In the Name of God

Journal of
Information Systems & Telecommunication
Vol. 9, Special Issue, December 2021

Research Institute for Information and Communication Technology
Iranian Association of Information and Communication Technology
Affiliated to: Academic Center for Education, Culture and Research (ACECR)

Print ISSN: 2322-1437
Online ISSN: 2345-2773
Publication License: 91/13216
Editorial Office Address: No.5, Saeedi Alley, Kalej Intersection., Enghelab Ave., Tehran, Iran,
P.O.Box: 13145-799 Tel: (+9821) 88930150 Fax: (+9821) 88930157
E-mail: info@jist.ir , infojist@gmail.com
URL: www.jist.ir

Indexed by:

- SCOPUS	www. Scopus.com
- Index Copernicus International	www.indexcopernicus.com
- Islamic World Science Citation Center (ISC)	www.isc.gov.ir
- Directory of open Access Journals	www.Doaj.org
- Scientific Information Database (SID)	www.sid.ir
- Regional Information Center for Science and Technology (RICeST)	www.ricest.ac.ir
- Iranian Magazines Databases	www.magiran.com

Publisher:
Iranian Academic Center for Education, Culture and Research **(ACECR)**

This Journal is published under scientific support of
Advanced Information Systems (AIS) Research Group and
Telecommunication Research Group, ICTRC

Acknowledgement

JIST Editorial-Board would like to gratefully appreciate the following distinguished referees for spending their valuable time and expertise in reviewing the manuscripts and their constructive suggestions, which had a great impact on the enhancement of this issue of the JIST Journal.

(A-Z)

- Abbasi, Mahdi, Bu Ali Sina University, Hamedan, Iran
- Bahadori, Farhad, Islamic University, Fasa Branch, Shiraz, Iran
- Ebadati, Omid Mahdi, Kharazmi University,Tehran, Iran
- Farzi, Saeid, K. N. Toosi University of Technology, Tehran, Iran
- Fathi, Amir, Urmia University, Urmia, Iran
- Ghaffari, Ali, Islamic Azad University, Tabriz Branch, Iran
- Kashef, Seyed Sadra, Urmia University, Urmia, Iran
- Kheirkhah, Fatemeh, ACECR, Tehran, Iran
- Khazaei, Mehdi, Kermanshah University of Technology, Kermanshah, Iran
- Maleki Sadr, Mohammad Amin, Mc Gill university, Montral, Canada
- Mir, Ali, Lorestan University, Iran
- Moradi, Gholamreza, Amirkabir University,Tehran, Iran
- Mohammadzadeh, Sajjad, University of Birjand, South Khorasan, Iran
- Mojarad, Morteza, Urmia University, Urmia, Iran
- Nangir, Mahdi, University of Tabriz, Tabriz, Iran
- Rasi, Habib, University of Technology, Shiraz, Iran
- Shirvani Moghaddam, Shahriar , Shahid Rajaee Teacher Training University, Tehran, Iran
- Vijayaragavan, S.P, Bharat University, Tamil Nadu, India
- Yaghmaee, Mohammad Hossein, Urmia University, West Azerbaijan, Iran

Special Issue: Telecommunication & Signal Processing

Table of Contents

Measurement and Analysis of Radiation Levels from Base Transceiver Station in Sambas

Fitri Imansyah*
Faculty of Engineering, Electrical Engineering Major, Universitas Tanjungpura, Indonesia
fitri.imansyah@ee.untan.ac.id
Leonardus Sandy Ade Putra
Faculty of Engineering, Electrical Engineering Major, Universitas Tanjungpura, Indonesia
leonardusandy@ee.untan.ac.id
Eka Kusumawardhani
Faculty of Engineering, Electrical Engineering Major, Universitas Tanjungpura, Indonesia
ekawardhani@ee.untan.ac.id

Received: 29/Aug/2021 Revised: 01/Nov/2021 Accepted: 07/Dec/2021

Abstract
The development of telecommunications in Indonesia until now has experienced a very significant increase and has become a significant need in communication. Many people use communication tools daily, causing many providers to set up Base Transceiver Stations (BTS) to reach their users to remote areas. BTS has a transmit power that can reach the destination area, but most people still do not know the level of radiation emitted and the health effects on the body. Therefore the International Commission for Non-Ionizing Radiation Protection (ICNIRP) has set a threshold level of safe radiation for the human body. Sambas is one of the cities in West Kalimantan which is the target for the development of BTS establishments by operators. This makes the surrounding community feel afraid of the health caused by radiation from the BTS. So it is necessary to do some research, socialize, measuring, and evaluate the level of radiation emitted from BTS, especially in residential areas. The research was conducted through several stages, including; data collection, data collection methods on variations in distance from BTS, results of radiation level measurements, and comparisons to the safe threshold value for radiation intensity that has been set by ICNIRP. The measurement results from 20 BTS in Sambas show that the radiation level from the BTS measured is still far from the safe radiation threshold that has been set by ICNIRP.

Keywords: Base Transceiver Stations; Radiation Level; ICNIRP; Telecommunication; Sambas.

1- Introduction

The field of telecommunications currently has an essential role in developing and progressing a nation's culture. Especially during the pandemic, every human being is trying to narrow the space and time boundaries to get information, communicate anywhere, anytime, and with anyone in the form of voice, data, images, and videos. In communicating, we can use a communication tool such as a cellphone that almost owned by everyone.

Communicating using mobile phones requires the help of signals emitted from BTS (Base Transceiver Station) for the data exchange process. The establishment of BTS has simultaneously been regulated by the Ministry of Communication and Informatics and a Regulation of the Minister of Public Works and Public Housing which requires operators to use towers simultaneously in order to be more efficient and maintain the aesthetics of the environment and urban planning itself [1], [2]. The establishment of BTS is currently overgrowing due to the increasing public need for telecommunication networks. However, this is inseparable from the health facts that arise due to Radiofrequency (RF).

Communities in the Sambas feel that the sentence of telecommunication towers has an impact on health, safety, and social equity. Health impacts related to radiation emission from electromagnetic fields from transmitters in telecommunication towers have people's concerns about the impacts they experience, such as fertility problems, sleep problems, dizziness, and nausea. In addition, the people who live around the BTS feel afraid that the BTS tower may collapse and hit the surrounding settlements.

Previous research on the clustering effects of electromagnetic field radiation generated by GSM signals has been carried out [3], and concluded that continuous exposure to radiation in high-intensity values causes severe effects on human health. Evaluation of the effect of radiation on the human body by considering the distance to the BTS [4], concluded that radiation exposure could be reduced by a greater distance from the radiation source. The limit of exposure to radiation emission

* Corresponding Author

distribution in the middle of Mossoro City is 10 MHz to 8 GHz wideband with interpolation technique [5]. Research conducted by Agostinho L and Marco A [6], stated that the most substantial radiation exposure was received from 0 meters to 182 meters. So, therefore, the supervision and legal requirements in the development of BTS construction are needed. The antenna used must meet the standards set by ICNIRP so that the negative impact of radiation can be avoided [7]. In May 2011, the International Agency for Research on Cancer (IARC) evaluated the cancer risk from RF radiation providing evidence of the risk of gliomas and acoustic neuromas in humans [8].

Given the problems that arise in the community, it is necessary to research BTS in Sambas to measure and identify the level of radiation generated by BTS. This measurement is needed to provide information to the public about radiation that is around their residence. Measurement of radiation generated from the BTS tower is expected to have a value below the threshold set in ICNIRP. ICNIRP [9], state that there is a difference between the standard magnitude of radiofrequency radiation exposure between the general public living in a public environment and field workers whose daily lives are more often in an environment with a high intensity of radiofrequency or Occupational [10]. According to the ICNIRP General Public, radiation exposure standards for the range 1 - 10.W.m-2 are safe for the public.

This study aims to obtain radiation level data at 20 BTS in Sambas. Measurements were made using a Spectran HF V4 Spectrum Analyzer to measure the radiation level in watts per meter squared. Measurements were made at every 25m interval from the BTS with the maximum distance measured was 150m. The results of this study will be socialized to the surrounding community, especially those who live around BTS. Then a comparison of the measurement data will be carried out against the ICNIRP standard to determine the feasibility of radiation emission in the community.

2- Research Methodology

This research is descriptive and conducted in the Sambas area. Sambas was chosen because of the large number of public complaints regarding the impact of BTS establishment on health, safety, etc. The method used is the measurement of the actual value in the field. This method was chosen because it is easier and affordable to do when compared to the computational method. In addition, selective frequency measurement is better done directly because of the default frequency range setting.

The measured BTS is selected and investigated to see the radiation level (Power Density) produced by radiofrequency, especially on GSM 900 and GSM 1800 in units of watts per meter squared. Measurements were carried out using the Spectran HF V4 Spectrum Analyzer at every 25 m distance from the BTS guardrail. Measurements were carried out by directing the antenna of the measuring instrument slowly towards the direction of the radiofrequency source until the maximum power density value was obtained. Measurements are also carried out at the height of 1.5 m from the ground to minimize RF Radiation reflection from the ground. The measurement results obtained are then recorded, analyzed, and evaluated.

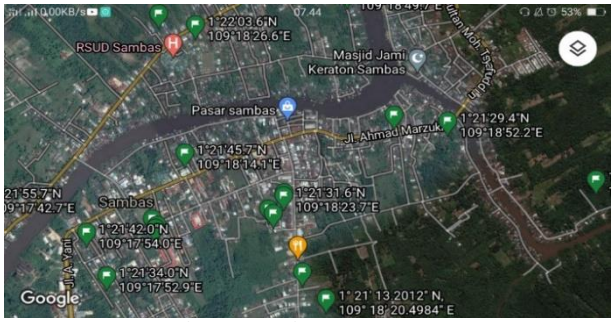


Fig. 1 Map of BTS Identification Locations in Sambas.

Figure 1 shows the distribution map of BTS sites in the Sambas. It can be seen that the towers scattered in the Sambas area are located in residential areas that need to be measured against the radiation generated.

2-1- Base Transceiver Station (BTS)

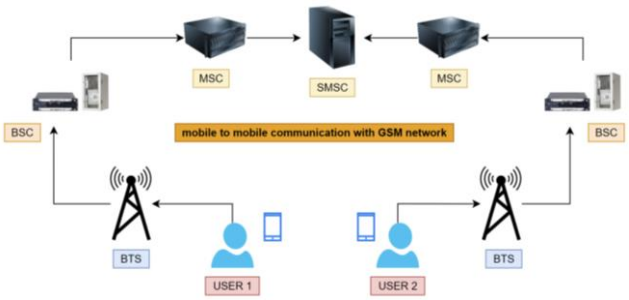


Fig. 2 Illustration of mobile phone network communication and BTS.

Base Transceiver Station is an essential part of the cellular telecommunications network that connects the network of a cellular telecommunications operator with its customers and is shown in Figure 3. BTS consists of three main parts, namely: Tower, Shelter, and Feeder. BTS shelter is a place to store telecommunications equipment. The BTS shelter serves as a device storage media that will be connected to a device center. There are various main and supporting components in the shelter, such as combiner, core module, power supply, fans, lights, and BTS shelter doors.

One beam range generated by a BTS can be called a Cell. Mobile communication is a modern communication that supports high mobility between users. It is then controlled

by a Base Station Controller (BSC) from several BTS, which is connected to a microwave or fiber-optic connection. Although the term BTS can be applied to any wireless communication standards, it is usually and generally associated with mobile communication technologies such as GSM and CDMA.

In this case, the BTS is part of the base station subsystem (BSS) development for the management system. It may also have equipment for encrypting and decrypting communications, spectrum filtering devices (bandpass filters), and antennas as components of the BTS to facilitate BTS functionality. Typically, a BTS will have multiple transceivers (TRXs), allowing it to serve several different frequencies and different cell sectors in the BTS. A BTS controller controls a base station via the base station control (BCF) function. These BCFs are implemented as discrete units or even incorporated into TRX in compact BTS. The BCF provides the operation and maintenance of connections with the network management system and manages the operating conditions of each TRX. The increase in the use of networks connected to BTS will affect the intensity of the electric field around it [11].

2-2- RF Radiation Exposure Standard

The effect of RF radiation on humans is still under endless debate, but the results of studies that have been carried out state that RF emitted from cellular phones [12], including BTS, need to be considered because they have a thermal effect that affects human health [13]. Meanwhile, according to the consensus of the International Scientific Community, it is stated that the energy from BTS is improbable to cause health risks if they do not make direct contact.

Depending on the wattage required and the intended use, BTS can deliver a few watts of power or more. Base station antennas are generally about 20 cm to 30 cm long with a length of 1 meter mounted on a tower with a heig ht of 15 to 50 meters from the ground. The antenna can transmit radio frequency with a very narrow type and perpendicular direction.

RF radiation exposure can be categorized into radiation exposure for workers and the general public [14]. In worker exposure, the person exposed as a consequence of their work is fully aware of the dangers of such exposure and the worker's radiation protection treatment. The International Commission Non-ionizing Radiation Protection (ICNIRP) has set limits for exposure to RF radiation as shown in the following Table 1:

Table 1: Radiation Exposure for the General Public by ICNIRP [14].

Frequency Range	**E-Field Strength** (V/m^{-1})	**H-Field Strength** (A/m^{-1})	**B-Field** (μT)	**Equivalent Plane Wave Power Density** $S_{eq}(Wm^{-2})$
Up to 1 Hz	-	$3.2x10^4$	$4x10^4$	-
1 - 8 Hz	10,000	$3.2x10^4/f^2$	$4x10^4/f^2$	-
8 - 25 Hz	10,000	$4{,}000/f$	$5{,}000/f$	-
0.025 - 0.8 kHz	$250/f$	$4/f$	5/f	-
0.8 - 3 kHz	$250/f$	5	6.25	-
3 - 150 kHz	87	5	6.25	-
0.15 - 1 MHz	87	$0.73/f$	0.92/f	-
1 - 10 MHz	$87/f^{1/2}$	$0.73/f$	0.92/f	-
10 - 400 MHz	28	0.073	0.092	2
400 - 2,000 MHz	$1.375f^{1/2}$	$0.0037f^{1/2}$	$0.0046f^{1/2}$	$f/200$
2 - 300 GHz	61	0.16	0.20	10

2-3- Effect of Radiation Below Standard

The increasing use of communication using cellular phones impacts the development of BTS construction adjacent to homes, schools, hospitals, and densely populated areas. The development that continues to grow raises public concerns about the safety of the population from the radiation it causes.

ICNIRP has set standards for exposure to RF radiation from BTS to minimize health problems due to radiation exposure. Health problems due to exposure to RF radiation cannot be felt instantly [15]. However, the longer a person is exposed to RF radiation, the more dangerous that person's health. Suppose residential areas are too close to BTS or do not follow standards. In that case, it can cause health problems such as headaches [16], [17], brain tumors and cancer [1 8] , and fetal disorders in pregnant women due to imperfect formation of deoxyribonucleic acid (DNA) [19]. Another thing states that if the location of residential areas is following standards, it will minimize health problems due to exposure to RF radiation. A study on cancer stated that no cancer-causing cells were found, and DNA damage was caused by exposure to low RF

radiation levels [20].

3- Result and Discussion

This study's measurement of BTS radiation levels was carried out at 20 BTS in Sambas. Radiation level measurements were carried out at intervals 25 meters from 0 to 150 meters. So, seven variations in distance describe the value of the radiation level emitted by BTS around residential areas. The measurements were carried out at two GSM frequencies, namely the 900 Mhz and 1800 Mhz frequencies. RF and EMF transmit power measurements were carried out using the Spectran HF V4 Spectrum Analyzer. The locations of 20 BTS in Sambas and the results of the measurement of transmit power at the GSM 900 frequency are shown in the following Table 2:

Table 2: Measurement results BTS radiation level at GSM 900 Mhz frequency.

BTS Number	BTS Location	Power Density ($W.m^{-2}$) measurement result in various distance (m)						
		0	25	50	75	100	125	150
1	Jl Kramat Dusun Lb Sari Rt 11 Rw 06 Desa Pendawan Kab Sambas.	0.0012	0.0041	0.0043	0.0057	0.0049	0.0035	0.0075
2	Kab Sambas Desa Pandawa.	0.0034	0.0032	0.0098	0.0016	0.0024	0.0074	0.0094
3	Jl. Tabrani, Lumbang, Kec. Sambas, Kabupaten Sambas, Kalimantan Barat 79462, Sambas.	0.0029	0.0030	0.0067	0.0039	0.0084	0.0056	0.0069
4	Jl Kramat Dusun Lb Sari Rt 11 Rw 06 Desa Pendawan, Sambas.	0.0014	0.0041	0.0061	0.0038	0.0041	0.0065	0.0079
5	Gapura Kec Sambas Kab Sambas.	0.0016	0.0042	0.0006	0.0051	0.0097	0.0053	0.0074
6	Jl. Dusun Sukaramai Rt013/Rw004, Kel. Dalam Kaum, Kec. Sambas, Kab. Sambas, Sambas.	0.0019	0.0029	0.0031	0.0047	0.0093	0.0033	0.0045
7	Jl. Pembangunan, Dn. Sukamantri Rt012/Rw003, Kel. Dalam Kaum, Kec. Sambas.	0.0030	0.0032	0.0022	0.0038	0.0059	0.0079	0.0071
8	Jl. Dagang Tim. Dalam Kaum Kecamtan Sambas Kabupaten Sambas.	0.0021	0.0031	0.0006	0.0036	0.0084	0.0011	0.0042
9	Jl. Gusti Hamzah, Kecamtan Sambas, Kabupaten Sambas.	0.0003	0.0039	0.0031	0.0026	0.0090	0.0031	0.0049
10	Jl. Pendidikan, Jagur, Kecamtan Sambas, Kabupaten Sambas.	0.0024	0.0044	0.0049	0.0017	0.0040	0.0057	0.0060
11	Jl. Sambas - Ledo, Dalam Kaum, Kecamtan Sambas, Kabupaten Sambas.	0.0020	0.0048	0.0084	0.0028	0.0057	0.0027	0.0048
12	Jl. Tabrani, Sunsung, Saing Rambi, Kecamtan Sambas, Kabupaten Sambas.	0.0026	0.0040	0.0061	0.0049	0.0062	0.0053	0.0028
13	Jl Raya Panji Anom	0.0006	0.0006	0.0058	0.0056	0.0069	0.0023	0.0071
14	Jl. Raya Ahmad Mardzuki Kacapuri Rt12/Rw06. Desa Sebayan. Kec. Sambas. Kab. Sambas. Kalimantan Barat	0.0032	0.0012	0.0053	0.0097	0.0046	0.0075	0.0056
15	Jl. Sukaramai Rt 03/ Rw 11. Desa Dalam Kaum. Kec. Sambas. Kab. Sambas.	0.0007	0.0016	0.0033	0.0015	0.0016	0.0059	0.0045
16	Jl. Raya Tabrani Sambas - Singkawang. Dusun Sunsung Rt 07/ Rw 03. Kel. Saing Rambi. Kec. Sambas.	0.0038	0.0032	0.0071	0.0025	0.0096	0.0019	0.0031
17	Jl Raya Sambas Semangau Kec Sambas.	0.0025	0.0039	0.0096	0.0071	0.0007	0.0032	0.0012
18	Jl. Sabas_Subah Dusun Sei Benuah Desa Lubuk Dagang Rt.15 Rw.07 Kec. Sambas Kab. Sambas	0.0030	0.0028	0.0014	0.0073	0.0075	0.0056	0.0075
19	Jl. Sungai Raya Dalam Ii_40. Sambas.	0.0033	0.0008	0.0039	0.0021	0.0098	0.0090	0.0027
20	Jl. Daya Nasional, Sambas.	0.0013	0.0067	0.0081	0.0052	0.0076	0.0049	0.0027

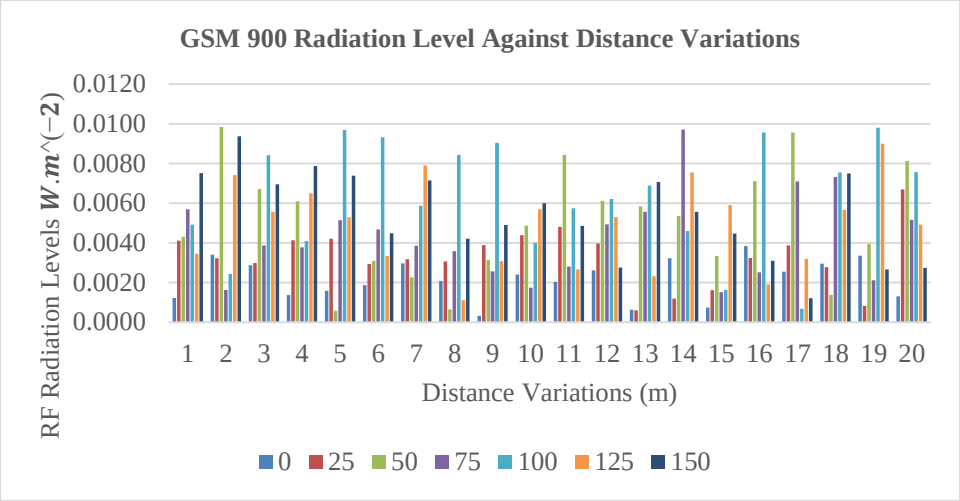


Fig. 3 GSM Radiation Level Frequency 900 to variations in measurement distance.

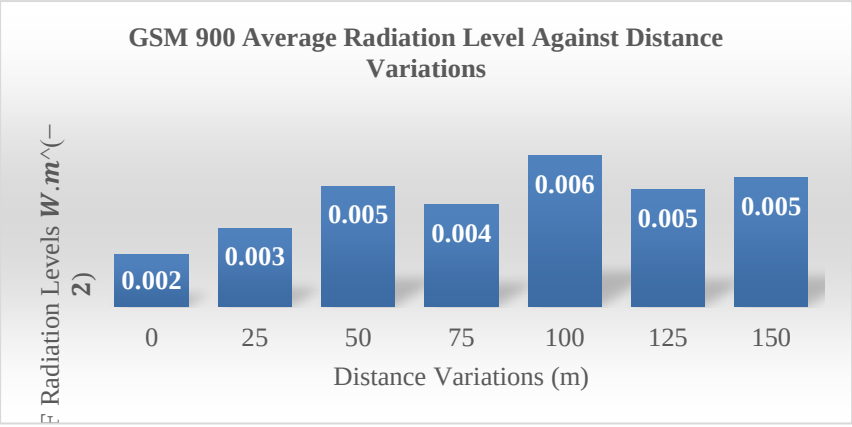


Fig. 4 The average radiation level of GSM 900 to the variation of the measurement distance.

Figure 3 and BTS 19 at 100 meters with a radiation level of 0.0098 W.m^{-2}. Then BTS 9 has the lowest radiation value of 0.0003 W.m^{-2} at 0 meters. Figure 4 shows the average radiation measurement at each distance variation. It can be seen that the lowest radiation frequency is at a distance of 0 meters with a value of 0.0002 W.m^{-2}. The highest average radiation value is at 100 meters which reaches 0.006 W.m^{-2}. The GSM 900 Mhz frequency has the advantage of covering a wider beam area. So that measurements at longer distances still produce higher transmit power.

Table 3 shows the results of measuring radiation levels at the GSM 1800 Mhz frequency against variations in the measurement distance. Figure 5 shows the results of the GSM frequency radiation level measurement at 1800 Mhz. BTS 16 has the highest radiation value of 0.007 W.m^{-2} at 100 m, while BTS 12 has the lowest radiation value of 0.0001 W.m^{-2} at 100 meters.

Table 3: Measurement results BTS radiation level at GSM 1800 Mhz frequency.

BTS Number	BTS Location	Power Density (W.m^{-2}) measurement result in various distance (m)						
		0	25	50	75	100	125	150
1	Jl Kramat Dusun Lb Sari Rt 11 Rw 06 Desa Pendawan Kab Sambas.	0.0021	0.0038	0.0011	0.0030	0.0013	0.0013	0.0044
2	Kab Sambas Desa Pandawa.	0.0062	0.0056	0.0038	0.0031	0.0022	0.0018	0.0026
3	Jl. Tabrani, Lumbang, Kec. Sambas, Kabupaten Sambas, Kalimantan Barat 79462, Sambas..	0.0036	0.0045	0.0034	0.0027	0.0047	0.0019	0.0023

4	Jl Kramat Dusun Lb Sari Rt 11 Rw 06 Desa Pendawan, Sambas.	0.0012	0.0011	0.0047	0.0047	0.0040	0.0005	0.0026
5	Gapura Kec Sambas Kab Sambas.	0.0048	0.0038	0.0044	0.0046	0.0002	0.0047	0.0022
6	Jl. Dusun Sukaramai Rt013/Rw004, Kel. Dalam Kaum, Kec. Sambas, Kab. Sambas, Sambas.	0.0021	0.0044	0.0029	0.0041	0.0016	0.0013	0.0015
7	Jl. Pembangunan, Dn. Sukamantri Rt012/Rw003, Kel. Dalam Kaum, Kec. Sambas.	0.0050	0.0055	0.0046	0.0037	0.0014	0.0012	0.0032
8	Jl. Dagang Tim. Dalam Kaum Kecamtan Sambas Kabupaten Sambas.	0.0047	0.0048	0.0039	0.0007	0.0035	0.0016	0.0032
9	Jl. Gusti Hamzah, Kecamtan Sambas, Kabupaten Sambas.	0.0038	0.0021	0.0049	0.0023	0.0041	0.0028	0.0046
10	Jl. Pendidikan, Jagur, Kecamtan Sambas, Kabupaten Sambas.	0.0016	0.0033	0.0043	0.0031	0.0017	0.0007	0.0013
11	Jl. Sambas - Ledo, Dalam Kaum, Kecamtan Sambas, Kabupaten Sambas.	0.0026	0.0015	0.0011	0.0042	0.0018	0.0006	0.0012
12	Jl. Tabrani, Sunsung, Saing Rambi, Kecamtan Sambas, Kabupaten Sambas.	0.0030	0.0012	0.0016	0.0015	0.0001	0.0013	0.0011
13	Jl Raya Panji Anom	0.0003	0.0013	0.0031	0.0031	0.0022	0.0017	0.0031
14	Jl. Raya Ahmad Mardzuki Kacapuri Rt12/Rw06. Desa Sebayan. Kec. Sambas. Kab. Sambas. Kalimantan Barat	0.0039	0.0060	0.0034	0.0013	0.0036	0.0027	0.0026
15	Jl. Sukaramai Rt 03/ Rw 11. Desa Dalam Kaum. Kec. Sambas. Kab. Sambas.	0.0044	0.0029	0.0050	0.0043	0.0023	0.0003	0.0026
16	Jl. Raya Tabrani Sambas - Singkawang. Dusun Sunsung Rt 07/ Rw 03. Kel. Saing Rambi. Kec. Sambas.	0.0010	0.0017	0.0021	0.0004	0.0070	0.0017	0.0031
17	Jl Raya Sambas Semangau Kec Sambas.	0.0023	0.0034	0.0036	0.0041	0.0034	0.0044	0.0030
18	Jl. Sabas_Subah Dusun Sei Benuah Desa Lubuk Dagang Rt.15 Rw.07 Kec. Sambas Kab. Sambas	0.0011	0.0039	0.0036	0.0045	0.0020	0.0031	0.0034
19	Jl. Sungai Raya Dalam Ii_40. Sambas.	0.0024	0.0033	0.0041	0.0017	0.0018	0.0031	0.0024
20	Jl. Daya Nasional, Sambas.	0.0050	0.0031	0.0026	0.0025	0.0023	0.0022	0.0048

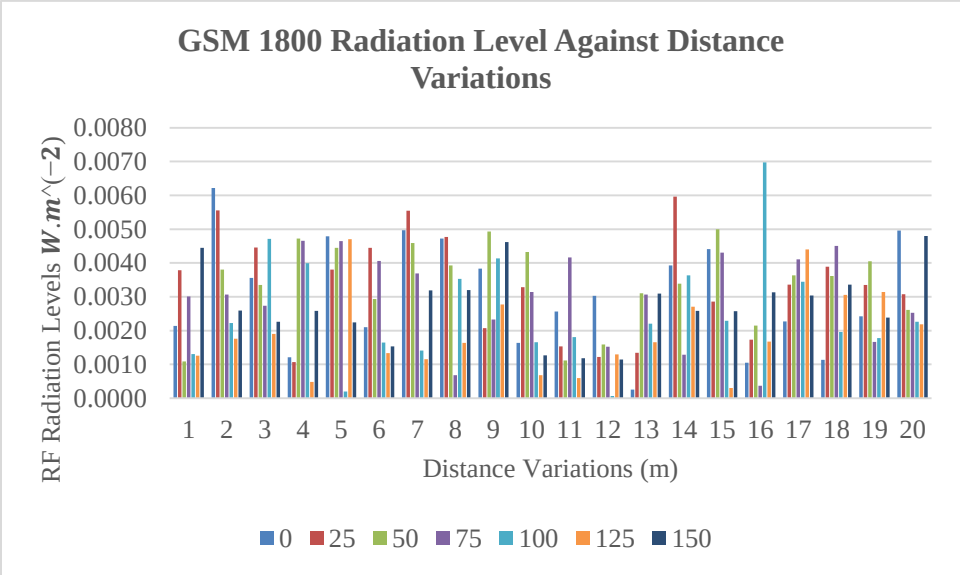


Fig. 5 GSM Radiation Level 1800 MHz Frequency against variations in measurement distance.

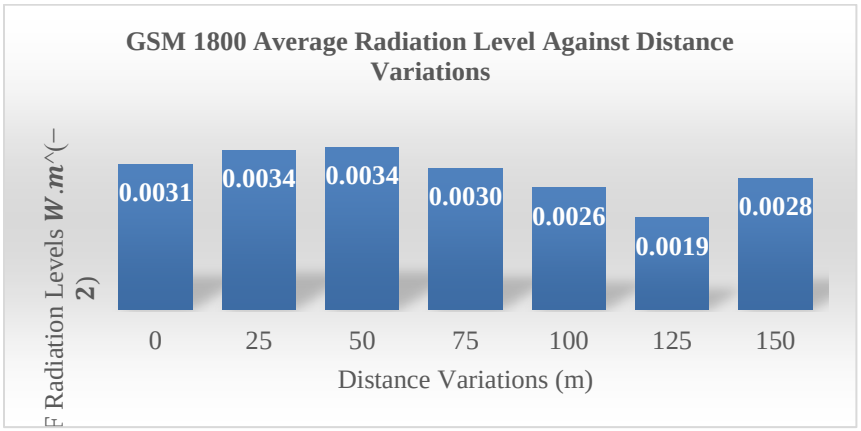


Fig. 6 The average radiation level of GSM 1800 to the variation of the measurement distance.

Figure 6 shows the average radiation measurement for each distance variation in the 1800 Mhz GSM frequency measurement. The measurements show that the transmit power of BTS has a relatively higher value at 0 meters to 75 meters because at the GSM 1800 Mhz frequency, the transmit power focuses more on the data capacity, which is distributed on the carrier frequency. So that the further the measurement distance, the value of BTS radiation decreases.

It can be seen that Figures 4 and Figure 6 have different measurement results caused by the characteristics of the GSM frequency itself and the measurement locations located in residential areas. The presence of obstacles such as buildings and other obstacles has affected the measurement results. However, in this study, measurements were made considering the time, the conditions of the measurement location, and the method of collecting data simultaneously [21], [22]. This step is done to determine the optimal measurement results and reduce the risk of measurement errors.

3-1- The Impact of BTS RF Radiation on Human Health

The negative impact of RF radiation on health is a significant issue developing in society. Health problems arise due to a lack of public knowledge about the RF emitted by BTS in residential areas supported by information from the public who do not know whether the RF from the BTS has a safe value for the human body. Ideally, the level of power density emitted by BTS is at 30m to 150m in densely populated residential areas and will experience a decrease in power at a distance of 200m [23]. According to a study conducted by Ayinmode B. and Idowu P. [24], there are reports that the effects of radiation can affect emotions, physical fatigue, skin problems, dizziness, nausea, headaches, and sleep disturbances. However, Kumar ruled out exposure to RF radiation as a cause of symptoms and health problems if the area has BTS transmit power which is still below the safe limit according to ICNIRP [25].

According to the ICNIRP General Public, radiation exposure standards for the range 1 - 10.W.m^{-2} are safe for the public. Based on the results of measurements that have been carried out at 20 BTS using distance variations, the radiation value in residential areas is still below the ICNIRP standard threshold. All BTS have transmitted power with different values but are still within safe limits for the surrounding population. Then the construction of the BTS tower has also been carried out with the Standard Operational Procedure (SOP) in terms of height, depth of grounding on the pole, and the materials used. So that the results of the BTS transmit power measurement in Sambas can be said to be safe for residents around the BTS and do not have severe problems in terms of health for humans and the security of the tower's location in residential areas.

4- Conclusions

Measurements on 20 BTS in Sambas have produced various RF radiation values. The measurement distance to the BTS significantly affects the results of the measurements. At the GSM frequency of 900 Mhz, the value of RF radiation has a higher value than the GSM frequency of 1800 Mhz because the frequency spectrum on GSM 900 Mhz is more substantial in transmitting power on frequency bands so that the transmission range is more expansive than GSM 1800 MHz. The measurement results have positively impacted residents in the BTS area who have been convinced that the radiation value is still within safe limits for humans. The establishment of a BTS

tower following the SOP is essential for the community to have a sense of security when living in the BTS tower area. Given the public's concerns about the health and safety conditions living in the BTS tower area, the radiation value measurement can be carried out periodically by both the provider and the relevant government agencies so as not to cause negative impacts on the community in the future.

Acknowledgments

We want to thank Sambas's Government for supporting this research activity. Then to LPPM Tanjungpura University, which has funded this research, and also to Universitas Tanjungpura for facilitating this research

References

[1] R. I. KOMINFO, "Peraturan Menteri Pekerjaan Umum Nomor 7 Tahun 2009 tentang Pedoman Pembangunan dan Penggunaan Bersama Menara Telekomunikasi." 2009.

[2] K. P. Republik Indonesia, "Peraturan Menteri Pekerjaan Umum Nomor 7 Tahun 2009 tentang Pedoman Pembangunan dan Penggunaan Bersama Menara Telekomunikasi." 2009.

