band and 5.7 % for the 5 GHz band, which meet the required bandwidth specification of WLAN standard. To reduce the mutual coupling and get high isolation between two dual-band antennas, we proposed the novel electromagnetic band gap (EBG) structures. When the EBG structure is employed, a -13dB and -30dB mutual coupling reduction is achieved at 2.4 and 5.2 GHz. It shows that the features of small size, uniplanar structure, good radiation characteristics and small mutual coupling are promising for multi-input multi-output (MIMO) applications.

Keywords: CPW, T-Shaped Antenna, WLAN, MIMO, Mutual Coupling

1. Introduction

Wireless communications, especially the wireless local area network (WLAN) communication, have evolved at astonishing rate during the last decade. Therefore, design of broad dual and multi-band antennas with low-profile, lightweight, flush mounted and single-feed to fit the limited equipment space of the WLAN device has gained increasing demands. Coplanar waveguide (CPW) transmission lines have been widely used as feeding networks with slot antennas. CPW lines have many useful design characteristics such as low radiation leakage, less dispersion, little dependence of the characteristic impedance on substrate height, and uniplanar configuration. They also allow easy mounting and integration with other microwave integrated circuits and RF frequency devices. Dual-band operations have become very important in wireless local area network (WLAN) applications. CPW-fed antennas for wireless communications have been discussed by many authors for dual-band operations.

The demand for high speed and high quality data transmission in communications has been rapidly increasing .This requires the effective utilization of the limited channel bandwidth. At the same time, multi-input multi-output (MIMO) systems have also received significant attention for their ability to increase the channel capacity.

A novel and simple printed dual-band with two different sizes T-shaped antenna which generates two separate resonant is proposed in [1].

Since WLAN antennas are usually required on broad bandwidth and small antenna size, researchers have studied monopole structures with dual resonant modes [2–4].

A new CPW-fed T-shaped monopole antenna with a trapeze form ground plane and two parasitic elements for WLAN/WiMAX dual mode operation is investigated in [5].

The CPW-fed G-shaped planar monopole antenna with dual band operation is a good choice for WLAN application [6]. For WLAN operations we can use the dual-band slot antenna with compact size [7]; this antenna can be easily integrated with other RF front-end circuits.

[8] explains a novel dual-band patch antenna on magnetic substrate for WLAN application. The dual-band operation is obtained by inserting a pair of L-shaped slots.

Using two-tapered patch with different slopes, a slot between them, modified feed and a slot in the ground plane, the impedance band width can be increased [9].

[*] Corresponding author

The mutual coupling or isolation between closely packed antenna elements is important in a number of applications. These include systems relying on array antennas and more recently multiple input multiple output (MIMO) wireless communication systems which rely on multiple antennas to offer increases in capacity without the need for additional power or spectrum, compared to conventional systems [10]. Achieving low mutual coupling between closely-packed antenna elements is difficult to achieve and has been well studied.

Printed antenna arrays suffer from relatively high level of mutual coupling between individual elements due to surface waves [11]. This becomes progressively worse with increasing frequency, dielectric constant, and substrate thickness.

While mutual coupling can be reduced by increasing the inter element spacing, this results in undesirable grating lobes. Mutual coupling can also be reduced by incorporating electromagnetic band gap (EBG) structures in between array elements. The electromagnetic-band gap (EBG) structures are periodical cells composed of metallic or dielectric elements. The major characteristic of EBG structures is to exhibit band gap feature in the suppression of surface-wave propagation. Electromagnetic band gap (EBG) structures show the characteristics of forbidding the propagation of electromagnetic waves in certain frequency range and this has been applied in antenna designs [11]. One of the applications is to reduce the mutual coupling between antennas [12], in which the EBG structure is put between two antenna elements and a much higher isolation is obtained.

In this paper, we propose a novel CPW-fed double T-Shaped antenna with simple structure and compact size. These antennas can be tuned to cover the 2.4/5.2 GHz WLAN bands. A multi feed 2-element planar antenna array with the same size is placed on a 78mm × 28.5mm substrate whose thickness is 0.5 mm and dielectric constant is 3.38.We also explore the effect of the proposed planar EBG structure in the mutual coupling reduction and suppression surface-wave between antenna arrays. The mutual coupling between adjacent elements with and without the EBG structure is simulated and compared.

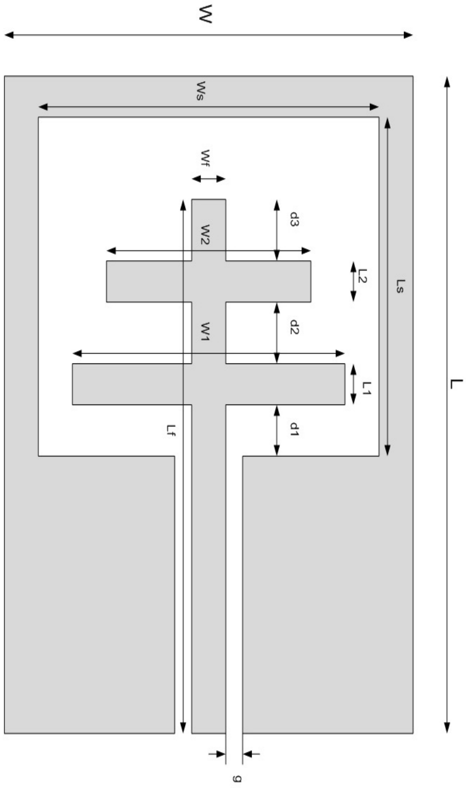


Figure 1. Geometry of proposed antenna.

Table 1. Double T-shaped antenna dimensions.

Parameters	Ws	Ls	W1	W2	L1	L2	d1	d2	d3
Value (mm)	23	18.5	19	15	3	5	5.2	1.9	1.5

2. Antenna Geometry

The geometry and parameters of the antenna with double T-shaped stub are shown in Figure 1. The antenna is etched on Rogers RO4003 substrate with a thickness of 20 mil (0.5 mm) and dielectric constant of 3.38. The size of the ground plane is W= 23.5mm × L=28.5 mm. The slot has a length Ls = 18.5 mm and width Ws = 23 mm.

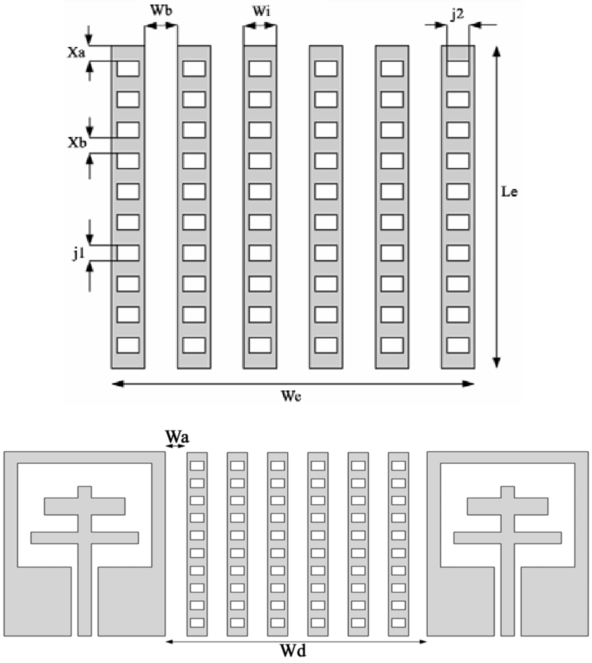
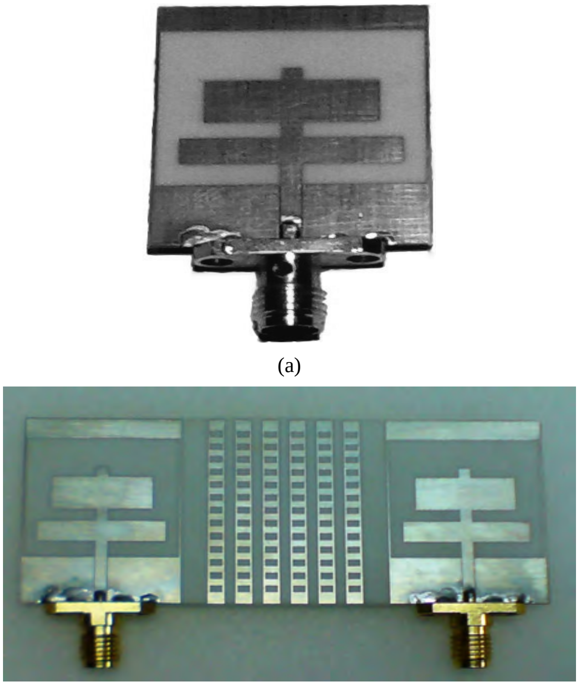


Figure 2. Geometry of EBG structure for coupling reduction.



(a)

(b)

Figure 3. Photograph of the antennas a) Single antenna b) Array antenna using EBG.

Table 2. EBG structure dimensions for Double T-shaped antenna.

Parameters	Xa	Xb	Wb	Wi	j1	j2
Value (mm)	1.75	1	1.6	2.5	1	1.8

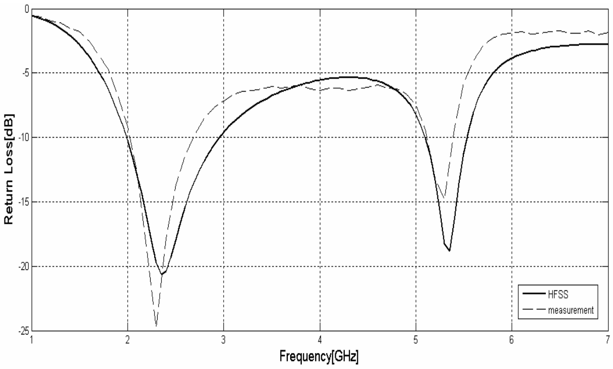


Figure 4. Simulated and measured return loss of the proposed antenna.

The antenna is excited by a 50Ω microstrip line with double T-shaped tuning stub. The width of the 50 Ω microstrip line is Wf =1.75 mm and the gap of the CPW line is g=0.1 mm. The double T-shaped tuning stub is located at the center of the slot, where the antenna is symmetrical along the center, x-axis. Detail dimensions and location of the T-shaped tuning stubs are d1= 2.5mm, d2=1.9mm, d3=1.5mm, L1=3mm, L2=5mm, W1=19mm and W2=15mm. A series of typical dimensions are presented in Table 1.

The proposed planar EBG structure with its geometrical parameters and the geometry of multi-feed 2-element planar antenna array is shown in Figure 2, the size of array is 78mm × 28.5mm, is placed on a substrate whose thickness is 20 mil (0.5 mm) and dielectric constant is 3.38, the distances between the adjacent elements are 31mm, is smaller than quarter wavelength.

The dimension of the EBG structure was 23 mm × 28.5 mm. The EBG structure is fabricated on a 0.5mm thick substrate with relative permittivity 3.38. In this case, 6 columns of metal are used, with an equivalent space 1.6mm, and each element of this metal consists of 13 rectangular slots (slits) etched on it. This structure will be low profile, light weighted and cheaper. The planar EBG structure was used in between a two-element planar antenna array to study its effect on antenna mutual coupling and surface wave. In fabrication, the distance between the antenna and the EBG surface is 4 mm. These dimensions are obtained after performing and optimization. The dimensions of EBG structure are presented in Table 2.

The photograph of the fabricated single and array antenna with EBG structure is shown in Figure 3.

3. Results

The antenna performance was investigated by simulation via HFSS software. Figure 4 shows the measured and simulated return losses of the proposed antenna.

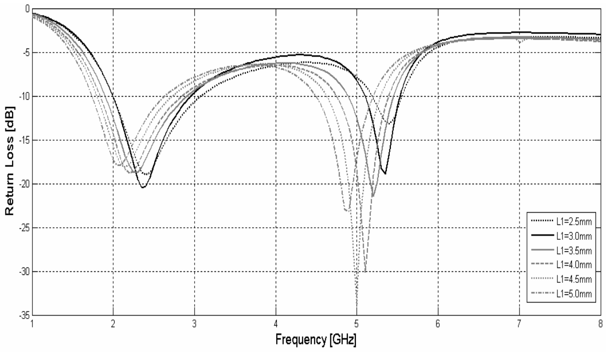


Figure 5. The effect on return loss due to the change of L2.

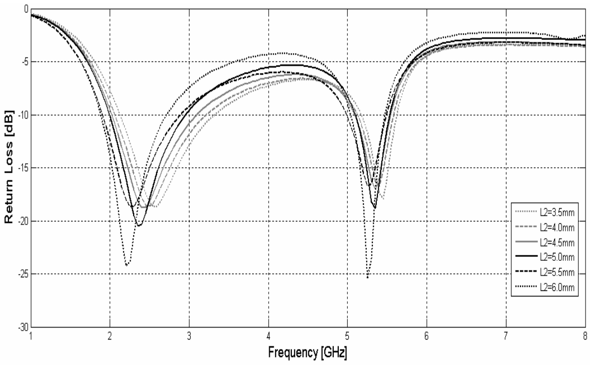


Figure 6. The effect on return loss due to the change of L2.

