

JOURNAL OF SCIENCE



SAKARYA UNIVERSITY

Sakarya University Journal of Science



SAKARYA
UNIVERSITY

e-issn: 2147-835X

Volume 23 Issue 1

FEBRUARY 2019

Sakarya University Journal of Science
Volume: 23 Issue: 1 February 2019
Editorial Boards

Editor-in-Chief

Emrah Dođan, Civil Engineering, Sakarya University (Turkey)
email: emrahd@sakarya.edu.tr

Editors

Ahmet ađatay ilingir, Mechanical Engineering, Sakarya University (Turkey)
email: cilingir@sakarya.edu.tr

Alparslan Serhat Demir, Industrial Engineering, Sakarya University (Turkey)
email: alparslanserhat@sakarya.edu.tr

Beytullah Eren, Environmental Engineering, Sakarya University (Turkey)
email: beren@sakarya.edu.tr

Cüneyt Bayılmış, Computer Science, Sakarya University (Turkey)
email: cbayilmis@sakarya.edu.tr

Ertan Bol, Civil Engineering, Sakarya University (Turkey)
email: ebol@sakarya.edu.tr

Kerem Küçük, Computer Science, Kocaeli University (Turkey)
email: kkucuk@kocaeli.edu.tr

Mehmet Nebiođlu, Chemistry, Sakarya University (Turkey)
email: nebioglu@sakarya.edu.tr

Meral Demirtaş, Meteorology Engineering, Ondokuz Mayıs University (Turkey)
email: mdemirtas@omu.edu.tr

Naci ađlar, Civil Engineering, Sakarya University (Turkey)
email: caglar@sakarya.edu.tr

Sadık Bađcı, Physics, Sakarya University (Turkey)
email: sbagci@sakarya.edu.tr

Serkan Zeren, Mechatronics Engineering, Kocaeli University (Turkey)
email: serkan.zeren@kocaeli.edu.tr

Şevket Gür, Mathematics, Sakarya University (Turkey)
email: sgur@sakarya.edu.tr

Zafer Barlas, Metallurgical and Materials Engineering, Sakarya University of Applied Sciences (Turkey)
email: barlas@sakarya.edu.tr

Editorial Board

Ahmet Aygün, Environmental Engineering, Bursa Technical University (Turkey)
email: ahmet.aygun@btu.edu.tr

Ali Çoruh, Physics, Sakarya University (Turkey)
email: coruh@sakarya.edu.tr

Ali Pınar, Civil Engineering, Boğaziçi University (Turkey)
email: pinara@boun.edu.tr

Ali Sarıbiyık, Civil Engineering, Sakarya University of Applied Sciences (Turkey)
email: alisaribiyik@sakarya.edu.tr

Azim Gökçe, Metallurgical and Materials Engineering, Sakarya University of Applied Sciences (Turkey)
email: azimg@sakarya.edu.tr

Barış Yüce, Industrial Engineering, University of Exeter (England)
email: b.yuce@exeter.ac.uk

Bengü Bayram, Mathematics, Balıkesir University (Turkey)
email: benguk@balikesir.edu.tr

Berrin Denizhan, Industrial Engineering, Sakarya University (Turkey)
email: denizhan@sakarya.edu.tr

Çiğdem Gündüz, Mathematics, Kocaeli University (Turkey)
email: caras@kocaeli.edu.tr

Dilek Angın, Food Engineering, Sakarya University (Turkey)
email: angin@sakarya.edu.tr

Emrah Evren Kara, Mathematics, Düzce University (Turkey)
email: eevrenkara@duzce.edu.tr

Emre Tabar, Physics, Sakarya University (Turkey)
email: etabar@sakarya.edu.tr

Erman Aslan, Mechanical Engineering, Istanbul University (Turkey)
email: erman.aslan@istanbul.edu.tr

Esra Erkuş Duman, Mathematics, Gazi University (Turkey)
email: eerkusduman@gmail.com

Giglou Abolfazl, Civil Engineering, Uidaho University (USA)
email: abolfazl@uidaho.edu

Habibullah Uzun, Environmental Engineering, Marmara University (Turkey)
email: habibullah.uzun@marmara.edu.tr

Halil Yiğit, Computer Science, Kocaeli University (Turkey)
email: halilyigit@kocaeli.edu.tr

Hüseyin Aksoy, Biology, Sakarya University (Turkey)
email: haksoy@sakarya.edu.tr

Ipek Güleç, Mathematics, Hacettepe University (Turkey)
email: ipek@hacettepe.edu.tr

İdris Cesur, Mechanical Engineering, Sakarya University of Applied Sciences (Turkey)
email: icesur@sakarya.edu.tr

İlkay Şişman, Chemistry, Sakarya University (Turkey)
email: isisman@sakarya.edu.tr

İnan Keskin, Civil Engineering, Karabük University (Turkey)
email: inankeskin@karabuk.edu.tr

Jamal Khatib, Civil Engineering, University of Wolverhampton (England)
email: j.m.khatib@wlv.ac.uk

Kadriye Ergün, Industrial Engineering, Balıkesir University (Turkey)
email: kergun@balikesir.edu.tr

Kasım Serbest, Mechatronics Engineering, Sakarya University of Applied Sciences (Turkey)
email: kserbest@subu.edu.tr

Mahmut Özacar, Chemistry, Sakarya University (Turkey)
email: mozacar@sakarya.edu.tr

Maryna Van De Venter, Biochemistry and Microbiology, Nelson Mandela University (Germany)
email: maryna.vandeventer@nmmu.ac.za

Mehmet Ozen, Mathematics, Sakarya University (Turkey)
email: ozen@sakarya.edu.tr

Murat Tuna, Chemistry, Sakarya University (Turkey)
email: tuna@sakarya.edu.tr

Murat Utkucu, Civil Engineering, Sakarya University (Turkey)
email: mutkucu@sakarya.edu.tr

Mustafa Alkan, Mathematics, Akdeniz University (Turkey)
email: alkan@akdeniz.edu.tr

Mustafa Yılmaz, Biology - Vegetative and Animal Production, Sakarya University of Applied Sciences (Turkey)
email: mustafayilmaz@sakarya.edu.tr

Nazan Deniz Yön, Biology, Sakarya University (Turkey)
email: ndyon@sakarya.edu.tr

Necati Olgun, Mathematics, Gaziantep University (Turkey)
email: olgun@gantep.edu.tr

Nesrin Güler, Statistics, Sakarya University (Turkey)
email: nesring@sakarya.edu.tr

Osman Kırtel, Civil Engineering, Sakarya University of Applied Sciences (Turkey)
email: okirtel@sakarya.edu.tr

Özer Uygun, Industrial Engineering, Sakarya University (Turkey)
email: ouygun@sakarya.edu.tr

Özgül Keleş, Metallurgical and Materials Engineering, Istanbul Technical University (Turkey)
email: ozgulkeles@itu.edu.tr

Öznur Özkan, Mathematics, Başkent University (Turkey)
email: oznur@baskent.edu.tr

Peter Claisse, Civil Engineering, Coventry University (England)
email: pete@claisse.info

Ramazan Meral, Biosystem Engineering, Bingöl University (Turkey)
email: rmeral@bingol.edu.tr

Salima Saib, Physics, University of M'sila (Algeria)
email: salima_saib@yahoo.fr

Sema Salur, Mathematics, Rochester University (England)
email: sema.salur@rochester.edu

Sezgin Aydın, Physics, Gazi University (Turkey)
email: sezginaydin@gazi.edu.tr

Sezgin Kaçar, Electrical and Electronics Engineering, Sakarya University of Applied Sciences (Turkey)
email: skacar@sakarya.edu.tr

Shivam Tripathi, Civil Engineering, Purdue University (USA)
email: shiva@iitk.ac.in

Tijen Talas Oğraş, Biology - Genetic Engineering, Tubitak (Turkey)
email: tijen.ogras@tubitak.gov.tr

Uğur Çalığülü, Metallurgical and Materials Engineering, Fırat University (Turkey)
email: ugurcaligulu@gmail.com

Vezer Ayhan, Mechanical Engineering, Sakarya University of Applied Sciences (Turkey)
email: vayhan@sakarya.edu.tr

SAKARYA UNIVERSITY JOURNAL OF SCIENCE
CONTENTS

Volume: 23 - Issue: 1 (FEBRUARY 2019)

RESEARCH ARTICLES

A Survey of Hybrid Main Memory Architectures	Zerrin Yıldız Çavdar, İsa Avcı, Murat Koca, Ahmet Sertbaş	1-15
Comparison and Analysis of Routing Protocols Using Riverbed Modeler	Mahdi Ali Warsame, Abdullah Sevin	16-21
Exploring Analytical Model to Performance Optimization for Mobile Application Using End-to-End Network Slicing in Cloud-Based Vehicular Networks	Yonal Kırsal	22-34
Assesment of Power Quality Disturbances For Grid Integration of PV Power Plants	Gökay Bayrak, Alper Yılmaz	35-42
Performance Comparison for Series and Parallel Modes of a Hybrid Electric Vehicle	Gökhan Canbolat, Halit Yaşar	43-50
Autonomous flight performance improvement of the morphing aerial robot by aerodynamic shape redesign	Harun Çelik, Tuğrul Oktay, Metin Uzun	51-65
Passive RFID Uplink and Downlink Link Budget and Comparison of ASK and BPSK Backscatter Modulations	Kazım Evcan	66-75
Numerical and experimental approach of various sectioned new concept of the crash-boxes to determine the reliability and crashworthiness of the vehicles during frontal impacts	İbrahim Kutay Yılmazçoban, Ömer Adanur, Ahmad Bakhtiyar, Aslı Ergün	76-84
A Modeling Study on Surface Roughness of Spinneret Mold Sections Machined By WEDM	Erdoğan Kanca, Volkan Cem Taşkın, Ali Günen	85-93
Secrecy Outage Probability of Modified TAS/Alamouti-STBC Schemes under Pilot Contamination Attacks	Ahmet Faruk Coşkun	94-105
Performance comparison and analysis of Linux block I/O schedulers on SSD	Yunus Ozen, Abdullah Yıldırım	106-112
Optimization of Surface Roughness of AISI 1040 Stainless Steel in Milling Process Using Taguchi Method	Neslihan Özsoy, Murat Özsoy	113-120
The influence of canola oil biodiesel on performance, combustion characteristics and exhaust emissions of a small diesel engine	Mehmet Şen	121-128



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: A Survey of Hybrid Main Memory Architectures

Authors: Zerrin Yıldız Çavdar, İsa Avcı, Murat Koca, Ahmet Sertbaş

Received: 2017-08-14 17:48:27

Revised: 2018-07-05 18:47:48

Accepted: 2018-07-11 15:06:57

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 1-15

How to cite

Zerrin Yıldız Çavdar, İsa Avcı, Murat Koca, Ahmet Sertbaş; (2019), A Survey of Hybrid Main Memory Architectures. Sakarya University Journal of Science, 23(1), 1-15, DOI: 10.16984/saufenbilder.334645

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/334645>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

A Survey of Hybrid Main Memory Architectures

Zerrin Yıldız Çavdar^{*1}, İsa Avcı², Murat Koca³, Ahmet Sertbaş⁴

ABSTRACT

Rapidly evolving technology, increased internet speed and capacity, and the widespread use of mobile technologies have increased the demands for faster applications and less power consumption of modern electronic systems. In modern electronic systems, RAM is as effective as CPU regarding performance and power consumption. Although DRAM is the most used types of main memory today, it has been insufficient in terms of provide increasing demands. One of the issues to be addressed is to improve DRAM in terms of performance and power consumption. Another study to address this increasing demand is the development of hybrid main memory architectures. Hybrid Main Memory is one of the most recent studies on RAM. In this research, we investigate hybrid main memory systems for a more efficient main memory architecture.