[3] Health Protection Agency, *Health Effects from Radiofrequency Electromagnetic Fields*. 2012.

[4] P. Bernardi *et al.*, "Evaluation of human absorption in the near field of a BTS antenna," *IEEE MTT-S Int. Microw. Symp. Dig.*, vol. 3, pp. 1449–1452, 2004.

[5] T. A. A. Santana *et al.*, "Measurement campaign on the electromagnetic environment in the central region of the City of Mossoro," *SBMO/IEEE MTT-S Int. Microw. Optoelectron. Conf. IMOC 2017*, vol. 2017-Janua, no. 303, pp. 1–4, 2017.

[6] A. Linhares, M. A. B. Terada, and A. J. M. Soares, "Estimating the location of maximum exposure to electromagnetic fields associated with a radiocommunication station," *J. Microwaves, Optoelectron. Electromagn. Appl.*, vol. 12, no. 1, pp. 141–157, 2013.

[7] M. Riederer, "EMF exposure due to GSM base stations: measurements and limits," pp. 402-405 Vol.1, 2008.

[8] IARC, "IARC classifies radiofrequency electromagnetic fields as possibly carcinogenic to humans," *World Heal. Organ.*, vol. 2008, no. May, pp. 1–6, 2011.

[9] R. Matthes, "EMF Safety Guidelines \r- The ICNIRP View - ," ITU *Work. Hum. Expo. Electromagn. F.*, no. May, 2013.

[10] E. A. Dialogue and E. Fields, "on Risks From," *Environ. Heal.*, pp. i–66, 2002.

[11] M. J. Schoemaker *et al.*, "Mobile phone use and risk of acoustic neuroma: Results of the Interphone case-control study in five North European countries," *Br. J. Cancer*, vol. 93, no. 7, pp. 842–848, 2005.

[12] J. Wilén *et al.*, "Electromagnetic Field Exposure and Health among RF Plastic Sealer Operators," *Bioelectromagnetics*, vol. 25, no. 1, pp. 5–15, 2004.

[13] N. Kwan-hoong and D. Ph, "Radiation, Mobile Phones, Base Stations and Your Health."

[14] G. Ziegelberger *et al.*, *Guidelines for limiting exposure to* electromagnetic *fields (100 kHz to 300 GHz)*, vol. 118, no. 5. 2020.

[15] S. Sector and O. F. Itu, "Guide on electromagnetic fields and health," vol. 1, 2020.

[16] M. E. van D. P. R. Repcholi, "Base stations and wireless networks: Exposures and health consequences," *Int. Work. Base Station. Wirel. Networks Expo. Heal. Consequences*, pp. 1–168, 2005.

[17] G. J. Rubin, G. Hahn, B. S. Everitt, A. J. Cleare, and S. Wessely, "Are some people sensitive to mobile phone signals? Within participants double blind randomised provocation study," *Br. Med. J.*, vol. 332, no. 7546, pp. 886–889, 2006.

[18] K. K. Kesari, M. H. Siddiqui, R. Meena, H. N. Verma, and S. Kumar, "Cell phone radiation exposure on brain and associated biological systems," *Indian J. Exp. Biol.*, vol. 51, no. 3, p. 187—200, Mar. 2013.

[19] P. A. Valberg, E. van Deventer, and M. H. Repacholi, "Workgroup report: Base stations and wireless networks - Radiofrequency (RF) exposures and health consequences," *Environ. Health Perspect.*, vol. 115, no. 3, pp. 416–424, 2007.

[20] IEEE, IEEE *Standard for Safety Levels With Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz*, vol. 2005, no. April. 2006.

[21] A. Linhares, A. J. M. Soares, and M. A. B. Terada, "Determination of Measurement Points in Urban Environments for Assessment of Maximum Exposure to EMF Associated with a Base Station," *Int. J. Antennas Propag.*, vol. 2014, 2014.

[22] R. Vadlamudi and D. Sriram Kumar, "97 Dual Band, Dual Slant ±45° Polarized 2 × 2 MIMO (8T 8R) Antenna Array with Low Mutual Coupling for A-LTE(4G) Band 41/42/43(5G) BTS Application," *2020 Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2020*, vol. 43, pp. 97–101, 2020.

[23] U. Bergqvist, G. Friedrich, Y. Hamnerius, and L. Martens, "Mobile Telecommunication Base Stations-Exposure to Electromagnetic Fields, Report of a Short Term Mission within COST 244bis," *Cost*, p. 77, 2000.

[24] B. O. Ayinmode and I. P. Farai, "Risks Associated with Low Level Radiofrequency Exposure at Close Proximities to Mobile Phone Base Stations," *Pacific J. Sci. Technol.*, vol. 13, no. 2, pp. 330–335, 2012.

[25] S. To and P. By, "REPORT ON Secretary , DOT , Delhi Prepared By," no. December, pp. 1–50, 2010.

Fitri Imansyah received the B.S. and M.Eng. degree in Electrical Engineering from Universitas Tanjungpura, Pontianak, Indonesia. Currently he is a Head of Telecommunication Laboratory and a lecture in Universitas Tanjungpura, Pontianak, Indonesia. His research interests include telecommunication, computer Science and energy.

Leonardus Sandy Ade Putra is a lecture from Universitas Tanjungpura. He received his Bachelor Degree in electrical engineering from Sanata Dharma University, Yogyakarta, Indonesia. Then he continued his master degree in telecommunication engineering from Universitas Diponegoro, Semarang, Indonesia. His research is focused on telecommunication network, IoT, image and video processing, and artificial intelligence.

Eka Kusumawardhani is working as a lecture in Faculty of engineering, Universitas Tanjungpura. She received her bachelor degree in Telecommunication Engineering from Telkom University, Indonesia in 2015, and in 2016 She continued her study and received her master degree in electrical – telecommunication engineering from Telkom University, Indonesia. She is conducting research activities in the areas of wireless communication system, massive MIMO, and signal processing.

Analytical Model to Create Proxy Server Sessions in Multimedia Networks

Mehdi Khazaei*
Department of Computer Engineering, Kermanshah University of Technology, Iran
m.khazaei@kut.ac.ir

Received: 29/Aug/2021 Revised: 13/Nov/2021 Accepted: 07/Dec/2021

Abstract
One of the most popular and widely applied protocols on multimedia networks is the Session Initiation Protocol (SIP) to create, modify, and terminate the sessions. SIP is the platform of Next Generation Networks (NGN). In this way, SIP should be able to respond to the needs of such a largely-used network. One of the major problems in SIP networks is overload. This challenge creates a sharp drop in quality of service for NGN users. In this regard, many studies have been conducted on the effectiveness of this protocol, especially under overload. A new analytical model is developed that prioritizes the SIP message processing. An analytical approach is proposed based on the Mean Value Analysis (MVA) algorithm in queue theory. Considering some appropriate assumptions customizing MVA as to implement this proposed model and to cope with the limitations of the MVA is highly essential. The output of the analytical model is compared with the standard SIP model obtained from the simulator and the results confirm that prioritizing original messages would enhance the SIP performance at different load conditions. Prioritization of original messages is advantageous, and outperforms the normal SIP. Nevertheless, prioritizing the repeated messages not only has no advantage, but also its performance is less than the normal SIP.

Keywords: Modeling; Prioritize; Queue theory; Proxy server; MVA algorithm.

1- Introduction

Today's world requires designing an efficient and scalable system, since the lack of these criteria may cause the system not to be implemented, so that the system is quickly eliminated from the competition. As the performance improvement after system designing may fail or associate with high costs, it is better to consider this issue during the design phase. Accordingly, it is necessary to develop a model of the system before construction that can describe a general view of the system to improve the performance of the system lifetime. Nevertheless, in practice, it is impossible to generate a model with all details and limitations of the actual system, and if possible, it is highly complex and expensive. Consequently, many critical details are mainly not considered in the modeling process that can affect the performance via either low effect or ineffectiveness. The details level of the model depends on the aim of the model that should not be more complex than the defined target.

The modeling process can be performed as simulation or analytic. In simulation method, a computer program mimics the various states of the system. However, the analytical method includes several mathematical and computational equations calculating performance parameters based on offered load. It should be noted the analytical model involves fewer details than the simulation method because it is implemented using some mathematical equations. The reasons to utilize the analytical modeling instead of the simulation method are: 1) simulators may depend on either time or event, so that non-convergence or very long convergence occurs by increasing the simulation time or load, and 2) if the parameters are not valued correctly, the results of the simulation are unrealistic. However, the analytical model calculates the performance parameters in a short time, and the results do not depend on the specific settings of the parameters. Ensuring the accuracy of the results of the analytical model, the output of the model is compared to the output of the simulation, which is highly useful.

In the queuing theory, the Markov model is utilized to describe the sequence of possible events, in which the probability of each event occurring depends only on the state of the previous event. In fact, the Markov model is a memory-less random process, which is the basis of most quantitative analysis methods. This model is formed based on the state diagram that is a powerful tool to describe systems. In addition, it can be visually understood by less-experienced people and is employed in many problems. It should be noted that as the Markov model often determines the interaction between the various components of the system, it could be utilized as a predictor of the system behavior. Nevertheless, the main challenge of the Markov model is its extreme sensitivity to

* Corresponding Author

the sudden increase in the state space. In some cases, as the number of requests increase, the state space becomes so large that solving its equations does not converge. Therefore, some alternative solution methods like Mean Value Analysis (MVA) have been proposed.

The MVA is an algorithm developed to solve various Markov models, which does not require both explicit and simultaneous solutions of a large number of equations. The MVA utilizes a recessive relation to obtain each state instead of solving simultaneous equations to find probabilities.

The Session Initiation Protocol (SIP) is applied in many multimedia networks, and is introduced as the kernel protocol for next-generation networks. SIP is a text-based protocol where both request and response are expressed as HTTP, standardized by IETF. In this protocol, the control and data transmission sections are separated. SIP contributes to the end points' coordination, finding session participants, and agreeing on the session features.

Many studies are done on the performance of the SIP servers under specific conditions like overload. Overload is one of the most unsolved issues in SIP networks. In general, when the offered load to the server exceed its processing capacity, overloading becomes inevitable and the retransmission leads to queues server overflow. Hence sessions creating is delayed and overload is increased. Overload problem solutions are presented within simulation or analytical models, which are often simulation dependent.

One approach to overcome the local overload is to prioritize processing of the incoming requests in the proxy servers, where retransmitted messages are reduced and the corresponding transaction is terminated faster [1-4]. The prioritization method can be combined with intelligent methods to reduce the unnecessary in retransmissions by giving high priority to retransmitted requests [5]. As to VLB-CAC [6], the Virtual load-balanced call admission controller (VLB-CAC) is introduced for the cloud-hosted SIP servers to control overload in SIP network. The VLB-CAC determines the optimal call admission rates and signaling paths based on a heuristic mathematical model. The proposed scheme is implemented in smart applications on virtual infrastructure as testbed. The new concepts of computer networks, OpenSIP^{Partial} and VLB-CAC are applied in overcoming the overload issue. As to the OpenSIP^{Partial}, the SIP network is upgraded based on SDN and NFV technologies [7, 8].

On the other, SIP proxy server is composed of the dynamic subsystems, each subsystem interacts with each other to manage session constantly. In addition, overload indicate the proxy server as dynamic and complex system. Specially, with the development of SIP networks and their role in communication, the number of SIP servers will increase, and they will be geographically distributed. Hence, simulation is slow and time-consuming method to model SIP server but an analytical method is the right tool.

A priority-based approach is modeled through M/D/1/R according to RFC 357 in [9, 10]. In this method, three normal (L), overload (H) and removal (R) areas are defined, the input load is controlled according to the number of requests in each area. Also, two types of loads are defined, the first has an exclusive priority over the second. In the overload area, only first type load is accepted, but no load is accepted in the rejection area. In this method, a random process is defined for each region, and the rapid probability distribution function of that region is obtained. At the end, the mean return time to the normal state is calculated by the distribution function [11].

Applying the prioritization and the Markov model, the proxy servers between each UAC and UAS are modeled through a series of M/M/1 open queues. Each proxy server consists of a number of parallel and series queues, each of which is associated with a specific request or response. In this method, each message enters one of these queues with a defined probability. Then, based on the Markov model, an analytical model is presented that assesses the performance of the SIP network by input load, service rate, and different delays. Finally, it has been shown to improve the performance in the use of memory and CPU [12, 13].

Appling the queue theory model, a hierarchical model is presented in [14] to assess the performance of the SIP server in normal and overload states, where, Petri Net networks are used to check the network performance. In this model, two types of load management are considered to reduce retransmissions. The present paper aims to prioritize resource planning and reservation, so that the performance of the degradation system is at a lower level.

As statistical analytical models, the SIP server is modeled by the queuing system and it is analyzed by numerical methods. In these method, two limited queues for invite and non-invite requests are considered, which the non-invite queue is a priority. For the non-invite queue, a threshold (L) is considered, and when the number of requests in the queue exceeds (L), the incoming requests are rejected in the invite queue. The switch between the two queues for processing is as exponential. The polling systems, along with two general and gradual policies, have also been utilized to process requests. In any gradual processing, a number of requests per queue are processed, and the rest of the requests are waited, until it is their turn to be processed. Evaluation indicates that the gradual method has better performance [15, 16].

Most of the non-statistical prioritization models are implemented through the recursive equations. The fluid-flow model of one server, and then tandem servers are obtained for infinite queue length [17]. In this research, the initial conditions of the queue are obtained, so that the server is not overloaded by prioritizing the messages. The fluid-flow model of tandem servers with a finite buffer size can be calculated [18].

The trapezoid topology to tandem servers and several user agents, providing a Liapanov function that can assess the probability of rejection of calls through the proxy servers to achieve stability [19]. For this purpose, the fluid-flow model of the trapezoid topology of the SIP server is obtained by giving higher priority to non-invite messages.
As to LB-CAC a call acceptance controller with load balancing is proposed in [20], where a linear programmed model is involved in estimating the acceptance rate and signaling pathways, to assure the CPU and memory allocations and acceptance rates optimization.

A. Innovations

Many solutions have been offered to overcome overload, one being the priority-based processing. There exists no article regarding an analytical model based on MVA with low cost computations. A new analytical model based on the MVA algorithm that models the SIP server performance under message prioritization, together with considerable details and accuracy is developed in this articles. For this purpose, the MVA algorithm must be customized to meet these proposed models requirements, something not done yet in this field. The output of this analytical model is compared with the same of the simulator to determine its reliability.
The article is organized as follows: The literature is reviewed in Sec. 2; the proxy server is modeled in Sec. 3; model is adapted to the MVA in Sec. 4; the results are analyzed in Sec. 5 and article is concluded is Sec. 6.

2- Literature Review

Since the purpose is modeled SIP server based on the MVA, needed the related concepts to be briefly explained.

2-1- MVA

The MVA has been developed for a broad class of useful Markov models that do not require the explicit solution to a large number of simultaneous equations. Instead of solving a set of simultaneous linear equations to find the steady state probability of being in each system state, from which performance metrics can be derived, MVA is a simple recursion. MVA calculates the performance metrics directly for a given number of requests, knowing only the performance metrics when the number of requests is reduced by one. The recursion is intuitive, efficient, and elegant. Its impact on the field of analytical performance evaluation has been huge. The basic MVA algorithm is quite powerful. It is applicable across a wide set of performance models. It has been the focus of much research and several extensions have been developed. These include: multi-class, load dependent servers and open and closed classes. The optimization parameters for each request are solved through MVA, with the following limitations [21]:

- MVA does not provide the steady-state probability on its own.
- MVA does not provide information on the transient and non-stationary state of the system.
- MVA does not model the state-dependent behavior.
- MVA is not applied in solving the non-multiplicative problems, like the Gaussian, uniform and stationary distributions, priority queues and multi-class FIFS queues. The approximate MVA techniques is applied to solve non-multiplicative cases.

2-2- SIP

The SIP protocol architecture contains two logical entities, namely the user agent and the server agent. The User Agents (UA) is divided into two groups of User Agent Client (UAC) is an applicant and User Agent Server (UAS) is a respondent. The proxy servers seek the user location and perform a routing operation to deliver messages to the desired user. SIP transaction is a request and all the relevant responses are exchanged between two adjacent components. Proxy servers are configured state-full or state-less depending on network conditions. In the state-full, the proxy server saves the data of each transaction to perform some operations like requests retransmission. In the state-less proxy, no operation log is recorded. Retransmission mechanism is utilized to prevent the loss of packets when the SIP is run on a non-reliable transmission layer protocol like UDP [22, 23]. The call establishment model between UACs and UASs as state-full proxy server is illustrated in Fig. (1).
The UACs start sending messages by delivering an invite messages (λ_{IO}), so that when this message reaches the server, the 100-Trying response ($\mu_{100\text{-}Trying}$) as confirming is sent to UACs, and the invite is passed to the other side of the session (μ_{IO}). When the invite message reaches by UAS, 180-Ringing message (λ_{180}) and 200-OK (λ_{200}) are sent by answering the call. Finally, a session is made when ACK (λ_{ACK}) is sent for every pair of 180-Ringing/200-OK ($\mu_{180\text{-}200}$). Afterward, the data are exchanged without passing through the proxy server. In terms of confirmation and retransmission, SIP transactions are classified as invite or non-invite invites. The different variables and parameters applied in this manuscript are tabulated in the Table 1.
When the invite request is transmitted, a timer is responsible to control the retransmission set to the initial value. Upon expiration of this timer, the request will be retransmitted again, thus, doubling the timer. This process continues until a response is received or the timer is expired with a default value, consequently, an invite request may be repeated for about six times. The λ_{IR} and λ_{IRS} symbolize the repeated invites retransmitted by the UACs or proxy servers. If no ACK is received the UAS will retransmitted a 200-OK message, that is, a 200-OK request can be retransmitted for about eleven times [22].

The μ_R and μ_{IR} symbolize the processing rate of responses and repeated invites, respectively.

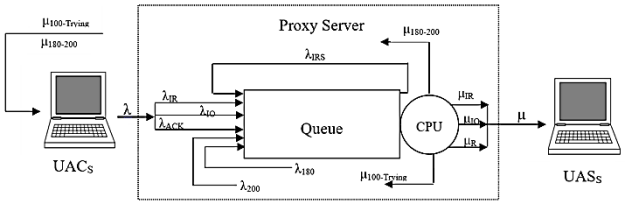


Fig. 1 The model of session create between the UAs

Table 1 The model parameters

Parameters	Description
λ	Offered load transmitted through UACs
λ_{IO}	Original invites rate transmitted through UACs
λ_{IR}	Repeated invites rate transmitted through UACs
λ_{IRS}	Repeated invites rate retransmitted through server
λ_{ACK}	ACK responses rate transmitted through UACs
λ_{200}	200-OK responses rate transmitted through UASs
λ_{180}	180 responses rate transmitted through UASs
μ_{IO}	Processing rate of original invites
μ_{IR}	Processing rate of repeated invites
μ_R	Processing rate of responses
μ	Processing rate of all requests
q_{IO}	Queue of original invites
T_j	Time to retransmit repeated messages
D_i	Sum of the times' means when a request receives service from the i^{th} source
V_i	Number of visits to the i^{th} source
S_i	Length of time that a request receives service from the i^{th} source
$D_{Cpu,k}^{Org}$	Demand service of original CPU applied through k^{th} category of invites (k= IO or IR)
$D_{Cpu,R}^{Org}$	Demand service of original CPU applied through responses
$D_{Cpu,R}$	Demand service of original CPU applied through responses
$D_{Cpu,R}^{shw\_y}$	Demand service of y^{th} Shadow CPU applied through responses
$D_{Cpu,k}^{shw\_y}$	Demand service of y^{th} Shadow CPU applied through k^{th} category of invites (k= IO or IR)
$D_{Cpu,k}$	Demand service of original CPU applied through k^{th} category of invites (k= IO or IR)
$U_{CPU,R}$	Original CPU utilization applied through responses
$U_{CPU,k}^{shw\_y}$	y^{th} Shadow CPU utilization applied through k^{th} category of invites (k= IO or IR)
$X_{0,R}$	Throughput of responses
$X_{0,k}$	Throughput of k^{th} category of invites (k= IO or IR)
$\alpha(j)$	Service rate coefficient
$\mu(j)$	Mean service rate of j requests
M	Coefficient determining the number of invites, repeated for the first time
H	Upper bound from which the number of processes remains constant
ω	Volume from which the service rate is constant
$P(j\|n)$	Probability of *j* occurrence when *n* requests are present throughout the network
N_{IO}	Number of original invites in the corresponding queue
N_{IR}	Number of repeated invites in the corresponding queue
N_R	Number of responses in the corresponding queue
$R_{IO}^{/}$	Average residence time for original invites
$R_{IR}^{/}$	Average residence time for repeated invites
$R_R^{/}$	Average residence time for responses
C	Maximum proxy server capacity

3- Proxy Server Modeling

The SIP proxy server model schemed in Fig. (1), is customized to implement prioritization, Fig. (2), where, the three separate queue of original invite requests, queue of repeated invite requests, and queue of responses are considered for transactions. In this modeling, the assumptions are as follows:

1- On proxy servers, UASs and UACs, the processing is compatible with the Poisson distribution [24, 25].
2- The transmitted original invites are compatible to the Poisson distribution through UACs [24, 25].
3- The UACs and UASs are high-powered systems to establish communication among subscribers by sending and receiving text-based SIP messages; message-processing delays are negligible and the responses are quickly returned.
4- In the proposed model, processing the responses queue is a priority, where, the relevant timers expire sooner, lead to a reduction of repeated requests and overload prevention. Once the queue is empty, it is time to process the queue with the next priority [5].
5- In this proposed model, the scheduling prioritization is exclusive. In request processing at low priority, if a request with a higher priority enters the relevant queue, the request processing with low priority stops, and the requests are processed with higher priority. The requests are processed in the lower priority queue when the higher priority queue is empty.
6- Because the three responses of 200-OK, 180-Rining, and ACK are similar in terms of CPU consumption and demand service, the messages within the queue of responses are considered as one class, consequently, the demand service is considered as D_R.
7- According to second assumption, the UASs quickly transmitted back responses for each invite. Timers are reset through processing returned responses, and no retransmission enters the queue through the server ($\lambda_{IRS}=0$). Therefore, original and repeated invites are only sent through the UACs.

4- Adapting the Model to the MVA Algorithm

The implementation of the proposed model through the MVA has some practical limitations, consequently, required compliance with this algorithm. The limitations consist of: 1) the adaptation of the input of the queues to the Poisson distribution, 2) prioritizations of queues and 3) processing the load dependence of the queues.

4-1- Adapting the Inputs to Poisson Distribution

As to the queue of responses, the third assumption states that, as soon as the original invite arrives at UAS, 200-OK

and 180-Ringing are returned immediately; consequently, the queue of responses is of exclusive priority, where the responses are transmitted to UACs and they return the ACK response without any delay, thus, no retransmission for the responses. The number of the 200-OK, 180-Ringing, and ACKs is equal to the number of original invites processed by the server obtained through Eq. (1).

$$\lambda_{ACK} = \lambda_{180} = \lambda_{200} = \mu_{IO} \quad (1)$$

The processing of the original invites in UAs is based on the Poisson distribution, accordingly, the entry rate of responses to the corresponding queue on the same distribution. As to second assumption, the input of the original invite queue must correspond to the Poisson distribution. In the repeated invite requests, it must be proved that the Poisson distribution is adapted. If success (p) is defined as not transmitting a repeated request to an original invite, then failure is defined as transmitting a repeated request to the same invite.

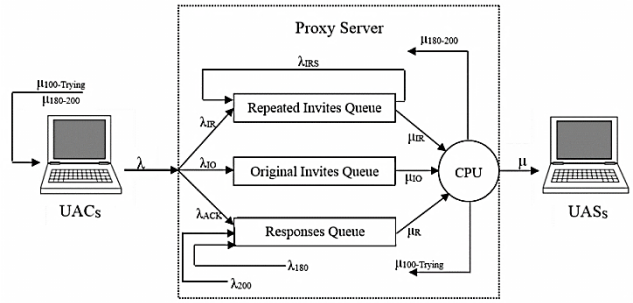


Fig. 2 Modeling the messages prioritization on the proxy server

P obtained through Eq. (2), where, p is percentage of invite requests, processed.

$$P = \frac{\mu_{IR} + \mu_{IO}}{\lambda_{IR} + \lambda_{IO}} \quad (2)$$

The number of repeated requests (x) from the total (n=λ_{IO}) original transmitted invites follow the binomial distribution obtained through Eq. (3).

$$\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{e^{-\mu}\mu^x}{x!} \;\; ; \;\; \mu = np \quad (3)$$

In binomial distribution, if the number of repetitions is very high and the probability of success is low, it can be estimated through the Poisson distribution. When the number of the invite requests transmitted by UACs is very high, it is not possible that the server will process an original invite successfully with no need for repetitions. Therefore, the number of the repeated invite requests follows the Poisson distribution.

Time is split in to discrete proportions to calculate the parameters, which allows discrete time modeling and makes it easier to understand and simulate the model. Smaller time proportions to obtaining more accurate results, while simulation time increases and vice versa. In the equations, the variables t and n are the time and time slices, respectively. The number of the repeated invite requests (λ_{IR}) is obtained through Eq. (4) [18].

$$\lambda_{IR}(n) = \sum_{j=1}^{6} \left[\lambda_{IO}(n - T_j) + q_{IO}(n - T_j) - \mu_{IO}(n - T_j, n)\right] \quad (4)$$

The length of the nth original invite queue (q_{IO}) is calculated through Eq. (5).

$$q_{IO}(n+1) = [q_{IO}(n) + \lambda_{IO}(n) - \mu_{IO}(n)] \quad (5)$$

The time to transmit repeated invites (T_j) is obtained through Eq. (6).

$$T_j = (2^j - 1)T_1 \qquad ; 1 \le j \le 6 \quad (6)$$

The first retransmitted is generated at T_1 = 0.5 sec and the last at maximum 64 * T_1 Sec [18].

4-2- Implementing the Priority of Queues

Because prioritization of the queues violates the multiplicative property of MVA, the approximate methods should be applied for this proposed model. The Stepwise Inclusion of Classes (SWIC) algorithm is run to implement the priority of queues, where the queue of responses never stops by the queue of invites unless it is empty. The invite queues see only the part of CPU that remains after processing the queue of responses. In this process, first, an estimation of CPU consumption by the queue of responses is made, which can be applied by modeling the queue of responses as the only queue of the system. Then, one of the invite queues is added, and modeling is made with two queues and an additional processor named the first shadow processor. Finally, the third queue with the second shadow processor is added to the model. The original invite queue takes precedence over the repeated invite, Fig. (3), while the opposite holds in Fig. (4).

The demand service (D_i) is the sum of the times when a request receives service from the ith source, Eq. (7).

$$D_i = V_i * S_i \quad (7)$$

D_i does not include the waiting time in the queue. S_i is the time of a request received service from the ith source for each visit, while V_i is the number of visits to the ith source to be made by the request.

The demand service of invite queues from the main CPU, and the demand service of response queue from the shadow CPUs are zero, respectively, Eqs. (8 and 9) [21].

$$D_{Cpu,k}^{Org} = 0 \qquad ; \;\; k = IO \; or \; IR \quad (8)$$

$$D_{Cpu,R}^{Shw_y} = 0 \qquad ; \;\; y = 1 \; or \; 2 \quad (9)$$

The variable k is defined as the IO for original invites and IR for repeated invite in all equations based on the intended priorities. Moreover, the variable y indicates the shadow processor number that is either one or two.

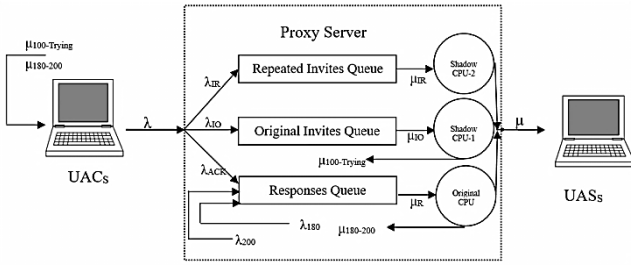


Fig. 3. Original invites queue has priority over the repeated invites queue

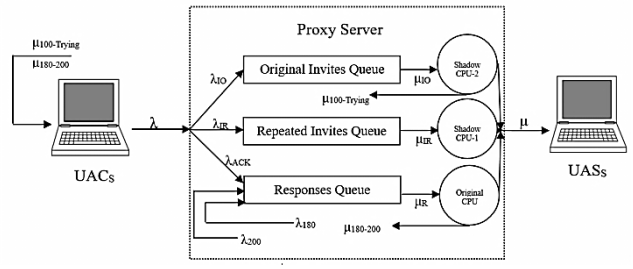


Fig. 4. Repeated invites queue has priority over the original invites queue

Because the queue of responses has the highest priority, the full CPU capacity is provided for responses, making the demand service for responses from the main CPU equal to the demand service expressed in Eq. (10).

$$D_{Cpu,R}^{org} = D_{Cpu,R} \tag{10}$$

The CPU utilization is obtained through Eq. (11).

$$U_{\mathrm{Cpu,R}} = \mathrm{D_{Cpu,R}} * \mathrm{X_{0,R}} \tag{11}$$

In the next step, the first shadow CPUs is added following the priority order. The demand service is then obtained through Eq. (12).

$$D_{Cpu,k}^{Shw\_y} = \frac{D_{Cpu,k}}{1 - \sum U_{high-priority}} \tag{12}$$

(e.g. $D_{Cpu,IO}^{Shw\_1}$ specifies the first shadow processor, allocated to the original invite queue, that is, the original invite requests take precedence over repeated invites, Fig. (3)). The CPU utilization of invite queues is obtained through Eq. (13) for the first and second shadow CPUs [21].

$$U_{Cpu,k}^{Shw\_y} = X_{0,k} * D_{Cpu,k}^{Shw\_y} \tag{13}$$

4-3- Request Processing

The request processing in each queue is performed in a load-independent or load-dependent manner. In the first, the mean service rate μ does not depend on the number of in-queue requests, while in the second, μ (n) is a function of load (n). In the model provided proxy server, the queue of responses and original invites is load independent, while in queues of repeated invite, processing is load-dependent. As to assumptions (3 and 4), the messages inside the response queue and the original invites queue are not repeated, consequently, they can be processed according to the capacity provided for the main and the corresponding shadow CPU. The number of original invites and responses in corresponding queue is obtained to modify Little's Law, Eqs. (14 and 15).

$$N_{IO} = \frac{U_{Cpu,IO}^{Shw\_y}}{1 - U_{Cpu,IO}^{Shw\_y}} \tag{14}$$

$$N_R = \frac{U_{Cpu,R}}{1 - U_{Cpu,R}} \tag{15}$$

The average residence time of the r[th] class is the total time spent by the same requests at the i[th] device in all visits. The average residence time of original invites and responses on corresponding CPUs is expressed as, Eqs. (16 and 17).

$$R_{IO}^{/} = \frac{D_{Cpu,IO}^{Shw\_y}}{1 - U_{Cpu,IO}^{Shw\_y}} \tag{16}$$

$$R_R^{/} = \frac{D_{Cpu,R}}{1 - U_{Cpu,R}} \tag{17}$$

In load-dependent processing, both the service rate and response time are a function of requests distribution, therefore, the MVA equations must be adjusted to fit this distribution. Therefore, instead of the queue length, a complete queue distribution is required. In the corresponding queue, the number of repeated invites is calculated through Eq. (18), provided that the stability condition ($\left[\frac{\mathrm{U}}{\alpha(\mathrm{w})}\right] < 1$) is met [21].

$$N_{IR} = p(0|\lambda_{IR}) * \left\{ \sum_{j=1}^{\omega} j \frac{U_{Cpu,IR}^{Shw\_y j}}{\beta(j)} + \frac{U_{Cpu,IR}^{Shw\_y \omega+1}}{\beta(\omega)*\alpha(\omega)} * \left[\frac{\frac{U_{Cpu,IR}^{Shw\_y}}{\alpha(\omega)} + (\omega+1)*\left(1 - \frac{U_{Cpu,IR}^{Shw\_y}}{\alpha(\omega)}\right)}{(1 - \frac{U_{Cpu,IR}^{Shw\_y}}{\alpha(\omega)})^2} \right] \right\} \tag{18}$$

$\mathrm{P(j|\lambda_{IR})}$ is the probability of j occurrence when λ_{IR} repeated invites enter to the queue, Eq. (19).

$$p(j|\lambda_{IR}) = \frac{D_{Cpu,IR}^{Shw\_y} * X_{0,IR}}{\alpha(j)} p(j-1|\lambda_{IR}), \;\; ; j = 1,2,\ldots.\lambda_{IR} \tag{19}$$

The probability $\mathrm{P}(0|\lambda_{\mathrm{IR}})$ is recursively obtained by resolving the $\mathrm{P}(j|\lambda_{\mathrm{IR}})$ and requiring all probabilities sum to one, Eq. (20).

$$p(0|\lambda_{IR}) = [\sum_{j=0}^{\omega} \frac{U_{Cpu,IR}^{Shw\_y j}}{\beta(j)} + \frac{[\alpha(\omega)]^{\omega}}{\beta(\omega)} * \frac{[\frac{U_{Cpu,IR}^{Shw\_y j}}{\alpha(\omega)}]^{\omega+1}}{1 - \frac{U_{Cpu,IR}^{Shw\_y j}}{\alpha(\omega)}}]^{-1} \tag{20}$$

α (j) is the coefficient of service rate, Eq. (21), directly related to the first repeated invites because they would terminate the state diagram of server transactions while processing the other repeated invites only waste the CPU.

$$\alpha(j) = \frac{\mu(j)}{\mu(1)} = \begin{cases} M * j & j \le H \\ H & j > H \end{cases} \tag{21}$$

M is a coefficient determining the number of the invites repeatedly delivered for the first time and is computed through Eq. (22).

$$M = \frac{\lambda_{IR}(j - T_1\} + q(j - T_1) - \mu_{IR}(j - T_1, j)}{\lambda_{IR}(j)} \quad (22)$$

However, service rate coefficient of repeated invites queue is limited to H, determining the upper bound where the throughput remains constant, Eq. (23).

$$H = X_{0,IR} = \frac{U_{Cpu,IR}^{Shw\_y}}{S_{IR}} = \frac{U_{Cpu,IR}^{Shw\_y}}{\frac{D_{Cpu,IR}^{Shw\_y}}{V_{IR}}} = \frac{V_{IR} * U_{Cpu,IR}^{Shw\_y}}{D_{Cpu,IR}^{Shw\_y}} \quad (23)$$

H is equal to shadow CPU throughput of repeated invites queue. W is the volume where the service rate is almost constant, that is, equal H. The β (j) is computed through Eq. (24) [21].

$$\beta(j) = \alpha(1) * \alpha(2) * \ldots * \alpha(j) \quad (24)$$

The average residence time for load-dependent processing of repeated invites is obtained through Eq. (25).

$$R_{IR}^{/} = \frac{N_{IR}}{\lambda_{IR}} \quad (25)$$

5- The Results Analysis

The C++ is applied to analyze the results. In this model, the server completes the C call per second (cps). The service time (S) for each session is equal to *1/C*. Each session consists of an original invite request and the three corresponding responses of 200-Ok, 180-Ringing and ACK, therefore, V is considered as 0.4 and 0.2 for the original invite and each one of the above responses. According to [26], the value of V for each repeated invite is more than the processing response and less than one original invite. Here, V is defined as 0.3 for each repeated invite. The volumes of Eqs. (7 and 10-13) parameters are tabulated in Table 2.