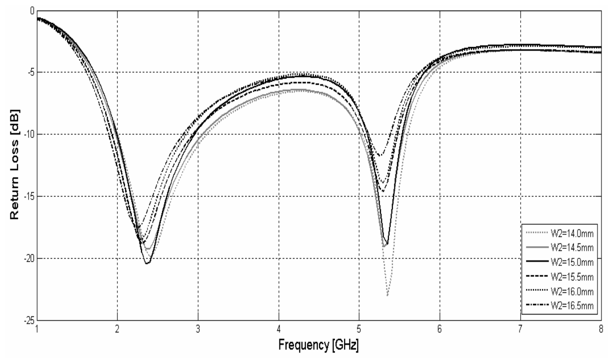
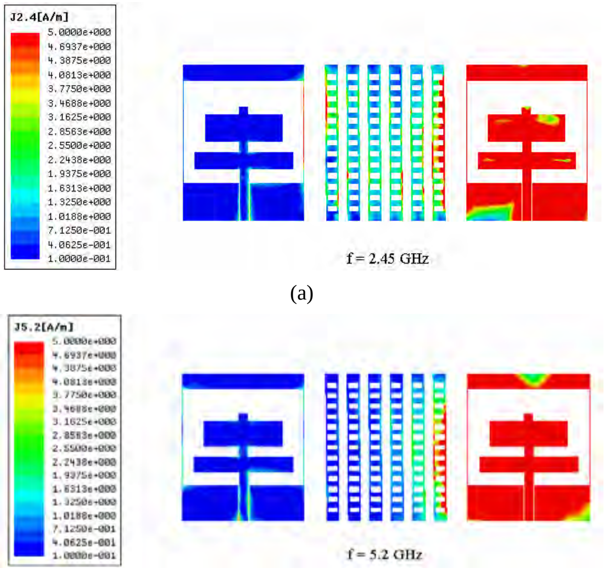


Figure 7. The effect on return loss due to the change of W2.



(a)

(b)

Figure 8. Electric current distribution over the substrate surface a) 2.45 GHz b) 5.2 GHz.

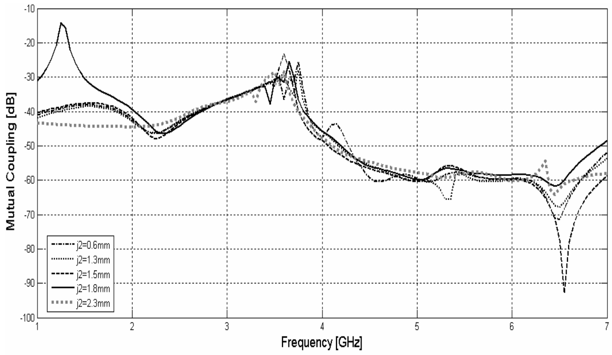


Figure 9. The effect on coupling due to the change of j2.

The obtained -10 dB return loss are 600 MHz (2.05–2.65GHz) and 500 MHz (5.1–5.4GHz), corresponding to an impedance bandwidth of 25.5% and 5.7 % with respect to the appropriate resonant frequencies. Obviously, the achieved bandwidths can cover the WLAN standards in the 2.4 GHz (2.4–2.484 GHz), 5.2 GHz (5.15–5.35 GHz) and 5.8 (5.725–5.825) GHz bands.

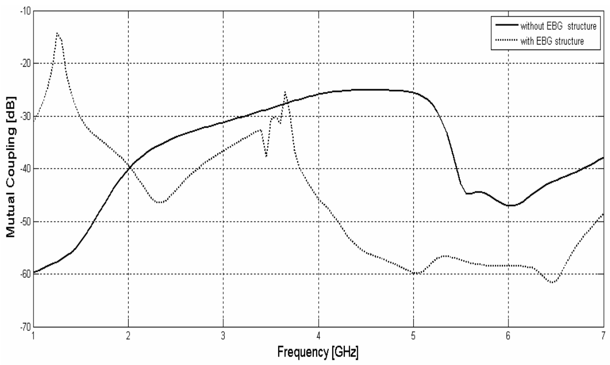


Figure 10. Mutual coupling with and without EBG structure.

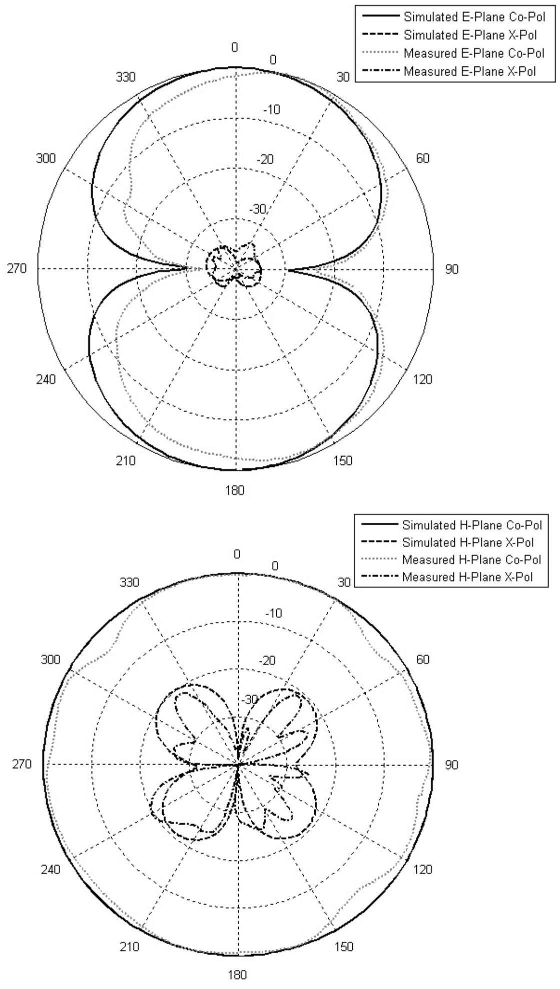


Figure 11. Radiation patterns of proposed antenna at 2.45 GHz.

In order to provide design criteria for this antenna, the effects of each geometrical parameter are analyzed. The antenna dimensions (L1, L2 and W2) and one parameter are changed at a time while the others are kept constant. Figure 5 to Figure 7 shows the effect of changing L1, L2 and W2, respectively.

In order to reduce coupling, an EBG structure is incorporated between two antennas. It is observed that the EBG structure demonstrates a band stop characteristic, which is enough to thoroughly eliminate the high order mode. The EBG structure not only successfully suppresses higher order modes, but also increases the impedance bandwidth. The planar EBG dimensions (j1 and j2) are chosen to be (1 and 1.5 mm), respectively, and one parameter is changed at a time while the others are kept constant.

Figure 8 shows the HFSS simulated current distributions of the 2*1 array antenna at 2.45, 5.2, GHz with EBG structures. EBG structures acts like a LC filter which rejects any undesired frequency range. This behavior is more studied in Figure 8. The electric current distribution over the substrate is plotted in presence of the EBG structures. Perusing currents distribution, it can be understood that with EBG structure, we decrease current distribution and reduce antenna's coupling.

Figure 9 shows the effect of changing j2. The rectangular slits etched in the metal are basically used to reduce mutual coupling between radiating elements, therefore the length of 1mm and width of 1.8 mm for slits in the metal is given best result of S12 with -46 dB and -58dB at 2.4 and 5.2 GHz.

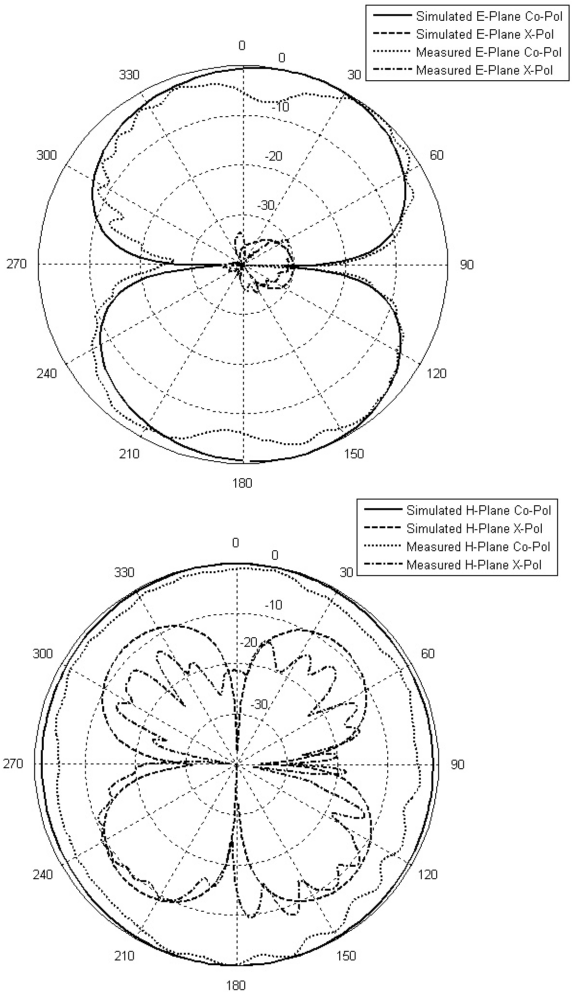


Figure 12: Radiation patterns of proposed antenna at 5.2 GHz.

The mutual coupling for the antenna arrays with and without the EBG structure is shown in Figure 10. It is evident that the proposed EBG structure can help achieve a significantly low level of mutual coupling. Measured mutual coupling (S21) with EBG structure are -46 dB and -58dB at 2.4 and 5.2 GHz. In addition, measured mutual coupling (S21) without EBG structure are -34 dB and -27dB at 2.4 and 5.2 GHz, (respectively.) This proves that the surface wave is suppressed.

The generation of surface waves decreases the antenna efficiency and degrades the antenna pattern. Since the EBG structure has already demonstrated its ability to suppress surface waves, the planar EBG structures are inserted between antenna arrays to reduce the mutual coupling.

The radiation characteristics of the proposed antenna have also been studied. Figure 11 to Figure 12 show the measured radiation patterns for the *E* and *H*-plane pattern including both Co- and Cross-polarization at 2.45 and 5.2 GHz for the proposed antenna.

4. Conclusions

A CPW-fed antenna with double T-shaped stub has been proposed for the 2.4/5.2 GHz dual-band WLAN operations. The antenna has characteristics of compact size, a simple structure, good omni directionality and the multi-feed 2-element planar antenna array formed with the compact CPW-fed antennas has satisfied input return loss bandwidth and good radiation patterns. In addition, a prototype of the EBG structure along with an array of two element antenna were fabricated and tested. The EBG structure is inserted between the antenna elements to reduce the mutual coupling and surface wave suppression. Therefore, it is an ideal option for MIMO applications in compact portable devices.

The measured impedance bandwidths are 25.5% at the lower frequency band of 2.4 GHz and 5.7% at the upper frequency band of 5 GHz for the desired bands.

5. References

[1] Y. L. Kuo and K. L. Wong, "Printed double-T monopole antenna for 2.4/5.2 GHz dual-band WLAN operations," IEEE Transactions on Antennas and Propagation, Vol. 51, No. 9, September 2003.

[2] L. Zhang, Y. C. Jiao, G. Zhao, Y. Song, and F. S. Zhang, "Broadband dual-band CPW-Fed closed rectangular ring monopole antenna with a vertical strip for WLAN operation," Microwave and Optical Technology Letters, Vol. 50, No. 7, pp. 1929–1931, 2008.

[3] C. M. Wu, "Dual-band CPW-fed cross-slot monopole antenna for WLAN operation," IET Microwave and Antennas Propagation. Vol. 1, No. 2, pp. 542–546, 2007.

[4] W. C. Liu, "Broadband dual-frequency CPW-fed antenna with a cross shaped feeding line for WLAN application," Microwave and Optical Technology Letters, Vol. 49, No. 7, pp. 1739–1744, 2007.

[5] Y. Y. Cui, Y. Q. Sun, H. C. Yang, and C. L. Ruan, "A new triple-band CPW-fed monopole antenna for WLAN and WiMAX applications," Progress in Electromagnetics Research M, Vol. 2, pp. 141–151, 2008.

[6] C. W. Liu, "Optimal design of dual band CPW-fed G-shaped monopole antenna for WLAN application," Progress in Electromagnetics Research, PIER 74, pp. 21–38, 2007.

[7] W. Ren, "Compact dual-band slot antenna for 2.4/5 GHz WLAN applications," Progress in Electromagnetics Research B, Vol. 8, pp. 319–327, 2008.