Keywords: Hybrid Main Memory, DRAM, Phase Change Memory/PCM, performance, energy saving

1. INTRODUCTION

With the rapid growth and widespread use of Internet speed and capacity, modern electronic systems that we use in many areas of our daily lives are expected to be faster. At the same time, the production of the mobile version of many devices has made the power consumption / battery life of the devices an important issue.

Performance and power consumption are two important factors in all modern electronic systems such as mobile phones, computers, smart home systems, and so on. It is desirable that the devices operate both very fast and with low power consumption (if mobile having long battery life). The biggest problem with developing electronic

systems is that speed performance and power consumption have a negative effect on each other. In other words, while efforts to increase the performance of systems often cause to more power consumption, on the other hand, the work done on power saving negatively affects the performance of the systems.

In a standard electronic system, CPU and RAM are the two fundamental components most affecting power consumption and performance. In order to achieve higher performance with low power consumption design goal in modern electronic systems the CPU and RAM related software and hardware enhancements have become the main interest in commercial and academic work. The studies show that RAMs are

* Corresponding Author: zerrinyildizcavdar@sehir.edu.tr

¹ Istanbul City University, Vocational School, Computer Technologies Department, Computer Programming, Istanbul, Turkey

² Turkish Airlines, Information Processing Unit, Istanbul, Turkey

³ Hakkari University, Vocational School of Health Services, Hakkari/Turkey

⁴ Istanbul University, Faculty of Engineering, Department of Computer Engineering, Istanbul, Turkey

highly efficient at high power consumption, so that in modern server systems, main memory contains almost 30-40% of total power consumption. [1].

At the present, DRAM is the most widely used main memory type. However, it is now difficult to meet the demand for rapidly evolving technology and increased performance. DRAM uses almost 20% to 40% of the energy consumed by an existing server system [2] [3]. This necessitates the creation of alternative main memory architectures (hybrid for example) for DRAM development. In the current studies, many hybrid main memory architectures are developed by using different memory types together.

Newest two commercial works on improving main memory performance and power consumption are Hybrid Memory Cube and IBM PRAM projects. First one is produced in 2011 by Micron and Samsung, and based on the principle that DRAM main memories are joined vertically. This makes it possible to use DRAM at more capacities without taking up more space. Technical details related to the study are available in [4] and [5]. The second one is developed by IBM in 2016. IBM scientists have worked on PCM (Phase Change Memory) memory and have been able to increase the data storage capacity from 1 bits per cell to 3 bit per cell [6].

In recent years, PCM has become a preferred solution to be used in hybrid main memory modules together with DRAM. DRAM memories are more advantageous than PCM memory types in write operation. PCM memories are more advantageous than DRAM in reading and stand-by operations because they do not need to refresh. Therefore, the Hybrid Main Memory architecture therefore focuses on the solutions used in combination with DRAM and PCM memories.

In this survey study, we are conducting a research on improving RAM technologies, which are very effective in terms of power consumption and performance. In the study, we investigate the academic works on hybrid main memory systems

in order to produce a more efficient main memory architecture. In the second section of our work, we briefly mention non-volatile main memory types. In the third section, we examine hybrid main memory architectures. We present architectures under the titles DRAM + PCM, DRAM + PRAM, DRAM + NVM, DRAM + Other Types. The fourth section of the paper conducts a performance review. The work in this section is presented under the titles of energy saving, performance analysis and endurance / lifetime. In these examinations, we have tried to compare techniques in different ways. Some of the studies have been addressed in more than one section, because they deal with multiple categories, such as energy saving and endurance. In the last section, we give conclusions of the work.

2. ELECTRONIC MEMORY TYPES

This section provides basic information about the types of memory that can be used as the main memory.

Memory types are electrically divided into Volatile (temporary) and Non-Volatile (permanent). We showed the memory types together in the following figure. The most common types of memory used in hybrid memory architectures are volatile DRAM and nonvolatile RAM. Also volatile DRAM is well known memory type, so we gave basic information only about Non-Volatile memories (2.1-2.4) below.

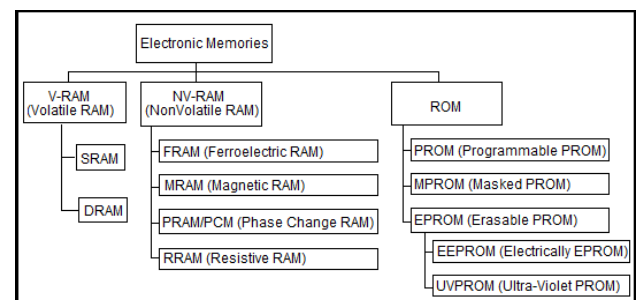


Figure 1: Electronic memory types

2.1. FRAM/FeRAM (Ferroelectric RAM)

FRAM is one of the memory types that are beginning to get attention in the hybrid main memory architectures, which is a low-voltage, non-volatile memory type with fast memory [7]. Ramtron International Corporation presented the first commercial FRAM in 1988 [8]. FRAM is an advanced non-volatile memory that marketed earlier than its counterparts like MRAM, PCRAM, ReRAM [9].

2.2. MRAM (Magnetic / Magnetoresistive RAM)

MRAM is a non-volatile RAM that stores information with electron spin. It works by turning on and off the magnetic moment, not by electricity to determine the write state. MRAM, which is highly ambitious against other RAM technologies due to its non-volatility and power efficiency, is a candidate for universal memory.

2.3. PRAM/PCM (Phase Change RAM)

PRAM is a type of phase change memory designed to protect the data even when power is lost. While reading data from the PRAM, the power consumption is much lower because no heating is required. There is also no energy requirement for PRAM refresh. While PCM is used in many sources to describe PRAM, phase change RAM, it is possible for PCM to serve outside of the main memory. Non-Volatile PCM uses the phase shifting property of chlorochemical glasses to store data (bit information) [10] [11].

2.4. RRAM (Resistive RAM)

The memory type, which is a type of a persistent / durable (nonvolatile) storage that functions via altering the exclusively designed solid non-conducting substance's impedance, is the resistive random access memory (RRAM). A RRAM embodies a memory resistor unit (memristor) of which shows a diverse level of resistance in accordance with the various voltages applied to it.

3. HYBRID MAIN MEMORY ARCHITECTURES

In this part, we examined the studies between the years 2000-2017 on the development of the hybrid main memory architecture. In Table 1, Group 1 and Group 2 are works using the same architecture (PRAM/PCM), listed them in different groups because of using the different terms in their architectural representations. In fact, both groups used Phase Change Memory. Group 3 and Group 4 are studies based on DRAM + NVM architecture, and DRAM + other memory types architectures respectively.

In the remainder of this section, we tried to examine some selected solutions, which include different hybrid main memory architecture models given in works listed on Table 1.

Table 1: Hybrid main memory architecture types

Group No	Suggested Hybrid Architectural Components	Studies Using This Architecture
1	DRAM + PCM	[12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26]
2	DRAM + PRAM	[27], [28], [29], [30], [31], [32], [33], [34], [35], [36]
3	DRAM + NVM	[37], [38], [39], [40], [41]
4	DRAM+Other Memory types	[42], [43], [44], [45], [46], [47], [48]

3.1. DRAM + PRAM/PCM Architectures

In this subsection, we studied on main memory architectures with Phase Change Memory and DRAM. Firstly, an energy-efficient main memory architecture, named PDRAM, using DRAM and PRAM together is suggested in Fig.2. In order to address the challenges of managing such an energy efficient system, the researchers anticipated that a hybrid (hardware + software) solution could be used. Because of the write

endurance problem of PRAM, in this work, a hardware solution is used to store the page-level write frequency information. On the software side, they proposed a page managing program using the writing frequency from hardware at the operating system level [27].

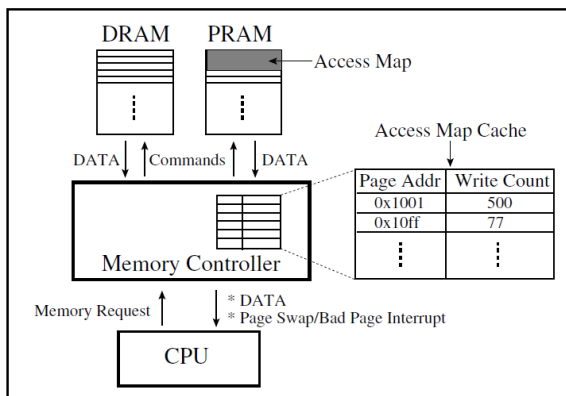


Figure 2: PDRAM diagram [27]

Unlike Dhiman and his friends' PDRAM named hardware and software based suggestion [27], Park and his colleagues implemented the memory management system only with a solution at the operating system level. In the study, the main memory management mechanism at OS (Operating System) level applied to DRAM+PRAM hybrid structure shown in Figure 3, the memory pages are divided into cold and hot and dynamically placed in the suitable memory areas.

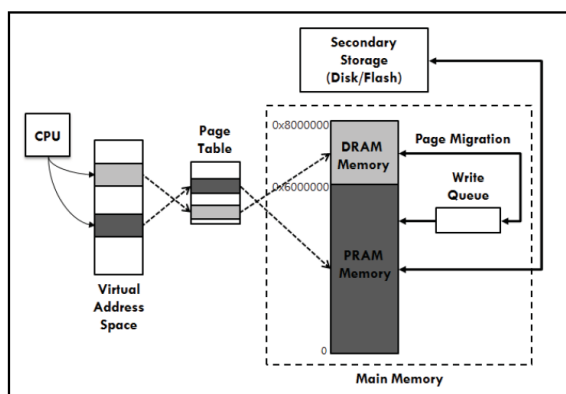


Figure 3: Hybrid main memory [28]

For this process, [28] used three methods called memory access monitoring, page switching and page assignment, cold pages are transferred to

PRAM with page transfer, hot pages are allocated to DRAM. They aimed to reduce energy consumption by this separation and by a page deletion structure that allows the DRAM to close when DRAM is not used [28].