The priority column identifies the next priority after response queue processing. If the priority is in the original invites queue, the method is named IO and if the same is in the repeated invites queue, named IR.

The N original invites produced based on Poisson distribution by accept and reject method. Goodput, the delay of a session creation, the mean number of original and repeated invites, the server utilization and fairness are considered as evaluation parameters. A session is completed when the ACK message is processed by the server and transmitted to UAS; accordingly, the Goodput is equal to the rate at which ACK messages exit the server. A session is considered successful, when ACK of corresponding invite is received less than 10 seconds [24].

The results obtained through the analytical model are compared with the results of the standard model, Fig. 1. The purpose of this comparison is to explore the main effect of prioritizing the messages of a session on the evaluation parameters. The NS-2 is applied to obtain the standard model volume of parameters. In the simulator, the SIP is implemented in accordance with RFC 3261. In this model, the classes are considered without priority, the processing is FIFO load-independent, and the server is state-full with a capacity of 100 cps. The UDP is applied as the transmission layer protocol. Also, the equivalent of the analytical model is implemented in the simulator, and compared with the evaluation parameters to verify the model validity and reliability.

The following two phases are considered in evaluation of the parameters: 1) all queues are empty at the beginning of evaluation, because there will be no overload until the mean input requests reaches the C and 2) twenty original invites are placed in the corresponding queue, in a sense that the server can experience the overload before reaching the C and determine how to return to normal status. The average of these two phases are repeated to achieve the required 95% confidence interval. In the charts, the offered load is the number of the calls per second, normalized by the C.

Table 2 Values of demands service and CPU utilization

Input Parameters	Values	Priority
$D_{Cpu,R}$	$0.2C$	--
$D_{Cpu,IO}$	$0.4C$	--
$D_{Cpu,IR}$	$0.3C$	--
$U_{Cpu,IO}^{Shw\_1}$	$X_{0,IO} * D_{Cpu,IO}^{Shw\_1}$	IO
$U_{Cpu,IR}^{Shw\_1}$	$X_{0,IR} * D_{Cpu,IR}^{Shw\_1}$	IR
$U_{Cpu,IO}^{Shw\_2}$	$X_{0,IO} * D_{Cpu,IO}^{Shw\_2}$	IR
$U_{Cpu,IR}^{Shw\_2}$	$X_{0,IR} * D_{Cpu,IR}^{Shw\_2}$	IO
$D_{Cpu,IO}^{Shw\_1}$	$\frac{D_{Cpu,IO}}{1 - U_R}$	IO
$D_{Cpu,IR}^{Shw\_1}$	$\frac{D_{Cpu,IR}}{1 - U_R}$	IR
$D_{Cpu,IO}^{Shw\_2}$	$\frac{D_{Cpu,IO}}{1 - (U_R + U_{IR}^{Shw\_1})}$	IR
$D_{Cpu,IR}^{Shw\_2}$	$\frac{D_{Cpu,IR}}{1 - (U_R + U_{IO}^{Shw\_1})}$	IO

5-1- The Sessions Good put and Delay

The Goodput and delay of the mechanisms under study are shown in Figs. (5 and 6). The efficiency of all mechanisms is almost the same when the offered load is less than C, Fig. (5). In the IR method close to C, the number of repeated invites increase, that is, the original invites processing is delayed due to priority, causing more than 10 seconds delay and unsuccessful sessions. This would lead to starvation of original invites, because, before the processing of repeated invites load reaches C, the already low Goodput reaches zero. When the offered load exceeds C, the IO method Goodput remains almost constant due to the priority of original invites. The Goodput of the normal mechanism becomes zero after a while because of an increase in the number of repeated invites, thus, a delay in the process of original invites.

5-2- The Average Number of the Original and Repeated Invite Requests on the Server

In the corresponding queue, the average number of the original invite requests is shown in Fig. (7). In all three methods, the mean number of original invites is the same as it approaches *C*. In the IO mechanism, this count is increased with a gentle slope by exceeding the *C* due to the priority and an increase in the number of original invites. In normal mechanism, the slope of the requests increase is sharper than that of the IO method, because as the number of the original invite requests increase, the repeated invite requests enter the FIFO queue, and delay original invites processing. Among the three mechanisms, the IR method has the highest number of unprocessed original invites in the queue. Because the offered load exceeds the *C*, the number of repeated requests increases, that is, the original invites queue cannot be processed yet, and the chart growth is severe.

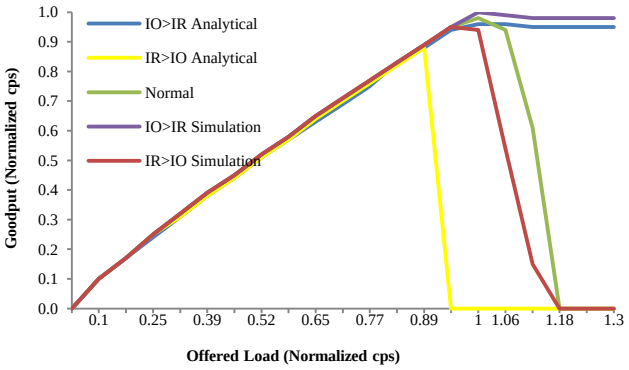


Fig. 5. Goodput of the mechanisms

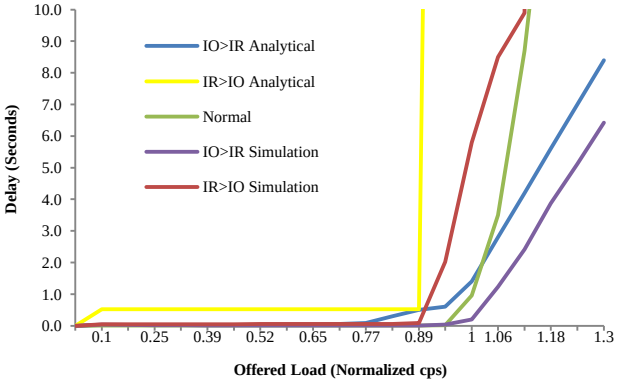


Fig. 6. Session delay of the mechanisms

In the corresponding queue, average number of repeated invites is shown in Fig. (8). The number of repeated requests in all three methods is almost the same before reaching the *C*. In IR, when the offered load exceeds *C*, the mean number of repeated requests increases. Because the repeated requests take precedence, and the original requests are processed with a long delay, the UACs constantly retransmit some repeated requests to the server, and this process continues. In the normal and IO methods, the mean number of repeated requests is lower, because the processing of the original requests took precedence, causing the UACs to transmit their corresponding repeated requests at a lower rate.

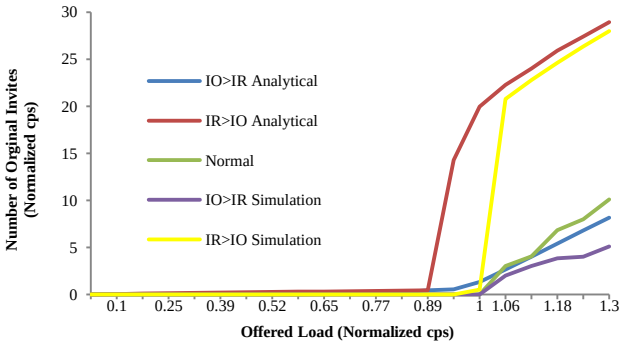


Fig. 7. Average number of original invites in the corresponding queue

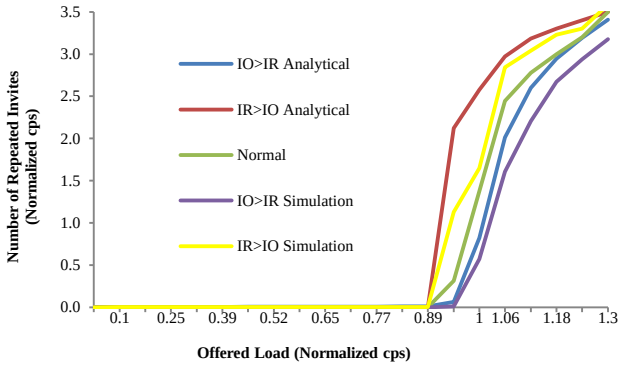


Fig. 8. Average number of repeated invites in the corresponding queue

5-3- Useful CPU Occupation

Processing of original invites is a useful task on the server CPU, while, the processing of repeated invites is a waste of capacity, respectively, Figs. (9 and 10). Among the subject mechanisms the one that makes better use of the capacity of server and reduce its waste outperforms others. The IO and IR mechanisms are such by increasing the offered load until it reaches the *C*, followed by the IR mechanism where the server capacity is wasted due to the processing more repeated requests than the IO mechanism. In normal mechanism, the CPU usage suddenly decreases because, likewise, the FIFO queue is filled by repeated invites after *C*.

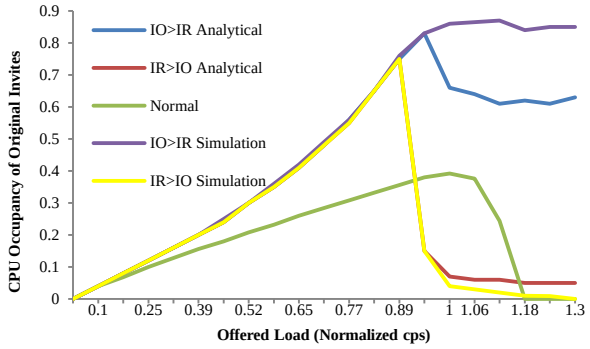


Fig. 9. Useful server occupation through the mechanisms

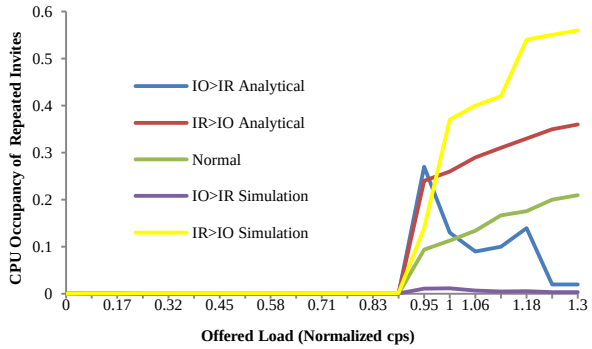


Fig. 10. Useless server occupation through the mechanisms

5-4- Fairness

The Jain's index, a statistical scattering measure is applied in measuring the average waiting time of original invites fairness, which is directly related to the number of calls made. This index with a volume within 0-1 is obtained through Eq. (26) [27].

$$J_w = \frac{(\sum_{i=1}^{m} w_i)^2}{k \sum_{i=1}^{m} w_i^2} \quad (26)$$

The closer the J_w rate to 1, the more the waiting time fairness and the opposite. W_i is the average waiting time of i^{th} offered load and K is the number of the applied offered load. The offered load range is within lower *C* to upper *C* range. The waiting time for M/M/1 queue is calculated through Eq. (27).

$$W_i = \frac{S * U_{CPU,k}^{Shw\_y}}{1 - U_{CPU,k}^{Shw\_y}} \quad (27)$$

The Jain's index of original invites waiting time with different priorities is shown in Fig. (11), when the priority is with IO queue, fairness is observed in the processing of original invites while priority to the IR queue acts contrariwise. Waiting time of original invites increases when IR queue is prioritized, and cause both the repeated invites retransmission and IO queue starvation. As observed in Fig. (5), there exists a direct relationship between the Goodput and the original invites waiting time fairness. In normal FIFO processing, waiting time of original invites increases for offered load up to C while less than C is acceptable.

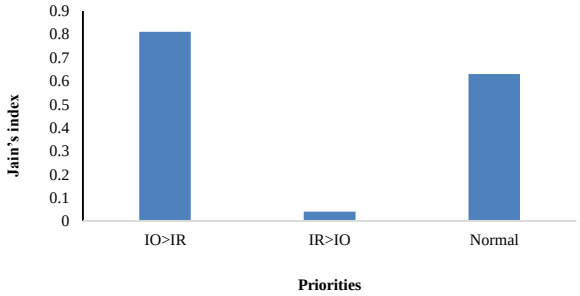


Fig. 11. Jain's index of waiting time of original invite

As to jitter, delay and package lost repeated processing of requests is not always in vain because this processing will not halt sessions. The fairness of repeated invites waiting time is shown in Fig. (12). In the process IR queue priority, is of high fairness and the same is observed in IO queue priority. In normal processing, fairness is almost the same for invites requests regardless of queue priority.

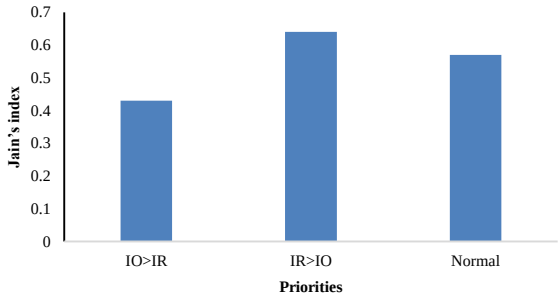


Fig. 12. Jain's index of waiting time of repeated invites

As can be deduced from the results, IO method is the most efficient, while, the IR method is more inefficient than the conventional SIP method and does not provide an appropriate response.

The results of the proposed analytical model are acceptable and close to the simulation results of Figs. (3 and 4). From comparison of the analytical and simulation results, it is deduced that the error percentage is acceptable and the results can be obtained with the time complexity of the O (n^2) where n is the number of original invites.

6- Conclusion and Future Works

Different messages are involved in creating the SIP session. If in the processing a message like invites is delayed, it would be retransmitted several times. In some cases, the number of these repeated messages are so big that no sessions are created or are delayed, where the result

is a considerable decline in the resource performance. The MVA algorithm of the queuing theory is customized to implement the proposed method. The question: at which priority should the messages involved in the session create be processed to cause less decline is assessed in this study. For this purpose, two scenarios are defined as follows: In the first scenario, priority is given to the original invites processing, because the rapid request processing and the prevention of retransmitting invites is evident. In the second scenario, priority is given to the repeated request processing, with the logic that this scenario would prevent retransmission. The results of these two scenarios are compared with the standard processing results of SIP messages, where priority is of no concern and are based on FIFS processing. As to the result, the mechanism for prioritizing original invite requests is more efficient than normal and IR mechanisms. The IR method performs poorly and does not provide acceptable answers. The simulation model confirms the analytical model reliability. As to the G/G/R queue and the corresponding algorithms can be replaced with M/M/1 discipline queue to model the proxy server. In multimedia networks, SIP tandem server is applied in establishing multimedia connection allowing the MVA algorithm to be run to model the consecutive proxy servers.

References

[1] D. Y. Yavas, I. Hokelek, and B. Gunsel, "Controlling SIP server overload with priority based request scheduling", International Conference on Computing, Networking and Communications (ICNC), 2015, pp. 510-514.

[2] I. Kuzminykh, "A combined LIFO-Priority algorithm for overload control of SIP server", International Conference on Modern Problems of Radio Engineering Telecommunications and Computer Science (TCSET), 2012, pp. 330-330.

[3] R. G. Garroppo, S. Giordano, S. Spagna, and S. Niccolini, "Queueing Strategies for Local Overload Control in SIP Server", In IEEE Global Telecommunications Conference, 2009, pp. 1-6.

[4] J. Lee and I. Joe, "An Overload Control Algorithm based on Priority Scheduling for SIP Proxy Server", Proceedings on the International Conference on Internet Computing, Athens 2012.

[5] K. K. Guduru and J. Usha, "Queuing strategies for self overload control in SIP servers", International Conference on Contemporary Computing and Informatics, 2014, pp. 1007-1011.

[6] A. Montazerolghaem, M. H. Yaghmaee, A. Leon-Garcia, M. Naghibzadeh, and F. Tashtarian, "A Load-Balanced Call Admission Controller for IMS Cloud Computing", IEEE Transactions on Network and Service Management, vol. 13, no. 4, pp. 806-822, 2016.

[7] A. Montazerolghaem, M. H. Y. Moghaddam, and A. Leon-Garcia, "OpenSIP: Toward Software-Defined SIP Networking", IEEE Transactions on Network and Service Management, vol. 15, no. 1, pp. 184-199, 2018.

[8] H. Kim and N. Feamster, "Improving network management with software defined networking", IEEE Communications Magazine, vol. 51, no. 2, pp. 114-119, 2013.

[9] P. Abaev, A. Pechinkin, and R. Razumchik, "Analysis of Queueing System with Constant Service Time for SIP Server Hop-by-Hop Overload Control", Berlin, Heidelberg, 2013, pp. 1-10: Springer Berlin Heidelberg.

[10] K. E. Samouylov, P. O. Abaev, Y. Gaidamaka, A. Pechinkin, and R. Razumchik, "Analytical Modelling And Simulation For Performance Evaluation Of SIP Server With Hysteretic Overload Control", In ECMS, 2014.

[11] Y. Gaidamaka, A. Pechinkin, R. Razumchik, K. E. Samouylov, and E. S. Sopin, "Analysis of an M|G|1|R queue with batch arrivals and two hysteretic overload control policies", International Journal of Applied Mathematics and Computer Science, vol. 24, pp. 519 - 534, 2014.

[12] V. K. Gurbani, L. J. Jagadeesan, and V. B. Mendiratta, "Characterizing session initiation protocol (SIP) network performance and reliability", presented at the Proceedings of the Second international conference on Service Availability, Berlin, Germany, 2005.

[13] S. V. Subramanian and R. Dutta, "A study of performance and scalability metrics of a SIP proxy server – a practical approach", Journal of Computer and System Sciences, vol. 77, no. 5, pp. 884-897, 2011/09/01/ 2011.

[14] G. Mishra, S. Dharmaraja, and S. Kar, "Performance analysis of SIP signaling network using hierarchical modeling", In Twentieth National Conference on Communications, 2014, pp. 1-5.

[15] S. Shorgin, K. Samouylov, Y. Gaidamaka, and S. Etezov, "Polling System with Threshold Control for Modeling of SIP Server under Overload", Cham, 2014, pp. 97-107: Springer International Publishing.

[16] Y. V. Gaidamaka, "Model with threshold control for analyzing a server with an SIP protocol in the overload mode", Automatic Control and Computer Sciences, vol. 47, no. 4, pp. 211-218, 2013/07/01 2013.

[17] Y. Hong, C. Huang, and J. Yan, "Modeling chaotic behavior of SIP retransmission mechanism", International Journal of Parallel, vol. Emergent and Distributed Systems, 02/01 2013.

[18] Y. Hong, C. Huang, and J. Yan, "Modeling and simulation of SIP tandem server with finite buffer", ACM Trans. Model. Comput. Simul., vol. 21, no. 2, p. Article 11, 2011.

[19] M. Jahanbakhsh, S. V. Azhari, and H. Nemati, "Lyapunov stability of SIP systems and its application to overload control", Computer Communications, vol. 103, pp. 1-17, 2017/05/01/ 2017.

[20] A. Montazerolghaem, M. H. Yaghmaee Moghaddam, and F. Tashtarian, "Overload Control in SIP Networks: A Heuristic Approach Based on Mathematical Optimization", 2015.

[21] V. A. F. A. Daniel A. Menascé, Lawrence W. Dowdy, Performance by Design: Computer Capacity Planning by Example. Prentice Hall PTR, 2004.

[22] M. Khazaei and N. Mozayani, "A dynamic distributed overload control mechanism in SIP networks with holonic multi-agent systems", Telecommunication Systems, vol. 63, 12/30 2015.

[23] M. Khazaei, "Occupancy Overload Control by Q-learning", Singapore, 2019, pp. 765-776: Springer Singapore.
[24] J. Wang, J. Liao, T. Li, J. Wang, J. Wang, and Q. Qi, "Probe-based end-to-end overload control for networks of SIP servers", Journal of Network and Computer Applications, vol. 41, pp. 114-125, 5// 2014.
[25] J. Liao, J. Wang, T. Li, J. Wang, J. Wang, and X. Zhu, "A distributed end-to-end overload control mechanism for networks of SIP servers", Comput. Netw., vol. 56, no. 12, pp. 2847-2868, 2012.
[26] C. Shen, H. Schulzrinne, and E. Nahum, "Session Initiation Protocol (SIP) Server Overload Control: Design and Evaluation", Berlin, Heidelberg, 2008, pp. 149-173: Springer Berlin Heidelberg.
[27] M. Khazaei and N. Mozayani, "Overload management with regard to fairness in session initiation protocol networks by holonic multiagent systems", International Journal of Network Management, vol. 27, no. 3, p. e1969, 2017.

Mehdi Khazaei received a B.Sc degree in computer Engineering (Computer Hardware) from Iran University of Science and Technology (Tehran, IRAN); M.Sc. and Ph.D degree in computer systems Architecture from Iran University of Science and Technology (Tehran, IRAN) in 2017. He is currently assistant professor in the School of Information Technology at Kermanshah University of Technology (Kermanshah, Iran).

SQP-based Power Allocation Strategy for Target Tracking in MIMO Radar Network with Widely Separated Antennas

Mohammad Akhondi Darzikolaei
Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology,Babol, Iran
M.akhondi@stu.nit.ac.ir
Mohammad Reza Karami Mollaei *
Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology,Babol, Iran
mkarami@nit.ac.ir
Maryam Najimi
Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran,Behshahr, Iran
m.najimi@mazust.ac.ir

Received: 22/Aug/2021 Revised: 06/Nov/2021 Accepted: 07/Dec/2021

Abstract
MIMO radar with widely separated antennas enhances detection and estimation resolution by utilizing the diversity of the propagation path. Each antenna of this type of radar can steer its beam independently towards any direction as an independent transmitter. However, the joint processing of signals for transmission and reception differs this radar from the multistatic radar. There are many resource optimization problems which improve the performance of MIMO radar. But power allocation is one of the most interesting resource optimization problems. The power allocation finds an optimum strategy to assign power to transmit antennas with the aim of minimizing the target tracking errors under specified transmit power constraints. In this study, the performance of power allocation for target tracking in MIMO radar with widely separated antennas is investigated. Therefore, a MIMO radar with distributed antennas is configured and a target motion model using the constant velocity (CV) method is modeled. Then Joint Cramer Rao bound (CRB) for target parameters (joint target position and velocity) estimation error is calculated. This is utilized as a power allocation problem objective function. Since the proposed power allocation problem is nonconvex. Therefore, a SQP-based power allocation algorithm is proposed to solve it. In simulation results, the performance of the proposed algorithm in various conditions such as a different number of antennas and antenna geometry configurations is examined. Results affirm the accuracy of the proposed algorithm.

Keywords: MIMO radar; Power allocation; SQP; Target tracking.

1- Introduction

The RADAR is a short form of Radio Detection And Ranging. Radar utilizes electromagnetic waves to detect, locate and measure the speed of reflected objects. It transmits the electromagnetic waves into space and receives the echo signals [1-2]. In recent years, radar network systems such as multi-static radars and multi-input multi-output (MIMO) radars have become an attractive and improtant problem [3]. Spatial diversity [4], waveform diversity [5], and multiplexing gain [6] over common monostatic radar [7] are some positive characteristics of networked radar systems. This structure of radars helps to raise tracking accuracy in multiple target tracking scenarios with radar, sonar, and video sensors [8]. A MIMO radar is a kind of radar structure that uses a combination of antennas as transmitter and receiver and each of them emits its own waveform apart from others [9]. Widely separated antennas and collocated antennas are two famous categories for MIMO radar. In collocated MIMO radar, the antennas are close to each other. The antennas in MIMO radar with widely separated antennas are far from each other. In other words, the transmit and receive antennas are located in a wide area. Therefore, the target is seen from different angles by antennas. In this type of radars, each receiver should receive all signals from all transmitters and then emit them to the central processor. This means that each receiver does not process or make a decision individually and sends its signals to the central processor to process all signals. This feature is the main difference between multi-static and MIMO radar with widely separated antennas. Power allocation is usually performed in radar networks to find the

* Corresponding Author

best strategy to assign power among various transmit antennas, aiming at minimizing the estimation error under specified transmit power constraints or its converse [10]. Power allocation is an important section of military operations in a hostile environment to obtain a low probability of interception [11]. Power allocation in radar networks is studied in last researches. For example, in [12] power allocation in widely separated multi-input multi-output (MIMO) radar for range-only target tracking is evaluated. This is performed by maximizing the Bayesian Fisher information matrix (B-FIM). B-FIM is derived for the predetermined signal model and then the problem is modeled as one cooperative game. [13] Shows the commercial application of power allocation in radar networks. In fact, it puts a cognitive radar network in an urban environment and this network tracks cars and vehicles. The power allocation problem for radar networks in a cooperative game-theoretic structure is considered in [14] to enhance the low probability of intercept (LPI) performance. In addition, by considering transmit power constraint and minimum signal to noise and interference ratio (SINR) for each radar, a cooperative Nash Bargaining power allocation game based on LPI is expressed. The radar network in [15] consists of unmodulated continuous wave (UCW) radars. This network utilizes a power allocation algorithm for Doppler-only target tracking. This algorithm minimizes the mean square error of target state estimation with a power budget constraint. [16] Investigates power allocation for radar networks to increase the performance of low probability of intercept. Two power allocation strategies are stated in this reference. One is for optimizing transmit power allocation with predetermined mutual information (MI) threshold and another is for finding optimal power allocation with minimum mean square error (MMSE) threshold. In [17], power allocation and target assignment in radar networks is mentioned to improve LPI performance. The target geometrical method is used to fuse information for measuring target localization by different radars. Authors in [18] propose joint antenna selection and power allocation for localization in distributed MIMO radar networks. The sensor management is performed by solving a constrained problem which is expressed to minimize the error in estimating target position, while it is constrained by transmitter number and power budget. [19] Talks about sensor selection in radar networks and for target tracking, the number of radars are selected. This problem is just for one target tracking. In this reference, sensor selection is performed by information theory. Joint transmitter and receiver selection in target tracking in distributed MIMO radar is described in [20]. Due to resource restriction in radars, it is necessary to select some radars in the MIMO radar network at each time and also keep the system performance in the best condition. So, lower bound PCRLB is u as an optimization criterion for an optimization problem in this literature. In [21], a joint power allocation and sensor selection algorithm for multi-target tracking in an LPI radar network with N monostatic radars is introduced. This algorithm can minimize the total transmitted power of a radar network based on predefined mutual information threshold between reflected signal and target impulse computed predictively with the estimation of the target state is needed for estimation of target parameters. [22] Describes the joint beam selection and power allocation strategies for multi-target tracking in collocated MIMO radar. Each radar works based on a multi-beam working mode, in which multiple simultaneous transmit beams are synthesized. This strategy applies an optimization technique to control the limited beam and power resource of each radar to obtain accurate target state estimation. Therefore, Bayesian Cramer Rao Lower Bound is extracted, normalized, and utilizes as the optimization criterion. To increase the system performance and resource utilization of widely separated MIMO radar, a joint resource allocation for velocity estimation problem in multi-target tracking is proposed in [23]. The authors chose one target as a key target and then examined their strategy by this target. They considered a Mean Square Estimation of velocity estimation of the key target as a minimization problem criterion. With limited resources and requirements for velocity estimation for targets, a joint optimization model with the selection of numbers of receivers and transmitters and allocation of transmit power and signal time is introduced. The Authors in [24] claims that since transmitters in MIMO radars with widely separated antennas emit waveforms with different powers and bandwidths, therefore these two parameters are limited. In this reference, they offer power allocation, bandwidth allocation, and joint power and bandwidth allocation problems. They compute Cramer Rao for target localization accuracy and utilize it as optimization criteria. In [25], a solution for joint beam and power scheduling in the netted Collocated MIMO radar systems for distributed multi-target tracking is suggested. An adaptive sensor scheduling integrated with power and bandwidth allocation is presented for centralized multiple target tracking in the netted collocated MIMO radar in [26].

By investigating the above references, we understand that power allocation for target tracking in MIMO radar with widely separated antennas is very essential and the performance of this problem should be improved. With our reviews, there are some challenges in the power allocation problem for target tracking in MIMO radar with widely separated antennas which are not investigated in other papers and we consider them in this paper.

1. In this paper, for the calculation of target tracking errors, joint target velocity and position estimation is used to improve tracking performance. Although someone works on this issue for MIMO radar, this is not used in power allocation problems in MIMO radar with widely separated antennas. For example, in [25] and [7], joint estimation is considered but certain power is determined for the transmit power and power allocation strategy is not performed. And also, in [25], it does not exactly specify its MIMO radar structure (distributed or collocated). In other researches about power allocation for target tracking in MIMO radar with widely separated antennas, joint estimation is not worked. For instance, in [12], range-only estimation for target tracking in MIMO radar with widely separated antennas is performed and it is not considered velocity. Whereas, in [23], velocity estimation is utilized to compute target tracking for power allocation problem in distributed MIMO radar. These are just some papers which we investigated and we concluded that joint estimation for target tracking in power allocation problem for MIMO radar with widely separated antennas is not used yet. Although it may be performed for collocated MIMO radar. Thus , joint target velocity and position estimation is the first novel idea for our power allocation strategy in MIMO radar with widely separated antennas.

2. Using random mathematical statistics for target RCS is another unique characteristic of this paper. Because in other researches, or a deterministic model for target RCS is considered [7, 24, 18, 22, 23], or if they supposed a random target RCS, they neglected it in their next calculations for simplicity [20, 21]. Therefore, none of the previous researches did contribute a power allocation strategy for target tracking in MIMO radar with widely separated antennas by considering random complex Gaussian target RCS. They usually neglected it or put a deterministic number instead of variance of Gaussian distribution. But in this paper, in all calculations, random complex Gaussian with random variance in different transmit-receive paths is considered. Note that assuming random RCS is necessary for MIMO radar with widely separated antennas. Because each antenna sees the target in a specified angle and target reflections in different transmit-receive paths are different with respect to each other.

The main contributions of this paper are as follows:

1. The system model for MIMO radar with widely separated antennas is introduced. Then CV model is considered for target motion. Considering a complex Gaussian random model for target RCS and using this feature in the next calculations, Cramer Rao bound for target parameters estimation error, is one of the prominent aspects of this paper.

2. Maximum likelihood (ML) estimation for unknown target parameters which were target position and velocity is calculated and then joint CRB for target position and velocity estimation is computed. Output of joint CRB is considered as an objective function for power allocation problem.

3. A power allocation strategy is formed. In fact, joint CRB for target position and velocity estimation function (target tracking errors) subject to some constraints such as limitation in total transmit power and transmit power of each transmit antenna is the power allocation problem of this paper. Our goal is to minimize tracking errors by using the mentioned constraints.

4. For solving the previous section problem, since it is nonconvex problem, SQP[1] based power allocation algorithm is proposed. This algorithm is formed based on the SQP algorithm and it can allot optimal power to each transmit antenna to satisfy the constraints in the problem. The rest of the paper is structured as follows:

The system model is mentioned in Section 2. Section3 exhibits the ML estimation calculations. Joint CRB for target parameters is computed in section4. Section 5 forms a power allocation problem to minimize target tracking error by considering total power limitation. And also, a proposed SQP-based algorithm is presented in this part to solve this problem. Simulation results are shown in section 6 and finally in part 7, concluding remarks are addressed.

2- System Model

Consider a MIMO radar with widely separated antennas with M transmitters and N receivers. Denote the location of mth transmitter in (x_m, y_m), where $m = 1,2, \ldots\ldots, M$ and the coordinates of nth receiver in (x_n, y_n), where $n = 1,2, \ldots\ldots, N$. Target is in initial location (x_q, y_q) with initial velocity of $(\dot{x}_q, \dot{y}_q)$. A set of low pass equivalent orthogonal waveforms,$s_m(t)$, is transmitted.

[1]sequential quadratic programming

($\int_{T_m}|s_m(t)|^2dt = 1$). period, effective bandwidth and transmit power of mth transmit waveform are shown as T_m,β_m, P_m. Target RCS corresponding to mnth path is modeled as a zero-mean complex Gaussian random variable $\xi_{mn}\sim\mathcal{CN}(0,\sigma_{mn}^2)$. Where σ_{mn}^2 is the variance of mnth path and it is known. Fig.1 shows the structure and location of antennas in MIMO radar with widely separated antennas with respect to the target.

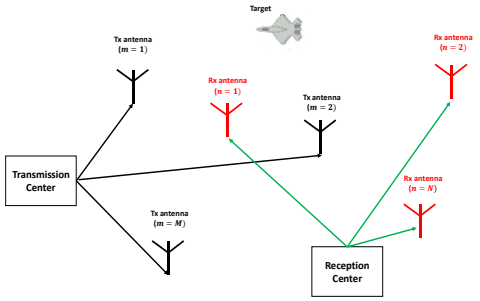


Fig. 1 configuration of a MIMO radar with widely separated antennas

We suppose the below assumption to simplify our problem.