[8] E. Wang, J. Zheng, and Y. Liu, "A novel dual-band patch antenna for WLAN communication," Progress in Electromagnetics Research C, Vol. 6, pp. 93–102, 2009.

[9] R. Zaker, C. Ghobadi, and J. Nourinia, "A modified microstrip-fed two-step tapered monopole antenna for UWB and WLAN applications," Progress in Electromagnetics Research, PIER 77, pp. 137–148, 2007.

[10] R. D. Murch and K. B. Letaief, "Antenna systems for broadband wireless access," IEEE Communication Magzine, Vol. 40, No. 4, pp. 76–83, April, 2002.

[11] F. Yang and Y. Rahmat-Samii, "Microstrip antennas integrated with electromagnetic band-gap (EBG) structures: A low mutual coupling design for array applications," IEEE Transactions Microwave Theory Techniques, Vol. 51, pp. 2936–2946, October, 2003.

[12] H. K. Xin, M. K. Matsugatani, J. Hacker, J. A. Higgins, M. Rosker, and M. Tanaka, "Mutual coupling reduction of low profile monopole antennas on high-impedance ground plane," Electronics Letters, Vol. 38, No. 16, pp. 849–850, August, 2002.

Int. J. Communications, Network and System Sciences, **2010, 3, 196-203**
doi:10.4236/ijcns.2010.32028 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Average Consensus in Networks of Multi-Agent with Multiple Time-Varying Delays

Tiecheng ZHANG, Hui YU
Institute of Nonlinear and Complex Systems, China Three Gorges University, Yichang, China
Email: *ztchongctgu@yahoo.cn*, *yuhui@ctgu.edu.cn*

Received November 19, 2009; *revised December* 15, 2009; *accepted January* 21, 2010

Abstract

The average consensus in undirected networks of multi-agent with both fixed and switching topology coupling multiple time-varying delays is studied. By using orthogonal transformation techniques, the original system can be turned into a reduced dimensional system and then LMI-based method can be applied conveniently. Convergence analysis is conducted by constructing Lyapunov-Krasovskii function. Sufficient conditions on average consensus problem with multiple time-varying delays in undirected networks are obtained via linear matrix inequality (LMI) techniques. In particular, the maximal admissible upper bound of time-varying delays can be easily obtained by solving several simple and feasible LMIs. Finally, simulation examples are given to demonstrate the effectiveness of the theoretical results.

Keywords: Average Consensus, Multi-Agent System, Multiple Time-Varying Delays, Linear Matrix Inequality

1. Introduction

Recently, more and more researchers have paid a great deal of attention on distributed coordinated control of networks of dynamic agents within the control community. Especially the consensus problem was discussed widely, which can be attributed to the broad applications of multi-agent systems in many areas, including cooperative control of unmanned air vehicles, formation control of multi-robot, flocking, swarming, distribution sensor fusion, attitude alignment and congestion control in communication networks, etc. In cooperative control of multi-agent system, a critical issue is to design appropriate protocol and algorithms such that all agents can reach a common consensus value. This problem is called consensus problem.

In the past decades, some theoretical results have been established in [1–11], to name a few. In [1], Vicsek *et al.* proposed a simple model but interesting discrete-time model of autonomous agents all moving in the plane with the same speed but with different headings. Simulation results provided in [1] show that all agents can eventually move in the same direction without centralized coordination. The first paper providing a theoretical explanation for these observed behaviors in Vicsek model is [2]. The theoretical results in [2] are extended to the case of directed graph by Ren *et al.* [3], in which matrix analysis and algebraic graph theory were used. Moreau [4] used a set-valued Lyapunov approach to study consensus problem with unidirectional time-dependent communication links. Saber *et al.* [5] discussed average consensus problem. When network communication is affected by time delay, the consensus problem is investigated in [5–7]. In [8], Saber provided a theoretical framework for analysis of consensus algorithms for multi-agent networked systems with an emphasis on the role of directed information flow, robustness to changes in network topology due to link/node failures, time-delays, and performance guarantees. Ren *et al.* provided a tutorial overview of information consensus in multivehicle cooperative control in [9]. Yu *et al.* [10] proposed weighted average consensus in direction networks and unidirectional networks with time-delay. Multi-vehicle consensus with time-varying reference state was discussed in [11]. Some other issues on consensus problem can be found in [12–16].

Currently, consensus problem for multi-agent networks with time delay was studied using linear matrix inequality method, for example [17–19] and [20]. In time delay systems of multi-agent, the network topology of multi-agent is a key factor in the analysis of stability of multi-agent system. Average consensus problem for multi-agent networks with both constant

Identify applicable sponsor/s here. *(sponsors)*

and time-varying delay has been extensively considered for the cases of multi-agent undirected and/or direction networks with fixed and/or switching topology [17,18] and [19]. However, the studies on consensus problem of multi-agent networks with multiple time-varying delays are still sparse. In this case, theoretical analysis is more challenging. In [20], the authors proposed a consensus protocol for multi-agent undirected network with multiple time-varying delays based on LMI method.

In this paper, we study the average consensus problem for continuous time undirected networks of multi-agent, coupling multiple time-varying delays. Because the closed-loop system matrix is singular and traditional LMI-based control theory is invalid. Therefore, theoretical analysis for this case is a challenging task. By using orthogonal transformation techniques, sufficient conditions based on LMI for multi-agent achieving average consensus is obtained. Compared to [20], a different approach is used in this paper. First of all, the original system is turned into a reduced dimensional system by orthogonal transformation; Secondly, via constructing a different Lyapunov-Krasovskii function, a sufficient condition expressed as LMI is proposed to guarantee all agents reach average consensus in fixed and switching network. Finally, the maximal admissible upper bound of multiple time-varying delays can be easily obtained by solving several simple and feasible LMIs.

This paper is organized as follows. Section 2 is the notation and formally states the problem. Section 3 contains our main results. Simulation results are presented in Section 4. The concluding remarks are made in Section 5.

2. Problem Statement and Preliminaries

In this section, we provide a brief introduction about algebraic graph theory [24] and state the problem.

2.1. Algebraic Graph Theory

Let $G(V,E,A)$ be a weighted undirected graph of order $n(n \ge 2)$, where $V=\{v_1,v_2,v_3\cdots,v_n\}$ is the set of nodes, $E \subseteq V \times V$ is the set of edges, and $A=\left[a_{ij}\right]\in\Re^{n\times n}$ is a weighted adjacency matrix. The node indexes belong to a finite index set $I=\{1,2,\cdots,n\}$.

In undirected graph, $e_{ij}=e_{ji}$. $e_{ij}\in E$ if and only if $a_{ij}>0$. Moreover, we assume $a_{ij}=0$ for all $i\in I$. A undirected graph is always connected. The set of neighbors of nodes v_i is denoted by

$$N_i=\left\{v_j\in V:\left(v_i,v_j\right)\in E\right\}.$$

The out-degree of node v_i is defined as follows: $\deg_{out}(v_i)=\sum_{i=1}^{n}a_{ij}$. The degree matrix of graph G is a diagonal matrix $D=\left[d_{ij}\right]$, where $d_{ij}=0$ for all $i\neq j$ and $d_{ii}=\deg_{out}(v_i)$. The Laplacian matrix associated with the graph is defined as

$$L=\left[l_{ij}\right]=D-A=\begin{cases}\sum_{j=1,j\neq i}^{n}a_{ij}, & j=i,\\ -a_{ij}, & j\neq i.\end{cases}$$

An important fact of L is that all the row sums of L are zero and thus $1_n=[1,1,\cdots,1]^T\in\Re^n$ is an eigenvector of L associated with the eigenvalue $\lambda=0$.

Lemma 1 [5]. If the undirected graph G is connected, then its Laplacian matrix L satisfies:

1) L is symmetric and $\operatorname{rank}(L)=n-1$;

2) zero is one eigenvalue of L, and 1_n^T and 1_n are the corresponding left and right eigenvector respectively, then, $1_n^T L=0$ and $L1_n=0$;

3) the rest $n-1$ eigenvalues are all positive and real.

2.2. Consensus Problem on Network

Consider a group of n agents with dynamics given by

$$\dot{x}_i(t)=u_i(t),\quad i\in I \tag{1}$$

where $x_i\in\Re^n$ is the state of the ith agent at time t, which might represent physical quantities such as attitude, position, temperature, voltage, and so on, and $u_i(t)\in\Re$ is the control input (or protocol) at time t.

We say protocol u_i asymptotically solves the consensus problem, if and only if the states of agents satisfy $\lim_{t\to\infty}\left\|x_i(t)-x_j(t)\right\|=0$, for all $i,j\in I$. Furthermore, if

$$\lim_{t\to\infty}x_i(t)=\frac{1}{n}\sum_{i=1}^{n}x_i(0)=Ave(x(0)),$$

We say protocol u_i asymptotically solves the average consensus problem.

2.3. Control Protocol for Time Delay

Let τ_{ij} notes the time delay for information communicated from agent j to agent i. Because of time delay, two difficult consensus protocols have been studied. One is

$$u_i=\sum_{v_j\in N_i}a_{ij}[x_j(t-\tau_{ij}(t))-x_i(t-\tau_{ij}(t))],\quad i\in I \tag{2}$$

and the other is

$$u_i = \sum_{v_j \in N_i} a_{ij}[x_j(t-\tau_{ij}(t)) - x_i(t)],\ \ i \in I. \qquad (3)$$

Literature [5] have taken the protocol in (2) with $\tau_{ij} = \tau$ into account, and obtained that it can asymptotically solves the average consensus problem in the fixed and undirected network topology if and only if $\tau \in \left[0, \pi/2\lambda_{\max}(L)\right]$. Some conclusions also were investigated in [16-18].

For consensus protocol in (3), communication delay only affects the agents who are transmitted. Literature [4] has established the consensus results in the directed and dynamically switching graph. Some results also appeared in [21–23].

In this paper, we are interested in discussing the average consensus problem in network of dynamic agents with fixed and switching topology coupling multiple time-varying delays, where the information passes through different edge with different time-varying delays. To solve such a problem, we assume the time delay satisfies $\tau_{ij} = \tau_{ji}$ in undirected graph, i.e., the delays in transmission from x_i to x_j and x_j to x_i coincide. So we use the following protocol:

$$u_i = \sum_{v_j \in N_i} a_{ij}[x_j(t-\tau_k(t)) - x_i(t-\tau_k(t))],\ \ i \in I \qquad (4)$$

With (4), (1) can be written in matrix form:

$$\dot{x}_i(t) = -\sum_{k=1}^{N} L_k x(t-\tau_k(t)), \qquad (5)$$

where $x(t) = \left(x_1(t), x_2(t), \cdots, x_n(t)\right)^T$, $N \le n(n-1)/2$, and $L_k = \left[l_{kij}\right]$ is the matrix defined by

$$l_{kij} = \begin{cases} -a_{ij} & i \ne j,\ \tau_k(t) = \tau_{ij}(t), \\ 0 & i \ne j,\ \tau_k(t) \ne \tau_{ij}(t), \\ \sum_{j=1}^{N} a_{ij} & i = j. \end{cases}$$

Because of A is symmetric and $\tau_{ij} = \tau_{ji}$ in the undirected network, we can obtain L_k is a symmetric and $\sum_{k=1}^{N} L_k = L \cdot$

The time-varying delay $\tau_k(t)$, $k = 1, 2, \cdots, N$ is assumed to satisfy the follow inequations:

$$0 \le \tau_k(t) < d_k,\ \dot{\tau}_k(t) \le h_k < h < 1,\ k = 1, 2, \cdots, N \qquad (6)$$

where d_k and h_k are constants. d and h are the upper bound of d_k and h_k, namely, $d = \max_{1\le k\le N}\{d_k\}$, $h = \max_{1\le k\le N}\{h_k\}$.

In the undirected network with switching topology, we address the follow hybrid system:

$$\dot{x}(t) = -\sum_{k=1}^{N} L_{ks} x(t-\tau_k(t)),\ \ s = \sigma(t) \in I_\Gamma. \qquad (7)$$

where the map $\sigma(t): [0, +\infty) \to I_\Gamma = \{1, 2, \cdots, M\}$ is a switching signal that determines the network topology, and $M \in Z^+$ denotes the total number of all possible switching undirected graphs. $\sum_{k=1}^{N} L_{ks} = L_s(G_s)$ is the Laplacian matrix of the graph $G_s(V_s, E_s, A_s)$ that belongs to a set $\Gamma = \{G_k : k \in I_\Gamma\}$, which is obviously finite.

Lemma 2 (Schur complement [25]). Let S_{11}, S_{12}, S_{22} be given symmetric matrices such that $S_{22} > 0$, then $\begin{bmatrix} S_{11} & S_{12} \\ S_{12}{}^T & -S_{22} \end{bmatrix} < 0 \ \Leftrightarrow S_{11} + S_{12}S_{22}{}^{-1}S_{12}{}^T < 0$.