Lee and his colleagues proposed a DRAM + PCM memory design, which increases performance without increasing energy consumption [12]. In addition to examining DRAM + PCM designs considering both performance and energy efficiency (Fig.4), a DRAM + PCM architecture is proposed that combines the energy efficiency achieved using DRAM as a chip cache and the performance levels accomplished when DRAM is used as a buffer of PCM. The architects have studied these approaches in terms of energy consumption and performance in their study of DRAM as a write cache and DRAM as cache memory. In the study, unlike previous studies, which emphasized the management of PCM memory, the idea of using DRAM memory in different forms was examined.

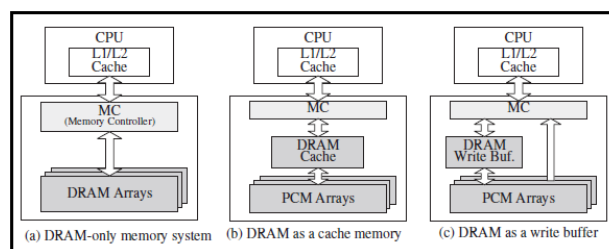


Figure 4: DRAM/PCM configuration [12]

In order to minimize the impact of DRAM's energy consumption, PRAM's write endurance and various limitations, [29] has worked on power control in DRAM + PRAM main memory, in Fig. 5.

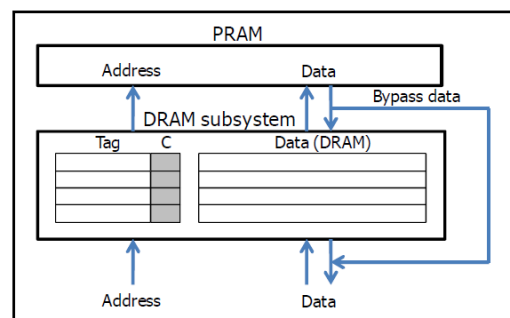


Figure 5: DRAM/PRAM memory [29]

[13] aimed enlarging the PCM's lifespan by reducing the writing number (Fig. 6), therefore they presented a caching scheme called CAR (Cache Address Remapping). In addition, RanCAR (Randomized CAR) also referred to a practical application. In this structure, RanCAR may reduce PCM write-back traffic by evenly distributing single cache's writes to different cache sets. Experiments with the M5 simulator have shown that CAR can decrease the $\sim 4600x$ DRAM cache miss ratio under particular attack and the PCM's lifespan can be extended to 13.8 years, with negligible performance loss.

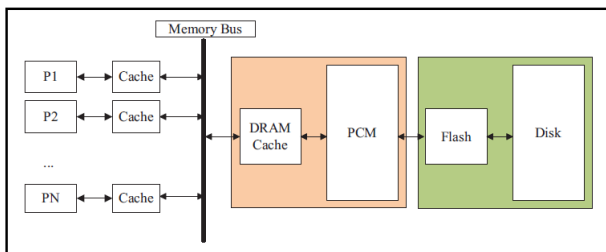


Figure 6: System architecture of DRAM/PCM [13]

An evaluation framework for improving performance in the DRAM + PRAM memory architecture, named the OPAMP is suggested in [30]. In such a structure that emphasizes that hybrid memory design must be done carefully, the frame aggregates the environmental parameters of the combined main memory and evaluates the best performance under relevant circumstances. Here the mentioned above study proposed a method of obtaining the most suitable value with the hybrid master memory profiling.

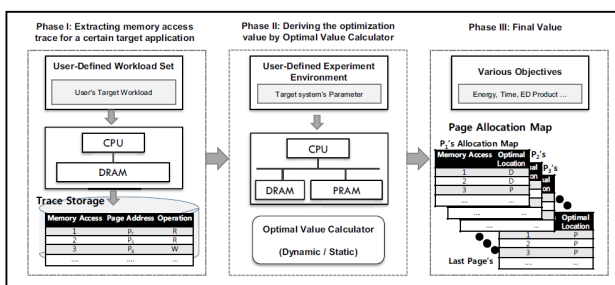


Figure 7: OPAMP process [30]

As an alternative to OPAMP process in [30], [16] proposed to design a new DRAM+PCM hybrid main memory architecture shown in Fig. 8 that takes the recently used metadata cache. Thus they

have set a different caching policy for the combined memory, which actuates the transition size causing the low access latency and low pass.

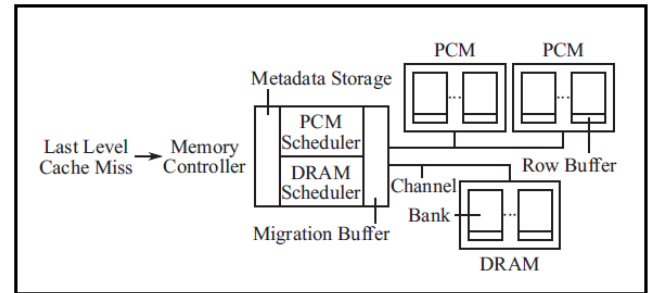


Figure 8: DRAM/PCM overall architecture [16]

Hu and his colleagues proposed SWL (Software Wear Leveling) algorithm and software techniques for correcting abrasions for prolonging the PCM life. They based on the architecture shown in Fig. 9 for their studies in the simulator environment. As a result of those studies, they could achieve a reduction about 80% of the number of writing by greedy algorithm and by ODA (Optical Data Allocation) algorithm a reduction of about 60% of memory access under 6% memory access overhead [17].

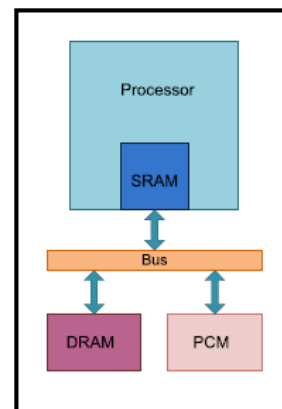


Figure 9: Hybrid memory architecture [17]

As similar to previous work focusing on software solutions, [31] examined the problem of task allocation. For to minimize energy consumption and memory size and extend the life of the hybrid memory, task allocation techniques have been studied in two stages, emphasizing PRAM's energy efficiency and DRAM write durability. Firstly, they designed ILP formulas to solve different objectives best. Then, they proposed 2

different heuristic algorithms, namely 3 polynomial time offline heuristics and 3 online heuristics. At the end of the studies, they conclude that the offline heuristics they offered performed better than the simple ones. Furthermore, when compared to ILP formulations, it has been shown that the offline exploratory methods offer similar solutions, but have lost much less time [31].

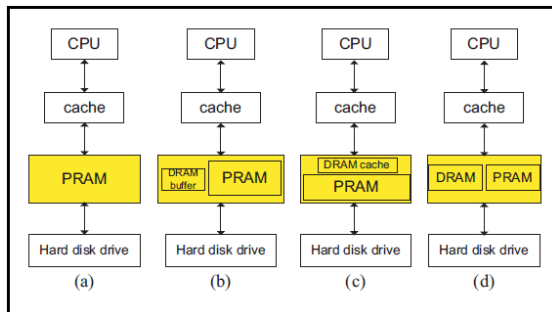


Figure 10: Steps from PRAM to hybrid DRAM + PRAM [31]

To increase the hybrid memory's reliability, Mao and his colleagues have set up three separate low cost ECC-based schemes in Fig. 11. In their work on a hybrid memory architecture [33], they used DRAM's PRAM buffer because it has some advantages like small standby power, speedy access time and big storage density. Fig.11(a) shows a conventional system in which only the PRAM's data is protected by ECC. For this reason, this system only runs if there are relatively low errors in DRAM. Systems 2 and 3 provide ECC protection for the PRAM as well as for the data in the DRAM cache (Figures 11(b) and 11(c)). In the second scheme, the same ECC unit protects the DRAM cache and the data in the PRAM, so the ECC scheme could be more powerful. System 3 uses other layouts for PRAM and DRAM so that 2 memories' error characteristics can be better identified [33].

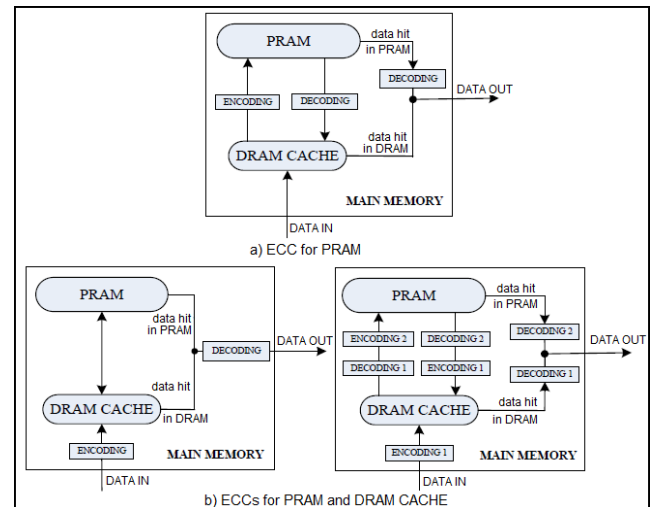


Figure 11: ECC for hybrid DRAM/PRAM [33]

For achieving high performance of DRAM + PRAM hybrid architecture, an architecture that can be accessed in parallel with memories (shown in Fig. 12) has been designed to support selective caching in DRAM to expand writing buffer. For resolving overmuch stall time the problem at the memory controller's data queue, they proposed caching data optionally, that incur the stall thereabout decreasing latency in memory access and ameliorating fairness. Their results presented that the DQSA (Data Queue Stall Aware) approach, achieves 21% better performance and 2.1 times melioration in fairness examined with the best of existing ways in a multi-care system consisting of collective GPU and CPUs [34].

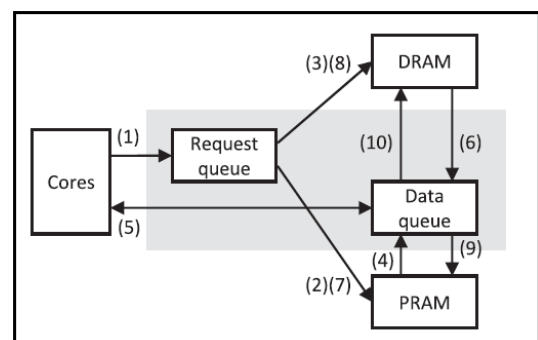


Figure 12: Hybrid main memory block diagram [34]

[35] introduced a main memory system based on PRAM including DRAM converter, CMT (Convert Management Table) and MLC(Multi Level Cell)/SLC(Single Level Cell) PRAM. This work showed that the DRAM converter, which

consisting AFSB(Aggressive Fetching Superblock Buffer) and SFB (Selective Filtering Buffer), could improve access latency and improves endurance. While the AFSB utilizes the regional location effectively / actively by bringing super blocks from the MLC PRAM, the SFB uses the temporary zone by adapting it. The CMT uses data classification management to ensure that pages are assigned and counting information is entered. The SLC PRAM takes on the task of extending the write buffer and the lifetime of the MLC PRAM to mask asymmetric write / read latencies [35].