1. w_{mn} (Noise of mnth path with a zero-mean complex Gaussian random variable and variance of σ_w^2) and ξ_{mn} in different paths are mutually independent.
2. Transmit waveforms are orthogonal.

$$\int_{-\infty}^{+\infty} s_m(t)\, s_{m'}^*(t)dt = \begin{cases}1 & if\ m = m' \\ 0 & if\ m \neq m'\end{cases} \quad (1)$$

This orthogonality also remains for time delays $\tau_m, \tau_{m'}$ and Doppler shifts $f_{dm'}$ and f_{dm} [26]:

$$\int_{-\infty}^{+\infty} s_m(t-\tau_m)\, s_{m'}^*(t-\tau_{m'})e^{j2\pi\left(f_{dm}-f_{dm'}\right)t}dt = \begin{cases}1 & if\ m = m' \\ 0 & if\ m \neq m'\end{cases} \quad (2)$$

3. Set $\sigma_w^2 = 1$ without loss of generality.
4. The antennas are adequately separated [27]. Therefore, each path provides an independent observation of the target and ξ_{mn} is independent for different m and n paths.

The time delay of (m,n)th channel in kth time slot is:

$$\tau_{mn,k} = \frac{d_{m,k}+d_{n,k}}{c} \quad (3)$$

Where,

$$d_{m,k} \triangleq \sqrt{\left(x_{q,k}-x_m\right)^2+\left(y_{q,k}-y_m\right)^2} \quad (4)$$

$$d_{n,k} \triangleq \sqrt{\left(x_{q,k}-x_n\right)^2+\left(y_{q,k}-y_n\right)^2}$$

In the above equations, c is light velocity. $d_{m,k}$ is distance from target and mth transmitter and $d_{n,k}$ is distance from a target and nth receiver.

With these assumptions, the received signal from mth transmit antenna at nth receive antenna at time k is given by:

$$r_{mn,k}(t) = \sqrt{\alpha_{mn,k}P_m}\ \xi_{mn,k} \quad (5)$$
$$s_m\left(t-\tau_{mn,k}\right)e^{j2\pi f_{mn,k}t} + w_{mn,k}(t)$$

In the above equation, $w_{mn,k}\sim\mathcal{CN}(0,\sigma_w^2)$ represents a zero-mean complex Gaussian noise with the variance of σ_w^2. Power variations due to path loss is shown as $\alpha_{mn,k} = \frac{1}{(4\pi)^3}\frac{1}{f_c^2}\frac{1}{d_{m,k}^2\, d_{n,k}^2}$. Where f_c is the carrier frequency.

Doppler frequency in mn path and time k is given by:

$$f_{mn,k} = \frac{\dot{x}_{q,k}(x_m-x_{q,k})+\dot{y}_{q,k}(y_m-y_{q,k})}{\lambda d_{m,k}} + \frac{\dot{x}_{q,k}(x_n-x_{q,k})+\dot{y}_{q,k}(y_n-y_{q,k})}{\lambda d_{n,k}} \quad (6)$$

λ is wavelength.

2-1- Target Dynamic Model

Target tracking in a MIMO Radar with widely separated antennas is the favorable problem of this paper. The target motion model is the constant velocity (CV). This model is as below[7]:

$$\boldsymbol{\theta}_{k+1} = \mathbf{F}_k\boldsymbol{\theta}_k + \boldsymbol{w}_k' \quad (7)$$

Target state vector is described as $\boldsymbol{\theta}_k = \left[x_{q,k}, \dot{x}_{q,k}, y_{q,k}, \dot{y}_{q,k}\right]^T$. This parameter is considered as unknown parameter in this paper and it should be estimated. The noise $\boldsymbol{w}_k'$ is a zero-mean Gaussian matrix as $\mathcal{N}(0,\boldsymbol{\Sigma}_k)$. Where $\mathbf{F}_k$ is state transition matrix and $\boldsymbol{\Sigma}_k$ is the covariance matrix [27]:

$$\mathbf{F}_k = \begin{bmatrix}1 & T & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & T\\ 0 & 0 & 0 & 1\end{bmatrix} \quad (8)$$

And

$$\boldsymbol{\Sigma}_k = \begin{bmatrix} \frac{T^2}{3} & \frac{T^2}{2} & 0 & 0 \\ \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^2}{3} & \frac{T^2}{2} \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix} l \tag{9}$$

Where T denotes sample intervals and l is the density of process noise.
In next section, we compute ML estimation of $\boldsymbol{\theta}_k$ to use it in calculation of joint CRB.

3- Maximum Likelihood Estimation

ML estimation of the unknown parameter ($\boldsymbol{\theta}_{\mathrm{k}}$) [28] can be achieved by testing of the likelihood ratio for two hypothesis pair, H_1 is corresponding to target existence hypothesis modeled in (5) and H_0 corresponding to noise only hypothesis:

$$\begin{aligned} &\boldsymbol{\Lambda}_{mn}\left(\boldsymbol{\theta}_k; r_{mn,k}(t)\right) \\ &= \frac{1}{\sigma_{mn}^2 P_m + 1} \\ &\times \exp\left\{\frac{\sigma_{mn}^2 P_m}{\sigma_{mn}^2 P_m + 1} \left| \int_{-\infty}^{+\infty} r_{mn,k}(t)\, s_m^*\big(t \right.\right. \\ &\left.\left. - \tau_{mn,k}\big) e^{-j2\pi f_{mn,k} t}\, dt \right|^2 \right\} \end{aligned} \tag{10}$$

Where $r_{mn,k}(t)$ denotes observation signal in nth receiver corresponding to mth transmitter.
Log-likelihood ratio based on (10) is as:

$$\begin{aligned} &\mathrm{L}_{mn}\big(\boldsymbol{\theta}_k; r_{mn}(t)\big) \\ &= Ln\, \Lambda_{mn}\left(\boldsymbol{\theta}_k; r_{mn,k}(t)\right) = \frac{\sigma_{mn}^2 P_m}{\sigma_{mn}^2 P_m + 1} \\ &\left| \int_{-\infty}^{+\infty} r_{mn,k}(t)\, s_m^*\big(t \right. \\ &\left. - \tau_{mn,k}\big) e^{-j2\pi f_{mn,k} t}\, dt \right|^2 + C_{mn} \end{aligned} \tag{11}$$

Where $C_{mn} = -\ln(\sigma_{mn}^2 P_m + 1)$ and it does not depend to $\boldsymbol{\theta}_k$.
According to problem assumptions, noise and reflection coefficients (RCS) are independent, the joint time delay and Doppler likelihood ratio term are obtained by:

$$\begin{aligned} &\Lambda_J\big(\boldsymbol{\theta}_k; \boldsymbol{r}(t)\big) \\ &= \prod_{m=1}^{M} \prod_{n=1}^{N} \Lambda_{mn}\left(\boldsymbol{\theta}_k; r_{mn,k}(t)\right) \end{aligned} \tag{12}$$

Where $\boldsymbol{r}(t)$ collects all observed signals from all antennas set and it is introduced as:

$$\boldsymbol{r}(t) = [r_{11}(t), r_{12}(t), \dots, r_{NM}(t)] \tag{13}$$

In the following part, joint Cramer Rao bound of target parameters is derived.

4- Joint Cramer Rao Bound

In this section, Bayesian Fisher Information and joint CRB for target position (x_q, y_q) and target velocity $(\dot{x}_q, \dot{y}_q)$ under our problem assumptions are provided.
Bayesian Fisher Information provides a lower bound on the target tracking Mean Square Error (MSE) for the estimation of target state. The Cramer Rao bound represents a lower bound on the accuracy of state estimates. If $\widehat{\boldsymbol{\theta}}_k$ be the estimation of target state, the Cramer-Rao bound is stated as [7]:

$$\mathbb{E}\left\{\left(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_{\mathrm{k}}\right)\left(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_{\mathrm{k}}\right)^T\right\} \geq \boldsymbol{J}_B^{-1}(\boldsymbol{\theta}_{\mathrm{k}}) \tag{14}$$

Where $\mathbb{E}$ is expectation operator and $\mathbf{J}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{k}})$ is the Bayesian Information Matrix (BIM).
According to [29], Bayesian Information Matrix for unknown parameter vector $\boldsymbol{\theta}_k$ is as:

$$\begin{aligned} \mathbf{J}_{\mathrm{B}}(\boldsymbol{\theta}_{\mathrm{k}}) = [\boldsymbol{\Sigma}_{\mathrm{k-1}} + \mathbf{F}\, \mathbf{J}_{\mathrm{B}}^{-1}(\boldsymbol{\theta}_{\mathrm{k-1}})\, \mathbf{F}^{\mathrm{T}}]^{-1} \\ + \mathbb{E}[\mathbf{J}_{\mathrm{D}}(\boldsymbol{\theta}_{\mathrm{k}})] \end{aligned} \tag{15}$$

Where $\mathbf{J}_{\mathrm{D}}$is Fisher Information Matrix (FIM).
The first step in CRB calculation is computing FIM which is a 4×4 matrix corresponding to the second derivatives of joint log-likelihood [26]:

$$\begin{aligned} &\mathbf{J}_{\mathrm{D}}(\boldsymbol{\theta}_{\mathrm{k}}) = \\ &\mathbb{E}_{\boldsymbol{r}(t);\boldsymbol{\theta}_k}\left\{\nabla_{\boldsymbol{\theta}_k} \ln \Lambda_J(\boldsymbol{r}(t); \boldsymbol{\theta}_k)[\nabla_{\boldsymbol{\theta}_k} \ln \Lambda_J(\boldsymbol{r}(t); \boldsymbol{\theta}_k)]^T\right\} \\ &= -\mathbb{E}_{\boldsymbol{r}(t);\boldsymbol{\theta}_k}\{\nabla_{\boldsymbol{\theta}_k}[\nabla_{\boldsymbol{\theta}_k} \ln \Lambda_J(\mathbf{r}(t); \boldsymbol{\theta}_k)]^T\} \end{aligned} \tag{16}$$

Where $\mathbb{E}$ denotes the expected value operator.
By considering (10) which is the function of $\tau_{mn,k}$ and $f_{mn,k}$, a new parameter as below is described:

$$\begin{aligned} &\boldsymbol{\vartheta}_k \\ &= \left[\tau_{11,k}, \tau_{12,k}, \dots, \tau_{NM,k}, f_{11,k}, f_{12,k}, \dots, f_{NM,k}\right]^T \end{aligned} \tag{17}$$

Based on Chain rule, a new FIM is introduced as:

$$\mathbf{J}_{\mathrm{D}}(\boldsymbol{\theta}_{\mathrm{k}}) = \left(\nabla_{\boldsymbol{\theta}_k}\boldsymbol{\vartheta}_k{}^{T}\right)\mathbf{J}_{\mathrm{D}}(\boldsymbol{\vartheta}_k)\left(\nabla_{\boldsymbol{\theta}_k}\boldsymbol{\vartheta}_k{}^{T}\right)^{T} \tag{18}$$

$$\nabla_{\boldsymbol{\theta}_k}\boldsymbol{\vartheta}_k{}^{T} = \begin{bmatrix} \frac{\partial\tau_{11,k}}{\partial x_{q,k}} & \frac{\partial\tau_{12,k}}{\partial x_{q,k}} & \dots & \frac{\partial\tau_{NM,k}}{\partial x_{q,k}} & \frac{\partial f_{11,k}}{\partial x_{q,k}} & \frac{\partial f_{12,k}}{\partial x_{q,k}} & \dots & \frac{\partial f_{NM,k}}{\partial x_{q,k}} \\ \frac{\partial\tau_{11,k}}{\partial y_{q,k}} & \frac{\partial\tau_{12,k}}{\partial y_{q,k}} & \dots & \frac{\partial\tau_{NM,k}}{\partial y_{q,k}} & \frac{\partial f_{11,k}}{\partial y_{q,k}} & \frac{\partial f_{12,k}}{\partial y_{q,k}} & \dots & \frac{\partial f_{NM,k}}{\partial y_{q,k}} \\ 0 & 0 & \dots & 0 & \frac{\partial f_{11,k}}{\partial \dot{x}_{q,k}} & \frac{\partial f_{12,k}}{\partial \dot{x}_{q,k}} & \dots & \frac{\partial f_{NM,k}}{\partial \dot{x}_{q,k}} \\ 0 & 0 & \dots & 0 & \frac{\partial f_{11,k}}{\partial \dot{y}_{q,k}} & \frac{\partial f_{12,k}}{\partial \dot{y}_{q,k}} & \dots & \frac{\partial f_{NM,k}}{\partial \dot{y}_{q,k}} \end{bmatrix}_{4\times 2NM} \tag{19}$$

Therefore, the above matrix parameters are defined as:

$$a_{mn,k} = \frac{\partial\tau_{mn,k}}{\partial x_{q,k}} = \frac{1}{c}\left(\frac{x_{q,k}-x_m}{d_{m,k}} + \frac{x_{q,k}-x_n}{d_{n,k}}\right) \tag{20}$$

$$b_{mn,k} = \frac{\partial\tau_{mn,k}}{\partial y_{q,k}} = \frac{1}{c}\left(\frac{y_{q,k}-y_m}{d_{m,k}} + \frac{y_{q,k}-y_n}{d_{n,k}}\right) \tag{21}$$

$$e_{mn,k} = \frac{\partial f_{mn,k}}{\partial x_{q,k}} = -\frac{\dot{x}_{q,k}}{\lambda}\left(\frac{1}{d_{m,k}}+\frac{1}{d_{n,k}}\right) + \frac{(x_m-x_{q,k})}{\lambda\,(d_{m,k})^3}\cdot\left[\dot{x}_{q,k}\,(x_m-x_{q,k}) + \dot{y}_{q,k}\,(y_m-y_{q,k})\right] + \frac{(x_n-x_{q,k})}{\lambda\,(d_{n,k})^3}\cdot\left[\dot{x}_{q,k}\,(x_n-x_{q,k}) + \dot{y}_{q,k}\,(y_n-y_{q,k})\right] \tag{22}$$

$$g_{mn,k} = \frac{\partial f_{mn,k}}{\partial y_{q,k}} = -\frac{\dot{y}_{q,k}}{\lambda}\left(\frac{1}{d_{m,k}}+\frac{1}{d_{n,k}}\right) + \frac{(y_m-y_{q,k})}{\lambda\,(d_{m,k})^3}\cdot\left[\dot{x}_{q,k}\,(x_m-x_{q,k}) + \dot{y}_{q,k}\,(y_m-y_{q,k})\right] + \frac{(y_n-y_{q,k})}{\lambda\,(d_{n,k})^3}\cdot\left[\dot{x}_{q,k}\,(x_n-x_{q,k}) + \dot{y}_{q,k}\,(y_n-y_{q,k})\right] \tag{23}$$

$$V_{mn,k} = \frac{\partial f_{mn,k}}{\partial \dot{x}_{q,k}} = \frac{x_m-x_{q,k}}{\lambda\ d_{m,k}} + \frac{x_n-x_{q,k}}{\lambda\ d_{n,k}} \tag{24}$$

$$q_{mn,k} = \frac{\partial f_{mn,k}}{\partial \dot{y}_{q,k}} = \frac{y_m-y_{q,k}}{\lambda\ d_{m,k}} + \frac{y_n-y_{q,k}}{\lambda\ d_{n,k}} \tag{25}$$

As shown in the above equations, the parameters of $a_{mn,k}$, $b_{mn,k}$,$e_{mn,k}$,$g_{mn,k}$,$V_{mn,k}$,$q_{mn,k}$ are determined by target position and velocity and c is light velocity.

If $\nabla_{\boldsymbol{\theta}_k}\boldsymbol{\vartheta}_k{}^{T}$ is divided into matrix blocks as:

First, by using (17) and $\boldsymbol{\theta}_{\mathrm{k}}$, we calculate $\left(\nabla_{\boldsymbol{\theta}_k}\boldsymbol{\vartheta}_k{}^{T}\right)$:

$$\nabla_{\boldsymbol{\theta}_k}\boldsymbol{\vartheta}_k{}^{T} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{0} & \mathbf{D}\end{bmatrix} \tag{26}$$

$\mathbf{A}$,$\mathbf{B}$,$\mathbf{D}$ are $2\times NM$ matrices and $\mathbf{0}$ is $2\times NM$ zero matrix.

$\mathbf{J}_{\mathrm{D}}(\boldsymbol{\vartheta}_k)$ is also obtained as:

$$\mathbf{J}_{\mathrm{D}}(\boldsymbol{\vartheta}_k) = -\mathbb{E}_{\mathbf{r}(t);\boldsymbol{\theta}_k}\{\nabla_{\boldsymbol{\vartheta}_k}[\nabla_{\boldsymbol{\vartheta}_k}\ln\Lambda_J(\mathbf{r}(t);\boldsymbol{\theta}_k)]^T\} \tag{27}$$

Where it is a $2NM\times 2NM$ matrix.

$\mathbf{J}_{\mathrm{D}}(\boldsymbol{\vartheta}_k)$ is represented as:

$$\mathbf{J}_{\mathrm{D}}(\boldsymbol{\vartheta}_k) = \begin{bmatrix}\mathbf{J}_{\mathrm{D}}^{UL} & \mathbf{J}_{\mathrm{D}}^{UR}\\ \mathbf{J}_{\mathrm{D}}^{LL} & \mathbf{J}_{\mathrm{D}}^{LR}\end{bmatrix} \tag{28}$$

Where $\mathbf{J}_{\mathrm{D}}^{UL}$, $\mathbf{J}_{\mathrm{D}}^{UR}$, $\mathbf{J}_{\mathrm{D}}^{LL}$,and $\mathbf{J}_{\mathrm{D}}^{LR}$ are $NM\times NM$ matrices. In fact, $\mathbf{J}_{\mathrm{D}}^{UL}$ includes all second-order derivatives terms with respect to $\tau_{mn,k}$ for all m and n in time slot k. $\mathbf{J}_{\mathrm{D}}^{UR}$, $\mathbf{J}_{\mathrm{D}}^{LL}$ contains second-order derivatives with respect to $\tau_{mn,k}$ and $f_{mn,k}$ for all m and n in time slot k. $\mathbf{J}_{\mathrm{D}}^{LR}$ contains second-order derivatives with respect to $f_{mn,k}$ for all m and n in time slot k.

Therefore,

$$\mathbf{J}_{\mathrm{D}}^{UL} = \boldsymbol{C}\odot(\boldsymbol{I}_N\otimes diag\,\{\epsilon_1,\epsilon_2,\dots,\epsilon_M\}) \tag{29}$$

Where $\boldsymbol{I}_N$ is an $N\times N$ identity matrix, $\odot$ is Hadmard product operator, and $\otimes$ is Kronecker product operator.

$$\boldsymbol{C} = 8\pi^2\ diag\,\{\frac{\sigma_{11}^4P_1^2}{\sigma_{11}^2P_1+1},\dots.,\frac{\sigma_{MN}^4P_M^2}{\sigma_{MN}^2P_M+1}\} \tag{30}$$

In fact, $\boldsymbol{C}$ is an $NM\times NM$ matrix.

$$\mathbf{J}_{\mathrm{D}}^{UR} = \mathbf{J}_{\mathrm{D}}^{LL} = \boldsymbol{C}\odot diag\,\{\gamma_{11,k},\gamma_{12,k},\dots,\gamma_{NM,k}\} \tag{31}$$

$$\mathbf{J}_{\mathrm{D}}^{LR} = \boldsymbol{C}\odot diag\,\{\eta_{11,k},\eta_{12,k},\dots,\eta_{NM,k}\} \tag{32}$$

Where ϵ_M, $\gamma_{mn,k}$ and $\eta_{mn,k}$ depend on receive waveform characteristics and they are defined as:

$$\epsilon_m = \int_{-\infty}^{+\infty} f^2 \, |S_m(f)|^2 \, df - \left| \int_{-\infty}^{+\infty} f \, |S_m(f)|^2 \, df \right|^2 \quad (33)$$

$$\gamma_{mn,k} = \frac{1}{2\pi} \mathcal{F}\{ \int_{-\infty}^{+\infty} t \, s_m^*(t - \tau_{mn,k}) \, \frac{\partial s_m(t - \tau_{mn,k})}{\partial \tau_{mn,k}} \, dt \} - \int_{-\infty}^{+\infty} f \, |S_m(f)|^2 \, df \, . \int_{-\infty}^{+\infty} t \, |s_m(t - \tau_{mn,k})|^2 \, dt \quad (34)$$

$$\eta_{mn,k} = \int_{-\infty}^{+\infty} t^2 \, |s_m(t - \tau_{mn,k})|^2 \, dt - \left| \int_{-\infty}^{+\infty} t \, |s_m(t - \tau_{mn,k})|^2 \, dt \right|^2 \quad (35)$$

Therefore, By considering the right-hand matrix as $H_{4\times4}$ which is the function of target position and velocity in each mn path and in different frames (although this matrix is not the function of transmit power), (36) can be reform as:

$$\mathbf{J}_D(\boldsymbol{\theta}_k) = \sum_{m=1}^{M} \sum_{n=1}^{N} 8\pi^2 \, \frac{\sigma_{mn}^4 P_m^2}{\sigma_{mn}^2 P_m + 1} \, . (H_{mn})_{4\times4} \quad (37)$$

In fact, $\mathbf{J}_D(\boldsymbol{\theta}_k)$ is a 4×4 matrix that is the function of transmit power.

Note that CRB for unknown parameter estimation is achieved by the diagonal elements of the inverse of FIM :

$$var\left(\hat{x}_{q,k}\right) \geq [\mathbf{J}_D^{-1}(\boldsymbol{\theta}_k)]_{1,1} \quad (38)$$

$$var\left(\hat{y}_{q,k}\right) \geq [\mathbf{J}_D^{-1}(\boldsymbol{\theta}_k)]_{2,2} \quad (39)$$

$$var\left(\hat{\dot{x}}_{q,k}\right) \geq [\mathbf{J}_D^{-1}(\boldsymbol{\theta}_k)]_{3,3} \quad (40)$$

$$var\left(\hat{\dot{y}}_{q,k}\right) \geq [\mathbf{J}_D^{-1}(\boldsymbol{\theta}_k)]_{4,4} \quad (41)$$

In above equations $var\left(\hat{x}_{q,k}\right)$,$var\left(\hat{y}_{q,k}\right)$, $var\left(\hat{\dot{x}}_{q,k}\right)$ and $var\left(\hat{\dot{y}}_{q,k}\right)$ respectively show the variance of target position in axis of x and y and also target velocity in axis of x and y in kth frame.In next section, our power allocation problem is formed.

In the above equations, $S_m(f)$ is a Fourier transform of $s_m(t)$. Therefore, (36) is obtained.

$$\mathbf{J}_D(\boldsymbol{\theta}_k) = \sum_{m=1}^{M} \sum_{n=1}^{N} 8\pi^2 \, \frac{\sigma_{mn}^4 P_m^2}{\sigma_{mn}^2 P_m + 1} \times \quad (36)$$

$$\begin{bmatrix} \epsilon_m a_{mn,k}^2 + 2\gamma_{mn,k} \, a_{mn,k} \, e_{mn,k} + \eta_{mn,k} e_{mn,k}^2 & (\epsilon_m a_{mn,k} + \gamma_{mn,k} e_{mn,k}) b_{mn,k} + (\gamma_{mn,k} a_{mn,k} + \eta_{mn,k} e_{mn,k}) g_{mn,k} & (\gamma_{mn,k} a_{mn,k} + \eta_{mn,k} e_{mn,k}) V_{mn,k} & (\gamma_{mn,k} a_{mn,k} + \eta_{mn,k} e_{mn,k}) q_{mn,k} \\ (\epsilon_m a_{mn,k} + \gamma_{mn,k} e_{mn,k}) b_{mn,k} + (\gamma_{mn,k} a_{mn,k} + \eta_{mn,k} e_{mn,k}) g_{mn,k} & \epsilon_m b_{mn,k}^2 + 2\gamma_{mn,k} \, b_{mn,k} \, g_{mn,k} + \eta_{mn,k} g_{mn,k}^2 & (\gamma_{mn,k} b_{mn,k} + \eta_{mn,k} g_{mn,k}) V_{mn,k} & (\gamma_{mn,k} b_{mn,k} + \eta_{mn,k} g_{mn,k}) q_{mn,k} \\ (\gamma_{mn,k} a_{mn,k} + \eta_{mn,k} e_{mn,k}) V_{mn,k} & (\gamma_{mn,k} b_{mn,k} + \eta_{mn,k} g_{mn,k}) V_{mn,k} & \eta_{mn,k} V_{mn,k}^2 & \eta_{mn,k} V_{mn,k} q_{mn,k} \\ (\gamma_{mn,k} a_{mn,k} + \eta_{mn,k} e_{mn,k}) q_{mn,k} & (\gamma_{mn,k} b_{mn,k} + \eta_{mn,k} g_{mn,k}) q_{mn,k} & \eta_{mn,k} V_{mn,k} q_{mn,k} & \eta_{mn,k} q_{mn,k}^2 \end{bmatrix}$$

5- Power Allocation Problem

In this section, we present power allocation strategy for a MIMO radar with widely separated antennas. In this problem and in the formulation of objective function, we consider a random RCS for target. This is rarely done in other papers and for simplicity, they consider deterministic reflection coefficients for target in MIMO radar with widely separated antennas. So this assumption make our main part of our objective function (equation (37)) complex.

5-1- Problem Formulation

In definition of our power allocation problem, we consider trace of CRB matrix as target tracking error and it is shown as:

$$\mathbb{F}(\boldsymbol{\theta}_k, \mathbf{P}) = trace(\boldsymbol{\Upsilon}_k \, \mathbf{J}_D^{-1}(\boldsymbol{\theta}_k)) \quad (42)$$

Where $\mathbf{P} = [P_1, \ldots, P_M]$ is a transmit power vector, $\boldsymbol{\Upsilon}_k$ is the normalization matrix and $\mathbb{F}$ is considered as the target tracking error.

Note that in equation (15), the first term (left-hand term) is approximately constant, therefore the second part, $\mathbf{J}_D(\boldsymbol{\theta}_k)$, is utilized for the calculation of the Another notable point is about diagonal elements of $\mathbf{J}_D^{-1}(\boldsymbol{\theta}_k)$, since the first and third diagonal elements are corresponding to the target position and second and fourth diagonal elements are corresponding to target velocity and their scales are different, so it is important to change them with no loss of generality to have the same scale. Fisher information matrix.Therefore we introduce $\boldsymbol{\Upsilon}_k$ as:

$$\boldsymbol{\Upsilon}_k = I_2 \otimes \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix} \quad (43)$$

Where I_2 is 2×2 identity matrix and $\otimes$ is the Kronecker product operator.

Therefore, the power allocation problem for target tracking in MIMO Radar with widely separated antennas by considering random radar cross-section (RCS) is defined as:

$$\min_{P_m} \mathbb{F}(\boldsymbol{\theta}_k, \mathbf{P}) \tag{44}$$

$$s.t. \sum_{m=1}^{M} P_m \le P_T \tag{45}$$

$$P_{min} \le P_m \le P_{max} \quad , m = 1,2, \dots, M \tag{46}$$

In the above problem, the first constraint represents that total transmit power is less than a predetermined value, P_T. In fact, it says that this value does not exceed P_T, because MIMO radar tries to use the least power for target tracking and it is not intercepted by other radars. The second constraint shows that each antenna also has power limitation.

5-2- Problem Solving

The problem (44) is nonlinear and it is nonconvex, So we propose the sequential quadratic programming (SQP) algorithm to solve it.

5-2-1 SQP Method

Sequential quadratic programming methods are famous to be affective for solving a series of related nonlinear optimization problems because of favorable hot and warm start properties—a solution for one problem is a good estimate of the solution of the next.

SQP applies Newton's method (or quasi-Newton methods) to directly solve the KKT conditions for the original problem. As a result, the accompanying subproblem turns out to be the minimization of a quadratic approximation to the Lagrangian function subject to a linear approximation to the constraints. Therefore, this kind of process is also known as a projected Lagrangian, or the Newton-Lagrange, approach. By its nature, this method produces both primal and dual (Lagrange multiplier) solutions.The procedure of SQP for constrained nonlinear problem as below [28]:

Minimize: $f(\boldsymbol{x})$ (47)

Subject to: $g_i(\boldsymbol{x}) \le 0, \quad i = 1, \dots, n_g$ (48)

$h_i(\boldsymbol{x})=0, \quad i = 1, \dots, n_h$ (49)

$x^l < \boldsymbol{x} < x^u$ (50)

Where, f is an objective function, g and h represent inequality and equality function and $f, g,$ and h are twice continuously differentiable. $\boldsymbol{x}$ shows the desirable variable matrix and it is limited by lower and upper bound x^l and x^u.

Given an iterate $(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)$, where $\boldsymbol{\lambda}^k$ and $\boldsymbol{\nu}^k \ge 0$ are the Lagrange multiplier estimates for the equality and inequality constraints, respectively, consider the following QP subproblem as a direct extension of QP$(\boldsymbol{x}^k, \boldsymbol{\lambda}^k)$:

$$\min_{\boldsymbol{d}^k} f(\boldsymbol{x}^k) + \nabla f(\boldsymbol{x}^k)^T \boldsymbol{d}^k + \frac{1}{2} {\boldsymbol{d}^k}^T \nabla^2_{xx} \mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\, \boldsymbol{d}^k \tag{51}$$

$$s.t.\ g_i(\boldsymbol{x}) + \nabla g_i(\boldsymbol{x}^k)^T \boldsymbol{d}^k = 0, \quad i = 1, \dots, n_g \tag{52}$$

$$h_i(\boldsymbol{x}) + \nabla h_i(\boldsymbol{x}^k)^T \boldsymbol{d}^k = 0, \quad i = 1, \dots, n_h \tag{53}$$

(51) is named as QP($\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k$) and $\mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k) = f(\boldsymbol{x}) + \boldsymbol{\nu}^T \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x})$.

Note that in above equations, for simplicity the constraint is not considered and also note that the KKT

Conditions for QP$(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)$ need that, in addition to primal feasibility, Lagrange multipliers $\boldsymbol{\lambda}^{k+1}$, $\boldsymbol{\nu}^{k+1}$ be found like that:

$$\nabla f(\boldsymbol{x}^k) + \nabla^2_{xx} \mathcal{L}(\boldsymbol{x}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k) \boldsymbol{d}^k + \nabla \boldsymbol{g}(\boldsymbol{x}^k)^T \boldsymbol{\nu}^{k+1} + \nabla \boldsymbol{h}(\boldsymbol{x}^k)^T \boldsymbol{\lambda}^{k+1} = 0 \tag{54}$$

$$[\boldsymbol{g}(\boldsymbol{x}^k) + \nabla \boldsymbol{g}(\boldsymbol{x}^k)^T \boldsymbol{d}^k]^T \boldsymbol{\nu}^{k+1} = 0 \tag{55}$$

With $\boldsymbol{\nu}^{k+1} \ge 0$ and $\boldsymbol{\lambda}^{k+1}$ unrestricted in sign. Clearly, if $\boldsymbol{d}^k = 0$, then $\boldsymbol{x}^k$ and $\boldsymbol{\lambda}^{k+1}$ and $\boldsymbol{\nu}^{k+1}$ yields KKT conditions to original problem. Otherwise, we should set $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{d}^k$ as before, increment k by 1and repeat the procedure.

5-2-2 Proposed SQP-based Power Allocation Algorithm

As mentioned before, Problem (44) subject to (45-46) is nonconvex and it should be solved by nonconvex suitable algorithms. In this section, we proposed SQP-based power allocation algorithm for our desired problem. In our scenario, problem (44) is our objective function and P_m is our desired variable . In fact, (44) is f in (47) and $\mathbf{P}$ in our problem is $\boldsymbol{x}$ in (47). By considering $\sum_{m=1}^{M} P_m - P_T \le 0$, we can claim that $\boldsymbol{g}(\boldsymbol{x})$ in (48) is equal to $\sum_{m=1}^{M} P_m - P_T$ in our problem.

There is not constraint (49) is our problem and $x^l = \boldsymbol{P_{min}}$ and $x^u = \boldsymbol{P_{max}}$.

Therefore, our proposed algorithm is based on previous section and it is proposed as:

Algorithm1: proposed SQP-based power allocation algorithm

Initialization step

Set the number of decision variables (Transmit Power ($\boldsymbol{P_m}$)), determine lower and upper bound of $\boldsymbol{P_m}(\boldsymbol{P_{min}}, \boldsymbol{P_{max}})$

Consider $\nabla^2_{xx} \mathcal{L}(\boldsymbol{x^k}, \boldsymbol{\lambda^k}, \boldsymbol{\nu^k})$ as $\boldsymbol{H^k}$ for simplicity.

Set $\boldsymbol{x} = \boldsymbol{P}$ (transmit power vector)

Choose an initial primal/dual point $(\boldsymbol{x^0}, \boldsymbol{\lambda^0}, \boldsymbol{\nu^0})$ with $\boldsymbol{\nu^0} \ge \mathbf{0}$

And a positive definite matrix $\boldsymbol{H^0}$.

Choose the relative tolerance ($\boldsymbol{\delta}$).
Set step size ($\boldsymbol{\alpha}$), $\boldsymbol{\alpha}$ is chosen to ensure the decrease in the objective function.
Let $\boldsymbol{k = 0}$ and go to main step.

Main step

1. Solve the quadratic subproblem QP$(\boldsymbol{x^k, \lambda^k, \nu^k})$, with $\boldsymbol{\nabla^2_{xx}\mathcal{L}(x^k, \lambda^k, \nu^k)}$ replaced by $\boldsymbol{H^k}$, to obtain a direction $\boldsymbol{d^k}$ along with a set of Lagrange multipliers $\boldsymbol{(\lambda^{k+1}, \nu^{k+1})}$.
2. if $\boldsymbol{d^k = 0}$, then $\boldsymbol{(x^k, \lambda^{k+1}, \nu^{k+1})}$ satisfies the KKT conditions for problem (44-46), stop.
3. if the previous section conditions are not satisfied, choose a new $\boldsymbol{\alpha_k}$, find $\boldsymbol{x^k = x^k + \alpha^k d^k}$. Update $\boldsymbol{H^k}$ to $\boldsymbol{H^{k+1}}$. Replace $\boldsymbol{k \leftarrow k + 1}$ and go to step 1.

Our proposed SQP-based power allocation algorithm flowchart is shown in Fig.2.