Lemma 3. For any Laplacian matrix $L \in \Re^{n\times n}$ of undirected connected network, there exists an orthogonal matrix W such that

$$W^T L W = \begin{bmatrix} \bar{L} & 0_{(n-1)\times 1} \\ 0_{1\times(n-1)} & 0 \end{bmatrix}$$

where the last column of matrix W is $1_n/\sqrt{n}$, $\bar{L} \in \Re^{(n-1)\times(n-1)}$.

Proof: Because $1_n{}^T$ and 1_n are the corresponding left and right eigenvector respectively. The proof of this lemma is straightforward.

3. Main Results

In this section, we provide the convergence analysis of the average consensus problem in undirected network with fixed and switching topology coupling multiple time-varying delays. Sufficient conditions expressed as LMI are presented for undirected networks of multi-agent with fixed and switching topology, respectively.

3.1. Networks with Fixed Topology

If the fixed communication topology $G(V, E, A)$ is kept connected, we have $1_n{}^T L_k = 0$, and $L_k 1_n = 0$, which imply $\sum_{i=1}^{n} \dot{x}_i = \sum_{i=1}^{n} u_i = 0$. Then, $\alpha = Ave(x(0))$ is an invariant quantity. Thus we have the decomposed equation $x(t) = \alpha 1_n + \delta(t)$, where $\delta = \left(\delta_1, \delta_2, \delta_3 \cdots, \delta_n\right)^T \in \Re^n$, satisfies $\sum_i \delta_i(t) = 0$, i.e., $1^T \delta = 0$. Then (5) is equivalent to

$$\dot{\delta}(t) = -\sum_{k=1}^{N} L_k \delta(t-\tau_k(t)) . \qquad (8)$$

By Lemma 3, we have

$$\dot{\delta}(t) = -WW^T \sum_{k=1}^{N} L_k WW^T \delta(t-\tau_k(t)) ,$$

then

$$W^T \dot{\delta}(t) = -W^T \sum_{k=1}^{N} L_k WW^T \delta(t-\tau_k(t)) .$$

Let $W^T\delta(t)=\left[\Delta^T(t),0\right]^T$, where $\Delta(t)\in\Re^{n-1}$. Then (8) can be transformed into the following equation:

$$\dot{\Delta}(t) = -\sum_{k=1}^{N} \bar{L}_k \Delta(t-\tau_k(t)) \qquad (9)$$

where $\bar{L}_k$ satisfies $W^T L_k W = \begin{bmatrix} \bar{L}_k & 0_{(n-1)\times 1} \\ 0_{1\times(n-1)} & 0 \end{bmatrix}$.

Lemma 4. If $\lim\limits_{t\to\infty}\left\|\Delta(t)\right\|=0$, then $\lim\limits_{t\to\infty}\left\|\delta(t)\right\|=0$.

Proof: From $W^T\delta(t)=\left[\Delta^T(t),0\right]^T$, we have $\delta(t)=W\left[\Delta^T(t),0\right]^T$. Therefore, when $\lim\limits_{t\to\infty}\left\|\Delta(t)\right\|=0$, we have $\lim\limits_{t\to\infty}\left\|\delta(t)\right\|=0$. This completes the proof.

In the following section, we will discuss the convergence of dynamical system (9), that is $\lim\limits_{t\to\infty}\Delta(t)=0$.

Theorem 1. Consider an undirected network of multi-agent with fixed topology coupling multiple time-varying delays satisfies (6). Assume the communication topology G is kept connected. Then, system (5) asymptotically solves the average consensus problem, if there exist positive definite matrices $P, Q_k, R_k \in \Re^{(n-1)\times(n-1)}$, satisfying

$$\begin{bmatrix} \Omega_1 & \sum\limits_{k=1}^{N} P\bar{L}_k & 0_{(n-1)\times(n-1)} \\ \sum\limits_{k=1}^{N} \bar{L}_k{}^T P & -\sum\limits_{k=1}^{N}\dfrac{R_k}{d_k} & 0_{(n-1)\times(n-1)} \\ 0_{(n-1)\times(n-1)} & 0_{(n-1)\times(n-1)} & \Omega_2 \end{bmatrix} < 0 \qquad (10)$$

where $\Omega_1 = -\sum\limits_{k=1}^{N}\left(P\bar{L}_k + \bar{L}_k{}^T P - Q_k\right)$,

$$\Omega_2 = diag[\tau_1(t)\bar{L}_1^T R_1 \bar{L}_1 - (1-h_1)Q_1, \cdots,$$
$$\tau_N(t)\bar{L}_N{}^T R_N \bar{L}_N - (1-h_N)Q_N].$$

Proof: Define a Lyapunov-Krasovskii function for system (9) as follows:

$$V(t) = V_1(t) + V_2(t) + V_3(t),$$

$$V_1(t) = \Delta^T(t) P \Delta(t),$$

$$V_2(t) = \sum_{k=1}^{N}\int_{t-\tau_k(t)}^{t} \Delta^T(s) Q_k \Delta(s) ds ,$$

$$V_3(t) = \sum_{k=1}^{N}\int_{-d_k}^{0}\int_{t+\theta}^{t} \dot{\Delta}^T(s) R_k \dot{\Delta}(s) ds d\theta$$

where P, Q_k and $R_k \in \Re^{(n-1)\times(n-1)}$ are positive definite matrix. Along the trajectory of system (9), we have

$$\dot{V}_1(t) = -2\sum_{k=1}^{N}\Delta^T(t)P\bar{L}_k\Delta(t-\tau_k(t)),$$

$$\dot{V}_2(t) = \sum_{k=1}^{N}\{\Delta^T(t)Q_k\Delta(t)$$
$$-(1-\dot{\tau}_k(t))\Delta^T(t-\tau_k(t))Q_k\Delta(t-\tau_k(t))\}$$
$$\le \sum_{k=1}^{N}\{\Delta^T(t)Q_k\Delta(t)$$
$$-(1-h_k)\Delta^T(t-\tau_k(t))Q_k\Delta(t-\tau_k(t))\},$$

$$\dot{V}_3(t) = \sum_{k=1}^{N}\{d_k\Delta^T(t-\tau_k(t))\bar{L}_k^T R_k \bar{L}_k \Delta(t-\tau_k(t))$$
$$-\int_{t-\tau_k}^{t}\dot{\Delta}^T(s)R_k\dot{\Delta}(s)ds\}.$$

By Newton-Leibniz formula

$$\Delta(t-\tau_k(t)) = \Delta(t) - \int_{t-\tau_k(t)}^{t}\dot{\Delta}(s)ds$$

and note that $2x^Ty \le x^TF^{-1}x + y^TFy$ hold for any appropriate positive definite matrix F, we have:

$$-2\sum_{k=1}^{N}\Delta^T(t)P\bar{L}_k\Delta(t-\tau_k(t))$$
$$= \sum_{k=1}^{N}\{-2\Delta^T(t)P\bar{L}_k\Delta(t) + \int_{t-\tau_k(t)}^{t} 2[\bar{L}_k{}^T P^T \Delta(t)]^T \dot{\Delta}(s)ds\}$$
$$\le \sum_{k=1}^{N}\{-2\Delta^T(t)P\bar{L}_k\Delta(t) + d_k\Delta^T(t)P\bar{L}_k \; R_k{}^{-1}\bar{L}_k{}^T P^T\Delta(t)$$
$$+\int_{t-\tau(k)}^{t}\dot{\Delta}^T(s)R\dot{\Delta}(s)ds\}.$$

Consequently,

$$\dot{V}(t) \le \sum_{k=1}^{N}\{-2\Delta^T(t)P\bar{L}_k\Delta(t) + d_k\Delta^T(t)P\bar{L}_k R_k{}^{-1}\bar{L}_k{}^T P^T\Delta(t)$$
$$+d_k\Delta^T(t-\tau_k(t))\bar{L}_k{}^T R_k\bar{L}_k\Delta(t-\tau_k(t)) + \Delta^T(t)Q_k\Delta(t)$$
$$-(1-h_k)\Delta^T(t-\tau_k(t))Q_k\Delta(t-\tau_k(t))\}$$

$$= \sum_{k=1}^{N}\{\Delta^T(t)\left(-P\bar{L}_k - \bar{L}_k{}^T P + Q_k + d_k P\bar{L}_k R_k{}^{-1}\bar{L}_k{}^T P\right)\Delta(t)$$
$$+\Delta^T(t-\tau_k(t))(d_k\bar{L}_k{}^T R_k\bar{L}_k - (1-h_k)Q_k)\Delta(t-\tau_k(t))\}.$$

Then, a sufficient condition for $\dot{V}(t) < 0$ is

$$\sum_{k=1}^{N}\{\Delta^T(t)\left(-P\bar{L}_k-\bar{L}_k^{\,T}P+Q_k+d_kP\bar{L}_kR_k^{-1}\bar{L}_k^{\,T}P\right)\Delta(t)\}<0, \tag{11}$$

and

$$\sum_{k=1}^{N}\{\Delta^T(t-\tau_k(t))(d_k\bar{L}_k^{\,T}R_k\bar{L}_k-(1-h_k)Q_k)\Delta(t-\tau_k(t))\}<0. \tag{12}$$

As a result, the matrix inequalities (11) hold, if and only if

$$\sum_{k=1}^{N}\left(-P\bar{L}_k-\bar{L}_k^{\,T}P+Q_k+d_kP\bar{L}_kR_k^{-1}\bar{L}_k^{\,T}P\right)<0 \tag{13}$$

Then, by Schur complement formula the matrix inequality (13) is equivalent to

$$\begin{bmatrix} \Omega_1 & \sum\limits_{k=1}^{N}P\bar{L}_k \\ \sum\limits_{k=1}^{N}\bar{L}_k^{\,T}P & -\sum\limits_{k=1}^{N}\dfrac{R_k}{d_k} \end{bmatrix}<0 \tag{14}$$

where Ω_1 is defined in (10).

Therefore, from Lemma 4, average consensus can be achieved if the matrix inequality (6) holds. This completes the proof.

3.2. Networks with Switching Topology

Considering the switching communication topology $G_s(s\in I_\Gamma)$ with system (7), we can obtain the following disagreement switching system:

$$\dot{\Delta}(t)=-\sum_{k=1}^{N}\bar{L}_{ks}\Delta(t-\tau_k(t)),\ s=\sigma(t)\in I_\Gamma, \tag{15}$$

where $\bar{L}_{ks}$ satisfies

$$W^TL_{ks}W=\begin{bmatrix} \bar{L}_{ks} & 0_{(n-1)\times 1} \\ 0_{1\times(n-1)} & 0 \end{bmatrix}.$$

Theorem 2. Consider a undirected network of multi-agent with switching topology coupling multiple time-varying delays satisfies (6). Assume the communication topology $G_s(s\in I_\Gamma)$ is kept connected. Then, system (7) asymptotically solves the average consensus problem, if there exist positive definite matrices $P,Q_k,R_k\in\Re^{(n-1)\times(n-1)}$, satisfying

$$\begin{bmatrix} \bar{\Omega}_1 & \sum\limits_{k=1}^{N}P\bar{L}_{ks} & 0_{(n-1)\times(n-1)} \\ \sum\limits_{k=1}^{N}\bar{L}_{ks}^{\,T}P & -\sum\limits_{k=1}^{N}\dfrac{R_k}{d_k} & 0_{n\times(n-1)} \\ 0_{(n-1)\times(n-1)} & 0_{(n-1)\times(n-1)} & \bar{\Omega}_2 \end{bmatrix}<0, \tag{16}$$

where $\bar{\Omega}_1=-\sum\limits_{k=1}^{N}\left(P\bar{L}_{ks}+\bar{L}_{ks}^{\,T}P-Q_k\right)$,

$$\bar{\Omega}_2=diag[\pi_1(t)\bar{L}_{1s}^{\,T}R_1\bar{L}_{1s}-(1-h_1)Q_1,\cdots,$$
$$\pi_N(t)\bar{L}_{Ns}^{\,T}R_N\bar{L}_{Ns}-(1-h_N)Q_N].$$

Proof: The proof of theorem 2 is similar to that of theorem 1, so it is omitted here.

Remark: When $h\le 1$ of the condition (6) is replaced with $h>1$, we can obtain corresponding results if choosing $V(t)=V_1(t)+V_3(t)$.