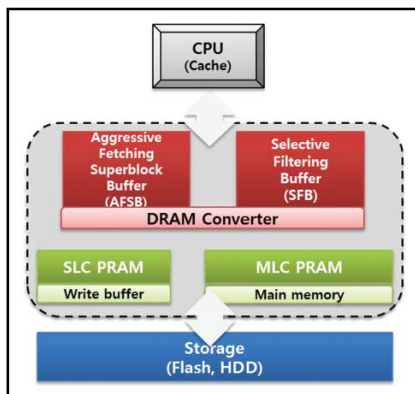


Figure 13: SLC/MLC PRAM + DRAM main memory [35]

For the sake of energy optimization of the PCM and DRAM combined main memory structure shown in Fig. 14, Wang and colleagues aimed to use the PCM completely to ensure the performance of real-time applications and reduce energy consumption [22]. This study proposed an optimal address-mapping algorithm for mapping an appropriate memory address for each address. First they calculated time cost and energy cost for every single address being founded on task types. Then they formulated an ILP model for the timeline issue on distinct memory types and given timing constraints. So they obtained an optimal solution. They concluded that the introduced approach could meaningfully decrease energy consumption at the minimum cost when analyzed by the conventional techniques [22].

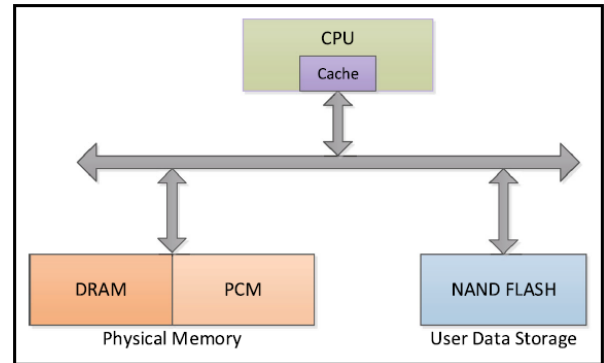


Figure 14: Hybrid DRAM/PCM memory in embedded system [22]

For real-time task-timeline problem for energy efficiency, [23] introduced static table-based and dynamic time lining algorithms for an extraordinary set of tasks. They also introduced 4 real-time programming algorithms based on RM (Rate-Monotonic) and EDF (Earliest Deadline First) timers for real-time embedded systems with a combined PCM-DRAM main memory (in Fig. 15) for a periodic task set. Finally, they have designed a dynamic-RM algorithm that takes advantage of the nearest idle time and a dynamic-EDF algorithm that recovers the entire available free time to enhance the results of static solutions. This architecture addresses DRAM and PCM consecutively in order to provide a direct CPU access to DRAM and PCM. At the end of this study, they presented the timing algorithms that reduce real time constraints and energy consumption [23].

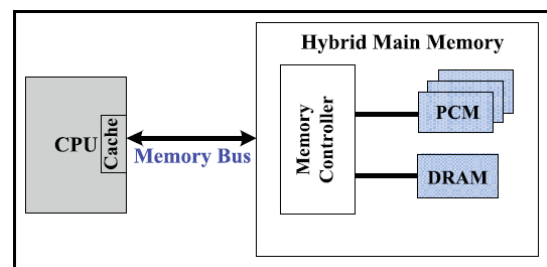


Figure 15: Hybrid DRAM/PCM memory [23]

In [36], Cai and his colleagues examined the task sharing issue for DRAM + PRAM combined main memory structure. To improve memory performance and decrease energy consumption of the memory subsystem shown in Fig.16, they assigned distinct memory devices for every task.

They have designed an ILP offline-ASA (Adaptive Space Allocation) algorithm to access ideal task distribution for an embedded system with a static periodic task set. Furthermore, they proposed an online-ASA algorithm for dynamic task set where accessions of tasks are unknown before [36].

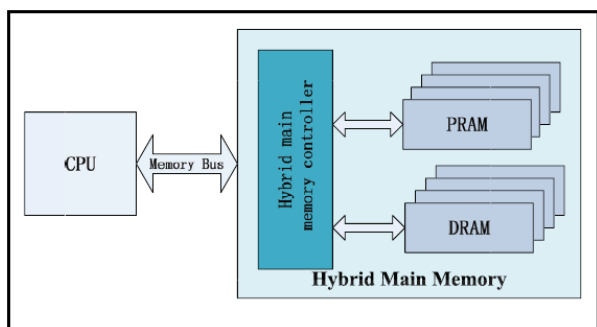


Figure 16: DRAM/PCM main memory [36]

The study in [25] described a multi-page concurrent transition (CMMP - Concurrent Migration of Multiple Pages), a new structure by using hardware and software to manage the construction shown in Fig. 17. In this work, it is indicated that CMMP carries multiple pages simultaneously independently the available memory bandwidth for programs significantly, brings a basic interface for the operating system to monitor memory access models. They also state that CMMP reduces the transfer bandwidth from PCM to DRAM by copying blocks at calls, decreases the bandwidth to the PCM from DRAM by preventing the blocks from being transferred back to the PCM [25].

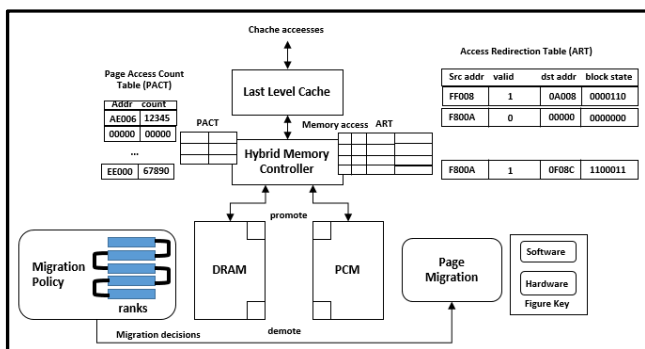


Figure 17: Overview of CMMP [25]

In [26] on power management of DRAM + PCM combined main memory (Fig. 18), it is aimed to

decrease energy consumption of combined memory and increase system efficiency. A new page-rated energy consumption strategy and a new data structure road saved the pages as local and global access data, classified according to the entry dates, and then reshaped by deciding which memory (DRAM or PCM) will be located according to the memory structure. Experimental studies on APG (Adaptive Page Grouping) and PDRAM (DRAM + PCM) in the Gem5 simulator environment have shown that they could reduce energy consumption and improve performance with their strategy [26].

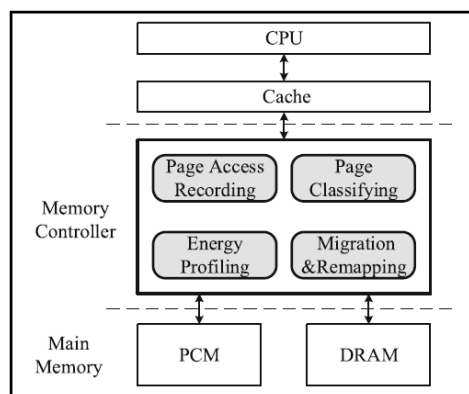


Figure 18: Hybrid DRAM/PCM architecture [26]

3.2. DRAM + NVM Architectures

In hybrid main memory systems, which are currently proposed as an ideal solution for main memory architecture, it is very important to balance the memory types used. Taking this balance into account, Knyagin and colleagues in [37] proposed Crystal, an analytical approach to resource partitioning at design time for hierarchical hybrid systems based on DRAM + NVM, modeled in Fig. 19. Crystal, which is rapidly identifies the most important quantities and trends of NVM for certain workloads and segmentation targets and obtains targeted design points for detailed evaluation. In this work, it has showed how to achieve system-level performance and energy efficiency using NVM with speed and energy consumption of NAND Flash in place of faster and more energy-efficient NVM like

PCM. Hybrid design calls the data to M2, if a requested page is not in M1 (if it is in M2). Each insertion (page entry from disk or a transition from M2) could be result in an evacuation sequence (transition to M2 or page output to disk) because of showing stationary state behavior of curve line. Sending disk only occurs dirty pages of non-memory programs [37].

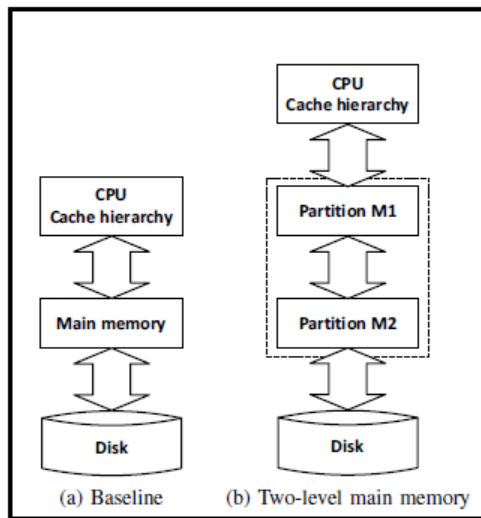


Figure 19: DRAM/NVM system [37]

In addition to the advantages of DRAM + NVM hybrid main memory, which is considered as an ideal next generation architecture, they have significant performance problems due to increased memory traffic, intensive data migration and lack of effective migration. For solving these problems in [40], it is developed a simulator named HMMSim shown in Fig. 20. In addition to introducing the properties of HMMSim, they observed that performance was improved with HMMSim in combined DRAM + NVM architecture in the probatory products. This work shows that hybrid main memory is a promising option with the right software support.

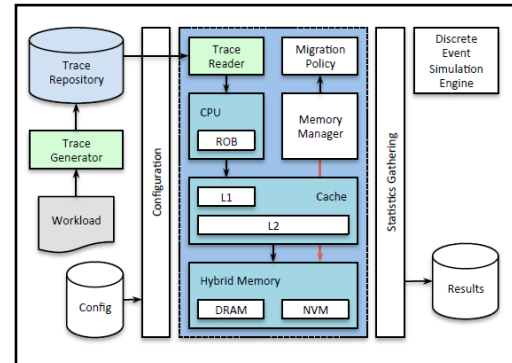


Figure 20: Overview of HMMSim [40]

In [41], Bock and his colleagues suggested the architecture in Fig. 21, which is composed of DRAM and NVM and is supported by the operating system as software. This architecture consists of one or more CPUs with special command and data L1 caching. Assume that requests from the CPU are tailed in L1 tails, all CPUs share an L2 cache, and a single L2 tail handles requests from all CPUs. In the study, working conditions such as application delay and page transition load of the hybrid memory managed by software were analyzed, and factors causing high cost in hybrid construction were determined. As a result of the study, it was seen that the main limiting factor was the delay in the NVM queue and better migration policies could be used to improve performance.

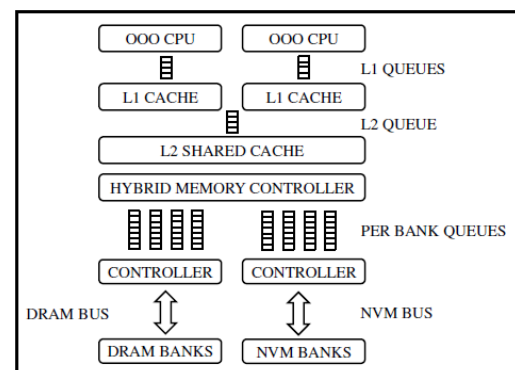


Figure 21: Limiting factors model [41]

3.3. DRAM + Other Memory Types

In regards of the hybrid main memory operations, although PRAM/PCM main memory is preferred well along with DRAM, there are some possible other hybrid solutions. In this section, we

examined other hybrid architectures related to these studies.