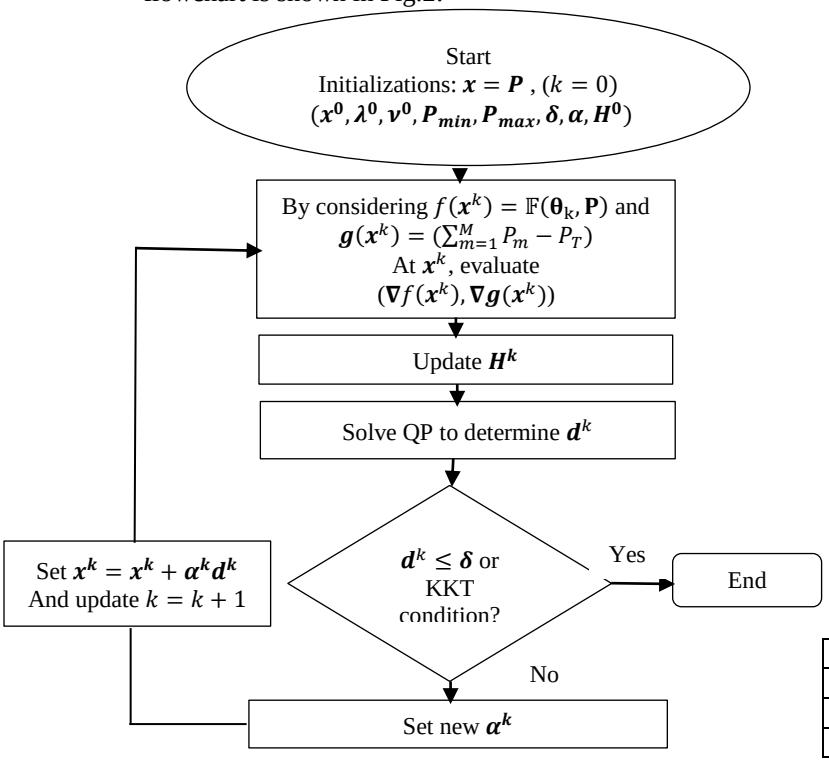


Fig. 2 the proposed SQP-based power allocation algorithm flowchart

6- Simulation Results

In this part, we consider some scenarios to investigate our proposed SQP-based power allocation algorithm. All simulations are executed in MATLAB. In this section, two various geometrical antenna deployment configurations are considered to illustrate the effect of antenna deployment on target tracking performance. These two configurations are shown in Fig. 3. Consider a MIMO radar with $M = 4$ and $N = 4$. Each antenna has 10 km distance from origin. Total power (P_T) is $10000\,watt$. Carrier frequency is $1GHz$. Target RCS variance, σ^2_{mn}, is random and it is different for any transmit-receive path (for simplicity, it is usually considered one in other papers). The power spectral density $l = 0.1$ and $T = 0.2s$. Target primary location and velocity in direction of x and y axis is equal to $[500m\ 1000m\ 50\frac{m}{s}\ \ 30\frac{m}{s}]$. Set $P_{min} = 0.02P_T$ (watt) and $P_{max} = 0.8P_T$.

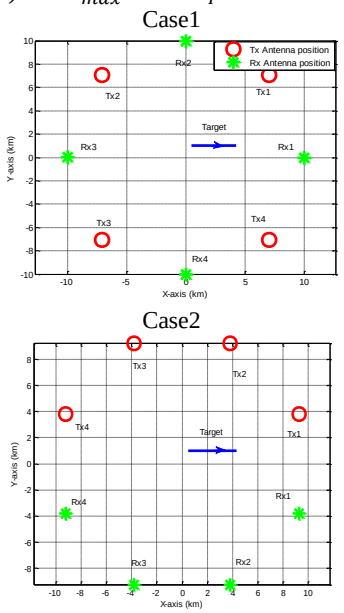


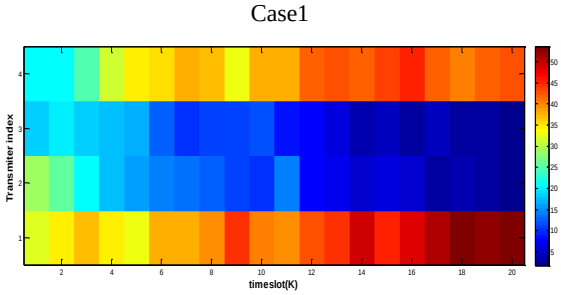
Fig. 3 MIMO radar antenna geometry deployment configurations $(M = 4, N = 4)$

The evaluation of proposed SQP-based power allocation algorithm is performed in these 2 Cases. Then this proposed algorithm is compared with uniform power allocation and Random power allocation and PSO-based algorithm. The values of the proposed SQP-based algorithm are shown in Table 1.

Table 1: proposed SQP-based parameters

parameter	value
δ	10^{-6}
α	10^{-6}
Maximum iteration	10000

Fig.4 shows the transmit power of each transmit antenna for two cases (Case1, Case2).Note that all results of this section are achieved by the proposed SQP-based power allocation algorithm)



Case2

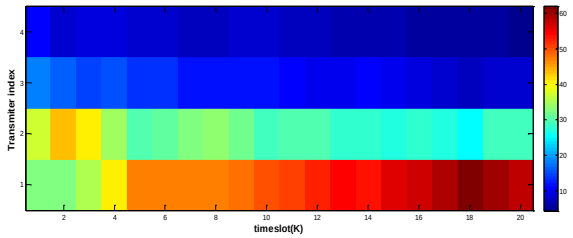


Fig. 4 Transmit power of each transmit antenna (%) for two cases with respect to timeslot

By attention to Fig.4, we can realize that by closing the target to the transmitter, more power is assigned to the transmitter. Therefore, in Case1, the transmitter 1 and 4, in Case2, the transmitter 1and 2, take more power to perform better target tracking performance.

Fig.5 exhibits that the proposed SQP-based power allocation algorithm has better performance than other strategies (Random and uniform power allocation) and PSO-based algorithm. it shows that the CRB of target tracking error is the least when the proposed SQP-based algorithm is used to perform the power allocation strategy. This priority is seen in all cases (Case1 (Fig.5 (a)), Case2 (Fig.5 (b))). Therefore, we can conclude that the proposed strategy has the best performance.

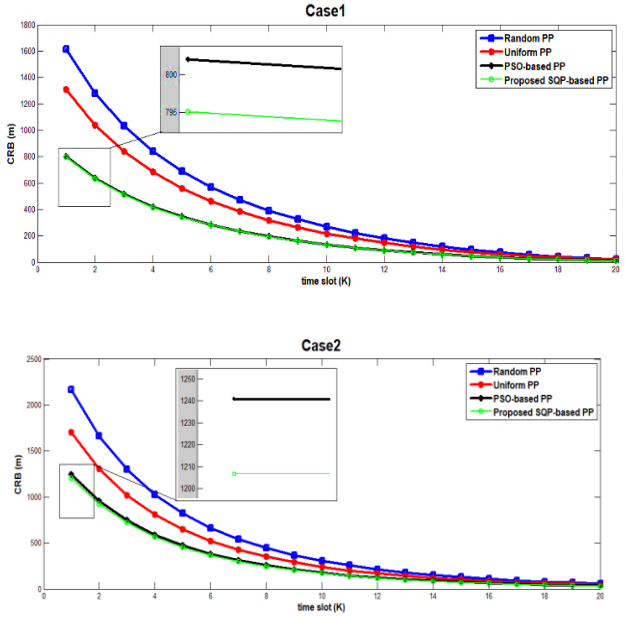


Fig.5. the comparison of the proposed SQP-based power allocation algorithm with uniform and random power allocation strategies in all two cases.

Table 2 illustrates the time duration which SQP-based and PSO-based algorithm need to run the simulations of our scenarios. Note that all the simulations are performed in system with intel(R) core(TM)i7-3612QM CPU @2.1GHz and 6 GB RAM.

Table 2: time cost of the proposed SQP-based and PSO-based algorithm

Cases	**SQP-based algorithm (s)**	**PSO-based algorithm (s)**	**Maximum iteration**
Case1	45.904317	505.183312	1000
Case2	45.512462	504.914239	1000

The above table shows that the proposed algorithm has more efficiency in real time scenarios. It is very crucial in target tracking scenarios because target is moving quickly in environment and the radar should track it simultaneously.

To show the impact of the number of antennas on MIMO radar target tracking performance, we consider another scenario, Case3, with $M = 6, N = 6$ for geometry deployment like Case1. Fig.6 demonstrates that increasing the antennas improves the performance of target tracking.

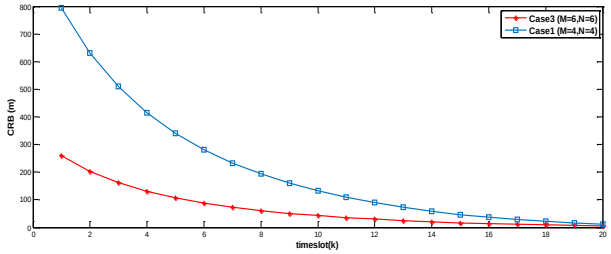


Fig.6. impact of the number of antennas on MIMO radar target tracking performance (with SQP-based algorithm)

To prove the use of the proposed tracking procedures, Fig.7 shows the target tracking RMSE and CRB in for Cases. The RMSE is computed as:

$$\mathrm{RMSE_k} = \sqrt{\frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} trace(\boldsymbol{\Upsilon}_k(\boldsymbol{\theta}_k - \widehat{\boldsymbol{\theta}}_k^j)(\boldsymbol{\theta}_k - \widehat{\boldsymbol{\theta}}_k^j)^T \boldsymbol{\Upsilon}_k^T)} \quad (56)$$

Where N_{MC} is the number of Monte Carlo trials and $\widehat{\boldsymbol{\theta}}_k^j$ denotes the state estimate of the target in the jth trial.

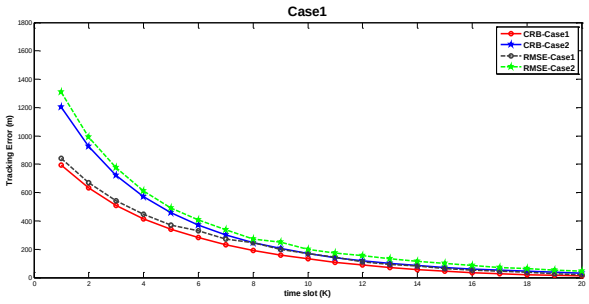


Fig.7. the evaluation of target tracking errors in two cases with proposed SQP-based target tracking algorithm

The proximity of CRB and RMSE results in Fig.7, exhibits that the proposed target tracking procedure is very near to actual conditions.

We consider random RCS model for target. Therefore we compare deterministic RCS and random RCS in performance of power allocation in target tracking in MIMO radar in Fig8. Fig 8 shows that by considering random RCS for power allocation strategy in MIMO radar, the target tracking error is decreased. This experiment is performed for Case1.

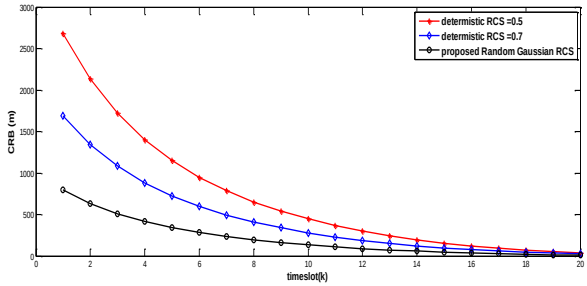


Fig.8. impact of deterministic and proposed random Gaussian RCS model on target tracking error

7- Conclusions

In this paper, a MIMO radar with widely separated antennas is considered to investigate power allocation strategy performance. A zero-mean random complex Gaussian with random variance in different transmit-receive paths is considered for target RCS. The simulation results demonstrate that this assumption enhances the MIMO radar with widely separated antennas performance. Using joint estimation of target position and velocity tracking error in power allocation problems for target tracking in MIMO radar with widely separated antennas is the another novel idea which is used in this paper.

In simulations, two different antenna geometry deployments with the different number of antennas (Case1 and 2) is considered to assess the proposed structure for power allocation problem. This paper demonstrates the power allocation problem to minimize the target tracking errors subject to the limitation in total transmit power and power of each transmit antenna using MIMO radar with widely separated antennas. Due to the nonconvexity of this problem, the SQP-based power allocation algorithm is proposed. The simulation experiments are performed in various conditions and simulation results exhibits the accuracy of the proposed algorithm.

Multi target tracking problem will be our future challenge to improve our work. In additions, we will add target RCS estimation to target state estimation vector to enhance target tracking performance in MIMO radar with widely separated antennas.

References

[1] M. A. Darzikolaei, A. Ebrahimzade, and E. Gholami, “Classification of radar clutters with artificial neural network,” in 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI), 2015, pp. 577–581.

[2] M. A. Darzikolaei, A. Ebrahimzade, and E. Gholami, “The Separation of Radar Clutters using Multi-Layer Perceptron,” Information Systems & Telecommunication, p. 41, 2017.

[3] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, “MIMO radar: An idea whose time has come,” in Proceedings of the 2004 IEEE Radar Conference (IEEE Cat. No. 04CH37509), 2004, pp. 71–78.

[4] E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik, and R. A. Valenzuela, “Spatial diversity in radars—Models and detection performance,” IEEE Transactions on signal processing, vol. 54, no. 3, pp. 823–838, 2006.

[5] J. Li and P. Stoica, “MIMO radar with colocated antennas,” IEEE Signal Processing Magazine, vol. 24, no. 5, pp. 106–114, 2007

[6] A. M. Haimovich, R. S. Blum, and L. J. Cimini, “MIMO radar with widely separated antennas,” IEEE Signal Processing Magazine, vol. 25, no. 1, pp. 116–129, 2007.

[7] M. Hai, “MIMO radar with widely separated antennas technology,” IEE Signal Magazine, vol. 26, no. 2, pp. 98–106, 2009.

[8] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation with applications to tracking and navigation: theory algorithms and software. John Wiley & Sons, 2004.

[9] A. Pakdaman and H. Bakhshi, “Separable transmit beampattern design for MIMO radars with planar colocated antennas,” AEU-International Journal of Electronics and Communications, vol. 89, pp. 153–159, 2018.

[10] M. Xie, W. Yi, T. Kirubarajan, and L. Kong, “Joint node selection and power allocation strategy for multitarget tracking in decentralized radar networks,” IEEE Transactions on Signal Processing, vol. 66, no. 3, pp. 729–743, 2017.

[11] H. Godrich, A. P. Petropulu, and H. V. Poor, “Power allocation strategies for target localization in distributed multiple-radar architectures,” IEEE Transactions on Signal Processing, vol. 59, no. 7, pp. 3226–3240, 2011.

[12] H. Chen, S. Ta, and B. Sun, “Cooperative game approach to power allocation for target tracking in distributed MIMO radar sensor networks,” IEEE Sensors Journal, vol. 15, no. 10, pp. 5423–5432, 2015.

[13] P. Chavali and A. Nehorai, “Scheduling and power allocation in a cognitive radar network for multiple-target tracking,” IEEE Transactions on Signal Processing, vol. 60, no. 2, pp. 715–729, 2011.

[14] C. Shi, S. Salous, F. Wang, and J. Zhou, “Power allocation for target detection in radar networks based on low probability of intercept: A cooperative game theoretical strategy,” Radio Science, vol. 52, no. 8, pp. 1030–1045, 2017.

[15] J. Yan, H. Liu, B. Jiu, and Z. Bao, "Power allocation algorithm for target tracking in unmodulated continuous wave radar network," IEEE sensors journal, vol. 15, no. 2, pp. 1098–1108, 2014.
[16] L. Wang, L. Wang, Y. Zeng, and M. Wang, "Jamming power allocation strategy for MIMO radar based on MMSE and mutual information," IET Radar, Sonar & Navigation, vol. 11, no. 7, pp. 1081–1089, 2017.
[17] S. M. H. Andargoli and J. Malekzadeh, "LPI radar network optimization based on geometrical measurement fusion," Optimization and Engineering, vol. 20, no. 1, pp. 119–150, 2019.
[18] B. Ma, H. Chen, B. Sun, and H. Xiao, "A joint scheme of antenna selection and power allocation for localization in MIMO radar sensor networks," IEEE communications letters, vol. 18, no. 12, pp. 2225–2228, 2014.
[19] X. Li, W. Yi, G. Cui, L. Kong, and X. Yang, "Radar selection for single-target tracking in radar networks," in 2015 IEEE Radar Conference (RadarCon), 2015, pp. 0545–0550.
[20] Y. Lu, Z. He, X. Zhang, and S. Liu, "Transmit and receive sensors joint selection for MIMO radar tracking based on PCRLB," in 2016 IEEE 13th International Conference on Signal Processing (ICSP), 2016, pp. 1551–1555.
[21] J. She, F. Wang, and J. Zhou, "A novel sensor selection and power allocation algorithm for multiple-target tracking in an LPI radar network," Sensors, vol. 16, no. 12, p. 2193, 2016.
[22] J. Yan, H. Liu, W. Pu, S. Zhou, Z. Liu, and Z. Bao, "Joint beam selection and power allocation for multiple target tracking in netted colocated MIMO radar system," IEEE Transactions on Signal Processing, vol. 64, no. 24, pp. 6417–6427, 2016.
[23] X. Song, N. Zheng, and T. Bai, "Resource allocation schemes for multiple targets tracking in distributed MIMO radar systems," International Journal of Antennas and Propagation, vol. 2017, 2017.
[24] N. Garcia, A. M. Haimovich, M. Coulon, and M. Lops, "Resource allocation in MIMO radar with multiple targets for non-coherent localization," IEEE Transactions on Signal Processing, vol. 62, no. 10, pp. 2656–2666, 2014.
[25] Yi, Wei, Ye Yuan, Reza Hoseinnezhad, and Lingjiang Kong. "Resource scheduling for distributed multi-target tracking in netted colocated MIMO radar systems." IEEE Transactions on Signal Processing 68 (2020): 1602-1617.
[26] Li, Zhengjie, Junwei Xie, Haowei Zhang, Houhong Xiang, and Zhaojian Zhang. "Adaptive sensor scheduling and resource allocation in netted collocated MIMO radar system for multi-target tracking." IEEE Access 8 (2020): 109976-109988.
[27] Q. He, R. S. Blum, and A. M. Haimovich, "Noncoherent MIMO radar for location and velocity estimation: More antennas means better performance," IEEE Transactions on Signal Processing, vol. 58, no. 7, pp. 3661–3680, 2010.
[28] V. Trees and L. Harry, Detection, Estimation, and Modulation Theory-Part l-Detection, Estimation, and Linear Modulation Theory. John Wiley & Sons New York, 2001.
[29] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localization accuracy gain in MIMO radar-based systems," IEEE Transactions on Information Theory, vol. 56, no. 6, pp. 2783–2803, 2010.
[30] D. Wassel, "Exploring novel designs of nlp solvers: architecture and implementation of worhp," PhD Thesis, Universität Bremen, 2013.

Mohammad Akhondi Darzikolaei received a B.Sc and M.Sc degree from the Babol Noshirvani University of Technology. He worked on radar clutters and radar signal processing. He is now a Telecommunication Ph.D. candidate at the Babol Noshirvani University of Technology. His research focuses on radar and sonar signal processing, speech processing and pattern recognition, and artificial neural networks.

Mohammad Reza Karami-Mollaei received the B.Sc in Electrical and Electronic Engineering in 1992, M.Sc of signal processing in 1994, and Ph.D in 1998 in Biomedical Engineering from I.N.P.L d'Nancy of France. He is now an associate professor with the Department of Electrical and Computer Engineering, Babol University of Technology. Since 1998 his research is in signal and speech processing. He teaches Digital Signal, Biomedical and speech processing at university. His research interests include Speech, Image and signal processing.

Maryam Najimi received her B.Sc in electronics from Sistanand Baloochestan University,Zahedan,Iran in 2004 and her M.Sc. in telecommunication systems engineering from K.N.Toosi University of Technology, Tehran, Iran, and Ph.D. degree in communication from Babol University of Technology, Mazandaran, Iran, in 2008 and 2014, respectively. She is currently an assistant professor with the department of electrical and computer engineering,University of Science and Technology, Behshahr, Iran. Her interests include Spectrum sensing in wireless cognitive sensor networks.

Reliability Analysis of the Joint LDPC Decoding Algorithms over the Multiple Access Channels

Mahdi Nangir *
Faculty of Electrical and Computer Engineering, University of Tabriz, Iran
nangir@tabrizu.ac.ir

Received: 11/Jun/2021 Revised: 24/Nov/2021 Accepted: 13/Dec/2021

Abstract
The joint Low Density Parity-Check (LDPC) decoding schemes iteratively decode the received data from multiple channels. Mostly, the available data in different channels are correlated and there is kind of dependency between the links or channels. In recent decades, the graph-based codes have been considered for the communication network scenarios. The performance of these codes is close to the existing theoretical bounds and their complexity is not high which cause the possibility of real world implementation and exploitation. The Multiple Access Channel (MAC) scenario with multiple senders which aim to send correlated data to a single receiver is considered. An analysis on the reliability of the Bit Error Rate (BER) performance of the Joint Sum-Product (JSP) decoding algorithm is presented for a two-link case, which can be extended to higher number of links. The effect of parameter variations on the BER performance is studied. These parameters include: the total number of iterations, the codeword length, the total number of rounds, and the coding rate in the JSP algorithm. An optimal value of the parameters is selected during the design procedure of a communication network by considering its limitations and complexity criterion. The JSP algorithm is a reliable scheme for jointly decoding of noisy binary data from different origins.

Keywords: Multi-Edge Type LDPC Codes; BER Performance; Joint SPA; Joint Channel Decoding Algorithm; Multiple Access Channel.

1- Introduction

Recently, wireless sensor networks (WSN) and radio access networks (RAN) have many applications in telecommunication industry. In these networks, numerous amount of correlated data have been generated and shared to finally send to a central destination. The receiver aims to decode multiple correlated data coming from multiple links or channels. A proper model for describing this scenario is the multiple access channel (MAC) model [1] which its theoretical bounds is derived in [2].

The central receiver in a MAC requires to jointly decode received data and reconstruct them by exploiting the correlation and similarity between them. Joint decoding of the data by employing graph-based codes has acceptable complexity and close performance to the existing theoretical bounds. Due to these advantages, it is practical and can be implemented in real-world applications. These schemes iteratively reconstruct corrupted data [3].

The successive decoding schemes are another efficient methods to decode multiple correlated data [4-6]. The performance of successive and joint decoding schemes are compared in [7]. Joint decoding scheme has close performance to the theoretical bounds. The joint sum-product algorithm (SPA) is an extended version of SPA [8] which is used in joint decoding scenarios in multi-terminal source coding problems [9] and multi-link channels such as the MAC. In the joint SPA there are various parameters which affects the bit error rate (BER) performance of the algorithm which aims to decode correlated data. In this work, we investigate the performance of the joint SPA with respect to the variation of key parameters which plays important role in the decoding process. In general, increasing the value of parameters cause to improve performance along with to increase the complexity [10]. This trade-off motivate us to study and analysis the BER performance of the joint SPA by changing parameters which provides a useful view to design decoder and to select parameters properly.

Capacity concept in communication networks is an extension of the channel capacity of a point-to-point communication system proposed by Shannon in 1948 and it determines the maximum possible rate under an arbitrary small BER value [11]. Practical implementations with finite code lengths almost have a small gap between rate values and capacity to achieve small enough values of BER. The gap value is a good measure for comparison of different decoding algorithms and schemes.

* Corresponding Author

The channel coding and source coding schemes and also their combinations are used in network scenarios such as the Wyner-Ziv problem [12], the Gelfand-Pinsker problem [13], and the CEO problem [14]. In this paper, we implement a joint SPA on the MAC and analysis its performance based on the variation of key parameters. This provides a reliability analysis for the algorithm over the MAC.
The joint SPA iteratively corrects errors which are occurred in the data by the noise of communication channel which is assumed to be an Additive White Gaussian Noise (AWGN) in our work. The LDPC codes is a strong and efficient group of error correction codes which can achieve the channel capacity and small BER values, when large enough code lengths are utilized [15]. In this work, we employ an extension of LDPC codes which are proper to network scenarios and called multi-edge type LDPC codes [16-17]. These graph-based codes are able to decode jointly and simultaneously the correlated data coming from separate links and channels.

A coding scheme is proposed for the two-user case of a Gaussian MAC based on LDPC codes [18]. Also, spatially coupled LDPC codes are utilized to multi-user detection over the MAC models to get advantages [19]. For the Multiple Access Relay Channels, a joint source-channel-network coding scheme is investigated in [20], which employs LDPC codes. A minimum error probability approach is studied in [21] to present a joint optimization method for two-user case in a non-orthogonal MAC. Besides the LDPC codes, other channel codes are employed for coding MACs. For instance, in [22] the polar codes which are strong channel codes are applied for the Gaussian MAC.

The purpose of a communication system and network designer in the practical implementation phase is to determine all parameters in the utilized algorithms. Our focus is on the key parameters of the joint decoding such as block length of codes, the total number of iterations and rounds in the joint SPA, and the coding rate. According to the criteria and expectations of the users with considering some aspects such as the complexity, one can choose the best possible value of the parameters by using our reported results and analyses. Throughout this paper, we present results based on a specific designed parity-check matrix H of a multi-edge type LDPC code, i.e. the reported results depend on the matrix H and its details. The performance of a reliable algorithm should not have tremendous variations by parameter changing. Finally, our presented results show that the joint SPA is a reliable algorithm for joint decoding of the MAC.

The main motivations of this research work can be listed as follows:

1- Graph-based codes are efficient tools in multi-user communication scenarios. Hence, we evaluate their performance in the MAC as an important case.

2- The SPA is generally for point-to-point communication, which is considered here for the case of multiple sender and encoder.

3- The BER performance of an efficient coding scheme should be reliable with respect to parameter variations, which is investigated in this paper.

The rest of this paper is organized as follows: In section II, the channel model and the joint decoding algorithm are proposed. Furthermore, the details of the joint SPA are provided in this section. Simulation results and discussions are presented in section III. Finally, section IV concludes this paper.

2- Problem Model and Decoding scheme

In this section, the problem model is provided and formulation of the joint SPA is given as an iterative message-passing algorithm as a joint decoding scheme. Prior to that we present the concept of the multi-edge type LDPC codes which the joint SPA is run by employing them.

2-1- MAC and its Limitations

The MAC is a many-to-one communication scenario which appears in real world applications such as the uplink transmission in wireless cellular and satellite communication systems. It also appears in computer networks when several information packets are sent to a single node via multiple links. A simple representation of the MAC is shown in Fig. 1, which N users aims to send their data to a single receiver.

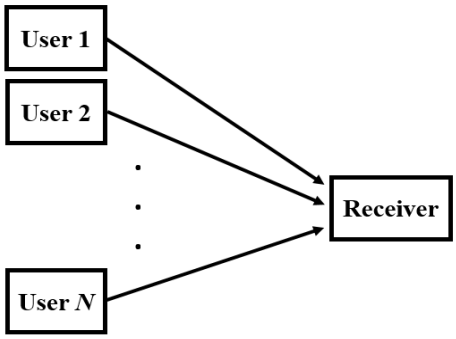


Fig. 1: Multiple Access Communication Channel Model.

A single-letter characterization of the capacity region of the MAC with only two channels are as follows:

$$\begin{aligned} R_1 &\le I(X_1;Y \mid X_2), \\ R_2 &\le I(X_2;Y \mid X_1), \\ R_1 + R_2 &\le I(X_1, X_2;Y), \end{aligned} \qquad (1)$$

where, X_i is the output of the encoder in the *i*-th channel. It is assumed that a message M_i is encoded to X_i. The codewords X_1 and X_2 are the inputs of the MAC and its output is Y. The probabilistic channel model is assumed to be $\Pr\{y \mid x_1, x_2\}$. The described model is depicted in the following figure.

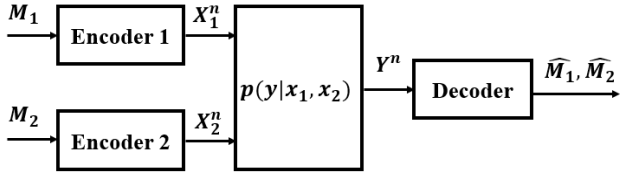


Fig. 2: Two-link case MAC.

We consider the binary erasure MAC in our implementations [11], in which X_1 and X_2 are binary and the channel output $Y = X_1 + X_2$ is ternary. In this case, the capacity region is completely determined [11]. We have the following capacity region that is plotted in Fig. 3. Hence, $H(Y \mid X_1, X_2) = 0$, we can write:

$$\begin{aligned} R_1 &\le I(X_1;Y \mid X_2) = H(Y \mid X_2) = 1, \\ R_2 &\le I(X_2;Y \mid X_1) = H(Y \mid X_1) = 1, \\ R_1 + R_2 &\le I(X_1, X_2;Y) = H(Y) = \frac{3}{2}. \end{aligned} \qquad (2)$$

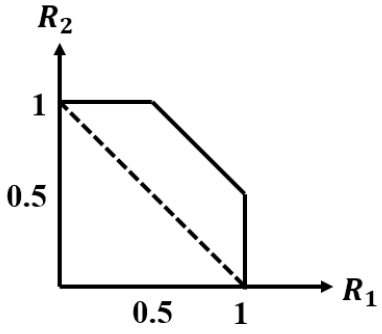


Fig. 3: The capacity region of the binary erasure MAC.

A joint decoding scheme such as the joint SPA wishes to recover the noisy corrupted data whit minimum possible BER and tolerable complexity.

2-2- Multi-Edge Type LDPC Codes

A conventional LDPC code consists of variable nodes and check nodes connected with each other with some edges that are in one-to-one correspondence with 1's in the parity check matrix *H*. For an LDPC code, the number of 1's are much less than the number of 0's in *H*. As an extension of the classic LDPC codes, we have the multi-edge type LDPC codes in which there are a single group of the variable nodes and several groups of the check nodes connected with the variable nodes. There is no connection between groups of the check nodes. Thus, for an *L*-edge type LDPC code, we have *L* parity check matrix H_i for $i = 1, 2, .., L$, each of which are related to a single LDPC code. They have shared variable nodes and separate check nodes. The total parity check matrix of the multi-edge type LDPC code are defined as follows:

$$H @ \begin{bmatrix} H_1 \\ H_2 \\ \mathrm{M} \\ H_L \end{bmatrix} \qquad (3)$$

For more illustration the Tanner graph representation of a two-edge type LDPC code is shown in Fig. 4. As it is seen, there are two types or two groups of the check nodes shown by squares, which are connected to the variable nodes shown by circles.

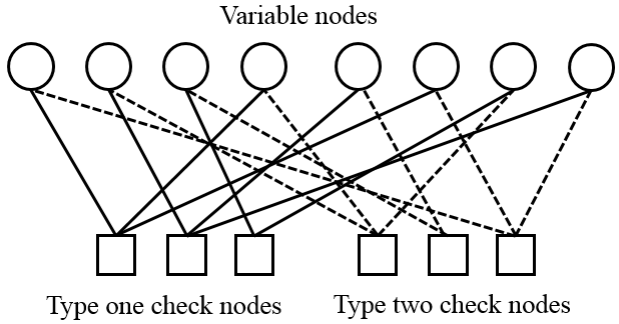


Fig. 4: Tanner graph of a two-edge type LDPC code.

The classical message passing algorithms for the channel decoding, e.g. the SPA, or for the source encoding, e.g. the Bias propagation consist of some iterations in which some messages are passed between the variable nodes and the check nodes of LDPC codes or LDGM [1] codes until reach to a stable condition. There exists a similar condition in the iterative joint decoding algorithms like the joint SPA in which some messages are passed between the variable nodes and a specific group of the check nodes in each

[1] Low Density Generator Matrix (LDGM) codes are the dual codes of the LDPC codes that are proper to the source coding scenarios in communication [23].

round within a specific number of iterations. Therefore, by utilizing the multi-edge type LDPC codes, we are able to implement joint MAC decoding algorithms iteratively.

Similar to the SPA, short length for the girth of the LDPC codes cause some problems in the decoding procedure [24]. The minimum required girth of Tanner graph of the LDPC codes is 6 to perform well, which is satisfied in our implementations.

Generally, for an *L*-edge type LDPC code, we can consider *L* rate values. If we want to achieve small values of the BER, then all *L* rates must be located inside of the capacity region. This condition should be considered in the design procedure of the multi-edge type LDPC codes and it obviously affects rate and code lengths. For instance, in the case of the binary erasure MAC, rates of the channel must be satisfied the inequalities in (1).

In general, a multi-edge type LDPC code is an extension of the classical LDPC code. The difference between them is that by considering all groups of check nodes together, unlike classical LDPC codes, the number of check nodes in multi-type LDPC codes can be more than number of variable nodes. It should be noted that a single group of check nodes along with the variable nodes is a classical LDPC code and it is employed in each link or channel for accomplishing a complete channel coding procedure.

2-3- The Joint SPA

The joint SPA is a modified version of the SPA [10]. It is applicable to joint decoding scenarios while the SPA is for point-to-point communication scenarios.

For the sake of simplicity, we describe a joint SPA for a two-link case over a two-edge type LDPC code. This algorithm consists of r rounds. For describing the decoding algorithm with more details, we use the following notations and assumptions:

- $q_{i,m}^{\text{out}}(q_{i,m}^{\text{in}})$ shows the LLR values passing from *m*-th edge and coming from (to) the *i*-th variable node in each iteration and each round and iteration of the algorithm.
- $d_{v,i}(d_{c,j})$ for $i=1,2,...,n$ shows the degree of the *i*-th variable node (*j*-th check node).
- $t_{j,m}^{\text{out}}(t_{j,m}^{\text{in}})$ shows the LLR values passing from *m*-th edge and coming from (to) the *j*-th check node in each iteration and each round of the algorithm.
- $\text{Rate}=\dfrac{\text{Total number of check nodes}}{\text{Total number of variable nodes}}$.
- The total number of rounds is $r=\max(r_1,r_2)$.
- The total number of iterations is $\mathrm{l}=l_1+l_2$.
- *n* is the codeword block length.

Clearly, the parity check matrix determines connections between the nodes and based on these connections LLR messages are passed between the check nodes and variable nodes.