4. Simulation

In this section, simulation examples will be given to validate the theoretical results obtained in the previous section. Consider a group of 10 agents labeled 1 through 10. Figure 1 shows four examples of undirected graph, which are all connected, and the corresponding adjacency matrices are limited to 0-1 matrices. Let the orthogonal matrix W is

$$\left[\begin{matrix} 0.0275 & -0.0303 & 0.3594 & -0.1500 & -0.4082 \\ 0.0275 & -0.0303 & 0.3594 & -0.1500 & -0.4082 \\ 0.0275 & -0.0303 & 0.3594 & -0.1500 & -0.8165 \\ 0.0594 & -0.0294 & 0.3782 & -0.1390 & -0.0000 \\ 0.0594 & -0.0294 & -0.2404 & -0.1390 & -0.0000 \\ -0.1323 & -0.5742 & 0.2482 & -0.5742 & -0.0000 \\ 0.4914 & 0.6506 & -0.2653 & -0.2557 & -0.0000 \\ -0.8162 & 0.3793 & -0.2234 & -0.1447 & -0.0000 \\ 0.2559 & -0.3149 & -0.4793 & -0.6714 & -0.0000 \\ 0.0000 & 0.0000 & -0.0000 & -0.1615 & -0.0000 \end{matrix}\right.$$

$$\left.\begin{matrix} 0.7071 & -0.0809 & 0.1749 & -0.2062 & 0.3162 \\ 0.7071 & -0.0809 & 0.1749 & -0.2062 & 0.3162 \\ 0.0000 & -0.0809 & 0.1749 & -0.2062 & 0.3162 \\ 0.0000 & -0.1310 & 0.6023 & -0.5946 & 0.3162 \\ 0.0000 & -0.1310 & 0.6694 & -0.5946 & 0.3162 \\ 0.0000 & -0.2061 & -0.1207 & -0.3231 & 0.3162 \\ 0.0000 & -0.1511 & -0.1290 & -0.2450 & 0.3162 \\ 0.0000 & -0.0820 & -0.1087 & -0.0230 & 0.3162 \\ 0.0000 & -0.0090 & -0.0206 & -0.0206 & 0.3162 \\ 0.0000 & -0.9348 & -0.0000 & -0.0000 & 0.3162 \end{matrix}\right]_{10\times 10}.$$

where the last column of matrix W is $1_n/\sqrt{10}$. For simplicity, we assume $N=1$ in undirected networks of multi-agent with both fixed and switching topology.

4.1. Examples of Networks with Fixed Topology

Consider an undirected network with fixed topology G_1 in Figure 1. Employing Theorem 1, we have:

1) for $h_1 = 0$, i.e., $\tau_1 \equiv d_1$, $d_1 \le 0.25$. Figure 2 shows the corresponding error system converges zero asymptotically.

2) for $h_1 = 0.5$, $d_1 \le 0.17$. Figure 3 shows the corresponding error system converges zero asymptotically.

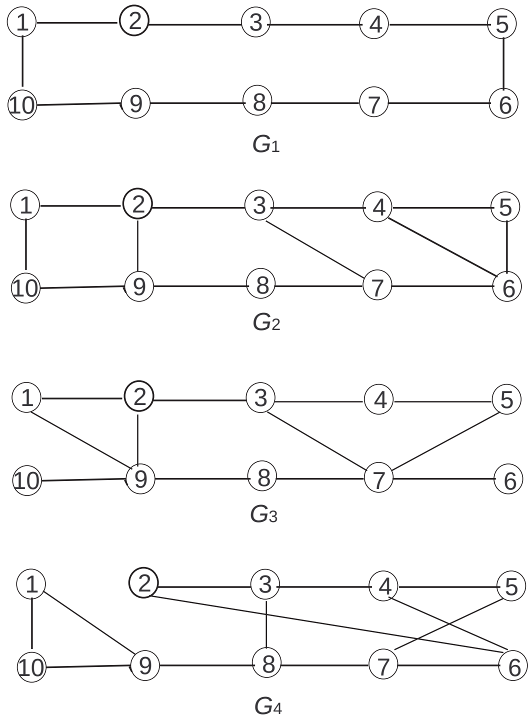


Figure 1. Examples of connected undirected graph.

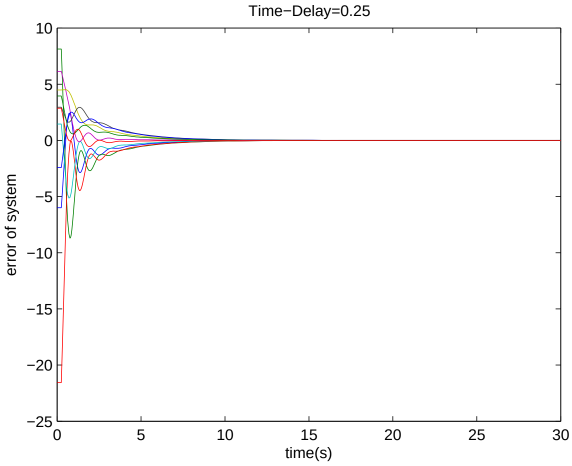


Figure 2. Error system with fixed topology and constant time-delay $d_1 = 0.25s$ with $h_1 = 0$ converges to zero asymptotically.

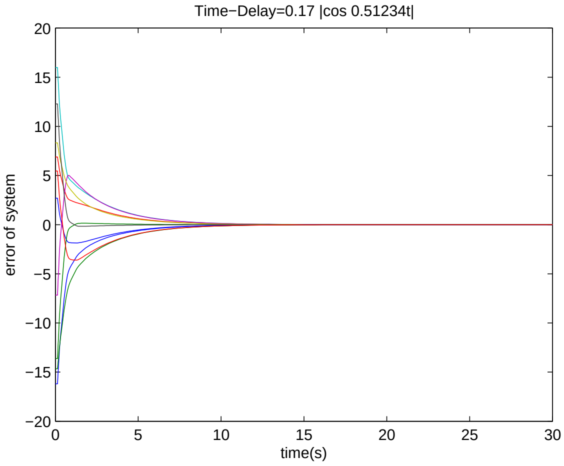


Figure 3. Error system with fixed topology and time-varying delay $\tau_1(t) = 0.17\left|\cos 0.51234t\right| s$ with $h_1 = 0.5$ converges to zero asymptotically.

4.2. Examples of Networks with Switching Topology

A finite automation with set of states $\{G_2, G_3, G_4\}$ is shown in Figure 4, which represents the discrete states of a network with switching topology and time delay as a hybrid system. It starts at the discrete state G_2 and switches every simulation time step to the next state according to the state machine in Figure 4. For multi-agent system with time-varying communication time delay, we take the derivative of delay $h_1 = 0.5$. From theorem 2, the feasible maximum delay bound of the system is $d_1 \le 0.12s$. And the corresponding feasible solution P, Q_1 and R_1 can be obtained by employing the LMI Tool box in Matlab.

Assume the time-varying delay of the error system is $\tau_1(t) = 0.12\left|\cos 0.51234t\right| s$. Figure 5 shows the corresponding error system converges zero asymptotically.

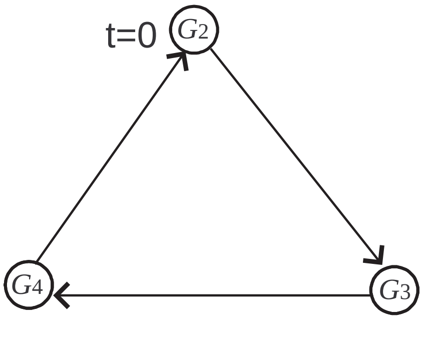


Figure 4. A finite automation with three states.

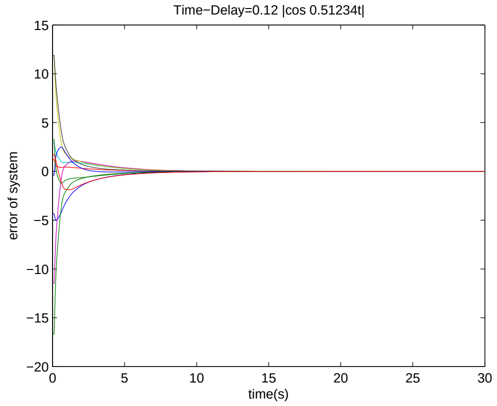


Figure 5. Error system with switching topology and time-varying delay $\tau_1(t) = 0.12\left|\cos 0.51234t\right| s$ converges to zero asymptotically.

5. Conclusions

This paper addresses an average consensus problem of multiagent systems. Undirected networks with fixed and switching network topology coupling multiple time-varying communication delays are considered in this paper. An orthogonal matrix is introduced and the original system is turned into a reduced dimensional system. At the same time, a Lyapunov-Krasovskii function is constructed in the stable analysis. Sufficient conditions in terms of LMI are given to guarantee the system reach average consensus. Moreover, numerical simulation examples are shown to verify the theoretical analysis.

6. Acknowledgment

This work was supported by National Natural Science Foundation of China under the grant 60604001 and 606-74081, Hubei Provincial Natural Science Foundation under the grant 2008CDB316 and 2008CDZ046, and Scientific Innovation Team Project of Hubei Provincial College under the grant T200809.

7. References

[1] T. Vicsek, A. Czirok, E. Benjacob, I. Cohen, *et al.*, “Novel type of phasetransition in a system of self-driven particles,” Physical Review Letters, Vol. 75, No. 6, pp. 1226–1229, 1995.

[2] A. Jadbabaie, J. Lin, and A. S. Morse, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” IEEE Transactions on Automatic Control, Vol. 48, No. 6, pp. 988–1001, 2003.

[3] W. Ren, and R. W. Beard, “Consensus seeking in multi-agent systems under dynamically changing interaction topologies,” IEEE Transactions on Automatic Control, Vol. 50, No. 5, pp. 655–661, 2005.

[4] L. Moreau, “Stability of multiagent systems with time-dependent communication links,” IEEE Transactions on Automatic Control, pp. 169–182, 2005.

[5] R. O. Saber, and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” IEEEE Transactions on Automatic Control, Vol. 49, No. 9, pp. 1520–1533, 2004.

[6] R. O. Saber and R. M. Murray, “Consensus protocols for networks of dynamic agents,” in Proceedings of American Control Conference, pp. 951–956, June 2003.

[7] Y. Tian and C. Liu, “Consensus of multi-agent systems with diverse input and communication delays,” IEEE Transactions on Automatic Control, Vol. 53, No. 9, pp. 2122–2128, October 2008.

[8] R. O. Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” in Proceedings of IEEE, Vol. 95, No. 1, pp. 215–233, January 2007.

[9] W. Ren, R. W. Beard, and E. M. Atkins, “Information consensus in multivehicle cooperative control,” IEEE Control Syst. Mag., Vol. 27, No. 2, pp. 71–82, April 2007.

[10] H. Yu, J. G. Jian and Y. J. Wang, “Weighted average consensus for directed networks of multi-agent,” Microcomputer Information, Vol. 23, No. 7, pp. 239–241, Augest 2007.

[11] W. Ren, “Multi-vehicle consensus with a time-varying reference state,” Systems & Control Letters, Vol. 56, No. 7-8, pp. 474–483, July 2007.

[12] D. P. Spanos, R. O. Saber, and R. M. Murray, “Dynamic consensus on mobile networks,” In IFAC World Congr., Prague, Czech Republic, 2005.

[13] Y. Hatano and M. Mesbahi, “Agreement over random networks,” IEEE Transactions on Automatic Control, Vol. 50, No. 11, pp. 1867–1872, November 2005.

[14] R. O. Saber, “Flocking for multi-agent dynamic systems: algorithms and theory,” IEEE Transactions on Automatic Control, Vol. 51, No. 3, pp. 401–420, March 2006.

[15] A. T. Salehi and A. Jadbabaie, “A necessary and sufficient condition for consensus over random networks,” IEEE Transactions on Automatic Control, vol. 53, No. 4, pp. 791–795, April 2008.

[16] G. M. Xie and L. Wang, “Consensus control for a class of networks of dynamic agents,” Robust Nonliner Control, Vol. 17, pp. 941–959, 2007.

[17] Y. G. Sun, L. Wang and G. M. Xie, “Average consensus in directed networks of dynamic agents with time-varying delays,” In Proceeding of IEEE Conference Decision & Control, Vol. 57, No. 2, pp. 3393–3398, December 2006.

[18] P. Lin and Y. Jia, “Average consensus in networks of multi-agents with both switching topology and coupling time-delay,” Physica A, Vol. 387, No. 1, pp. 303–313, January 2008.

[19] P. Lin, Y. Jia and L. Li, “Distributed robust H∞consensus control in directed networks of agents with time-delay,” Systems & Control Letters, Vol. 57, No. 8, pp. 643–653, Augest 2008.

[20] Y. G. Sun, L. Wang and G. M. Xie, “Average consensus

in networks of dynamic agents with switching topologies and multiple time-varying delays," Systems & Control Letters, Vol. 57, No. 2, pp. 175–183, February 2008.

[21] D. Angeli and P. A. Bliman, "Stability of leaderless discrete-time multiagent systems," Math. Control Signs Systems, Vol. 18, pp. 293–322, 2006.