3.3.1. DRAM + SCM

In a review of the use of SCM (Storage Class Memory) memory in computer systems, in [46] Kwon suggested using a combination of DRAM and SCM memory (Fig. 22) for a combined memory system. Although SCM is one of the memory types having the optimal memory and storage capabilities, it is not widely used because it is only available in small capacity. He pointed out that the operating system have an considerable role in alleviating the imperfections of the SCM and in using the SCM as a working memory. The most common difficulty of a SCM+DRAM combined memory is to estimate the access order of the data locks, for placing the hot data in DRAM during the write operation, and to place the cold write data in the SCM. He noted that in combined memory systems with the SCM + DRAM, it is necessary to investigate how to fit the data between the two memory moments in order to mitigate the imperfections of the SCM and use the persistence of the SCM. He also emphasized that the delay problem for both DRAM and SCM.

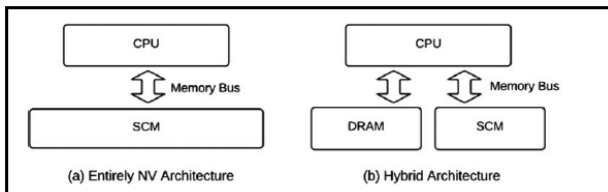


Figure 22: SCM architecture [46]

A study in [48] comparing the architecture of a big DRAM + SCM (small) given with the name of RAHMM (Retention-Aware Hybrid Main Memory) to BSLD (a Big SCM + a Little DRAM hybrid architecture) was performed. In this work, Jing and his colleagues used a small SCM to change the DRAM tail section's data, thereby providing for less refreshing of the DRAM and thus less energy consumption. They have also proposed a HBS (Hidden Buffer Strategy) in order to improve the writing performance and solve endurance problem.

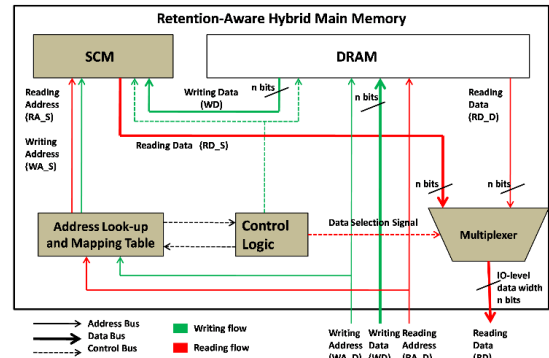


Figure 23: RAHMM architecture [48]

3.3.2. DRAM + PM

In [47], the researchers investigated the influence of hardware stationed page replacement on a combined main memory. The guided model, fixed page swap activity was evaluated by using the suggestions model shown in Fig. 24.

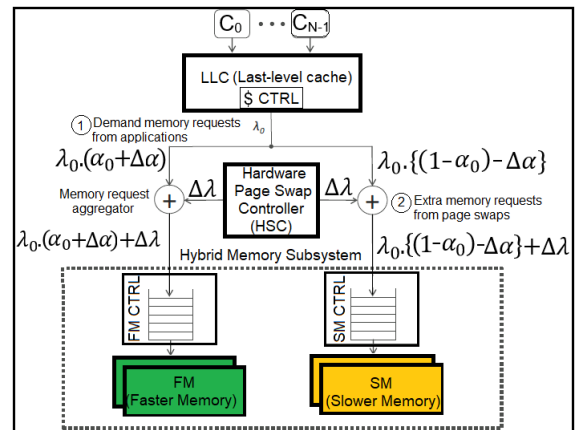


Figure 24: DRAM+PM in hybrid structure [47]

3.3.3. Other Main Memory Architectures

In [42], for performance modeling and analysis of new computer architects, an analysis by using a VMM (Virtual Machine Monitor) was performed for combined main memories includes DRAM + other memory type (DRAM + NAND,) shown in Fig. 25. In the study, the hybrid structure mainly consists of reinforced DRAM and a slower second level memory. In [43] using a customized VMM for performance analysis of hybrid main memory, it has been observed that only certain workloads are more suitable for these structures.

The study confirmed that performance evaluation with VMM is a valid and useful technique.

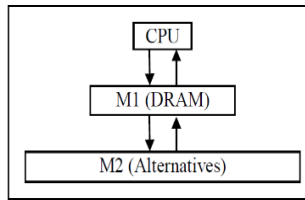


Figure 25: Alternative hybrid options [42]

In [44], a new combined main memory structure with multi-page cache was suggested as shown in Fig.26. In this architecture, Dai and his colleagues proposed GFDP (Global File Data Block Placement) algorithm for placement on file data block issue. They also developed an ILP model for the placement on file data-block issue [44].

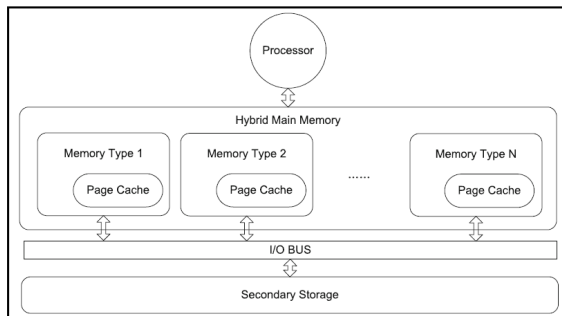


Figure 26: Hybrid memory options [44]

4. PERFORMANCE ANALYSIS

In this part, we examined combined main memory architectures through comparison in terms of energy saving rate, performance rate and endurance/lifetime rate. While the studies shown in Table 2 share mathematical values and measurement results, others interpret them in general, not give numerical results.

Apart from the studies given in Table 2; [31], [37], [18], [39], [22], [23], [26] have indicated that they have improved energy saving; [42], [14], [15], [37], [19], [41], [40], [21], [26] stated that they could provide performance improvement with the solution they proposed ; [28], [38], [35] have pointed out that they achieved positive improvements in durability, in their works.

Table 2: Performance analysis

Group	Paper	Energy Saving Rate	Perf. Rate	Endurance / Life Time Rate	Method Explanation /
1	[12]	-	%42	-	Chip cache
1	[16]	%18	-	-	A new caching policy
1	[24]	%11.7	%4.2	-	Refree
1	[25]	%29	%14	-	Concurrent Migration of Multiple Pages)
2	[27]	%30	-	-	Hybrid (hardware + software)
2	[28]	%35	-	-	Hot and cold pages method
2	[29]	%23.5 - %94.7	-	-	- Runtime-adaptive time out control - DRAM bypass - Keep dirty data cleaner
2	[30]	%20	-	-	OPAMP
2	[32]	%49	-	%88	ILP and graphic model for DSP systems
2	[34]	-	%21	-	Data Queue Stall Aware
2	[36]	%27	-	-	ASA (Adaptive Space Allocation)
4	[47]	-	%13 - %28.9	-	Instruction Per Cycle perform. with fixed page swap
4	[48]	%45	%30	-	Retention-Aware Hybrid Main Memory

5. CONCLUSION

As DRAM has begun to widespread use to meet increasing demand, studies on new main memory

solutions continue increasingly. It is desirable to use a second / more memory with DRAM to meet the increasing demand without giving up the advantages of DRAM. Instead of using PCM as an alternative to DRAM, adding PCM memory to DRAM (which is faster than other types of memory in writing) is seen as the most preferred method to provide read and stand-by improvements. Due to factors such as the fact that PCM memory does not need refreshing and that there are some very successful studies on eliminating the negativities such as lifetime, PCM memory is preferred to other memory types in hybrid main memory architectures, We expect that a hybrid main memory with DRAM and PCM memory and managed with a hybrid solution consisting of software + hardware can be produce very successful results in terms of endurance, performance and energy saving.

In this survey paper, the combined/hybrid main memory architectures are examined in order to show that combined/hybrid main memory solutions are promising. As well known, the use of non-volatile PRAM in conjunction with DRAM is one of the most ambitious solutions for combined/hybrid main memory operations. So, in the future, it is expected that the main memory technology can be improved much more with new works regarding the hybrid main memory architectures.

REFERENCES

- [1] L. A. Barroso ve U. Hölzle, «The Case for Energy-Proportional Computing,» *IEEE Computer Society*, pp. 33-37, 2007.
- [2] C. Lefurgy, K. Rajamani, F. Rawson, W. Felter, M. Kistler and T. W. Keller, "Energy Management for Commercial Servers," *IEEE Computer Society*, pp. 39-48, 2003.
- [3] A. N. Udipi, N. Muralimanohar, N. Chatterjee, R. Balasubramonian, A. Davis and N. P. Jouppi, "Rethinking DRAM design and organization for energy-constrained multi-cores," in *ISCA '10 Proceedings of the 37th annual international symposium on Computer architecture*, Saint-Malo, 2010.
- [4] Micron, "Micron Technology, Inc. - Hybrid Memory Cube | Memory and Storage," 16 04 2017. [Online]. Available: <https://www.micron.com/products/hybrid-memory-cube>.
- [5] H. M. C. Consortium, "Hybrid Memory Cube Consortium - Home," 16 04 2017. [Online]. Available: <http://hybridmemorycube.org/>.
- [6] A. Sammons and C. Sciacca, "IBM New room," 17 05 2016. [Online]. Available: <https://www-03.ibm.com/press/us/en/pressrelease/49746.wss>.
- [7] S. Bagheri, A. A. Asadi, W. Kinsner and N. Sepehri, "Ferroelectric random access memory (FRAM) fatigue test with Arduino and Raspberry Pi," in *2016 IEEE International Conference on Electro Information Technology (EIT)* , Grand Forks, 2016.
- [8] Cypress, "Microcontrollers, Connectivity, Memory Solutions," 22 05 2017. [Online]. Available: <http://www.cypress.com>.
- [9] T. Eshita, W. Wang, K. Nakamura, S. Mihara, H. Saito, Y. Hikosaka, K. Inoue, S. Kawashima, H. Yamaguchi and K. Nomura, "Development of ferroelectric RAM (FRAM) for mass production," in *Applications of Ferroelectrics, International Workshop on Acoustic Transduction Materials and Devices & Workshop on Piezoresponse Force Microscopy (ISAF/IWATMD/PFM), 2014 Joint IEEE International Symposium on the*, State College, PA, 2014.
- [10] N. Yamada, E. Ohno, K. Nishiuchi, N. Akahira and M. Takao, "RAPID-PHASE TRANSITIONS OF GETE-SB2 TE3 PSEUDOINARY AMORPHOUS THIN-FILMS FOR AN OPTICAL DISK MEMORY," *AMER INST PHYSICS*, pp. 2849-2856, 1991.