First, the received syndromes are located in the check nodes of the two-edge type LDPC code. The first one is located in the type one check nodes and the second one is located in the type two check nodes, respectively. The proposed algorithm have r rounds, where $r=\max(r_1,r_2)$. Each round contains of $\mathrm{l}=l_1+l_2$ iterations, where l_1 iterations are done in the first LDPC code and l_2 in the second one. Note that the information in the first and the second links are reconstructed in r_1 and r_2 iterations, respectively.

In the end of the algorithm, the reconstructed bit can be found in the variable nodes after a sufficient number of rounds.

At the beginning of the algorithm, the initial LLR values of the variable nodes are randomly determined. During iterations in each round, their values are changed and updated until reaching to the final iteration or the LLR values converge.

Encoding and Decoding Procedure:

Two information sequences are encoded by two LDPC codes with parity check matrices of H_1 and H_2 in the first and second links. Therefore, two LDPC codewords are sent to the receiver through a noisy channel which causes some errors in the bit streams.

In the receiver, a joint decoder receives the noisy codewords and calculate their syndromes at once. Then, these syndromes are employed to reconstruct information sequences of the links by the joint SPA.

By multiplying the information bit of length k_i with the generator matrix G_i with the size of $k_i \times n$, a codeword C_i is obtained, for *i*=1,2. C_i is a valid codeword if and only if $C_i H_i^T = 0$. These two codewords are sent via a noisy channel to the decoder side. Decoder receives $C_i' = C_i + n_i$, where n_i is a random Bernoulli noise with parameter p_i. Then, it calculates syndromes $S_i = C_i' H_i^T \neq 0$ and starts to accomplish the joint SPA to recover LDPC codewords and information packets.

Without loss of generality, we suppose $p_1 \leq p_2$. Under this assumption, we expect that the codeword C_1 is decoded and appeared in the variable nodes after r_1 rounds. Similarly, that the codeword C_2 is decoded and appeared

in the variable nodes after r_2 rounds and obviously we have $r_1 \le r_2$.

Here, we present the update equations of the joint SPA. The initial values of LLRs are as follows:

$$q_{i,o} = \log \frac{\Pr[v_i = 0 \mid v_i']}{\Pr[v_i = 1 \mid v_i']} \qquad i = 1, 2, ..., n \qquad (4)$$

where, v_i' is the bit value of the *i*-th variable node in the two-edge type LDPC code. Note that, $q_{i,o}$ is equal to $(1-2v_i')\log\frac{1-p_1}{p_1}$ for the messages toward the type one check nodes; and it is $(1-2v_i')\log\frac{1-p_2}{p_2}$ for the messages toward the type two check nodes.

During an iteration, the LLR message that is sent from the *i*-th variable node to a check node using *m*-th edge is as follows:

$$q_{i,m}^{\text{out}} = q_{i,0} + \sum_{j=1, j\neq m}^{d_{v,i}} q_{i,j}^{\text{in}} \qquad (5)$$

for $m = 1, 2, ..., d_{v,i}$ and $i = 1, 2, ..., n$. In the first iteration, we set $q_{i,j}^{\text{in}} = 0$. In each iteration the LLR value of $q_{i,m}^{\text{out}}$ is allocated to the related $t_{j,\pi(i,m,j)}^{\text{in}}$ based on the connections of the two-edge type LDPC code. These LLR values are used in the check node update equations and $\pi(i,m,j)$ is the number of edge [1] connecting *i*-th variable node to the *j*-th check node through *m*-th edge from variable nodes side. The update equations engaged in the *j*-th check node of the two-edge type LDPC code are as follows:

$$\tanh\left(\frac{t_{j,m}^{\text{out}}}{2}\right) = (1-2s_j)\prod_{i=1, i\neq m}^{d_{c,j}} \tanh\left(\frac{t_{i,m}^{\text{in}}}{2}\right) \qquad (6)$$

where, $m = 1, 2, ..., d_{c,j}$ and s_j is bit value of the *j*-th check node. Note that it can belong to either S_1 or S_2. We set $q_{i,m}^{\text{in}} = t_{j,\pi(i,m,j)}^{\text{out}}$ and continue the algorithm with update equations. Finally, after running a sufficient number of iterations, we can estimate the codewords and decide about the binary values of variable nodes based on the following rule:

$$\hat{v}_i = \begin{cases} 0, & \text{if} \quad q_{i,0} + \sum_{m=1}^{d_{v,i}} q_{i,m} \ge 0 \\ 1, & \text{if} \quad q_{i,0} + \sum_{m=1}^{d_{v,i}} q_{i,m} < 0 \end{cases} \qquad (7)$$

[1] From check nodes side.

Based on (7), we decode the codewords that have been sent from the different links of the MAC. Note that, this decoding is in the category of the error correction algorithms.

3- Implementation Results and Analyses

In this section, we present implementation results which reports the performance of the algorithm for different values of the important parameters. As it is known, the BER value can represent the efficiency and power of a scheme in channel decoding scenarios. All reported BER values are obtained by averaging over 50 tests when random binary information sources are generated.

We focus on some key parameters which play important role in the decoding procedure. They include: the number of iterations in the algorithm, the number of rounds in the algorithm, the code length, and the coding rate. We compare the results in various cases and present the BER performance versus different values of the signal to noise ratio. We compute the BER in our implementation results using the following formula:

$$\text{BER}_1 = \text{E}\left[\frac{1}{n} d_{\text{Ham}}(C_1, \hat{V})\right] \qquad (8)$$

after r_1 rounds when C_1 is decoded. Similarly, C_2 is decoded after r_2 rounds and BER is equal to:

$$\text{BER}_2 = \text{E}\left[\frac{1}{n} d_{\text{Ham}}(C_2, \hat{V})\right] \qquad (9)$$

In (8) and (9), $\text{E}[.]$ is expectation, $d_{\text{Ham}}(.,.)$ shows the Hamming distance between two binary sequences, and $\hat{V} = (\hat{v}_1, \hat{v}_2, ..., \hat{v}_n)$ is obtained from (7). We define

$$\text{BER} = \max(\text{BER}_1, \text{BER}_2) \qquad (10)$$

In what follows, the simulation and implementation results are reported which shows the BER performance of a MAC scenario with binary sources which is equipped with the joint SPA for jointly decoding of noisy information sources employing a two-edge type LDPC code. Generally, we know that the BER value decreases by increasing the signal to noise ratio.

In Fig. 5, it is shown that the BER performance of the algorithm improves by increasing the number of maximum number of iterations in the SPA in each round. Note that we have used a same parity-check matrix *H* with the rate of 1. The code length is considered to be 10^4 and the total number of rounds is 25. Increasing the number of iterations cause to increase the accuracy of the convergence in the joint SPA similar to the conventional SPA.

Furthermore, it is seen that by increasing the total number of iterations more and more, the BER performance does

not improve significantly. Therefore, the sensitivity of the joint SPA over a MAC with respect to the total number of iterations is high for small number of iterations and vice versa.

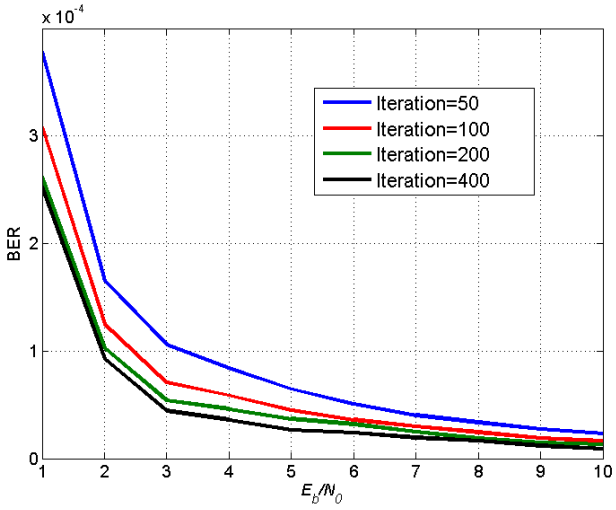


Fig. 5: The BER performance of the joint SPA for the MAC for different number of iterations.

In fig 6, we propose the BER performance of the joint SPA with respect to the variation of the codeword length *n*. The total number of iterations is 100. The total number of rounds is assumed to be 25 and for simplicity we consider rate is equal to be 1 in this implementation. As it is seen, increasing codeword length cause to reduce the BER value. This observation exactly happens in the point-to-point channel coding scenarios in telecommunication. It should be noted that the complexity increases with growing the codeword length *n*.

A communication system designer can choose appropriate and reliable parameters of the message-passing algorithms and establish a trade-off between the acceptable BER performance and the complexity. Acceptable performance and complexity strongly depend on the application and the communication scenario. We accept complexity increasing in some real world applications that high security or high reliability is required. In some other applications, we can employ not so large codeword lengths.

We defined the total number of rounds which is shown by *r*. As it is depicted in Fig. 7, by increasing the total number of rounds, the BER performance of the algorithm improves. This is somehow predictable, because by increasing the total number of rounds the accuracy of the convergence in a message-passing algorithm, such as the joint SPA, increases for a fixed number of iterations and fixed codeword length and rate. Simulation results that are presented in Fig. 7 are obtained for 100 number of iterations. Furthermore, the codeword length is 10^4 and the rate is equal to be 1. The comparison will be simplified when all parameters except one of them are fixed.

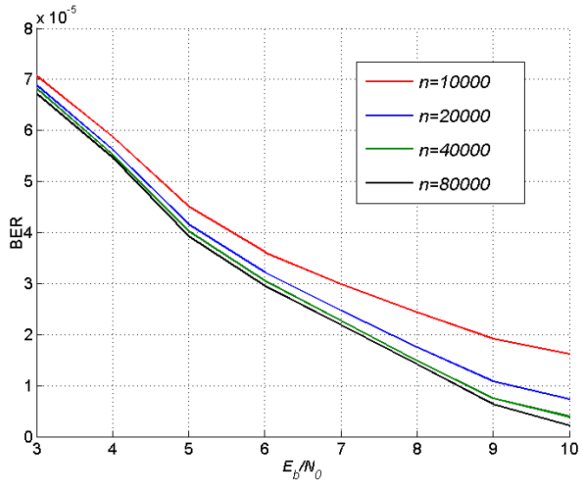


Fig. 6: The BER performance of the joint SPA for the MAC for different code lengths (*n*).

The running time of an algorithm is another factor which should be considered. The running time increases by growing the total number of iterations and rounds in the joint SPA.

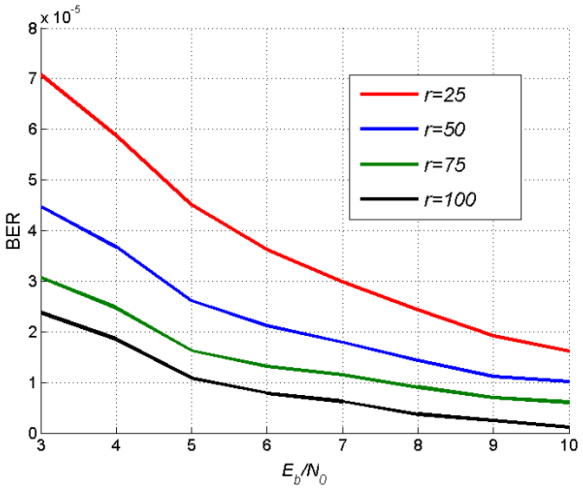


Fig. 7: The BER performance of the joint SPA for the MAC for different number of rounds (*r*).

Finally, we show the BER performance of the joint SPA with respect to variation of rate values in Fig. 8. We defined the rate as the ratio of the total number of check nodes to the total number of variable nodes in a multi-edge type LDPC code. In this case, *n* is equal to 10^4, the total number of iterations is 100, and *r* is equal to 25.

Based on the results presented in Fig. 8, the BER performance of the joint SPA improves by increasing the value of the rate, i.e. by growing the total number of check nodes when the total number of variable nodes is assumed to be fixed.

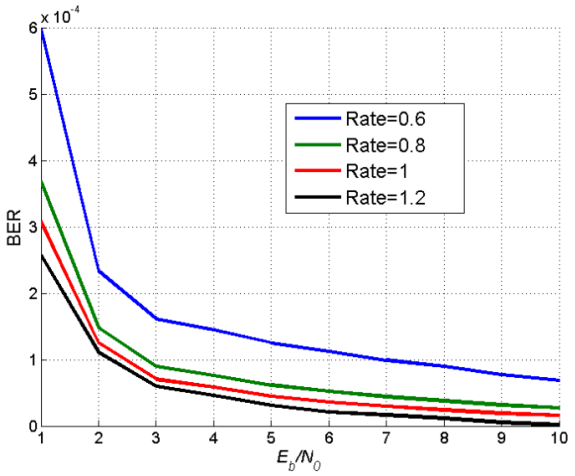


Fig. 8: The BER performance of the joint SPA for the MAC for different values of rate.

In message-passing algorithms over the graph-based codes, the number of check nodes shows the independent linear equations should be satisfied on the set of variable nodes. Thus, it is concluded that by growing the number of check nodes, the number of conditions over the variable nodes that should be satisfied increases and this cause to reach more accuracy of convergence in the algorithm. In other words, by increasing the total number of check nodes, we achieve low BER values.

Based on the results which are presented in this section can report the BER value of the joint SPA for some other values of the key parameters by using linear interpolation technique. The obtained results from the linear interpolation are not exact, but they can provide a good view about the performance of the message-passing algorithm. Estimation of the performance in these message-passing algorithms can be done with an arbitrary precision by finding more BER curves versus the signal to noise ratio for various parameters.

4- Conclusions

In this paper, a new scheme for the MAC decoding presented. Graph-based codes are employed to design the joint SPA for reconstructing binary data which are encoded by using multi-edge type LDPC codes. The performance of the proposed scheme is evaluated form the BER point of view. As a communication system designer, we analysis the reliability of the joint SPA with respect to some key parameters including the total number of iterations, the codeword length, the total number of rounds, and the coding rate. We present the BER values of the joint decoding versus different values of the signal to noise ratio. These results helps a system designer to choose appropriate values of the parameters according to the desirable complexity-performance trade-off.

Appendix

The degree distribution of the parity-check matrices are given here. These degree distributions are utilized in the simulations and implementations of the section 3. As it is known, for a single degree distribution, there is not a unique parity-check matrix and there a lot. We construct matrices by utilizing the progressive edge growth method with some approximations. We employ the degree distributions reported in [25].

We employ the following degree distribution for the coding rate of 0.3:

$$\lambda(x) = 0.1392x + 0.2007x^2 + 0.2522x^6 + 0.0134x^{11} + 0.172x^{17} + 0.0424x^{31} + 0.0855x^{41} + 0.0953x^{49}$$

and

$$\rho(x) = 0.3x^{16} + 0.7x^{17}$$

We employ the following degree distribution for the coding rate of 0.4:

$$\lambda(x) = 0.0665x + 0.3059x^2 + 0.0885x^6 + +0.0028x^7 + 0.0933x^8 + 0.027x^9 + 0.0288x^{10} + 0.0067x^{11} + 0.0622x^{13} + 0.013x^{14} + 0.0041x^{15} + 0.0023x^{16} + 0.0016x^{17} + 0.0016x^{18} + 0.0014x^{19} + 0.0013x^{20} + 0.001x^{21} + 0.001x^{37} + 0.003x^{38} + 0.281x^{39}$$

and

$$\rho(x) = x^{13}$$

We employ the following degree distribution for the coding rate of 0.5:

$$\lambda(x) = 0.1528x + 0.2825x^2 + 0.0062x^3 + 0.5586x^{19}$$

and

$$\rho(x) = x^9$$

We employ the following degree distribution for the coding rate of 0.6:

$$\lambda(x) = 0.2365x + 0.3137x^3 + 0.0714x^4 + 0.1707x^{10} + 0.0004x^{11} + 0.0002x^{18} + 0.0002x^{19} + 0.0002x^{20} + 0.0003x^{21} + 0.0007x^{22} + 0.0183x^{23} + +0.1854x^{24} + 0.0016x^{25}$$

and

$$\rho(x) = 0.9x^6 + 0.1x^7$$

References

[1] E. C. Van Der Meulen, "The discrete memoryless channel with two senders and one receiver", In Proc. IEEE Int. Symp. Information Theory (ISIT), September 1971, pp. 78.

[2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers", Annals of probability, Vol. 2, No. 5, 1974, pp. 805-814.

[3] A. Khandekar, "Graph-based codes and iterative decoding", Ph.D. thesis, California Institute of Technology, USA, 2002.

[4] J. Scarlett, A. Martinez, and A. G. Fabregas, "Mismatched multi-letter successive decoding for the multiple-access channel", IEEE Transactions on Information Theory, Vol. 64, No. 4, 2017, pp. 2253-2266.

[5] P. Li, X. Jian, F. Wang, S. Fu, and Z. Zhang, "Theoretical throughput analysis for massive random access with spatial successive decoding", IEEE Transactions on Vehicular Technology, Vol. 69, No. 7, 2020, pp. 7998-8002.

[6] M. Nangir, R. Asvadi, J. Chen, M. Ahmadian-Attari, and T. Matsumoto, "Successive Wyner-Ziv Coding for the Binary CEO Problem under Logarithmic Loss", IEEE Transactions on Communications, Vol. 67, No .11, 2019, pp. 7512-7525.

[7] M. Nangir, J. Pourrostam, J. M. Niya, and B. M. Tazehkand, "Comparison Between the Joint and Successive Decoding Schemes for the Binary CEO Problem", In IEEE 28th Iranian Conference on Electrical Engineering (ICEE), August 2020, pp. 1-5.

[8] S. Papaharalabos and P. T. Mathiopoulos, "Simplified sum-product algorithm for decoding LDPC codes with optimal performance," Electronics Letters, Vol. 45, No. 2, January 2009, pp. 116-117.

[9] M. Nangir, R. Asvadi, M. Ahmadian-Attari, and J. Chen, "Analysis and code design for the binary CEO problem under logarithmic loss", IEEE Transactions on Communications, Vol. 66, No. 12, 2018, pp. 6003-6014.

[10] H. Khodaei Jooshin, and M. Nangir, "Reliability Analysis of the Sum-Product Decoding Algorithm for the PSK Modulation Scheme", Journal of Information Systems and Telecommunication (JIST), Vol. 8, No. 3, 2020, pp. 167-174.

[11] A. El Gamal, and Y. H. Kim, Network information theory, Cambridge University Press, 2011.

[12] M. Nangir, R. Asvadi, M. Ahmadian-Attari, and J. Chen, "Successive Wyner-Ziv coding for the binary CEO problem under log-loss", In IEEE 29th Biennial Symposium on Communications (BSC), Canada, 2018, pp. 1-5.

[13] Z. Goldfeld, and H. H. Permuter, "A Useful Analogy Between Wiretap and Gelfand-Pinsker Channels", In IEEE International Symposium on Information Theory (ISIT), 2018, pp. 121-125.

[14] M. Nangir, R. Asvadi, M. Ahmadian-Attari, and J. Chen, 'Binary CEO problem under log-loss with BSC test-channel model", In IEEE 29th Biennial Symposium on Communications (BSC), 2018, pp. 1-5.

[15] F. Vatta, A. Soranzo, M. Comisso, G. Buttazzoni, and F. Babich, "Performance study of a class of irregular LDPC codes through low complexity bounds on their belief-propagation decoding thresholds", In IEEE AEIT International Annual Conference, 2019, pp. 1-6.

[16] S. Jayasooriya, M. Shirvani moghaddam, L. Ong, G. Lechner, and S. J. Johnson, "A new density evolution approximation for LDPC and multi-edge type LDPC codes", IEEE Transactions on Communications, Vol. 64, No. 10, 2016, pp. 4044-4056.

[17] S. Jeong, and J. Ha, "On the Design of Multi-Edge Type Low-Density Parity-Check Codes", IEEE Transactions on Communications, Vol. 67, No. 10, 2019, pp. 6652-6667.

[18] J. Du, L. Zhou, L. Yang, S. Peng and J. Yuan, "A New LDPC Coded Scheme for Two-User Gaussian Multiple Access Channels," in IEEE Communications Letters, vol. 22, no. 1, pp. 21-24, Jan. 2018.

[19] S. Cammerer, X. Wang, Y. Ma and S. t. Brink, "Spatially Coupled LDPC Codes and the Multiple Access Channel," 2019 53rd Annual Conference on Information Sciences and Systems (CISS), 2019, pp. 1-6.

[20] M. B. Abdessalem, A. Zribi, T. Matsumoto and A. Bouallègue, "LDPC-based Joint Source-Channel-Network Coding for the Multiple Access Relay Channel," 2018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM), 2018, pp. 1-6.

[21] B. K. Ng and C. Lam, "Joint Power and Modulation Optimization in Two-User Non-Orthogonal Multiple Access Channels: A Minimum Error Probability Approach," in IEEE Transactions on Vehicular Technology, vol. 67, no. 11, pp. 10693-10703, Nov. 2018.

[22] M. Ebada, S. Cammerer, A. Elkelesh, M. Geiselhart and S. t. Brink, "Iterative Detection and Decoding of Finite-Length Polar Codes in Gaussian Multiple Access Channels," 2020 54th Asilomar Conference on Signals, Systems, and Computers, 2020, pp. 683-688.

[23] Z. Sun, M. Shao, J. Chen, K. M. Wong, and X. Wu, "Achieving the rate-distortion bound with low-density generator matrix codes", IEEE Transactions on Communications, Vol. 58, No. 6, 2010, pp. 1643-1653.

[24] T. Richardson, and R. Urbanke, Modern coding theory, Cambridge University Press, 2008.

[25] D. H. Schonberg, "Practical distributed source coding and its application to the compression of encrypted data", Ph.D. thesis, University of California, Berkeley, USA, 2007.

Mahdi Nangir received the B.Sc degree with first rank in Electrical Engineering from University of Tabriz and M.Sc. degree in Communication System Engineering from Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively. He received the Ph.D. degree from K. N. Toosi University of Technology, Tehran, Iran, in 2018. He was a finalist of the Mathematics Olympiad and owner of bronze medal in 2005 from Young Scholar Club. He received high ranks in Iranian National Electrical Engineering Student Olympiads of 2009 and 2010. In 2017, he joined McMaster University, Hamilton, Ontario, Canada as a research visiting student. He is now an assistant professor in Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran. His research interests include coding and information theory, distributed source coding, data compression algorithms and optimization.

Remote Sensing Image Registration based on a Geometrical Model Matching

Zahra Hossein-Nejad
Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.
hoseinnejad.zahra@yahoo.com
Hamed Agahi*
Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.
agahi@iaushiraz.ac.ir
Azar Mahmoodzadeh
Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.
mahmoodzadeh@iaushiraz.ac.ir

Received: 31/Aug/2021 Revised: 24/Nov/2021 Accepted: 07/Dec/2021

Abstract
Remote sensing image registration is the method of aligning two images from the same scene taken under different imaging circumstances containing different times, angles, or sensors. Scale-invariant feature transform (SIFT) is one of the most common matching methods previously used in the remote sensing image registration. The defects of SIFT are the large number of mismatches and high execution time due to the high dimensions of classical SIFT descriptor. These drawbacks reduce the efficiency of the SIFT algorithm. To enhance the performance of the remote sensing image registration, this paper proposes an approach consisting of three different steps. At first, the keypoints of both reference and second images are extracted using SIFT algorithm. Then, to increase the speed of the algorithm and accuracy of the matching, the SIFT descriptor with the vector length of 64 is used for keypoints description. Finally, a new method has been proposed for the image matching. The proposed matching method is based on calculating the distances of keypoints and their transformed points. Simulation results of applying the proposed method to some standard databases demonstrated the superiority of this approach compared with some other existing methods, according to the root mean square error (RMSE), precision and running time criteria.

Keywords: SIFT; Matching Method; Remote Sensing Image Registration; Transformation Model.

1- Introduction

Image registration is a keystone in programmed remote sensing image analysis such as change detection, image synthesis and image mosaic [1, 2]. Remote sensing image registration (RSIR) is the procedure for finding geometric transformation models between multiple remote sensing images, taken at different times or viewpoints or by different cameras or sensors [3, 4]. In general, registration methods of remote sensing images are categorized in intensity-based and feature-based approaches [3, 5, 6]. The former techniques use the intensity distribution in the masks of the same sizes. To this end, using similarity arrays, parameters of the geometric transformation model are tuned. Common similarity indices for RS images include mutual information (MI) [7] and phase correlation [8]. Accordingly, intensity-based techniques encounter various challenges [9, 10]. One main challenge is that for images with large intensity differences, overlapping regions in the images become limited [11]. Additionally, the run-time of these approaches is high, since they make use of the whole content in the image [12]. In [3, 13] several registration methods are compared, confirming that the feature-based techniques attain higher performance indices for remote sensing images registration. Feature-based approaches find salient features of the images; and afterwards they compute the matching and all parameters of the transformation [14]. Harris [15], scale-invariant feature transform (SIFT) [16], and speed-up robust feature (SURF) [17] are typically used in the feature-based RSIR. The SIFT algorithm is robust against scale and rotation changes and also intensity variations, affine distortion and noise [18]. Although these advantages made this algorithm significant in the registration process, the complex nature of remote sensing images has resulted in many mismatches [13, 19, 20].

Several studies have been conducted on the enhancement of the matching accuracy of the RS images in the feature extraction and descriptor computation steps. In [21], SIFT was adapted for multi-modal remote sensing image registration. This method considered a threshold value of 0.08 to remove keypoints with low contrast, and an 8×8

window to create a descriptor for keypoints. In [20] the AB-SIFT descriptor is proposed to improve the matching of RS images. In this method, the features are identified using the Hessian. Then for each feature, one direction is considered according to the SIFT method. Finally, the proposed AB-SIFT descriptor and the nearest neighbor-based matching method are used to match the images. In [13] the uniform robust SIFT (URSIFT) algorithm was introduced, in which extracted features were made uniform in terms of spatial and scale distribution. Via this method, the features were distributed uniformly in the image and then, the nearest neighbor-based matching is used to match the remote sensing images. In [22] an effective SIFT-based mode-seeking algorithm was proposed in order to register the remote sensing images. To eliminate the outliers, the authors used mode seeking of scale, rotational difference, and vertical and horizontal shifts between all SIFT keypoints. In [23], the scale restriction SIFT (SRSIFT) was introduced that eliminated a number of mismatches by gradient orientation modification and scale restriction criteria. The main challenge in the SRSIFT algorithm is high execution time due to the high dimensions of its descriptor. These problems led to interference in the process of remote sensing image registration. Although these improvements have increased the accuracy of the correspondence, on the one hand, there are still a number of incorrect correspondences in the SIFT algorithm. Indeed, in some applications, including object identification, even an incorrect correspondence interferes with this process. On the other hand, all the reviewed methods have increased the accuracy of the comparisons by reducing the total number of correspondences, while the total number of correspondences is very effective in subsequent processes. Nowadays, creating matching methods which can match the maximum keypoints and create the most correct matches has been one of the important challenges.

In this paper, first the keypoints are identified using the SIFT. The first innovation of this paper is the use of a 64-dimensional SIFT descriptor to describe keypoints in remote sensing images, which improves the speed and increases the matching accuracy compared to the classic 128-dimensional SIFT descriptor. The NNDR matching method is one of the common methods of local matching, which has problems, such as increasing the number of local incorrect matches and matching multiple errors. Next innovation, a global matching method is introduced to solve these problems. The proposed matching method is based on calculating the distances of keypoints and their transformed ones. The proposed matching method does not create multiple matches and local incorrect matches. The proposed matching method was able to match the maximum keypoints and maximized the correct matches.

The remainder of article is organized as follows. Section 2 presents the NNDR matching method and its problems. Section 3 presented the proposed method. In section 4 results are described and the paper is concluded in section 5.

2- NNDR Matching Method and Its Problems

In this section, the NNDR matching method is described briefly, and then the disadvantages of this algorithm in the remote sensing image registration are reported.

2-1- NNDR Matching Method

The matching operation is done using descriptors of each feature. First the Euclidean distance between descriptors in both images is computed and then an appropriate nearest neighbor distance ratio (NNDR) criterion is used. For each feature descriptor (FD_i) in the first image, its Euclidean distance to the first nearest (SD_j) and the second nearest (SD_m) neighbors in the second image are calculated, and their ratio is computed according to (1) [24]. If this ratio is smaller than a pre-determined threshold value T_e, the matching is performed. The value of T_e is set at 0.8 [16].

$$\frac{\|FD_i - SD_j\|}{\|FD_i - SD_m\|} < T_e \quad (1)$$

2-2- NNDR Matching Method Problems

The NNDR matching method is a common method, which one of its drawbacks is creating multiple matches, as shown in the schematic Fig. 1. In the figure, the NNDR method has matched the keypoints of p_3, p_4, and p_5 to a same point at q_4 (i.e., multiple matches). In this illustration, matches done for keypoints p_5 and p_3 are incorrect. On the other hand, the NNDR is a local matching method; thus, it cannot generate an accurate model for the whole image. For example, the point p_5 is located in the vicinity of point p_4 which is another weakness of this matching method; that causes p_5 to be mistaken for q_4. Multiple matches and local matchings increase incorrect correspondences and decrease matching accuracy; the importance of which leads to the proposal of a novel matching method based on a global transformation model.

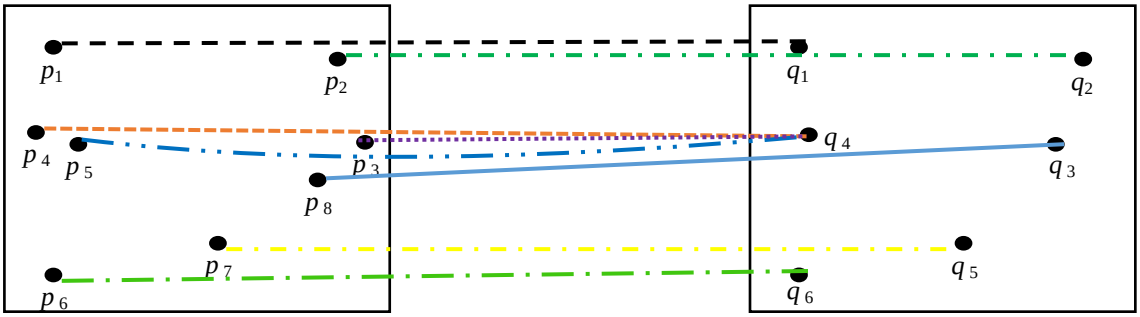


Fig. 1 The NNDR based matching method

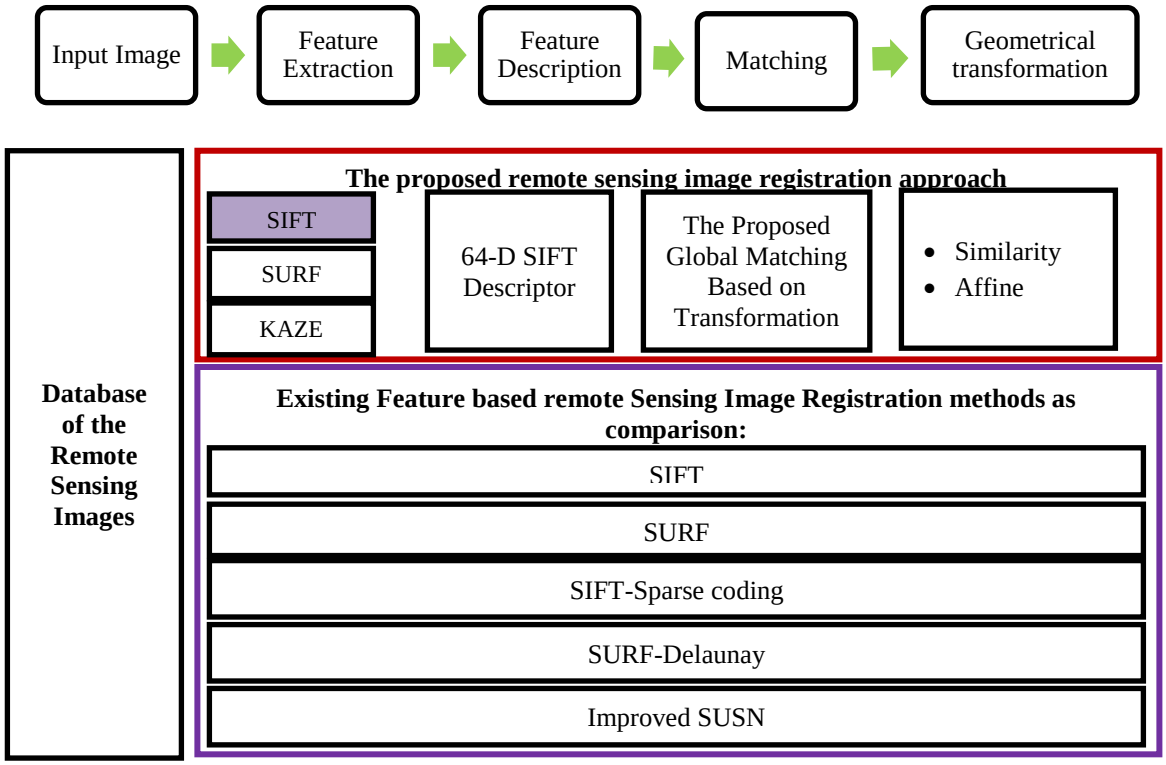


Fig. 2 Remote sensing image registration system

3- The proposed Method

This paper focuses on the feature-based registration method for RS images. Fig. 2 shows the proposed system containing four main phases including (1) feature extraction using the SIFT, (2) descriptors creation using 64-D SIFT descriptor, (3) matching process using the proposed method, and (4) making the transformation model. In this matching approach, other feature extraction algorithms can also be replaced for different circumstances, for example the improved versions of SIFT, Fast, KAZE and etc.

3-1- Feature Extraction

Features of images are identified by the SIFT method. Details of the SIFT are referred to [16].

3-2- 64-D SIFT Descriptor Creation

Once the keypoints are extracted using the SIFT algorithm, the next phase is to create a descriptor as a tool for finding matches between images. In this paper, the 64-D SIFT

descriptor is used instead of the 128-D one, using a smaller window [25]; as shown in Fig. 3. In this figure, the keypoint is represented by a blue circle, whose eight neighborhoods are considered according to each keypoint around it. Then, eight directions are found for each neighborhood and a 64-D descriptor is created. This descriptor window increases the matching accuracy and reduces the run-time.