[22] A. Papachristodoulou and A. Jadbabaie, "Synchronization in oscillator networks: Switching topologies and non-homogeneous delays," In Proceedings of the 44th IEEE Conference Decision and Control and the European Control Conference, pp. 5692–5697, December 2005.

[23] A. Papachristodoulou and A. Jadbabaie, "Synchronization in oscillator networks with heterogeneous delays, switching topologies and nonlinear dynamics," In Proceedings of the 45th IEEE Conference Decision and Control, pp. 4307–4312, December 2006.

[24] N. Biggs, "Algebraic Graph Theory," Cambridge University Press, Cambridge, U. K., 1974. 6

[25] R. A. Hornn and C. R. Johnson, "Matrix Analysis," Cambridge University Press, Cambridge, New York, 1985.

Int. J. Communications, Network and System Sciences, **2010, 3, 204-207**
doi:10.4236/ijcns.2010.32029 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Geometry Aspects and Experimental Results of a Printed Dipole Antenna

Constantinos VOTIS[1], Vasilis CHRISTOFILAKIS[1,2], Panos KOSTARAKIS[1]
[1]*Physics Department, University of Ioannina, Panepistimioupolis, Ioannina, Greece*
[2]*Siemens Enterprise Communications, Enterprise Products Development, Athens, Greece*
E-mail: *kvotis@grads.uoi.gr, basilios.christofilakis@siemens-enterprise.com, kostarakis@uoi.gr*
Received November 2, 2009; *revised December* 10, 2009; *accepted January* 12, 2010

Abstract

Detail experimental measurements of a 2.4 GHz printed dipole antenna for wireless communication systems is presented and discussed. A group of printed dipoles with integrated balun have been designed and constructed on a dielectric substrate. This paper is based on modifications of the known printed dipole architecture. The corresponding printed dipole antennas have differences on their forms that are provided by two essential geometry parameters. The first parameter *l* is related to the bend on microstrip line that feeds the dipole and the second *w* corresponds to the form of the dipole's gap. The impact of these parameters on reflection coefficient and radiation pattern of antenna has been investigated. The corresponding measured results indicate that the return loss and radiation pattern of a printed dipole antenna are independent of the *w* parameter. Instead, variations in the value of the *l* parameter in the dipole's structure affect the form of the corresponding return loss. These observations are very important and provide interesting considerations on affecting design and construction of antenna elements at frequency range of 2.4 GHz.

Keywords: Printed Dipole, Scattering Parameters, Radiation Pattern

1. Introduction

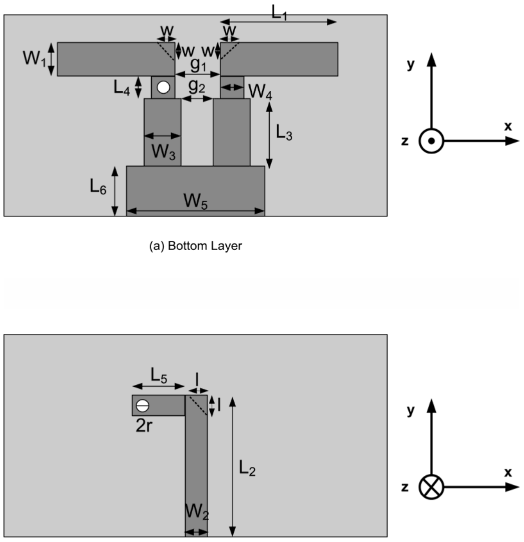
Modern wireless communications offer higher bit rates and efficient quality of services. The majority of the equipment used today introduces requirements for better performance and lower cost. Antennas with quite small sizes, low profiles and versatile features represent interesting solutions that provide modern wireless applications. The printed dipole antenna with integrated balun is widely used as a radiation element on communication systems because of its omni-directional features, narrowband character and simple structure [1–4]. This type of antenna because of its small size can be integrated on the same PCB with other electronics circuits and devices. For the same reason, it can also be used as element on antenna array architecture. The last feature is very interesting and attractive in MIMO modern wireless systems. This printed dipole architecture offers versatile characteristics for design and implementation of antenna arrays on both ends of a MIMO wireless system.

In the present paper, we will study and discuss the effect of the variation of the two geometrical parameters (*l*, *w*) of the printed dipole antenna structure. The first corresponds to a discontinuity on microstrip line of printed dipole and the second is related to the discontinuity in the gap. Details of structure concept and design process are presented in Section 2; the experimental results for return loss and radiation pattern for each of the printed dipoles are presented and discussed in Section 3. The paper concludes in Section 4.

2. Design and Structure Aspects

As mentioned above, the proposed analysis is based on geometrical characteristics of a prototype printed dipole antenna with integrated balun. This kind of printed dipole antenna is considered for use in many applications [1–3]. In our study the geometrical parameters of the printed dipole antenna were modified to achieve better performance in the frequency range of 2.4 GHz. This modified design and the corresponding parameters are shown in Figure 1 while the values summarized in Table 1.

Identify applicable sponsor/s here. *(sponsors)*



(a) Bottom Layer

(b) Top Layer

Figure 1. Geometry of printed dipole.

Table 1. Printed dipole dimensions.

Parameter	Values

Dipole strips	L1 = 20.8 mm
	W1 = 6 mm
	g1 = 3 mm

Microstrip Balun	L2 = 32 mm
	L3 = 16 mm
	L4 = 3 mm
	L5 = 3 mm
	W2 = 2 mm
	W3 = 5 mm
	W4 = 3 mm
	g2 = 1 mm

Via radius	r = 0.375 mm

Ground plane	L6 = 12 mm
	W5 = 17 mm

Side of microstrip bend	l variable (0 mm – 3 mm)
Side of dipole’s arms in the gap	w variable (0 mm – 3 mm)

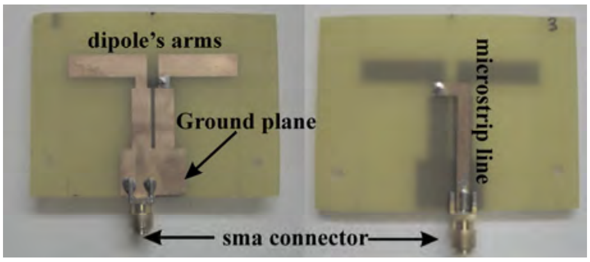


Figure 2. Prototype printed dipole antenna: Top Layer (left) –Bottom Layer (right).

An Fr-4 substrate with thickness of 1.5 mm and permittivity ε_r =4.4 has been used for the fabrication of the dipoles. Figure 2 shows the top and bottom layer for the one of them. It also presents the dipole’s arms and gap, the balun, the ground plane and the microstrip line that interface the dipole with the coaxial feed line via sma connector.

From this Figure, we can also see the right angle at the microstrip line and the other two right angles at the dipole’s gap. It is known that the presence of right angles in conductors cause discontinuities that leads to degradation in circuit performance [5]. Microwave theory suggests that these angles introduce parasitic reactances which can lead to phase and amplitude errors, input and output mismatch and possibly spurious coupling [5–7]. In order to reduce this effect it is proposed to modify these discontinuities directly, by mitering the conductor. Our investigation and the experimental measurements show the effect of mitering these discontinuities. At first, a prototype printed dipole antenna with unaffected geometrical parameters has been designed and constructed. Secondly, we constructed and measured six different printed dipoles. Three of them had *w = 0 mm* and different *l* values (1 mm, 2 mm, 3 mm) and the other three dipoles had *l = 0 mm* and different values of *w* (1 mm, 2 mm, 3 mm). All these seven dipoles we constructed, the unaffected one and the mitered ones were measured in an anechoic environment. Figures 3 and 4 show a printed dipole for *l = 2 mm* and *w = 0* mm and for *l = 0 mm* and *w = 3 mm*, respectively. The aim of this study is to investigate the return loss coefficient and radiation pattern in each of these seven dipole’s forms. The next section discusses the obtained results and presents the significant observations.

3. Results and Discussion

The return loss of the prototype dipole and the six different modified printed dipole antenna we constructed are measured using a Network Analyzer. These results are shown in two Figures. The first (Figure 5) corresponds to *l* parameter’s variations keeping the *w* parameter equal to zero. The second (Figure 6) shows the return loss curves where *w* parameter varies but the *l* parameter equals to zero. In both figures we can see the return loss curve that belongs to the prototype printed dipole (*l* and *w equal to 0 mm*).

From these curves, it seems that this dipole antenna design has a resonance point at 2.4 GHz with 500 MHz –

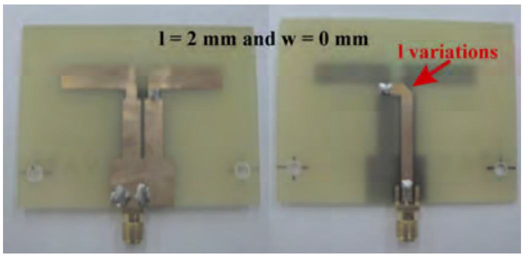


Figure 3. Printed dipole antenna for *l = 2 mm* and *w = 0 mm* Top Layer (left) – Bottom Layer (right).

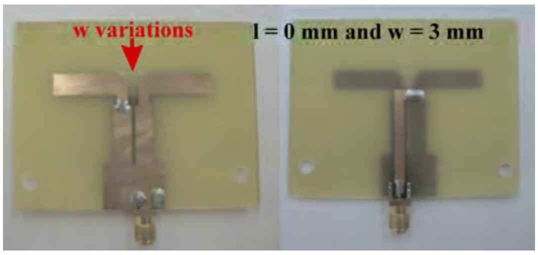


Figure 4. Printed dipole antenna for *l = 0 mm* and *w = 3 mm* Top Layer (left) – Bottom Layer (right).

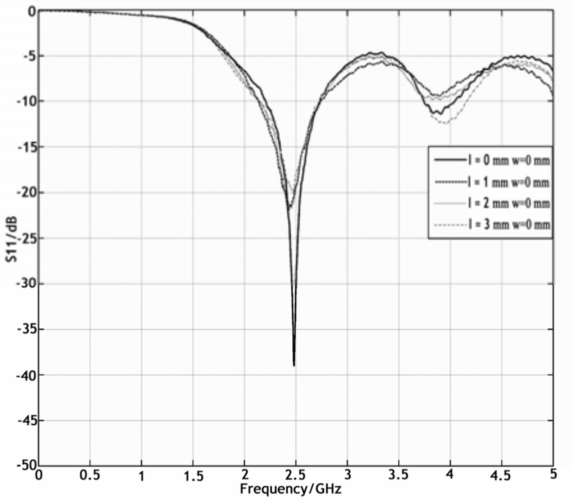


Figure 5. Measured return loss of printed dipole for each value of *l* parameter and *w = 0* mm.

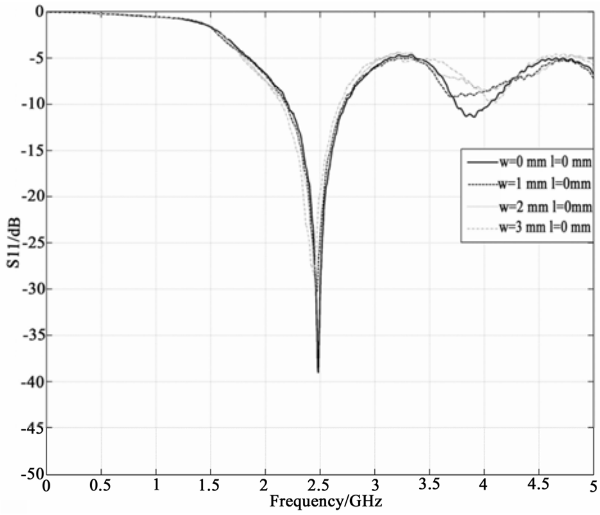
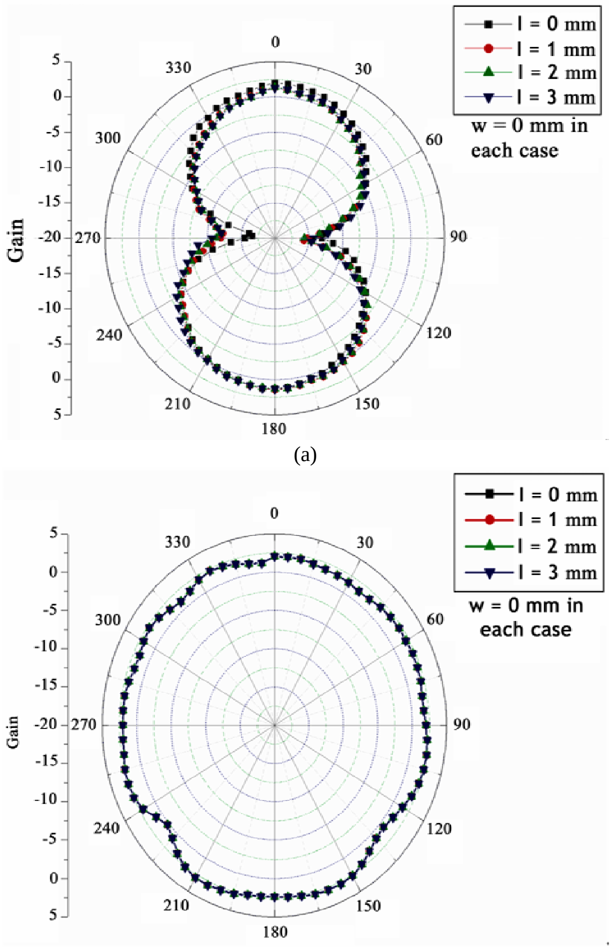


Figure 6. Measured return loss of printed dipole for each value of *w* parameter and *l = 0 mm.*

10 dB bandwidth. The last frequency range has center frequency close to 2.4 GHz which is the frequency value that the return loss is minimized. For these frequencies the corresponding values of return loss are smaller or equal to -10 dB. From Figure 5, it is obvious that as the value of *l* parameter increases, the form of the corresponding return loss curve changes and becomes more flat at the resonance frequency range. On the other hand, the value of *w* parameter does not affect the form of the return loss curve. Each of these seven forms of printed dipole antenna has quite similar return loss curves and introduces narrowband operation at the frequency range of 2.4 GHz. Moreover, for a wireless application that requires design and construction of many identical printed dipoles, it is recommended to choose *l* parameter equals to *2 mm* and *w* parameter equals to *0 mm* for better performance. As it can be seen from Figure 5 the above investigation ensures that the printed dipole antennas will have quite identical return loss curves and performance as elements in an antenna array configuration.