- [11] J. Tominaga, T. Kikukawa, M. Takahashi and R. Phillips, "Structure of the optical phase change memory alloy, Ag-V-In-Sb-Te, determined by optical spectroscopy and electron diffraction," *AMER INST PHYSICS*, pp. 3214-3218, 1997.
- [12] H. G. Lee, S. Baek, C. Nicopoulos and J. Kim, "An Energy- and Performance-Aware DRAM Cache Architecture for Hybrid DRAM/PCM Main Memory Systems," in *2011 IEEE 29th International Conference on Computer Design (ICCD)*, Amherst, MA, 2011.
- [13] G. Wu, H. Zhang, Y. Dong and J. Hu, "CAR: Securing PCM Main Memory System with Cache Address Remapping," in *2012 IEEE 18th International Conference on Parallel and Distributed Systems*, Singapore, 2012.
- [14] L. Ramos and R. Bianchini, "Exploiting Phase-Change Memory in Cooperative Caches," in *2012 IEEE 24th International Symposium on Computer Architecture and High Performance Computing*, New York, NY, 2012.
- [15] S. Kwon, D. Kim, Y. Kim, S. Yoo and S. Lee, "A Case Study on the Application of Real Phase-Change RAM to Main Memory Subsystem," in *2012 Design, Automation & Test in Europe Conference & Exhibition (DATE) Design*, Dresden, 2012.
- [16] J. Meza, J. Chang, H. Yoon, O. Mutlu and P. Ranganathan, "Enabling Efficient and Scalable Hybrid Memories Using Fine-Granularity DRAM Cache Management," *IEEE Computer Architecture Letters*, pp. 61-64, 2012.
- [17] J. Hu, Q. Zhuge, C. J. Xue, W.-C. Tseng and E. H.-M. Sha, "Software enabled wear-leveling for hybrid PCM main memory on embedded systems," in *2013 Design, Automation & Test in Europe Conference & Exhibition (DATE) Design*, Grenoble, 2013.
- [18] Z. Wang, Z. Gu and Z. Shao, "Optimized Allocation of Data Variables to PCM/DRAM-based Hybrid Main Memory for Real-Time Embedded Systems," *IEEE Embedded Systems Letters*, pp. 61-64, 2014.
- [19] L. Ramos and R. Bianchini, "Robust performance in hybrid-memory cooperative caches," *Parallel Computing*, p. 514-525, 2014.
- [20] J. Hu, M. Xie, C. Pan, C. J. Xue, Q. Zhuge and E. H.-M. Sha, "Low Overhead Software Wear Leveling for Hybrid PCM + DRAM Main Memory on Embedded Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 654-663, 2014.
- [21] K. Kavi, S. Pianelli, G. Pisano, G. Regina and M. Ignatowski, "Memory organizations for 3D-DRAMs and PCMs in processor memory hierarchy," *Journal of Systems Architecture*, pp. 539-552, 2015.
- [22] G. Wang, Y. Guan, Y. Wang and Z. Shao, "Energy-aware assignment and scheduling for hybrid main memory in embedded systems," *Computing. March 2016*, p. 279-301, 2016.
- [23] Z. Zhang, Z. Jia, P. Liu and L. Ju, "Energy Efficient Real-Time Task Scheduling for Embedded Systems with Hybrid Main Memory," *Journal of Signal Processing Systems*, p. 69-89, 2016.
- [24] B. Pourshirazi and Z. Zhu, "Refree: A Refresh-Free Hybrid DRAM/PCM Main Memory System," in *2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Chicago, IL, 2016.
- [25] S. Bock, B. R. Childers, R. Melhem and D. Moss'è, "Concurrent Migration of Multiple Pages in Software-Managed Hybrid Main Memory," in *2016 IEEE 34th International Conference on Computer Design (ICCD)*, Scottsdale, AZ, 2016.
- [26] J. Zhang, X. Liao, H. Jin, D. Liu, L. Lin and K. Zhao, "An Optimal Page-Level Power Management Strategy in PCM-DRAM Hybrid Memory," *International Journal of Parallel Programming*, pp. 4-16, 2017.

- [27] G. Dhiman, R. Ayoub and T. Rosing, "PDRAM: A Hybrid PRAM and DRAM Main Memory System," in *2009 46th ACM/IEEE Design Automation Conference Design Automation Conference*, San Francisco, CA, 2009.
- [28] Y. Park, D.-J. Shin, S. K. Park and K. H. Park, "Power-Aware Memory Management for Hybrid Main Memory," in *The 2nd International Conference on Next Generation Information Technology Next Generation Information Technology (ICNIT)*, Gyeongju, Korea (South), 2011.
- [29] H. Park, S. Yoo and S. Lee, "Power Management of Hybrid DRAM/PRAM-Based Main Memory," in *2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC)*, San Diego, CA, 2011.
- [30] J.-H. Choi, S.-M. Kim, C. Kim, K.-W. Park and K. H. Park, "OPAMP: Evaluation Framework for Optimal Page Allocation of Hybrid Main Memory Architecture," in *2012 IEEE 18th International Conference on Parallel and Distributed Systems Parallel and Distributed Systems*, Singapore, 2012.
- [31] W. Tian, Y. Zhao, L. Shi, Q. Li, J. Li, C. J. Xue, M. Li and E. Chen, "Task Allocation on Nonvolatile-Memory-Based Hybrid Main Memory," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 1271-1284, 2013.
- [32] T. Liu, Y. Zhao, C. J. Xue and M. Li, "Power-Aware Variable Partitioning for DSPs With Hybrid PRAM and DRAM Main Memory," *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, pp. 3509-3520, 2013.
- [33] M. Mao, C. Yang, Z. Xu, Y. Cao and C. Chakrabarti, "Low cost ECC schemes for improving the reliability of DRAM+ PRAM MAIN memory systems," in *Signal Processing Systems (SiPS), 2014 IEEE Workshop on*, Belfast, 2014.
- [34] D. Kim, S. Yoo and S. Lee, "Hybrid Main Memory for High Bandwidth Multi-Core System," *IEEE TRANSACTIONS ON MULTI-SCALE COMPUTING SYSTEMS*, pp. 138-149, 2015.
- [35] S.-I. Jang, S.-K. Yoon, K. Park, G.-H. Park and S.-D. Kim, "Data Classification Management with its Interfacing Structure for Hybrid SLC/MLC PRAM Main Memory," *COMPUTER JOURNAL*, pp. 2852-2863, 2015.
- [36] X. Cai, L. Ju, X. Li, Z. Zhang and Z. Jia, "Energy efficient task allocation for hybrid main memory architecture," *Journal of Systems Architecture*, pp. 11-22, 2016.
- [37] D. Knyagin, G. N. Gaydadjiev and S. Per, "Crystal: A Design-Time Resource Partitioning Method for Hybrid Main Memory," in *Parallel Processing (ICPP), 2014 43rd International Conference on*, Minneapolis MN, 2014.
- [38] G. Nakagawa and S. Oikawat, "Language Runtime Support for NVM/DRAM Hybrid Main Memory," in *2014 IEEE COOL Chips XVII (COOL Chips)*, Yokohama, 2014.
- [39] A. Hassan, H. Vandierendonck and D. S. Nikolopoulos, "Energy-Efficient Hybrid DRAM/NVM Main Memory," in *International Conference on Parallel Architecture and Compilation*, San Francisco, CA, 2015.
- [40] S. Bock, B. R. Childers, R. Melhem and D. Moss'e, "HMMSim: A Simulator for Hardware-Software Co-Design of Hybrid Main Memory," in *2015 IEEE International Conference on Grey Systems & Intelligent Services (GSIS)*, Leicester, United Kingdom, 2015.
- [41] S. Bock, B. R. Childers, R. Melhem and D. Mosse, "Characterizing the Overhead of Software-Managed Hybrid Main Memory," in *IEEE 23rd International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, Atlanta, GA, 2015.

- [42] D. Ye, A. Pavuluri, C. A. Waldspurger, B. Tsang, B. Rychlik and S. Woo, "Prototyping a Hybrid Main Memory Using a Virtual Machine Monitor," in *IEEE International Conference on Computer Design Computer Design*, Lake Tahoe, CA, 2008.
- [43] J. Stevens, P. Tschirhart, M.-T. Chang, I. Bhati, P. Enns, J. Greensky, Z. Chisti, S.-L. Lu and B. Jacob, "An Integrated Simulation Infrastructure For The Entire Memory Hierarchy: Cache, Dram, Nonvolatile Memory, And Disk," *Intel Technology Journal*, pp. 184-200, 2013.
- [44] P. Dai, Q. Zhuge, X. Chen, W. Jiang and E. H.-M. Sha, "Effective file data-block placement for different types of page cache on hybrid main memory architectures," *DESIGN AUTOMATION FOR EMBEDDED SYSTEMS*, pp. 485-506, 2013.
- [45] Z. Chen, Y. Lu, N. Xiao and F. Liu, "A hybrid memory built by SSD and DRAM to support in-memory Big Data analytics," *KNOWLEDGE AND INFORMATION SYSTEMS*, pp. 335-354, 2015.
- [46] J. B. Kwon, "Exploiting Storage Class Memory for Future Computer Systems: A Review," *IETE Technical Review*, pp. 218-226, 2015.
- [47] J.-Y. Jung and R. Melhem, "Empirical, Analytical Study of Hardware-based Page Swap in Hybrid Main Memory System," in *2016 28th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, Los Angeles, CA, 2016.
- [48] W. Jing, K. Yang, Y. Lin, B. Lee, S. Yoon, Y. Ye, Y. Du and B. Chen, "Retention-Aware Hybrid Main Memory (RAHMM): Big DRAM and Little SCM," *IEEE Transactions on Computers*, pp. 912-918, 2017.
- [49] S.-I. Jang, C.-G. Kim and S.-D. Kim, "An Efficient DRAM Converter for Non-Volatile Based Main Memory," in *IT Convergence and Security*, Pyeong Chang, Korea, 2012.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Comparison and Analysis of Routing Protocols Using Riverbed Modeler

Authors: Mahdi Ali Warsame, Abdullah Sevin

Received: 2018-07-24 16:31:39

Revised: 2018-07-29 04:44:49

Accepted: 2018-07-30 11:53:16

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 16-22

How to cite

Mahdi Ali Warsame, Abdullah Sevin; (2019), Comparison and Analysis of Routing Protocols Using Riverbed Modeler. Sakarya University Journal of Science, 23(1), 16-22, DOI: 10.16984/saufenbilder.447345

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/447345>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Comparison and Analysis of Routing Protocols Using Riverbed Modeler

Mahdi Ali Warsame^{*1}, Abdullah Sevin¹

ABSTRACT

In recent years several routing protocols have been proposed to be used on different types of networks. Some comparative studies were carried out to analyze and compare the performance of several routing protocols over Local Area Network (LAN). In this work, RIP, EIGRP, OSPF routing protocols were analyzed, compared and simulated over Riverbed Modeler. The performance of different protocols are compared using Riverbed Modeler tool in which metrics like delay, throughput, and convergence time are measured. The behavior of the protocols over different network topologies and different data rates were analyzed.