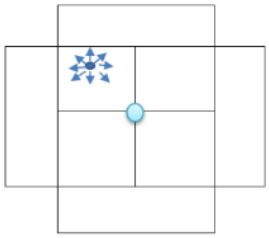


Fig. 3 a 64-D SIFT descriptor

3-3- Matching Process using the Proposed Global Matching

In this sub-section, to solve the problems of the NNDR matching method due to its locality, a new global approach is proposed. In this method, each keypoint in the first image is checked with all the keypoints in the second image, and then the point that has a better fit than the whole set is selected to be match to the desired candidate. This global matching method is based on the Euclidean distances and the transformation model, which is shown in Fig. 4 and described next.

- First, the Euclidean distance between each keypoint in the first image and all keypoints in the second image is computed, and this operation is performed for all keypoints in the first image, to create an $m \times n$ matrix. Here, m and n are the numbers of keypoints in the first and second images respectively.
- In each row of this matrix, the smallest Euclidean distance is considered as the candidate matching of that keypoints in the first and second images. For example, in Fig. 4, the pairs p_1-q_1, p_2-q_2, p_3-q_3, p_4-q_3, p_5-q_5, p_6-q_6 and p_7-q_4 are considered as candidate matches.
- To compute the transformation parameters and the model, three pairs of candidates with the least Euclidean distance, while simultaneously far away each other at least by three pixels are selected. In Fig. 4, the pairs p_1-q_1, p_3-q_3 and p_6-q_6 are selected to compute the parameters and the model. It is noted that the pair p_4-q_3 is not chosen since the distance of p_4 to p_3 is less than three pixels.
- The transformation model is computed from the mentioned three pairs of candidate points, as in
- (2). Here, θ represents the model parameters.

$$(x',y')=T\left((x,y),\theta\right) \quad (2)$$

- The distance between each candidate matching point and its transformed point is computed in the second image. If this distance is less than a specified threshold, the candidate matching point will be approved, otherwise it will not be matched according to (3).

$$dis\left(p_i',q_j\right)=\sqrt{(x_i'-x_j)^2+(y_i'-y_j)^2}$$
$$dis(p_i',q_j)<T \quad (3)$$

In (3), $dis(p_i',q_j)$ is the distance between the candidate point q_j and the transformed point p_i'. The threshold value T is simply set at 1.

3-4- Transformation Model Estimation

Several transformation models including affine, similarity, and the thin-plate spline (TPS) are utilized for the geometric distortion between images [26]. The transformation model type is selected according to the number of matches. Similarity [27] and affine distortion [28] are applied in the system of this work. The similarity transformation is chosen where the matches are two and the affine one is selected when this number is in the range of 3-6. Finally, the transformation parameters are computed by the least-square method [29].

4- Results

In this part, the functioning of the proposed method was evaluated using the database in [13]. This database contains five sets of remote sensing images with different texture conditions, such as urban areas and natural landscapes (see Table 1). To perform a fair evaluation, a comparison is made with the classical SIFT-NNDR matching [16], SURF-Delaunay triangulation matching(SURF-DTM) [30] , SURF-NNDR matching [17], SIFT-Sparse Coding(SIFT-SC) [31] and improved SUSAN [32]. The tests were executed on a PC with Intel® Core™ i7 processor and 16 GB RAM memory, implemented in MATLAB®2018A software. Evaluation criteria and the results of the tests are reported next.

4-1- Evaluation Criteria

To assess the functioning of the proposed method, four assessment indices including RMSE, SITMMR [33], SITMMC [33], and precision are used, as formulated respectively in (4-7). In (4), (x_i,y_i) and (x_i',y_i') are the coordinates of the i^{th} matched points pair, and NB_{TM} is the total number of matches. Also, in (5-7), NB_{CM} and NB_{FM} denote the number of true and false matches, respectively. A system with high SITMMC and precision, and also low

SITMMR and RMSE is appropriate for the registration task.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - x_i'\right)^2 + \left(y_i - y_i'\right)^2}{NB_{TM}}} \quad (4)$$

$$SITMMR = \frac{NB_{FM} + 1}{NB_{TM}} \quad (5)$$

$$SITMMC = \frac{NB_{CM} - 1}{NB_{TM}} \quad (6)$$

$$precision = \frac{NB_{CM}}{NB_{TM}} \quad (7)$$

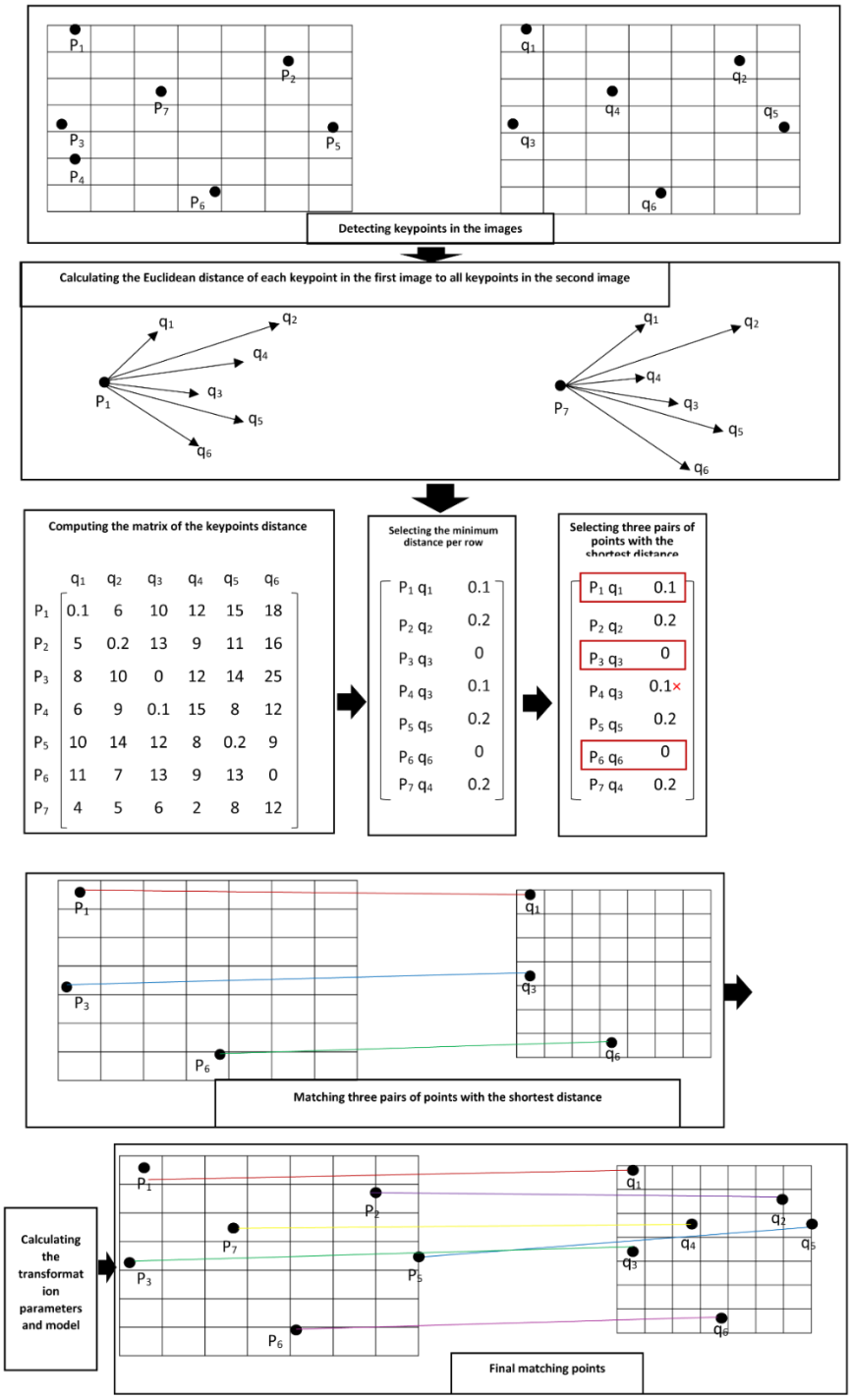


Fig. 4 Flowchart of the proposed matching method

Table 1: Properties of the datasets of the remote sensing images

No.	Satellite	Spectral Mode	Image Size	Pixel Size	Bits Per Pixel	Acquisition Date	Location
1	SPOT4	Multispectral Band: 1	611×1235	20	8	2001	China-Hangzhou
	Landsat TM	Multispectral Band: 3	648×1230	30	8	2004	China-Hangzhou
2	SPOT5	Panchromatic	1311×1215	10	8	2006	Canada-Halifax
	Landsat ETM^{+}	Multispectral Band: 3	440×410	30	8	1999	Canada-Halifax
3	IRS-1C	Panchromatic	1346×1135	5	8	1998	Iran-Tehran
	SPOT4	Panchromatic	700×590	10	8	1996	Iran-Tehran
4	IKONOS	Panchromatic	1288×1085	1	11	2001	Iran-Shiraz
	IRS-P6	Panchromatic	651×548	2.5	10	2006	Iran-Shiraz
5	IKONOS	Panchromatic	2335×1987	1	11	2004	Iran-Tehran
	IRS-1C	Panchromatic	567×482	5	8	1998	Iran-Tehran

4-2- Experimental Setup

Four experiments were performed to assess the functioning of the proposed method in the remote sensing image registration task. In the first experiment, the functioning of the matching phase is examined. In the second experiment, the registration process is evaluated. Then, images artificially contaminated by different speckle noises were used. Finally, the proposed matching method was evaluated on other detectors such as SURF, FAST and KAZE.

4-2-1- Experiments for the Matching Process

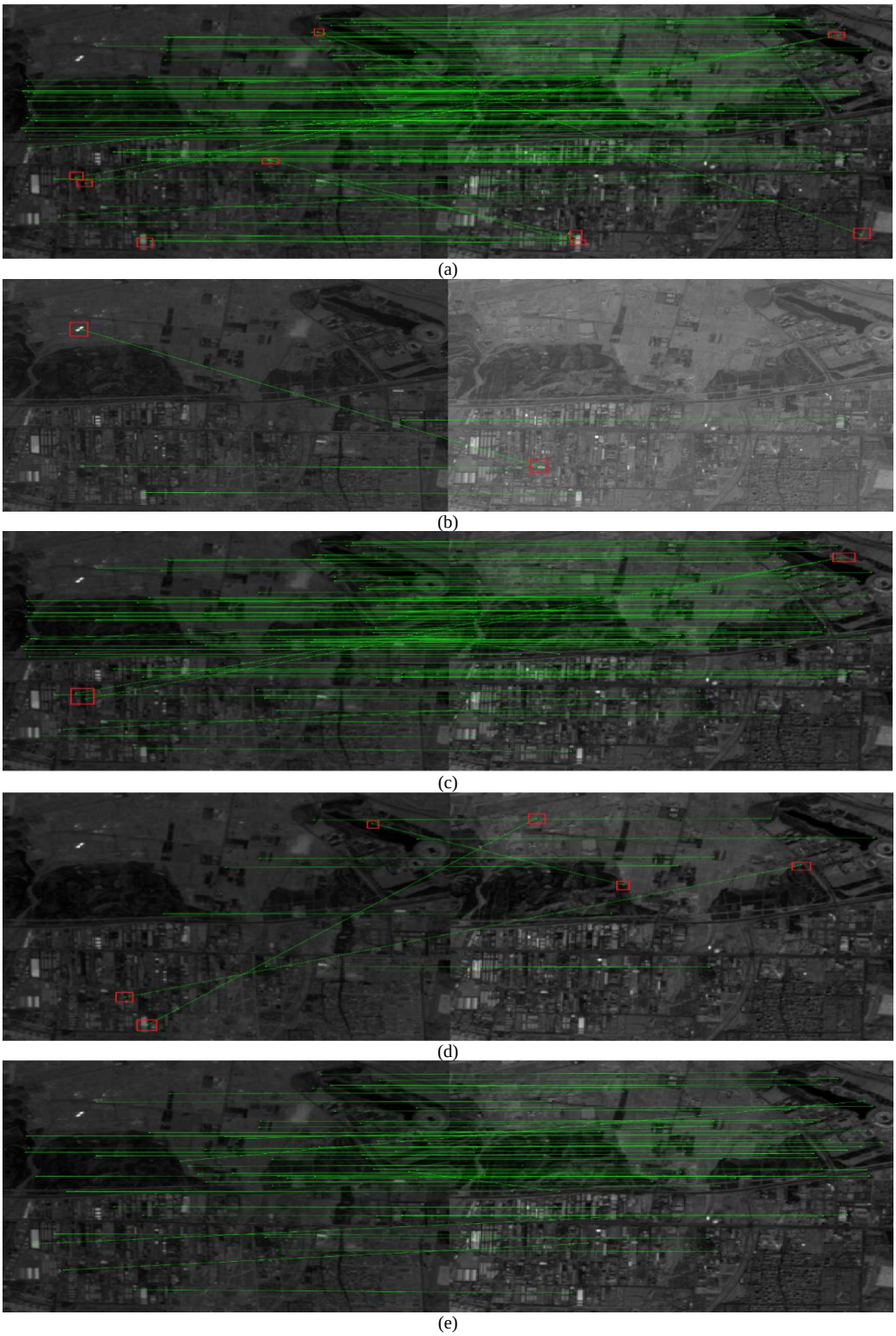
In this experiment, the database [13] was used to assess the functioning of the proposed method for the matching process; the results of which are shown in Fig. 5 and Table 2. The number of matchs in SURF-NNDR and SURF-DTM is low. In SIFT-NNDR and SIFT-SC, the number of matches is good, but there are a lot of incorrect matches. In the proposed method, compared to other methods, the number of matches is good and there is no wrong match in Fig.5. As it is clear, the proposed method had higher performance than the classic SIFT-NNDR, SURF-DTM, SURF-NNDR and SIFT-SC. In the SIFT and SURF matching algorithms, using the method based on the NNDR, the total number of matches is reduced and the number of incorrect matches is increased. However, the proposed method increased the number of correct matches and reduced the number of mismatches. Also, the results reported in Table 2 confirmed that our approach revealed higher performance rates with respect to the others. This reflects that this approach outperformed the classic algorithms (SIFT-NNDR matching, SURF-NNDR matching, SURF-DTM, Improved SUSAN) in terms of the mentioned performance criteria. As shown in Table 2, the run-time of our approach was higher than that of the SURF; but the run-time of the proposed method was shorter than the SIFT-SC and SURF-DT algorithms; since the proposed method uses a smaller-length descriptor.

4-2-2- Experiments for the Registration Process

In this test, four pairs of remote sensing images [13] are used and the functioning of the registration methods is investigated; the results are shown in Fig. 6 and Table 3. The black areas in Fig. 6 represent the incorrect registrations. These areas in the SIFT-NNDR matching and SURF-DTM algorithms were due to several incorrect matches. While, the proposed method, the registration process performed well thanks to the improvement of the matching algorithm. Based on the RMSE (as one of the important indices in the registration process), our approach was shown to do well.

4-2-3- Experiments for the Images with Simulated Noise

In this experiment, simulated remote sensing images with different speckle noises were used to evaluate the performance of the proposed method. We added a multiplicative noise to any remote sensing image I using the relation $IB = B * I$; in which B denotes a speckle noise with mean $E(B) = 1$ and different variances $v(B) = 0, 0.2, 0.4, \ldots, 2$. The results are shown in Fig. 7. From Fig. 7, it is deduced that the number of correct matches decreased with the increase of the noise variance in both SIFT-NNDR matching and the proposed method. However, the values for our approach were higher than those for the basic SIFT-NNDR matching. This validates the higher robustness of the proposed method against the speckle noise. Based on the results of this experiment and due to the persistent existence of speckle noises in natural remote sensing images, our approach is expected to work well for matching these images. Also, this technique could be generalized to other types of images such as SAR images, which are usually corrupted by speckle noises.



(a)

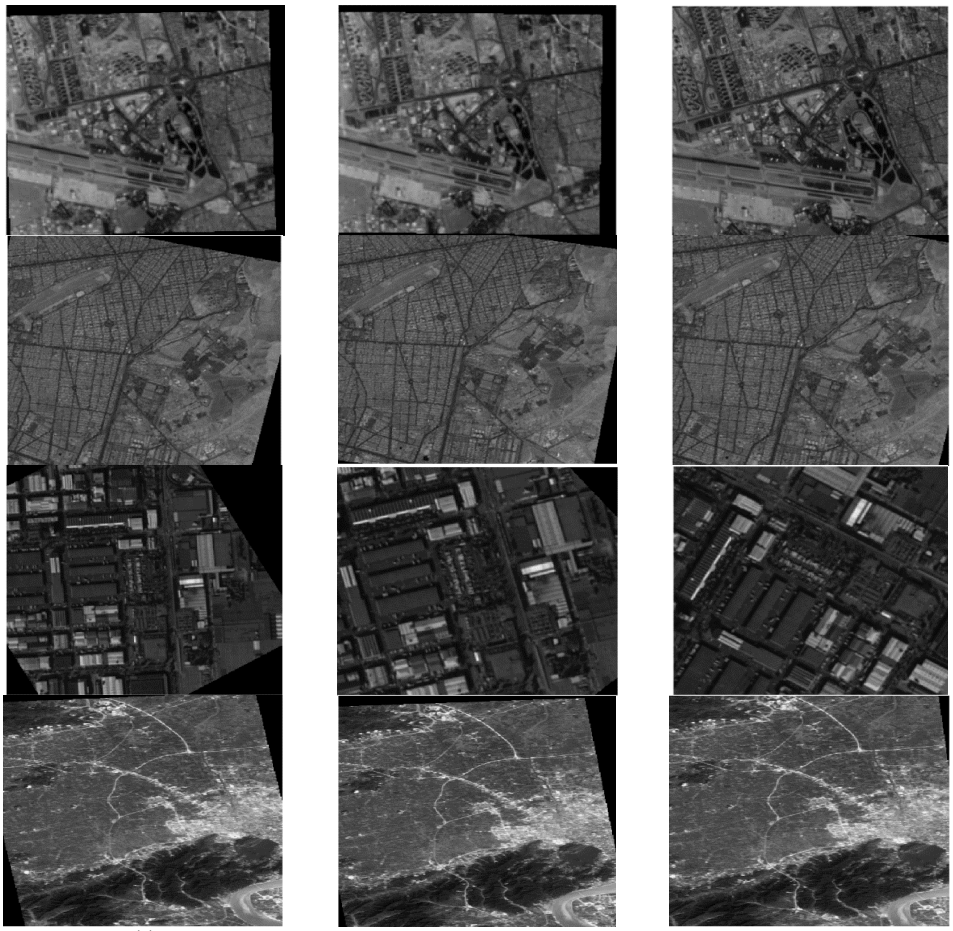
(b)

(c)

(d)

(e)

Fig. 5 Results of the remote sensing image matching; (a) the classic SIFT- NNDR matching [16]; (b) SURF-NNDR matching [17]; (c) SIFT-SC [31] (d) the SURF-DTM [23]; (e) the proposed method. The squares in this figure show the mismatches.



(a) (b) (c)

Fig. 6 Remote sensing image registration. Labels are set for the subfigure columns. (a) the SIFT-NNDR matching [16]; (b) the SURF- DTM[23]; (c) the proposed method

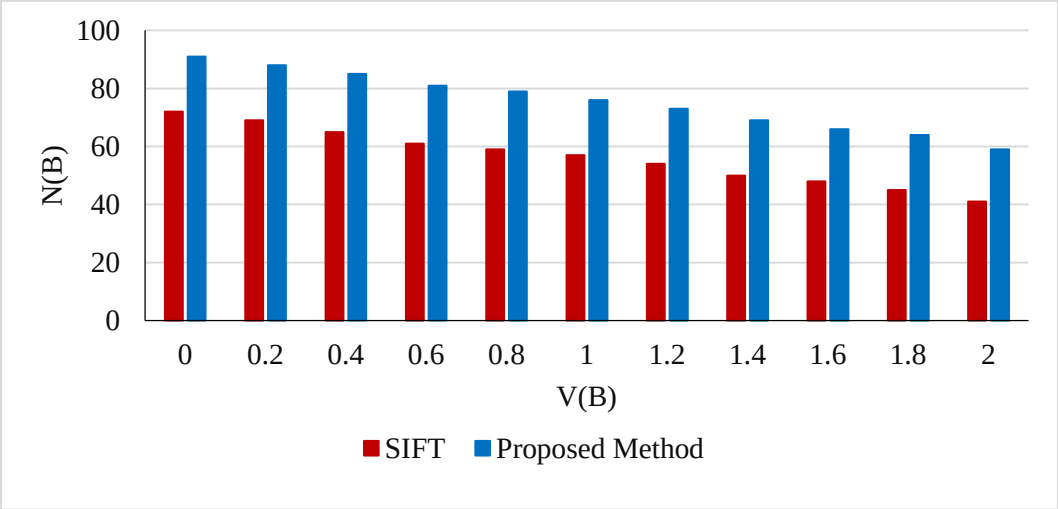


Fig. 7 Comparison of the performance of the SIFT-NNDR matching and proposed method for noisy images. The vertical and horizontal axis represent the number of correct matches and noise variances, respectively.

Table 2: Performance results of the proposed method and other methods

Method	SITMMC	SITMMR	Precision	Time (s)
SIFT **[16]**	0.69	0.31	0.69	0.52
SURF **[17]**	0.66	0.34	0.54	0.36
SURF- DTM **[30]**	0.55	0.44	0.66	0.68
Imp- SUSAN **[32]**	0.63	0.36	0.72	0.62
SIFT-SC **[31]**	0.74	0.26	0.79	0.72
The proposed method	0.90	0.09	0.92	0.65

Table 3: RMSE results of the classic methods and the proposed method

Method	RMSE
SIFT-NNDR matching **[16]**	11.43
SURF-NNDR matching **[17]**	12.02
SURF-DTM **[30]**	10.98
RKEM-SIFT **[18]**	9.82
Improved SUSAN **[32]**	8.46
SIFT-SC **[31]**	7.38
The proposed method	6.26

Table 4: Results of the proposed matching method on other detectors

Method	Precision
SURF-proposed global matching	0.73
SURF-NNDR matching	0.52
FAST- proposed global matching	0.59
FAST-NNDR matching	0.48
KAZE- proposed global matching	0.65
KAZE-NNDR matching	0.57

4.2.1 Applying the proposed Matching Method to Other Detectors

In this experiment, the functioning of the proposed matching method on other feature extraction algorithms such as SURF, FAST, KAZE was evaluated and the results are shown in Table 4. As shown in Table 4, performance of the proposed matching method is more effective on SURF, FAST and KAZE identifiers. Precision in SURF-proposed global matching and KAZE-proposed global matching is better than FAST-proposed global matching.

5- Conclusion

In this paper, a novel approach to the remote sensing image registration process was introduced using a combination of the SIFT method, the 64-D SIFT descriptor and the proposed global matching method. Initially, using the SIFT algorithm, keypoints are extracted from images and a 64-D SIFT descriptor is used to describe keypoints. Finally, a new method based on the Euclidean distances of keypoints and the transformation model is proposed to improve matching. The main goals of the proposed method were to improve the matching performance criteria and the run-time. For the future work, improvement of the next processes including remote sensing image mosaicking could be considered.

6- References

[1]A. Wong and D. Clausi, "ARRSI: automatic registration of remote-sensing images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, pp. 1483-1493, 2007.

[2]W.-J. Lee and S.-J. Oh, "Remote sensing image registration using equivariance features," in *2021 International Conference on Information Networking (ICOIN)*, 2021, pp. 776-781.

[3]Z. Hossein-Nejad and M. Nasri, "A New Method in Image Matching Based on Spatial Relationships in Multi-Sensor Remote Sensing Images " *Iranian Remote Sensing & GIS*, pp. 73-94, 2018.

[4]Y. Liu, H. Cao, Y. Zhao, Q. He, Y. Yang, L. Wang, *et al.*, "A Remote sensing image registration algorithm based on multiple constraints and a variational Bayesian framework," *Remote Sensing Letters*, vol. 12, pp. 296-305, 2021.

[5]Z. Hossein-Nejad and M. Nasri, "A Review on Image Registration Methods, Concepts and applications," *Journal of Machine Vision and Image Processing*, pp. 39-67, 2017.

[6]W. Lee, D. Sim, and S.-J. Oh, "A CNN-Based High-Accuracy Registration for Remote Sensing Images," *Remote Sensing*, vol. 13, p. 1482, 2021.

[7]H.-M. Chen, M. K. Arora, and P. K. Varshney, "Mutual information-based image registration for remote sensing data," *International Journal of Remote Sensing*, vol. 24, pp. 3701-3706, 2003.

[8]X. Xie, Y. Zhang, X. Ling, and X. Wang, "A novel extended phase correlation algorithm based on Log-Gabor filtering for multimodal remote sensing image registration," *International Journal of Remote Sensing*, vol. 40, pp. 5429-5453, 2019.

[9]M. I. Patel, V. K. Thakar, and S. K. Shah, "Image registration of satellite images with varying illumination level using HOG descriptor based SURF," *Procedia computer science*, vol. 93, pp. 382-388, 2016.

[10] P. Schwind, S. Suri, P. Reinartz, and A. Siebert, "Applicability of the SIFT operator to geometric SAR image registration," *International Journal of Remote Sensing*, vol. 31, pp. 1959-1980, 2010.

[11] G. Sreeja and O. Saraniya, "A Comparative Study on Image Registration Techniques for SAR Images," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2019, pp. 947-953.

[12] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 2941-2958, 2017.

[13] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for

optical remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 4516-4527, 2011.

[14] Z. Hossein-Nejad and M. Nasri, "A-RANSAC: Adaptive random sample consensus method in multimodal retinal image registration," *Biomedical Signal Processing and Control*, vol. 45, pp. 325-338, 2018.

[15] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey vision conference*, 1988, p. 50.

[16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, pp. 91-110, 2004.

[17] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol. 110, pp. 346-359, 2008.

[18] Z. Hossein-Nejad and M. Nasri, "RKEM: Redundant Keypoint Elimination Method in Image Registration," *IET Image Processing*, vol. 11, pp. 273-284, 2017.

[19] A. Sedaghat and N. Mohammadi, "High-resolution image registration based on improved SURF detector and localized GTM," *International Journal of Remote Sensing*, vol. 40, pp. 2576-2601, 2019.

[20] A. Sedaghat and H. Ebadi, "Remote sensing image matching based on adaptive binning SIFT descriptor," *IEEE transactions on geoscience remote Sensing Technology and Application*, vol. 53, pp. 5283-5293, 2015.

[21] M. Hasan, M. R. Pickering, and X. Jia, "Modified SIFT for multi-modal remote sensing image registration," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, 2012, pp. 2348-2351.

[22] B. Kupfer, N. S. Netanyahu, and I. Shimshoni, "An Efficient SIFT-Based Mode-Seeking Algorithm for Sub-Pixel Registration of Remotely Sensed Images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 12, pp. 379-383, 2015.

[23] Z. Yi, C. Zhiguo, and X. Yang, "Multi-spectral remote image registration based on SIFT," *Electronics Letters*, vol. 44, pp. 107-108, 2008.

[24] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, pp. 1615-1630, 2005.

[25] N. Y. Khan, B. McCane, and G. Wyvill, "SIFT and SURF performance evaluation against various image deformations on benchmark dataset," in *2011 International Conference on Digital Image Computing: Techniques and Applications*, 2011, pp. 501-506.

[26] R. Bouchiha and K. Besbes, "Comparison of local descriptors for automatic remote sensing image registration," *Signal, Image and Video Processing*, vol. 9, pp. 463-469, 2015.

[27] M. Deshmukh and U. Bhosle, "A survey of image registration," *International Journal of Image Processing (IJIP)*, vol. 5, p. 245, 2011.

[28] M. Hasan, X. Jia, A. Robles-Kelly, J. Zhou, and M. R. Pickering, "Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, 2010, pp. 1011-1014.

[29] S. Saxena and R. K. Singh, "A survey of recent and classical image registration methods," *International journal of signal processing, image processing and pattern recognition*, vol. 7, pp. 167-176, 2014.

[30] S. Chen, S. Zheng, Z. Xu, C. Guo, and X. Ma, "AN IMPROVED IMAGE MATCHING METHOD BASED ON SURF ALGORITHM," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, 2018.

[31] P. Etezadifar and H. Farsi, "A New Sample Consensus Based on Sparse Coding for Improved Matching of SIFT Features on Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[32] W. He and X. Deng, "A modified SUSAN corner detection algorithm based on adaptive gradient threshold for remote sensing image," in *2010 International Conference on Optoelectronics and Image Processing*, 2010, pp. 40-43.

[33] Z. Hossein-Nejad and M. Nasri, "An adaptive image registration method based on SIFT features and RANSAC transform," *Computers & Electrical Engineering*, vol. 62, pp. 524-537, 2017.

Zahra Hossein-Nejad received the B.Sc and M.Sc degrees in Electrical Engineering from the Islamic Azad University, Jahrom, and Sirjan branches, Iran, in 2013 and 2016, respectively. She is a Ph.D. student in Electrical Engineering at Shiraz University, Shiraz, Iran. Her research interests include image processing and computer vision.

Hamed Agahi received B.Sc, M.Sc and Ph.D degrees in Electrical Engineering from University of Shiraz, Amirkabir University of Technology and University of Tehran, Iran, in 2005, 2008 and 2013, respectively. From 2009, he was with the Islamic Azad University, Shiraz Branch, Shiraz, Iran. His research interests include pattern recognition, image and signal processing, and fault detection and diagnosis applications.

Azar Mahmoodzadeh received B.Sc, M.Sc and Ph.D degrees in Electrical Engineering from University of Shiraz, University of Shahed and University of Yazd, Iran, in 2005, 2008 and 2013, respectively. From 2009, she was with the Islamic Azad University, Shiraz Branch, Shiraz, Iran. Her research interests include pattern recognition and image and signal processing.

Error Reconciliation based on Integer Linear Programming in Quantum Key Distribution

Zahra Eskandari*
Department of Computer Engineering, Quchan University of Technology, Quchan, Iran
z.eskandari@qiet.ac.ir
Mohammad Rezaee
Department of Computer Engineering, Quchan University of Technology, Quchan, Iran
rezaee@qiet.ac.ir

Received: 14/Sep/2021 Revised: 24/Nov/2021 Accepted: 12/Dec/2021

Abstract
Quantum telecommunication has received a lot of attention today by providing unconditional security because of the inherent nature of quantum channels based on the no-cloning theorem. In this mode of communication, first, the key is sent through a quantum channel that is resistant to eavesdropping, and then secure communication is established using the exchanged key. Due to the inevitability of noise, the received key needs to be distilled. One of the vital steps in key distillation is named key reconciliation which corrects the occurred errors in the key. Different solutions have been presented for this issue, with different efficiency and success rate. One of the most notable works is LDPC decoding which has higher efficiency compared to the others, but unfortunately, this method does not work well in the codes with a high rate. In this paper, we present an approach to correct the errors in the high rate LDPC code-based reconciliation algorithm. The proposed algorithm utilizes Integer Linear Programming to model the error correction problem to an optimization problem and solve it. Testing the proposed approach through simulation, we show it has high efficiency in high rate LDPC codes as well as a higher success rate compared with the LDPC decoding method - belief propagation – in a reasonable time.

Keywords: Key reconciliation algorithm; error correction; LDPC codes; Belief Propagation; Integer Linear Programming.

1- Introduction

Quantum key distribution protocols [1,2] share a secure key between two remote parties and then establish secure communication between them using two channels; quantum channel and public channel. Quantum channel is used to carry qubits which include secret key information and after the key agreement between the two parties, Alice (A) and Bob (B), they use the secret key to establish secure communication in the public channel. It is unfortunate that the key establishment process always takes along with errors because of channel noise, device imperfection [3, 4], or the presence of eavesdropper (Eve) [5]. Thus, after transmitting the key via the quantum channel, A and B use the public channel to estimate the amount of Quantum Bit Error Rate (QBER) and correct errors in the key to establish symmetric key on both sides of the connection.

To do this, A send a key, and then at the side B, 1- the key bits are sifted, then 2- B estimates errors in sifted key bits, QBER, using the error estimation approaches [6-9]. Comparing the estimated QBER with a determined threshold, B decides that the channel is affected by channel or device noises or the presence of Eve. If QBER is higher than the determined threshold, because of the no-cloning theorem [10], it shows that Eve spoofs the connection and the key is not safe anymore. Otherwise, the next step is begun, 3- B uses a reconciliation algorithm [11-15,45-47] to correct the errors and then uses some privacy amplification methods [16,17] to remove the disclosed information along the reconciliation step.

Key reconciliation is important because using an efficient approach with minimum information leakage, in addition to increase secure key generation rate, it impacts on the security of the communication between two sides. The first error reconciliation was the BBBSS protocol [11] which uses some passes to exchange raw key subsets and check the parity of the blocks to determine and correct the errors in the subset. This approach was then improved by [12] as the Cascade algorithm. Other common approaches based on the BBBSS algorithm are Furukawa–Yamazaki [13] and Winnow protocol [14] which uses a Hamming code to reduce the number of errors.

Currently, LDPC (low-density parity check) [18] method has been widely used in this subject [15, 48, 49] as the Belief propagation algorithm (BP) [19], also known as sum-product algorithm, was used to correct the errors. This approach has attracted a lot of attention because it works more efficiently rather than others [20]. It should be

*Corresponding Author

mentioned in comparison to Cascade and Winnow, LDPC provides lower communication overhead and also it can reconcile errors at a higher rate than those.

Using artificial neural networks for error correction was introduced in [21]. The work uses the mutual synchronization of artificial neural networks to correct errors in the sifted key after the transmission in the quantum channel. Two sides create the neural networks based on the keys that they have. After the mutual learning process, they correct all errors and can use the key. But, this approach suffers from high processing time and source consumption so it has not been investigated in higher code length.