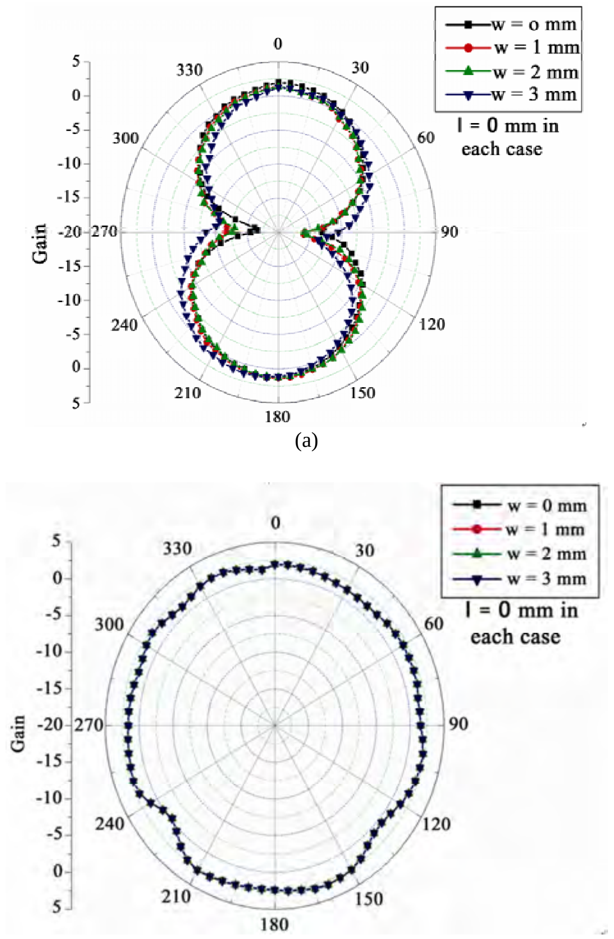
For deeper analysis on this topic, experimental measurements on radiation pattern of these antennas have also been made. Measurements were carried out in a RF anechoic chamber using a calibrated measuring system. In particular, Figure 7 shows the measurements of radiation pattern in E- plane and in H – plane for each dipole



(a)

(b)

Figure 7. Radiation pattern of dipole for each value of *l* parameter (a) E – plane, (b) H -plane.



(a)

(b)

Figure 8. Radiation pattern of dipole for each value of *w* parameter (a) E – plane, (b) H – plane.

with *w* parameter equals to 0 mm and *l* parameter equals to integer values that ranging from *0 mm* to *3 mm*,. Figure 8 shows the corresponding results for each dipole with *l* parameter equals to *0 mm* and *w* parameter's integer values ranging from *0 mm* to *3 mm*. All these dipole structures introduce radiation characteristics that correspond to a fundamental dipole antenna [6,7]. Each of them has a measured peak gain that equals to quite 2 dBi and introduces omni-directional features. Quite small variations on these curves are on the limits of measurements' accuracy. For this reason, it can be observed that the radiation characteristics of the printed dipole antenna are not affected by the variations on *l* and *w* geometrical parameters. Therefore, the radiation diagrams of them are independent of the *l* and *w* parameters.

4. Conclusions

A number of printed dipole antennas with integrated balun are constructed and studied in terms of return loss and radiation pattern. Each of them has a defined form and geometry. Starting from a dipole antenna we mitered the angles introducing the parameters *l* and *w* that we varied. Experimental measurements on return loss provide the obtained results. These are quite similar and also introduce a resonance point at frequency range of 2.4 GHz with narrow resonance bandwidth. The form of this resonance range is affected only by the *l* parameter. The radiation pattern of these dipoles is also investigated. The corresponding radiation diagrams are independent of these geometrical parameters (*l*, *w*) and are similar to that of the fundamental dipole. These observations on printed dipole architecture are very crucial for wireless communication engineering and antenna design. This is because they introduce the ability of constructing a group of identical dipoles choosing an appropriate value of *l* parameter (*l =2 mm*) with quite identical resonance and radiation characteristics.

5. Acknowledgment

This research project (PENED) is co-financed by E.U.-European Social Fund (80%) and the Greek Ministry of Development-GSRT (20%).

6. References

[1] D. Edward and D. Rees, "A broadband printed dipole with integrated balun," Microwave J, 1987, pp. 339–344.

[2] N. Michishita, H. Arai, M. Nakano, T. Satoh, and T. Matsuoka, "FDTD analysis for printed dipole antenna with balun," Asia Pacific Microwave Conference, 2000, pp. 739–742.

[3] G. S. Hilton, C. J. Railton, G. J. Ball, A. L. Hume, and M. Dean, "Finite–difference time–domain analysis of a printed dipole antenna," 19th Int. IEEE Antennas and Propagation Conference, 1995, pp. 72–75.

[4] H. R. Chuang and L. C. Kuo, "3-D FDTD design analysis of a 2.4 GHz polarization – diversity printed dipole antenna with integrated balun and polarization– switching circuit for wlan and wireless communication application," IEEE Transactions on Microwave Theory and Techniques, Vol. 51, No. 2, 2003.

[5] D. M. Pozar, Microwave Engineering, Wiley, 1998.

[6] C. A. Balanis, Antenna Theory Analysis and Design, Wiley, 1997.

[7] R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, Microstrip Antenna Design Handbook, Artec House, 2001.

Int. J. Communications, Network and System Sciences, 2010, 3, 208-211
doi:10.4236/ijcns.2010.32030 Published Online February 2010 (http://www.SciRP.org/journal/ijcns/).

Scientific Research

Two Algorithms of Power Loading for MCM System in Nakagami Fading Channel

Lev GOLDFELD, Vladimir LYANDRES, Lior TAKO
Department of Electrical and Computer Engineering, Communication Laboratory
Ben-Gurion University of the Negev, Beer-Sheva, Israel
Email: *lyandres@ee.bgu.ac.il*

Received November 19, 2009; *revised December* 22, 2009; *accepted January* 17, 2010

Abstract

We consider two non-iterative algorithms of adaptive power loading for multicarrier modulation (MCM) system, The first one minimizes the average power of the system transmitter and ensures the preset average bit-error rate, while the second reduces the average transmitting power subject to the given values of demanded bit-error rate and of the outage probability. The algorithms may be used for power-efficient management of the up-link in cellular communication, where mobile terminals use rechargeable batteries, or of the downlink in satellite communication with solar power source of a transponder. We present performance analysis of the adaptive MCM systems supported by computer simulation for the case of the *m*-Nakagami fading and additive white Gaussian noise in the forward and backward channels. Evaluation of the power gain of the proposed strategies and its comparison with uniform power loading shows that the gain depends on the fading depth and average signal to noise ratio in the system sub-channels.

Keywords: MCM, Fading Channel, Adaptive Power Loading

1. Introduction

Data transmission over bandwidth-limited channel using a multicarrier modulation (MCM) has attracted a lot of interest, due to its ability to achieve relatively good reception quality in spite of frequency selective fading. If the channel state information (CSI) is available at the transmitter side, the MCM system performance can be significantly improved by means of adaptive power loading (APL) [1–5].

In this paper, we propose two non-iterative APL strategies. The first strategy, relevant to the case of continuous data transmission, is to minimize the average transmitting power required for the given value the average bit-error rate (ABER). The purpose of the second strategy, more suitable for packet transmission, is to search for a value of the average transmitting power which is sufficient for given values of ABER *and* of the outage probability. The proposed strategies might be useful for power management of the uplink cellular communication channel, since mobile terminals use rechargeable batteries, or of the downlink from a satellite transponder with solar power sources.

At first, we assume that the communication channel is time-invariant, i.e. it has a "frozen" frequency-selective transfer function and suffers only from additive Gaussian noise. Under the assumption of the perfect channel state information we evaluate average power gain (APG) due to adaptation as a function of the Signal-to-Noise Ratio averaged over the system sub channels (ASNR). Although this premise is too optimistic, it is widely used in the literature while evaluating the upper bound for performance of the adaptive MCM system.

2. Algorithms of Adaptive Power Loading

We assume that the channel amplitude transfer function is well staircase approximated, i.e. it may be presented as a set of N equal-bandwidth sub-channels, each with the same noise power σ^2 and with a constant random gain $K_n,\ n \in [1, N]$ known to the transmitter precisely. The ABER of the MCM system is defined as

$$\overline{P}_{er\,N} = \frac{1}{N}\sum_{n=1}^{N} p_{er\,n}\left(\gamma_n P_{Tr\,n}\right) \tag{1}$$

where $p_{er\,n}$ is the bit-error rate in the n-th sub channel

$$\gamma_n = K_n^2 \big/ \sigma^2 \tag{2}$$

is the normalized partial SNR and $P_{Tr\,n}$ is the power loaded in the *n*th sub channel.

2.1. The First Algorithm

We define the normalized partial BER in the AMCM system as

$$b_n = \frac{p_{er\,n}}{N \cdot \overline{P}_{er\,N}} \tag{3}$$

It is evident that

$$\sum_{n=1}^{N} b_n = 1 \tag{4}$$

and the total transmitted power is written now as

$$P_{Tr\,N} = \sum_{n=1}^{N} P_{Tr\,n}(b_n) \tag{5}$$

The search of the optimal power partition vector

$$\mathbf{P}_{Tr\,N\,opt} = \left\{P_{Tr\,n\,opt}\right\}_{n=1}^{N}$$

can be formulated as minimization of (5) subject to the required value of ABER $\overline{P}_{er\,N} = const$ and can be found by the LaGrange method, i.e. solving the system of equations:

$$\frac{\partial\left[\Phi(\vec{b})\right]}{\partial(b_n)} = 0, \quad n = 1,\ldots,N\,, \tag{6}$$

where

$$\Phi(\vec{b}) = P_{Tr\,N}(\vec{b}) - \lambda\left(\sum_{n=1}^{N} b_n - 1\right) \tag{7}$$

and λ is the LaGrange multiplier.