Keywords: Routing, Protocols, Riverbed, RIP, LAN Networks, OSPF, EIGRP.

INTRODUCTION

Computers interconnect with each other through multi-hop link in Local Area Networks(LAN). Routing has great significance in the circumstances of transmitting the data from source to destination. Research colleagues have done several researches of routing protocols over different types of networks, one of such research is carried out to analyze and compare the function of most common routing protocols such as RIP, EIGRP and OSPF. The aim of the paper is to make a deep analysis of the common routing

protocols using Riverbed Modeler by implementing over different network types with different data rates. As a consequence of this work we can determine routing protocols in terms of the suitability of common protocols over different network types. This way we can recommend the type of routing protocol should be used in a defined network type. Routing technique is critical part of every network types. Today most of enterprises prefer usage of LAN over their internal infrastructures for security reasons. Generally Routing is classified into

* Corresponding Author: mahdi.abdulle@ogr.sakarya.edu.tr

¹ Sakarya University, Department of Computer Science, Sakarya, Turkey

parts as static and dynamic routing. [1], [2]. Static routing is performed manually while in dynamic routing is done by algorithms. In this paper, LAN is designed and various routing protocols were implemented over LAN in order to test and compare the performance of above routing protocols. According to the behavior of the protocol over different network types it can be implemented to the suitable network type.

LITERATURE SURVEY

Different studies over routing protocols have been carried out and some of those shows below; Nurul I. and Wilford G. Have offered some ways of how to increase the performance e.g by incrementing the buffer size. Wireless Local Area Network (LAN) is evaluated with using high priority traffic in [2]. The work of Krishna Gorantala in [3], tries to improves data transmissions over large network without any connection damage. A. Ahmad et.al. [4] Have carried out a realistic study over two routing protocols (AODV and DSR) on a university campus, they have used a quantative metrics like throughput, delay and receiving traffic over a given scenarios to analyze the performance of the routing protocols. Some other routing protocols were proposed in these works. [5] [6] [7] [8] [9] [10] [11] [12] but a few comparative study were carried out. The focus of our work is to make comparison between routing protocols over LAN network using Riverbed modeler, metrics like delay, router convergence duration, and throughput were measured for different data rates and topologies.

ROUTING PROTOCOLS

Every network has a goal of sharing information between its nodes which resides on the network. It's is simple to send/receive data to a node inside a network, but it gets a little complicated when it comes to sharing an information to an outside network. When a network wants to communicate with another network routing protocol is necessary. Routing is fundamental for systems to communicate with outside networks. The duty of every router is to forward the coming packets to their suitable destinations. To achieve this, routers need routing protocols, that's why routing protocols are crucial for every network. In order to do routing every router should synchronize its routing table with its associative neighbors and also throughout the network, after that a network topology is formed. Various routing protocols are described below.

Routing Information Protocol

RIP is a distance vector protocol which uses single distance metric as 'Hop Count' for minimum hop. RIP prevents wrong routing of information using a limit of 15 hubs (intermediate nodes). Despite the weakness and limited size of RIP, it offers a great advantages such us using UDP and port number of 520 [12]. Distance vectors were exchanged every 30 seconds which is called Advertisement and each advertisements route to up to 25 destination in networks. The protocol routed in 4.3 BSD Unix distribution.

Enhanced Interior Gateway Protocol

EIGRP is one of hybrid routing protocols it composed of the properties of distance vector and link-state routing protocols. The relationships forms by EIGRP known as adjacencies. With routes in the identical autonomous system-as, by interchanging hello packets. Routes share their routing information just after the formation of adjacencies. The multicast address of 224.0.0.10 is used by EIGRP to propagate its hello packets. For protocol packets delivery reliability purpose EIGRP uses Reliable transport protocol. Unlike other routing protocols EIGRP uses multiple metrics to determine optimal route to destination. Bandwidth (K1), Load (K2), delay of the line (K3), reliability (K4) and Maximum Transmission Unit-MTU (K5), But as default EIGRP uses just bandwidth and delay as metrics. EIGRP is one of classless routing protocols which supports VLSM. [13].

Open Shortest Path First Protocol

OSPF is based on Dijkstra's algorithm and it has the function of calculating the minimum path tree of the routes. One of the main pros of the protocol is to detect errors by itself and finally broadcast domain is provided by multicast addressing in OSPF [12]. By default every router has a built in information of Link-state protocol, with the advantage of this OSPF makes up its topology. Forming the topology it helps OSPF to determine routing decisions. Like EIGRP, OSPF supports VLSM.

Intermediate System to Intermediate System

IS-IS, is usually a preferred routing protocol in large-scale networks. IS-IS is working with reference to the OSI reference model which makes it different. With IS-IS the data is transported with the Connectionless Network Service (CLNS) standard. Hierarchical / scaling can be achieved by separating wide domains into sub-domains (areas or sub-domains). However, each router can only belong to one area. There is another important detail in transmission between ISs. IS-IS packages are encapsulated in Layer 2 without being subjected to OSI Layer 3 encapsulation [14].

PERFORMANCE METRICS

There are different metrics to analyze routing protocols over Riverbed Modeler. Delay, throughput, and convergence duration as performance metrics can give us deep view of how a given routing protocol over different network topology behaves and we can retrieve form that whether the protocol is suitable for the topology or not, with this we will also determine There are different metrics to analyze routing protocols over Riverbed Modeler. Delay, throughput, and convergence duration as performance metrics can give us deep view of how a given routing protocol over different network topology behaves and we can retrieve form that whether the protocol is suitable for the topology or not, with this we will also determine the performance of the protocols, over a given data rates to see how it adopts with different data

rates (increasing) as (1, 2, 4, 8) Mbps.

Performance metrics are detailed below:

End-to-End delay or E-to-E delay is the average of successfully completed packets from one source to destination over a network.

Throughput is the average of successfully delivered packets (messages) per unit of time (seconds) through communication channel. In computer networks, throughput is measured in bits per second and some situations in data packets per second.

Convergence duration is the time which a group of routers reach the state of convergence by creating routing tables after the convergence each router gets a map of the topology it resides from there each router decides which packet should be sent in which route. Optimally the routing protocols must have fast convergence time. It is calculated with the rate of second.

Table 1. Network Parameters

<i>Component</i>	<i>Model</i>
<i>Nodes</i>	<i>Ethernet workstation</i>
<i>Routers</i>	<i>Cisco 7200</i>
<i>Link</i>	<i>100Base T, DDS3, DDS1</i>
<i>Protocol</i>	<i>EIGRP, RIP, OSPF</i>
<i>Switch</i>	<i>Ethernet 16 switch</i>
<i>Data rate</i>	<i>1,2,4,8 Mbps</i>
<i>Packet rate</i>	<i>1 Packet/s</i>

SIMULATION SETUP AND EVALUATION

In this paper, the different network topologies (Star and Mesh) has been modeled by using Riverbed modeler and the performances of various routing protocols in different data rates (1,2,4,8) Mbps has been analyzed. Attributes of the network design, network topologies, LAN parameters with different protocols and traffic

configuration parameters are used in the scenarios. Results of the simulation shown in below section and also traffic and LAN parameters are shown in Table 1. The performance as delay, throughput and convergence duration for various protocols were chosen and analysis is performed.

ANALYSIS AND RESULTS

The goal of this paper is to compare routing protocols and evaluate the performance of the protocols using the Riverbed modeler in different situations and in different topologies. we created two separate topologies (star and mesh) on the Riverbed Modeler, and we created 12 scenarios in each topology. We evaluated the RIP, OSPF and EIGRP routing protocols to evaluate as follows; each protocol evaluated 8 different scenarios with 4 different data rates on each topology for delay time, throughput, convergence time, and protocol traffic, and as a result, we obtained the following graphs and summarized comparison of protocols is given in Table 2.

Table 2. General comparison of the routing protocols

	EIGRP	RIP	OSPF	IS-IS
Algorithm	DUAL	Bellman-Ford	Dijkstra	Dijkstra
Type	Hybrid	Distance Vector	Link- state	Link-state
Interior/Exterior	Interior	Interior	Interior	Interior
Classful	Yes	Yes	No	No
Metric	Bandwidth/Delay	Hop Count	Cost	Cost
Convergence Time	Very Low	High	Low	Low
Updates	Any changes	Full table	Any Changes	Any changes
Protocol And Port	IP, 88	UDP, 520	IP, 89	Ethernet/PPP, No Port

Figures 1 and 2 illustrates that RIP, and OSPF have more End-to-End delay than EIGRP. Also it's notable that from 1 to 4 Mbps the performance looks same which means at low bandwidth RIP and OSPF generates less delays than high bandwidth.

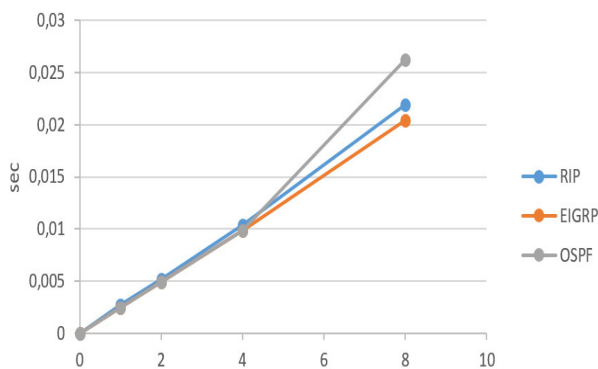


Figure 1. End to End Delay (Star Topology)

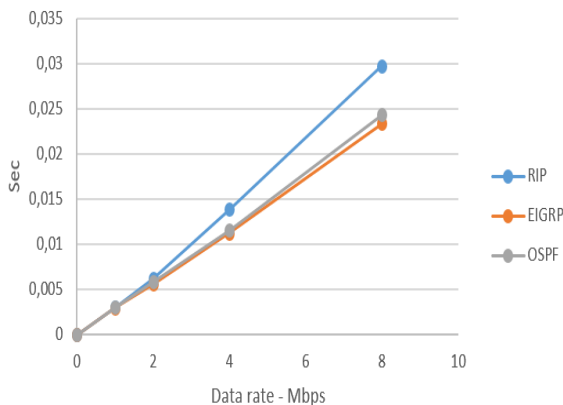


Figure 2. End to End Delay (Mesh Topology)

Throughput is the total amount of sent/received bits- per second, RIP demonstrates major decrease of throughputs when it comes topology type. As it shows on the graph RIP transferred more than 1200 bits at Star Topology but just less than 100 bits for Mesh topology. As we know RIP sends update messages every 30 seconds as default.