In this paper, we propose an efficient approach to reconcile errors in quantum keys distribution. As mentioned earlier, by comparing reconciliation efficiency of the three most common reconciliation algorithms, LDPC, Cascade and Winnow, it shows that the efficiency of the LDPC based reconciliation algorithm is superior to the two other in most of the QBERs. So, by considering this fact, we focus on LDPC codes approach. But as discussed later, the reconciliation based on the decoding of these codes suffers from lack of efficiency in codes with high code rate.

Indeed, the code rate has a direct influence on the amount of disclosed information in the reconciliation process. In higher rates that benefit from less leakage, LDPC approach or more specifically, BP decoding approach does not work with enough success. Focusing mentioned problem, in this paper, we decided to propose an approach to correct the errors in high rate codes. To do this we utilize Integer Linear Programming (ILP) approach.

It is noteworthy that the (Mixed) (Integer) Linear programming approaches have already been used to decoding the LDPC codes [22-25], but as it known, it is the first time that an Integer Linear Programming (ILP) model is utilized to reconcile the key in Quantum key distribution and more specifically key reconciliation algorithms. Furthermore, compared to mentioned works, the way we model the problem here is completely different in the number of variables and constraints which has a direct impact on the complexity of solving the ILP problem.

The rest of the paper proceeds as follows: in section 2 we review briefly the required preliminaries, LDPC coding concepts and ILP basics. In section 3, a detailed description of the proposed approach based on ILP is presented. The experimental results are evaluated, discussed and compared in section 4. Finally, section 5 concludes the work.

2- Preliminaries

In this section, we describe the basic concepts of this study. First, a brief overview of the LDPC codes is given, and then the concepts of (Mixed) (Integer) Linear Programming are discussed.

2-1- LDPC Codes

Low-density parity check (LDPC) codes were first introduced by Gallager in 1962 [18] as a method of transmitting a message over a noisy channel with error correction capability. Later, significant attentions were drawn to LDPC code due to its near-Shannon performance [26, 27]. The decoders for LDPC codes are based on Belief propagation (BP) algorithm and its variants [28-30]. However, BP decoding usually suffers from decreasing the success rate in presence of high error rates.

LDPC codes (n, k) can be considered as a k-dimensional subspace of $\{0,1\}^n$ which represented by a generator matrix G whose rows span code C and a parity check matrix H whose rows span $C^{\perp}$ - i.e., $c \in C$ if and only if $cH^T = 0$, where $m = n \times (1 - r)$. In LDPC codes, n and m are defined as the code and codeword length, correspondingly. The parameter r is as code rate in range $[0, 1]$ which defines the correcting power and efficiency.

When the sender wants to send vector a through noisy channel, instead of sending raw data, to have correction chance, considering G as generator matrix, she calculates codeword b as

$$b = a \times G \tag{1}$$

The symbol $\times$ corresponds to matrix (vector) multiplication in modulo 2 arithmetic.

At the channel end, the receiver receives the codeword c and use H as parity check matrix to verify that the received codeword c is error-free or not. If $c = b$, it means the error syndrome

$$S = H \times c^T \tag{2}$$

is equal to a zero vector. Otherwise, the non-zero elements of S can provide that the channel was noisy and there are some positions in the received codeword which are affected by some errors. In conditions that the occurred errors do not exceed the correction capacity of the channel code, the decoding process can correct the errors.

Decoding algorithms for LDPC codes are called message passing algorithms [19, 28-30], and work iteratively. These algorithms work based on Tanner graph [31], considering LDPC code as a graph consisting of message and check nodes corresponding to columns and the rows of the H, respectively. The reason for their name is that at each round of the algorithms, messages are passed from

message nodes to check nodes, and from check nodes back to message nodes, iteratively until the state of the graph converge to a valid codeword.
What is important about BP algorithm is that the iteration of decoding is ended when all parity check equations are satisfied and a valid codeword has been found. In some cyclic graphs, or when errors are higher than they can be decoded correctly, more iterations do not change the state of the graph and errors cannot be corrected. So, decoding fails without finding the correct codeword as the output.
For more details about LDPC codes, the readers can refer to [18, 19].

2-2- Optimization Approaches

Combinatorial optimization deals with the problem of minimizing or maximizing a function of several variables subject to some constraints. The constraints can be in the form of equality or inequality, linear or nonlinear. If the constraints are linear, the feasible region is a convex polyhedron, which has a global minimum, and the solving methods can converge to this optimal point. But if the constraints are nonlinear, solving the problem has some complexity and requires more provisions [32].
Linear programming (LP) problem is defined as follows:

$$\min\{\mathrm{Cx} \mid \mathrm{Ax} \leq \mathrm{b}, \mathrm{x} \in \mathrm{R}^{\mathrm{n}}\} \quad (3)$$

The $Ax \leq b$ inequalities determine the feasible region which is bounded by the system of constraints, including the possible values of the variables that satisfy all of the constraints. The aim of linear programming is to find the best solution to a problem by maximizing or minimizing the objective function Cx. As shown in equation (3), the objective function and all constraints are linear. It must be mentioned that the complexity of solving the LP problem is polynomial [33].
If the variables are integer, we have Integer Linear Programming (ILP) problem as follows:

$$\min\{\mathrm{Cx} \mid \mathrm{Ax} \leq \mathrm{b}, \mathrm{x} \in \mathrm{Z}^{\mathrm{n}}\} \quad (4)$$

As the variables are discrete in these problems, the solving methods differ and consequently the complexity of solving is affected such that in general, the time complexity of solving these problems is exponential [33].
However, some works have been done and some improvements have been achieved. A prominent one is LP relaxation which by analyzing the problem constraints, some conversions could be performed [31] or some new constraints may be added to the problem [34] and convert the ILP problem to LP one with polynomial solving time which is practical way to solve large size problem.
In addition to this naïve approach, Branch-and-bound [35] and cutting-plane [36] methods have been principle tools for solving ILP models in recent times. Both of them deal with the models by solving a sequence of LP problems by simplex methods [33]. It will be clear that only finitely many LP problems need to be solved in principle.
It is noteworthy that because of great industrial interest in optimization applications, there exist many well-develop solvers such as IBM ILOG Cplex [40], Gurobi [41], and so forth which employ proper techniques to solve large problems with reasonable complexity in terms of time and space.

3- Proposed Error Correction Approach based on ILP Model

In this section, we present the detail of the proposed approach to correct errors in the sifted key based on the Integer Linear Programming approach.
Suppose A and B prepare H and share it. A after choosing the key X_A, calculates syndrome S_A as equation (2) and send them to B via quantum and public channel, respectively.
After B obtains the sifted key, X_B and syndrome of A, S_A, through mentioned channels, because of the presence of the noise, indeed he received a noisy version of X_A:

$$X_B = X_A \oplus e \quad (5)$$

Where e is the noise vector and $\oplus$ used for summation operation in module 2.
Then B calculates the syndrome of X_B as equation (2), S_B using H and X_B. Comparing S_A and S_B, he gets some information about error occurrence. In a more precise view, he has:

$$S_B = H \times X_B = H \times (X_A \oplus e) = H \times X_A \oplus H \times e = S_A \oplus H \times e \quad (6)$$

Which equals to:

$$H \times e = S_A \oplus S_B \quad (7)$$

Where the symbol $\times$ and $\oplus$ corresponds to matrix multiplication and summation operation in modulo 2 arithmetic.
B should determines e vector as it satisfies equation (7). Indeed along with decoding approaches aim to find the nearest codeword to the received one, B should determines e vector with minimum weight. In the other word, B searches for a codeword with minimum Hamming distance to received codeword to decode it.
So B can model the problem to an optimization problem aiming to find an error vector with minimum weight to satisfy equation (7). To do this, he defines variables, constraints, and objective function based on the nature of the original problem. After modeling the problem, using

solving tools, B solves it to find the solution and determine the error vector e. Defining e as:

$$e = (e_1, e_2, \dots, e_n) | e_i = \{0,1\}, i = 1, \dots, n \quad (8)$$

Where e_i value 1 or 0 corresponds to the presence or absence of noise at the position i in the received key, X_B. As B tries to find nearest codeword to the received key X_B which equally means he should find e vector with minimum weight. So, the objective function is defined as follow:

$$\min \sum_{i=1}^{n} e_i \quad (9)$$
$$e_i = \{0,1\}, i = 1, \dots, n$$

To find the feasible region for possible values of noise vector, it should be mentioned that the noise vector should be satisfied by the equation (7). Each row of the $H(H_j)$ behaves like a parity check equation that must be satisfied. Given the equation (7), the variable b, as a m binary vector, is defined as:

$$b = S_A \oplus S_B \quad (10)$$

Therefore, the constraints could be shown as:

$$H_j \times e = b_j, \quad j = 1, \dots, m \quad (11)$$

More precisely, the constraints are as follows:

$$\oplus_{i=1}^{n} H_{ji} * e_i = b_j, \quad j = 1, \dots, m \quad (12)$$

Where $*$ used for multiplication operation over integer numbers and by $\oplus_{i=1}^{n} H_{ji} * e_i$, we mean $H_{j1} * e_1 \oplus H_{j2} * e_2 \oplus \dots \oplus H_{jn} * e_n$.

Thus the model is as follow:

$$\min \sum_{i=1}^{n} e_i \quad (13)$$
$$\oplus_{i=1}^{n} H_{ji} * e_i = b_j, \quad j = 1, \dots, m$$
$$e_i = \{0,1\}, i = 1, \dots, n$$

As seen in the problem, the variables are binary and the constraints are using module 2 multiplication operation. So, here we convert the module 2 constraints to constraints over Z by doing as follows: in [42], The Integer Adapted Standard Conversion Method (IASC) was proposed to apply to Boolean polynomials.

This method is based on the ability of presenting an equation modulo 2 as an equation over Z and works as follows: considering a Boolean polynomial f over $F2$ and assume this polynomial as a polynomial g over the integers by replacing XOR by addition. All solutions of the Boolean equation $f = 0$ will yield a multiple of 2 when plugged into g. Thus, for $x \in \{x | f(x) = 0\}$ it holds that $g(x) = 2 \cdot k$. Then we obtain an integer equation by subtracting a multiple of 2 from g: $g - 2 \cdot k = 0$ where $0 \leq k$.

As an example, consider the Boolean equation

$$x1 + x2 + x3 + x4 = 0 \quad (14)$$

If we evaluate the corresponding real-valued polynomial $y1 + y2 + y3 + y4 = 0$ for all solutions of (14), we get 0, 2, 4 as results. That means that a solution of (14) is a solution to the following equation over the integers

$$y1 + y2 + y3 + y4 - 2.k = 0 \quad (15)$$

where $k \in \{0, 1, 2\}$ and $yi \in \{0, 1\}$ for $i = 1 \dots 4$.

The number of variables per equation is increased only by one compared to the Boolean polynomial.

ILP formulation over Z: In order to conquer Boolean constraints, we use the IASC conversion method to convert the model to ILP one. Indeed, we convert summation in binary to integer space, so we define new variables as:

$$k_j, \quad k_j \in Z^+, j = 1 \dots m \quad (16)$$

Thus, the proposed ILP model for error correction problem is as follows:

$$\min \sum_{i=1}^{n} e_i$$
$$\oplus_{i=1}^{n} H_{ji} * e_i - 2 * k_j = b_j, \quad j = 1, \dots, m$$
$$e_i = \{0,1\}, i = 1, \dots, n,$$
$$k_j \in Z^+, j = 1 \dots m \quad (17)$$

Computational and Space Complexity Analysis: The proposed ILP model, has $n + m$ variables, m bits for error vector and n integer variables, and m linear constraints corresponds to the rows of parity check matrix and n integer constrains for error vector bits. As mentioned before, solving this type of problems is known to be NP-hard in general [51]. It should be noted that decoding methods and more specifically BP algorithms are known as NP-hard problems [50] and in addition, the performance of these algorithms depend on the iteration numbers.

Though LDPC codes were constructed using a sparse Tanner graph, so the corresponding generator and parity check matrices were sparse, too. Such as in this work, we model the error correction in LDPC codes as ILP, so the sparsity property of H parity matrix causes the sparsity in the constraints set and we have a sparse ILP. Recently some works [37, 38] have been done on Sparse Integer Linear Programming (SILP), the case that the coefficient

matrix is sparse. It was shown that SILP can be solved in polynomial time in such problems [39].

4- Evaluation of the Results

To evaluate the proposed ILP approach for error correction, in this section, we provide the detailed comparisons of the original BP algorithm, Multi-matrix BP (MBP) [8] and ILP proposed approach in different parameters; efficiency, success rate, and speed as three criteria for judging a key reconciliation algorithm. All simulation data are generated by random scenarios. In Multi-matrix BP (MBP) [8] approach, in each iteration multiple matrices were employed to generate more useful information in error correction. As claimed by the authors the iteration number falls and the convergence speed increases. Indeed cycles which appear in one matrix and reduce the success rate, could be eliminated by other matrices.

For comparisons, we use 4 pool standard LDPC codes [43, 44] with code length $n = 1944$ with different code rates $r = \{0.5, 0.6667, 0.75, 0.8333\}$.

To evaluate the performance of ILP proposed approach, we compare it with recent approaches to correct errors, BP and MBP. To do this, we generate different sets of keys at QBERs in range [1,1.7] step by 0.1. At a given QBER, we generate 100 random scenarios of keys, perform each approach on the keys, and present the results as the average over the random scenarios for each QBER. The simulation parameters were presented in table 1.

Table 1: simulation parameters

Parameter	***Value***
Code length	1944
Code rates	0.5, 0.6777, 0.75, 0.8333
QBER range	[1,1.7]
Number of random keys	100

It should be mentioned that the ILP, BP and MBP approaches were implemented in Python programming language and all the experiments were done on Intel (R) core (TM) i7-9750H CPU @ 2.60 GHz with 16 GB memory.

4-1- Efficiency of LDPC-based Approach

As the most important factor in reconciliation algorithms, we can name reconciliation efficiency f which shows the relation of the amount of information B obtains to the minimum amount of information he needs for correcting all errors that theoretically calculated. Therefore, to ensure that he can correct all errors, f must be greater than or equal to 1, i.e. $f \geq 1$. In theory, $f = 1$ happens when LDPC code tends to be infinite in length and no cycles in structure, which can reach the Shannon Limit [9]. So in practice, $f > 1$. The reconciliation efficiency f which implies the efficiency and security of a reconciliation strategy, is calculated as:

$$f = \frac{m}{n.h(e)} \tag{17}$$

where m and n are the numbers of check nodes and variable nodes of the corresponding Tanner graph of LDPC code, equivalent by number of rows and columns of parity check matrix H, e is the result of Quantum Bit Error Rate estimation, and $h(e)$ is the Shannon binary entropy represented as:

$$h(e) = -e \log_2 e - (1-e) \log_2 (1-e) \tag{18}$$

In figure (1), we present the efficiency of the LDPC codes for different code rates in QBERs in range $[1, 1.7]$. As seen, by increasing the QBER, the efficiency of LDPC codes in all code rate scenarios was decreased. But, in higher code rate, the efficiency of LDPC codes gets closer to 1.

Indeed under same value of n, in a high code rate compared to lower code rate, m is smaller which means the sender needs to send syndrome with smaller length. So based on equation (17), the f value gets closer to 1 meaning that it needs to disclose a lower amount of data.

So by using LDPC codes with high code rates in key reconciliation algorithms, we can achieve better efficiency as well as a higher secure key generation rate.

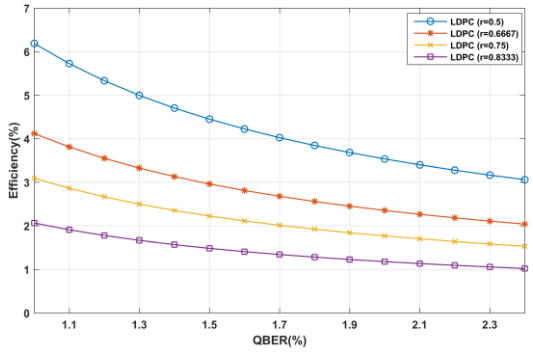


Fig. 1 Comparison of reconciliation efficiency of LDPC codes with different code rate values at different QBER.

4-2- Success Rate

The success rate of reconciliation algorithms shows the number of the successful scenarios that the reconciliation algorithm can correct the errors. Unfortunately, the success rate of the BP algorithm may be relatively impacted by cycles which means if LDPC codes are not cycle-free [31], it cannot get to the corrected answer even by consuming more time or with running the message passing in more iterations. So in these situations, the success rate of the BP is decreased.

In higher code rate, because of the lower number of check nodes in comparison to message nodes in Tanner graph of LDPC codes, the probability of successful decoding and correcting the errors is decreased, so as a result, the success rate is decreased too.
As shown in figure (2), in the code rate $r = 0.6667$, the BP algorithm works well and it could decode all the corrupted received keys even in high QBER values, and it achieve 100% success rate. But, in a higher code rate $r = 0.8333$ which the number of check nodes, m is decreased compared to $r = 0.6667$, the BP falls in wrong decoding or fails to reach the result. So the success rate of BP decreased drastically by increasing the QBERs.
So we can conclude that the BP algorithm is not proper enough to correct errors in high rate LDPC codes and we should look forward to a more successful approach to correct errors in such codes.

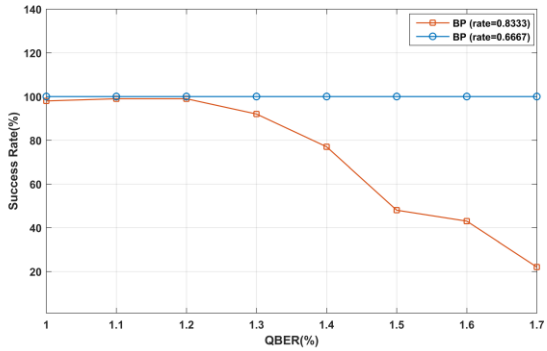


Fig. 2 Comparison of success rate of BP algorithm with different code rate values at different QBER.

About ILP-based error correction proposed algorithm, in low code rate $r = 0.6667$, the proposed approach can correct the errors in all scenarios in reasonable time and it has a success rate 100% for all QBERs in range $[1, 1.7]$ as well as the BP at code rate $r = 0.6667$.

But in contrast to BP algorithm, as shown in figure (3), at code rate $r = 0.8333$, the ILP proposed approach has more success in comparison to the BP and MBP algorithms. Indeed by analyzing the results, it reveals that in all scenarios which BP gets the results, the ILP proposed approach gets the correct result too, and in addition, in most of the failed scenarios by BP, ILP approach can solve the problem and correct the errors successfully. In fact we can conclude that in such problems that the cycles could degrade the performance of the decoding, ILP approach can step forward and correct errors.

4-3- Reconciliation Speed

For key reconciliation, the convergence speed is calculated as the required time to correct the errors. Here we evaluate the convergence speed of BP, MBP and ILP proposed algorithms in error correction of the LDPC codes with different code rate $r = \{0.6667, 0.8333\}$ by calculating the consumed time to perform error correction under different QBERs. At a given QBER, we generate 100 sets of keys, perform each algorithm on the keys, and calculate the average amount of the required time for error correction. The results are shown in figure (4) and figure (5) for different code rates. Clearly, under different code rates, the algorithms spend reasonable time for correcting errors. As shown in figure (4) in code rate $r = 0.6667$, by increasing the QBER, the speed of the correction algorithm grows slightly. Indeed, enough number of check nodes in comparison with message nodes causes the algorithms to obtain the solution and correct the errors in the low time. As seen in figure (4) and (5), since MBP did the decoding computation for some matrices, even it achieves less iteration number, it consumes much more time in comparison to BP with only one matrix.

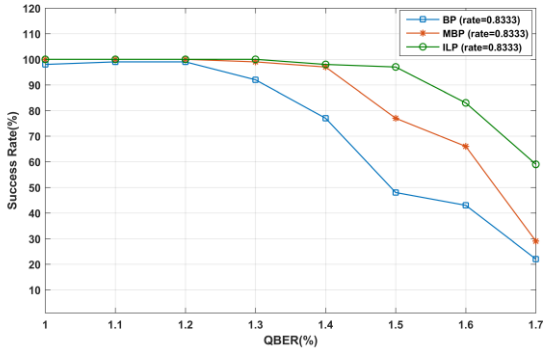


Fig. 3 Comparison of success rate of BP algorithm and Proposed ILP model in error correction of LDPC code with r=0.8333 at different QBER.

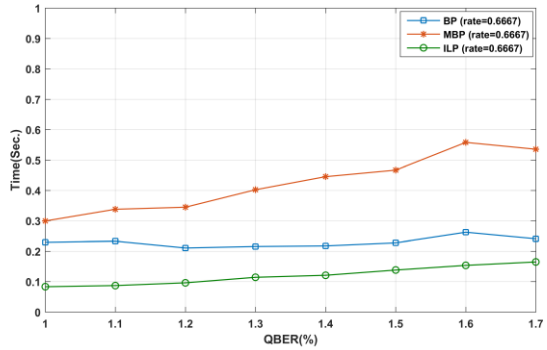


Fig. 4 Comparison of speed of BP algorithm and Proposed ILP model in error correction of LDPC code with r=0.6667 at different QBER.

But as seen in figure (5), which is correspond to the code rate $r = 0.8333$ with the lower number of check nodes at the same number of message nodes compared to the previous scenario, determining the right codeword requires more time in both BP approaches and ILP proposed algorithms. So the consumed time is increased compared to the code rate $r = 0.6667$.

It is notable that the solving time of the ILP problem is less than the required time for BP algorithms in both of the mentioned code rates. The increase in the consumed time of ILP at the high values of QBER in the figure (5) is caused by the fact that in most of the scenarios with high QBER that the BP algorithm fails, the ILP purposed model continue and solve the problem. considering the difference in the success rate of these algorithms presented at figure (3) can justify the results in the figure (5).

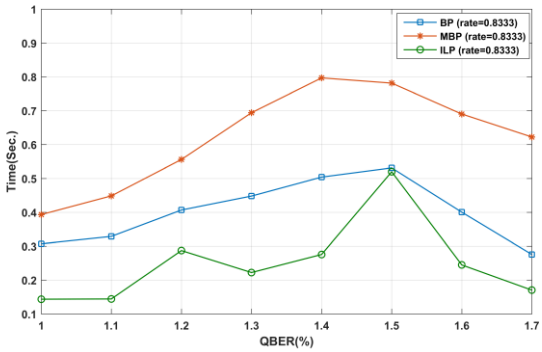


Fig. 5 Comparison of speed of BP algorithm and Proposed ILP model in error correction of LDPC code with r=0.8333 at different QBER.

5- Conclusion and Future Works

Quantum key distribution protocols share a secure key to using the Quantum channel. Because of the unavoidability of the existence of noise, distillation algorithms are necessary to purify the key. As one step in the distillation process, key reconciliation has the responsibility to correct errors of the key efficient manner.

Comparing reconciliation efficiency of the reconciliation algorithms, LDPC code-based reconciliation algorithms have revealed the higher efficiency but as the code rate grows, the success rate of the most used decoding algorithm, belief propagation, decreased considerably. Besides the fact that in such codes, the amount of disclosed information was decreased. So to use this helpful feature of high rate LDPC codes, we have to overcome this problem. In this paper, focusing on high rate LDPC codes, we propose an approach to correct the errors in such codes. The proposed approach utilizes Integer Linear Programming (ILP) approaches. To do this, we model the error correction problem, by defining the variables, constraints, and objective function corresponding to the reconciliation algorithm aim. Then to have more efficient modeling, we convert the binary model to a model over Z. So the final ILP model is defined over Z and it has sparsity property, which led to having an efficient model in terms of time and space solving complexity to have with reasonable solving time.

Finally, by evaluating the proposed algorithm at the crucial parameters for judging the efficiency of the reconciliation algorithms, our approach is superior to the BP algorithm in high rate codes regarding success rate and reconciliation speed in different rates and different QBERs.

As future work, by considering scalability the presented optimization model can be improved, so it can be utilized more efficiently in problems with more variables and constraints. By utilizing this improvement it can be employed to perform error correction in codes with longer length.

References

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, “Quantum cryptography,” Rev. Mod. Phys. 74(1), 145–195 (2002).

[2] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, “The security of practical quantum key distribution,” Rev. Mod. Phys. 81(3), 1301–1350 (2009).

[3] H. Weier, H. Krauss, M. Rau, M. Fuerst, S. Nauerth, and H. Weinfurter, “Quantum eavesdropping without interception: an attack exploiting the dead time of single photon detectors,” New J. Phys. 13(7), 073024 (2011).

[4] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs, “Device calibration impacts security of quantum key distribution,” Phys. Rev. Lett. 107(11), 110501 (2011).

[5] C. H. Bennet and G. Brassard, “Quantum cryptography: public key distribution and coin tossing,” in Proceedings of the IEEE International Conference on Computers Systems and Signal Processing (IEEE, 1984), pp. 175–179

[6] X.B. Wang, “Beating the photon-number-splitting attack in practical quantum cryptography,” Phys. Rev. Lett. 94(23), 230503 (2005).

[7] P. Treeviriyanupab, T. Phromsaard, C.M. Zhang, M. Li, P. Sangwongngam, T. S. N. Ayutaya, N. Songneam, R. Rattanatamma, C. Ingkavet, W. Sanor, W. Chen, Z.F. Han, and K. Sripimanwat, “Rate-adaptive reconciliation and its estimator for quantum bit error rate,” in Proceedings of International Symposium on Communications and Information Technologies (IEEE, 2014), pp. 351–355.

[8] C. Gao, J. Dong, G. Yu, L. Chen, Multi-matrix error estimation and reconciliation for quantum key distribution. Optics Express. (2019). 27. 14545. 10.1364/OE.27.014545.

[9] C. Gao, Y. Guo, D. Jiang, L. Chen, Multi-matrix rate-compatible reconciliation for quantum key distribution. ArXiv(2020)., abs/2001.01074.

[10] Wootters, W.K., Zurek, W.H.: A single quantum cannot be cloned. Nature 299, 802–803 (1982)

[11] Bennett, C.H., Bessette, F., Brassard, G., Salvail, L., Smolin, J.: Experimental quantum cryptography. J. Cryptol. 5, 3–28 (1992)

[12] Brassard, G., Salvail, L.: Secret-Key Reconciliation by Public Discussion, pp. 410–423. Springer, Berlin (1994)

[13] Furukawa, E., Yamazaki, K.: Application of existing perfect code to secret key reconciliation. In: Proceedings of International Symposium on Communication and Information Technologies, pp. 397– 400 (2001)

[14] Buttler, W.T., Lamoreaux, S.K., Torgerson, J.R., Nickel, G.H., Donahue, C.H., Peterson, C.G.: Fast, efficient error

reconciliation for quantum cryptography. Phys. Rev. A 67, 052303 (2003)

[15] E. Kiktenko, A. Malyshev, A. Bozhedarov, N. Pozhar, M. Anufriev, and A. Fedorov, "Error estimation at the information reconciliation stage of quantum key distribution," J. Russ. Laser Res. 39(6), 558–567 (2018).

[16] C. H. Bennett, G. Brassard, and J.M. Robert, "Privacy amplification by public discussion," SIAM J. Comput. 17(2), 210–229 (1988).

[17] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," IEEE Trans. Inf. Theory 41(6), 1915–1923 (1995).

[18] R. G. Gallager, Low Density Parity-Check Codes. MIT Press, Cambridge, MA, 1963.

[19] S. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," IEEE Trans. Inf. Theory 47(2), 657–670 (2001).

[20] Mehic M., Niemiec M., Siljak H., Voznak M. (2020) Error Reconciliation in Quantum Key Distribution Protocols. In: Ulidowski I., Lanese I., Schultz U., Ferreira C. (eds) Reversible Computation: Extending Horizons of Computing. RC 2020. Lecture Notes in Computer Science, vol 12070. Springer, Cham.

[21] Niemiec, M. Error correction in quantum cryptography based on artificial neural networks. Quantum Inf Process 18, 174 (2019). https://doi.org/10.1007/s11128-019-2296-4

[22] J. Feldman, "Decoding Error-Correcting Codes via Linear Programming". PhD thesis, M.I.T., Cambridge, MA, 2003

[23] K. Yang, X. Wang, and J. Feldman, "A new linear programming approach to decoding linear block codes," IEEE Trans. Inf. Theory, vol. 54, no. 3, pp. 1061–1072, Mar. 2008.

[24] H. Wei and A. H. Banihashemi, "An iterative check polytope projec tion algorithm for ADMM-based LP decoding of LDPC codes," IEEE Commun. Lett., vol. 22, no. 1, pp. 29–32, Jan. 2018.

[25] J. Bai, Y. C, Wang, and F. C. M. Lau, "Minimum-polytope-based linear programming decoder for LDPC Codes via ADMM approach", IEEE Wireless Commun. Lett., vol. 8, no. 4, pp. 1032-1035, Aug. 2019

[26] D. J. C. MacKay and R. M. Neal, "Good codes based on very sparse matrices," in Cryptography and Coding, ser. Lecture Notes in Computer Science, C. Boyd, Ed. Heidelberg/Berlin: Springer, 1995, vol. 1025, pp. 100-111.

[27] D. J. C. MacKay and R. M. Neal, "Near Shannon-limit performance of low density parity check codes," Electron. Lett., vol. 33, no. 6, pp. 457- 458, Mar. 1997.

[28] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," IEEE Trans. Inf. Theory, vol. 47, no. 2, pp. 498-519, Feb. 2001. 16

[29] W. Ryan and S. Lin, Channel Codes: Classical and Modern. Cambridge University Press, 2009.

[30] T. Richardson and R. Urbanke, Modern Coding Theory. Cambridge University Press, 2008.

[31] R. M. Tanner, "A recursive approach to low complexity codes," IEEE Trans. Inf. Theory 27(5), 533–547 (1981).

[32] L. A. Wolsey and G. L. Nemhauser: Integer and Combinatorial Optimization Wiley-Interscience, November 1999.

[33] Clovis C. Gonzaga, On the Complexity of Linear Programming, Resenhas IME-USP 1995, Vol. 2, No. 2, 197-207.

[34] EgonBalas, Sebastián Ceria, Gérard Cornuéjols: A lift-and-project cutting plane algorithm for mixed 0–1 programs, Mathematical Programming, Volume 58, January 1993, pp 295-324

[35] H. Land , A. G. Doig, An Automatic Method of Solving Discrete Linear Programming Problems, July 1960, Econometrica 28(3):497-520

[36] Ralph Gomory, Outline of an Algorithm for Integer Solutions to Linear Programs, September 1958, Bulletin of the American Mathematical Society 64(5):275-278

[37] Pritchard, D., Chakrabarty, D. Approximability of Sparse Integer Programs. Algorithmica 61, 75–93 (2011). https://doi.org/10.1007/s00453-010-9431-z

[38]Andres Iroume, SPARSITY IN INTEGER PROGRAMMING. PhD thesis, Georgia Institute of Technology, 2017

[39] Koutecký, Martin; Levin, Asaf; Onn, Shmuel (2018). A Parameterized Strongly Polynomial Algorithm for Block Structured Integer Programs. Michael Wagner: 14 pages. arXiv:1802.05859. doi:10.4230/LIPICS.ICALP.2018.85. S2CID 3336201.

[40]www.ibm.com/software/commerce/optimization/cplex-optimizer/

[41] www.gurobi.com/

[42] J. Borghoff, Mixed-integer Linear Programming in the Analysis of Trivium and Ktantan, IACR Cryptol. ePrint Arch. 2012.

[43] EEE Standard for Information Technology—Local and Metropoli tan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Spec ifications Amendment 5: Enhancements for Higher Throughput, IEEE Standard 802.11n-2009, Oct. 2009

[44] Elkouss, D., Martinez-Mateo, J. & Martin, V. Untainted Puncturing for Irregular Low-Density Parity-Check Codes. IEEE Wireless Communications Letters 1, 585–588 (2012).

[45] Guo, D., He, C., Guo, T. et al. Comprehensive high-speed reconciliation for continuous-variable quantum key distribution. Quantum Inf Process 19, 320 (2020).

[46] E. O. Kiktenko, A. O. Malyshev and A. K. Fedorov, "Blind Information Reconciliation With Polar Codes for Quantum Key Distribution," in IEEE Communications Letters, vol. 25, no. 1, pp. 79-83, Jan. 2021.

[47] Liu, Z., Wu, Z. & Huang, A. Blind information reconciliation with variable step sizes for quantum key distribution. Sci Rep 10, 171 (2020). https://doi.org/10.1038/s41598-019-56637-y

[48] K. Zhang, X. -Q. Jiang, Y. Feng, R. Qiu and E. Bai, "High Efficiency Continuous-Variable Quantum Key Distribution Based on Quasi-Cyclic LDPC Codes," 2020 5th International Conference on Communication, Image and Signal Processing (CCISP), 2020, pp. 38-42, doi: 10.1109/CCISP51026.2020.9273490.

[49] B. Bilash, B. K. Park, C. Hoon Park and S. -W. Han, "Error-Correction Method Based on LDPC for Quantum Key Distribution Systems," 2020 International Conference on Information and Communication Technology Convergence

(ICTC), 2020, pp. 151-153, doi: 10.1109/ICTC49870.2020.9289451.
[50] Georgios Papachristoudis, John W. Fisher, Adaptive Belief Propagation, 32th International Conference on Machine Learning, Lille, France, 2015. JMLR: W&CP volume 37.
[51] Daniel Lokshtanov, New Methods in Parameterized Algorithms and Complexity, Dissertation for the degree of Philosophiae Doctor (PhD) University of Bergen Norway April 2009.

Zahra Eskandari received the B.S degree in Computer Engineering from Kharazmi University, Tehran, Iran, in 2006. She received her M.S. and PhD. degrees in Computer Engineering from Ferdowsi University of Mashhad, Iran, in 2008 and 2020, respectively. She was with the cybersecurity section at DTU compute, Denmark as a visiting researcher from July 2016 to March 2017. She is a full-time Assistant-Professor in the Department of Computer Engineering at Quchan University of Technology, Iran. Her research interests include security and cryptography. She is particularly interested in algebraic cryptanalysis and optimization approaches.

Mohammad Rezaee received the PhD degree in computer engineering from Ferdowsi University of Mashhad, Mashhad, Iran, in 2019. Currently, he is an assistant professor at the Computer Engineering Department, Quchan University of Technology. His research interests include Smart Grid Communication, and Optimization of Communication Networks.