Let us consider, for example, the solution to this problem for the case of MPSK modulation in sub channels, for which BER in the *n*th sub channel is [7]

$$p_{er\,n} = \frac{2Q\left(\sqrt{2\log_2 M \cdot \gamma_n P_{Tr\,n}}\,\sin \pi/M\right)}{\log_2 M} \tag{8}$$

It is well known that Q-function may be approximated as

$$Q(x) = a\exp\left(-x^2/a\right) \tag{9}$$

where a is the coefficient dependent on x^2. From (8) and (9), we obtain the following expression for the partial BER

$$p_{er\,n} \approx \alpha_1 \exp\left(-\alpha_2\gamma_n P_{Tr\,n}\right) \tag{10}$$

and using a simple transformation of (1) together with (3), the total transmitter power may be written as

$$P_{Tr\,N} = -\sum\nolimits_{n=1}^{N} \frac{\ln\left(N\overline{P}_{erN} b_n / \alpha_1\right)}{\alpha_2\gamma_n} \tag{11}$$

where

$$\alpha_1 = 2a/\log_2 M \tag{12}$$

and

$$\alpha_2 = \frac{2\log_2 M \sin^2(\pi/M)}{a} \tag{13}$$

Substituting (11) into (6) and searching for $\min P_{Tr\,N}$ subject to $\overline{P}_{er\,N} = const$, we obtain the optimal transmitter power partition

$$P_{Tr\,n\,opt} = -(\alpha_2\gamma_n)^{-1} \ln\left[\frac{N\overline{P}_{er\,N}}{\alpha_1\gamma_n}\left(\sum\nolimits_{k=1}^{N} 1/\gamma_k\right)^{-1}\right] \tag{14}$$

The minimal total transmitting power $P_{Tr\,N\,opt}$ providing the given value of ABER of the MCM system is written as

$$P_{Tr\,N\,opt} = \sum_{n=1}^{N} -\frac{1}{\alpha_2 \cdot \gamma_n} \ln\left(\frac{N \cdot \overline{P}_{er\,N}}{\alpha_1} \frac{1/\gamma_n}{\sum\limits_{k=1}^{N} 1/\gamma_k}\right) \tag{15}$$

2.2. The Second Algorithm

Let us assume that certain time invariant bounded value of the ABER is required

$$P_{er\,N} = \frac{1}{N}\sum_{n=1}^{N} p_{er\,n}\left(\gamma_n P_{Tr\,n}\right) \le P_{er\,bound} \tag{16}$$

and the average transmitting power is to be minimized. For the non adaptive MCM system the condition is reached for any value $K_n \ge K_{\min}$ if the equal power $P_{Tr\,n}$ loaded into any sub-channels satisfies the following condition

$$\gamma_{\min} = P^{-1}_{er\,bound}(\gamma) = \frac{K^2_{\min} P_{Tr\,n}}{\sigma_n^2} \tag{17}$$

To achieve the requested value of BER in all sub channels of the AMCM system, the total power of the transmitter $P_{Tr\,N}\left(\overline{K}_n\right)$ should be distributed between sub channels in conformity with the following rule

$$\begin{aligned} P_{Tr\,n} &= P^{(\max)}_{Tr\,N} = \frac{\left(K_{\min}/K_n\right)^2}{N} P^{(\max)}_{Tr\,N} \quad for\ K_n \ge K_{\min} \\ P_{Tr\,n} &= P^{(test)}_{Tr\,n} = \frac{P^{(\max)}_{Tr\,N}}{N} \quad for\ K_n < K_{\min} \end{aligned} \tag{18}$$

Subject to the equity of the total transmitting powers in the adaptive and in the system with uniform power loading of the sub-channels. If the condition $K_n \geq K_{\min}$ is not satisfied for certain sub-channels, they are blocked, and instead a test signal for the sub-channels quality control is transferred[1].

Communication reliability can be defined as the probability of this condition being satisfied in all sub-channels of the system

$$R_N = (R_n)^N = \left(\int_{K_{\min}}^{\infty} f_K(K)dK \right)^N = 1 - P_{out} \qquad (19)$$

where $K_{\min}$ it the minimal value of the sub channel gain, that provides the required value of the BER for given $P_{Tr\,N}$ and P_{out} is outage probability.

3. Performance of the Adaptive MCM System in Fading Channel

In this Section, we consider the performance of the proposed systems in a time-varying channel. We assume that fading processes in the sub channels are mutually independent and have the same *m*-Nakagami probability density function (PDF) [7] of the sub channels gain

$$f_{K_i}(K_i) = \frac{2m^m}{\Gamma(m)K_0^{2m}} K_i^{2m-1} \exp\left(-\frac{mK_i^2}{K_0^2} \right) \qquad (20)$$

where $m \geq 0.5$ and $K_0^2 > 0$ are parameters and $\Gamma(\cdot)$ is the Gamma function. The PDF (20) is able to span the entire range from the deepest one-sided Gaussian fading to no-fading conditions and it is widely used to fit experimental data obtained on a variety of propagation paths from HF sky wave transmission to UHF urban radio propagation and VHF satellite links [7].

Let us define the Average Power Gain (APG) of the AMCM system in fading channel as

$$\eta_N(dB) = 10\lg\left(P_{Tr\,N}^{(U)} \big/ \overline{P}_{Tr\,N} \right) \qquad (21)$$

where $P_{Tr\,N}^{(U)}$ is the total transmitted power of MCM-U system and $\overline{P}_{Tr\,N}$ is the average transmitting power of the system with APL. Using Equations (15), (18), (19) and (20) we obtain, after certain transformations, the final expressions for APG. For the optimal APL

$$\eta_N^{(1)} = \frac{2\gamma_0 \log_2 M \cdot \sin^2(\pi/M)}{a\left[\frac{2^{N-1} m^{0.5mN+1}}{\Gamma^N(m)} A_1 - \frac{m^m}{\Gamma(m)} \ln \frac{N\overline{P}_{er\,N}}{2a} A_2 \right]} \qquad (22)$$

where

$$A_1 = \int_0^\infty y^{0.5m-1} \ln(1+y) \left\{ \int_0^t t^{0.5mN-1} I_{m-2}\left(2\sqrt{mty}\right) K_m^{N-1}\left(2\sqrt{mt}\right) dt \right\} dy$$

$$A_2 = NT(m+1)(m)^{-(m+1)} \qquad (23)$$

$I_k(\cdot)$ is the modified Bessel function of the k-th order and $K_m(\cdot)$is the McDonald function of the m-th order.

For the second APL strategy APG is written as

$$\eta_N^{(2)} = \frac{\Gamma(m)}{m\gamma_0^2 \Gamma\left(m-1, m\gamma_0^2\right) + \gamma\left(m, m\gamma_0^2\right)} \qquad (24)$$

where $\gamma_0^2 = \frac{K_{\min}^2}{K_0^2}$, $\Gamma(x, y)$ is incomplete gamma function and $\gamma(x, y)$ is alternative incomplete gamma function. For this case the equation for the outage probability P_{out} may be obtained from (17) and (18):

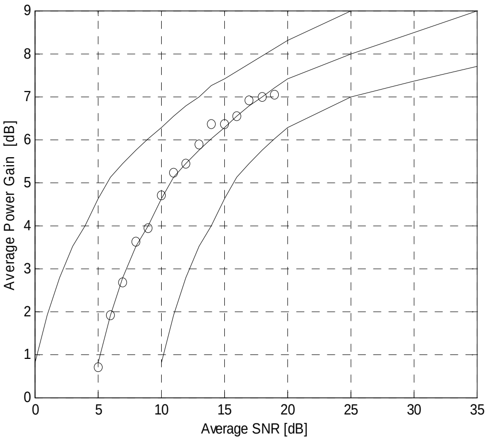


Figure 1. Average power gain of the first APL algorithm as a function of the average SNR.

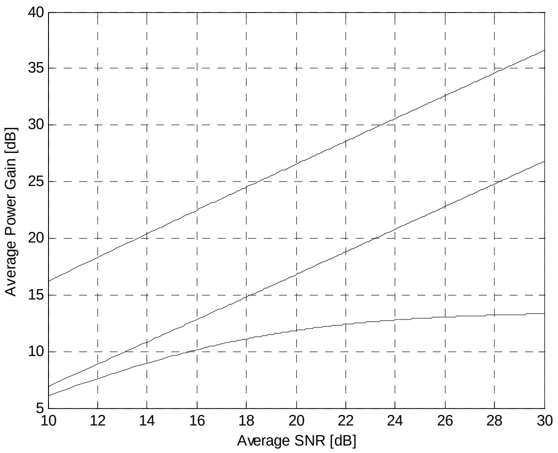


Figure 2. Average power gain of second APL algorithm as a function of the average SNR.

[1] This algorithm was mentioned under the name: Truncated Channel Inversion Power Loading Algorithm in [7]. In [5] the result of its use for minimization of error probability, and in [7] for maximization of the system capacity was investigated.

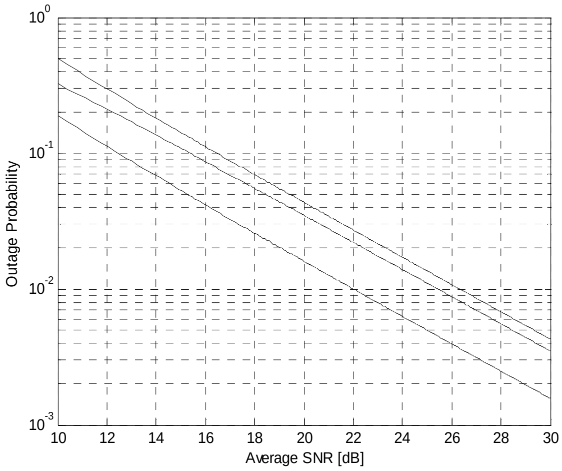


Figure 3. Outage probability of the second APL algorithm as a function of the average SNR.

$$P_{out} = 1 - \left[\frac{\Gamma\left(m, m K_{\min}^2 / K_{02}\right)}{\Gamma(m)} \right]^N \qquad (25)$$

The results of simulation and calculation of APG as a function of average normalized SNR is presented in Figure 1 for the first APL algorithm and in Figure 2 for the second one (N=16, m=0.75; 1.; 2). The results indicate that the gain increases, with increasing of average normalized SNR in sub channels (at the expense of increase $K_0^2 = E\left\{K_n^2\right\}$). It is necessary to underline here, that as the depth of fading increases (lower *m*) the APG becomes more significant.

When comparing the gain values for the proposed strategies it is necessary to take into account that optimal APL algorithm (15)) provides minimum of average transmitting power at the same size of ABER that was obtained in MCM system with uniform power distribution (MCM-U). On the other hand the algorithm (18)) reduces the average power for a given values of BER and outage probability. Outage probability P_{out} as a function of average normalized SNR is presented in Figure 3, from where it follows that P_{out} decreases with increasing of average normalized SNR and increases for more deeply of fading (lower value of *m*).

4. Conclusions

In this letter, we investigate low-complexity (non-iterative) algorithms of power-adaptation of MCM system in slow *m*-Nakagami fading channel. The first of the algorithms is relevant to continuous data transmission and minimizes the average power of the system transmitter subject to the required value of the average BER. The second algorithm searches for such value of the transmitter power that is sufficient for the given values of the average BER and of the system outage probability. The gain in transmitting power increases with increasing of the fading depth, and of the average SNR. In sub-Rayleigh and Rayleigh channels it can exceed 10 dB.

5. References

[1] L. Goldfield, V. Lyandres and D. Wulich, “Minimum BER power loading for OFDM in fading channel,” IEEE Transaction on communications, Vol. 49, pp. 14–18, January 2001.

[2] L. Goldfield, “Adaptive OFDM system in fading channel,” European Transactions on Telecommunications, Vol. 14, pp. 293–300, February 2002.

[3] N.Y. Ermolova and B. Makarevich, “Power and subcarrier allocation algorithms for OFDM systems,” in Proc. IEEE 16-th International Symposium on Personal, Indoor and Mobile Radio Communications, pp.352–356, 2005.

[4] L. Goldfeld and V. Lyandres “Minimum power loading strategy for multicarrier system in the capacity constrained frequency selective channel,” Electronic Letters, Vol. 41, No. 25, pp. 1386–1388, December 2005.

[5] C. Mutti, D. Dahlhaus and T. Hunziker, “Optimal power loading for multiple-input single-output OFDM systems with bit level interleaving,” IEEE Transaction on Wireless Communications, Vol. 5, pp. 1886–1895, May 2006.

[6] A. Goldsmith, “Wireless Communications”, Cambridge University Press, 2005.

Call for Papers

International Journal of

Communications, Network and System Sciences (IJCNS)

ISSN 1913-3715 (Print)   ISSN 1913-3723 (Online)

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

IJCNS is an international refereed journal dedicated to the latest advancement of communications and network technologies. The goal of this journal is to keep a record of the state-of-the-art research and promote the research work in these fast moving areas.

Editors-in-Chief

Prof. Huaibei Zhou	Advanced Research Center for Sci. & Tech., Wuhan University, China
Prof. Tom Hou	Department of Electrical and Computer Engineering, Virginia Tech., USA

Subject Coverage

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

- MIMO and OFDM technologies
- UWB technologies
- Wave propagation and antenna design
- Signal processing and channel modeling
- Coding, detection and modulation
- 3G and 4G technologies
- Sensor networks
- Ad Hoc and mesh networks
- Network protocol, QoS and congestion control
- Efficient MAC and resource management protocols
- Simulation and optimization tools
- Network security

We are also interested in:

- Short reports—Discussion corner of the journal:

  2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data.
- Book reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

http://www.scirp.org/journal/ijcns   ijcns@scirp.org

***Int. J. Communications, Network and System Sciences*, 2010, 3, 107-211**
Published Online February 2010 in SciRes (http://www.SciRP.org/journal/ijcns/).

TABLE OF CONTENTS

Volume 3 Number 2 **February 2010**

ISSN: 1913-3715

9771913371005 15