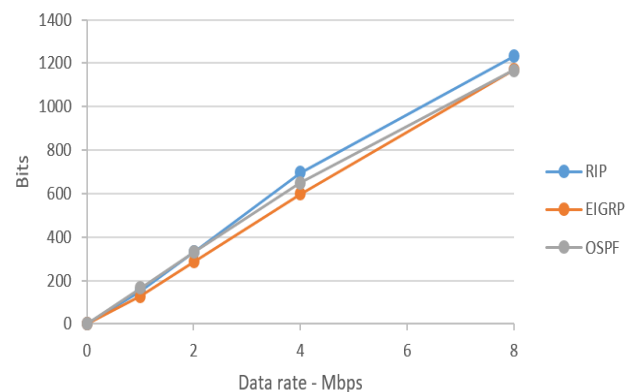


Figure 3. Throughput (Star Topology)

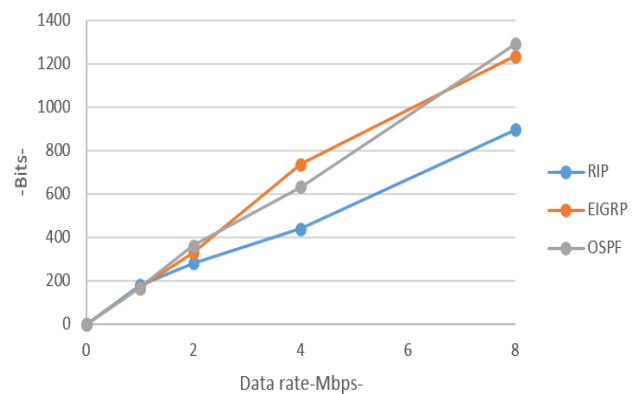


Figure 4. Throughput (Mesh Topology)

Convergence duration is the time that a cluster of routers achieve the situation of convergence. The routing protocols should have low convergence rate. Figure 6 and 7 shows the convergence duration of the routing protocols over topologies (star and mesh). The figures gives us a clear view of how EIGRP converges faster than other protocols over both topologies with less than a second. Thus we can say that EIGRP has a fast convergence rate.

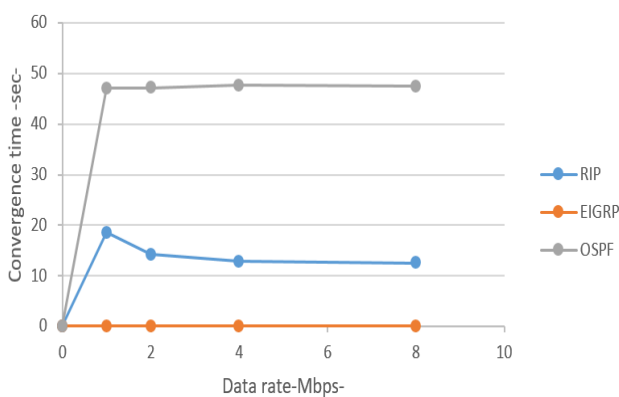


Figure 5. Convergence duration (Mesh Topology)

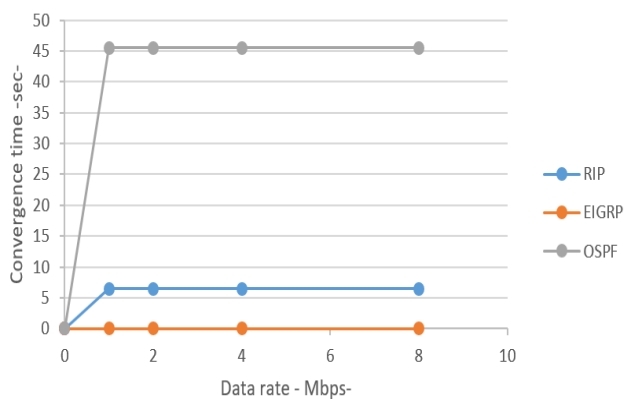


Figure 6. Convergence duration (Star Topology)

CONCLUSION

In this paper, common routing protocols were analyzed and compared each other according to end-to end delay, throughput and convergence duration using Riverbed Modeler, in advance to that the behavior of EIGRP, RIP and OSPF routing

protocols were studied over different network topologies (Star, Mesh) and data rates (1,2,4,8 Mbps). According to our findings we can say that CISCO's proprietary protocol EIGRP showed best performance over both topologies but it has a drawback when the routing traffic which EIGRP generated more routing traffic according to RIP and OSPF that results extra delay to the network. Also RIP demonstrated poor performance to the experiment according to its long convergence duration which causes long delays as it shown to the end-to end delay graph. OSPF protocol gave good performance at 1 to 2 Mbps data rates but it has more delays at 4 to 8 Mbps data rates which shows that OSPF can be more successful at low bandwidth networks other than high bandwidth networks. We can recommend to the high bandwidth networks to use EIGRP because of its less delay and quick convergence duration with high throughput.

REFERENCES

- [1] M. S. a. B. M. Z. Bojković, in *Current Developments towards the 4G Wireless System,* Proceedings of International Conference TELSIKS, Niš, Serbia, September 2005.
- [2] N. I. S. & W. G. Lol, *A Study of MANET Routing Protocols: Joint Node Density, Packet Length and Mobility,* Vols. 978-1-4244-7755-5/10/\$26.00, no. IEEE, pp. Page no. 515-520, ©2010.

- [3] K. Gorantala, in *Routing Protocols in Mobile Ad Hoc Networks*, Sweden, Umeå University, Department of Computing Science, June 2006., p. Master Thesis.
- [4] A. Ahmad, Sh. Husain, M. Chand, S.C. Gupta, R. Gowri, "Simulation Study for Performance Comparison," *International Journal of Information and Communication Engineering*, vol. Vol:4, p. No:10, 2010.
- [5] IEEE Computer Society LAN MAN Standards Committee, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," The Institute of Electrical and Electronics, New York, IEEE Std 802.11-1997.
- [6] P. Jacquet, P. Muhlethaler, A. Qayyum, Optimized Link State Routing Protocol, Internet Draft, draft-ietf-manet-olsr-00.txt, November 18, 1998.
- [7] M. Jiang, J. Li and Y. C. Tay, "Cluster Based Routing Protocol (CBRP) Functional specification," draft-ietf-manet-spec-00.txt, August 1998.
- [8] D.B. Johnson, & D.A. Maltz, "Dynamic source routing in ad hoc Wireless Networks," in *Mobile Computing*, H. K. Tomasz Imielinski, Ed., Kluwer Academic Publishers, pp. 153-181.
- [9] D.B. Johnson, & D.A. Maltz, "Protocols for adaptive wireless and mobile computing," in *IEEE Personal Communications*, February 1996.
- [10] V.D. Park and M.S. Corson, "A performance comparison of the Temporally-Ordered Routing Algorithm and Ideal Link-state routing," in *IEEE Symposium on Computers and Communication*, June 1998.
- [11] I. F. & G. Todorean, "Network performance evaluation for RIP, OSPF and EIGRP routing protocols," in *International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, , 2013.
- [12] A. I. & S. Khan, "Performance Evaluation of Real Time Applications for RIP, OSPF and EIGRP for flapping links using OPNET Modeler," *International Journal of Computer Networks and Communications Security*, vol. 3, no. 1, pp. 16-26, 2015.
- [13] A. Balchunas, "Enhanced Interior Gateway Routing Protocol," [Online]. Available: <http://www.routeralley.com/guides/eigrp.pdf>. [Accessed 27 April 2018].
- [14] Baraković, S., & Baraković, J.x, "Comparative performance evaluation of Mobile Ad Hoc routing protocols.," in *In MIPRO, Proceedings of the 33rd International Convention*, pp. 518-523, 2010.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Exploring Analytical Model to Performance Optimization for Mobile Application Using End-to-End Network Slicing in Cloud-Based Vehicular Networks

Authors: Yonal Kırşal

Received: 2018-03-12 13:19:11

Revised: 2018-05-26 19:09:52

Accepted: 2018-07-31 15:08:44

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 23-34

How to cite

Yonal Kırşal; (2019), Exploring Analytical Model to Performance Optimization for Mobile Application Using End-to-End Network Slicing in Cloud-Based Vehicular Networks. Sakarya University Journal of Science, 23(1), 23-34, DOI:

10.16984/saufenbilder.404414

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/404414>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Exploring Analytical Model to Performance Optimization for Mobile Application Using End-to-End Network Slicing in Cloud-Based Vehicular Networks

Yönel Kırsal*¹

ABSTRACT

Today, many large usages of cloud-based vehicular networks and applications have rapidly increased. This rapid increase causes the requirement of systems to be reliable to share their resources without delay in order to ensure a better quality of service (QoS) for mobile users. Hence, network slicing is considered one of the key concepts to enhance QoS in 5G networks. At present, new architectures attempt to provide support for end-to-end server quality mechanisms. A key mechanism of network slicing supported by such modern architectures is able to either handover to better network or migrate services closer to the users as they move around. This can be done by advanced handover and server localization techniques. These sorts of advanced handover and server localization help to maintain the QoS for mobile application in heterogeneous environments. In order to obtain QoS measurements and get the network conditions in a specific area, a cloud-based vehicular network slicing management framework is proposed using an analytical modeling approach. The analytical model results obtained consider real scenarios from the Middlesex University vehicular ad-hoc networks (VANET) testbed. Using this framework, the mobile users will make a decision on which situation is better suited to obtain the service based on the latencies as well as queuing capacities of the networks.

Keywords: network slicing, QoS management, analytical modelling, advanced handover, server localisation, vehicular ad-hoc network

1. INTRODUCTION

The next generation wireless network (e.g., 5G) will need to support new demands from a wide variety of users, machines, industries, governments and other organizations. 5G has several heterogeneous requirements from the application perspective, as well as from an architectural perspective [1,2]. From an application perspective, several Internet of Things (IoT) applications present many strengths and challenges, in terms of power, scalability and latency requirements. At the same time, 5G also needs to be highly secure whilst supporting mobility. Mobile subscribers demand uninterrupted connection from anywhere and

anytime as the mobile users move between networks [3]. From an architectural point of view, the new top-down network architecture should be service centric. They fundamentally expose software-defined networking (SDN) and network function virtualization (NFV), where both connectivity services and highly contextualized services should be provided to end users [4]. Hence, in order to meet these requirements, several emerging architectures are under consideration. One of the most important points is network slicing, which allows you to slice the network in a top-down application manner [5]. Efficient network slicing can be achieved by advanced handover and server localization techniques [6,7]. Handovers can be classified according to general characteristics such as network-based/client-based and hard/soft

* Corresponding Author

¹ European University of Lefke, Faculty of Engineering, Department of Electronics and Communication Engineering, Lefke, North Cyprus TR-10 Mersin, Turkey, E-mail:ykirsal@eul.edu.tr

handovers. Handovers can be further classified as advanced handover, dealing with the different mechanisms and inputs [8]. In order to improve the QoS, the mobile users may change their network operator based on the level of performance, or are forced to change the network point of attachment (PoA) [8,9,10]. Proactive handover is a type of advanced handover; and attempts to get the condition of the available networks in close proximity at a specific location before the handovers occur. Mobile users can calculate the time before vertical handover (TBVH) [11], using proactive handover policies [12]. Knowledge-based and model-based handovers policies are two types of proactive handover currently being developed. The knowledge-based policies allow mobile users to measure the signal strength of available networks over a given area beforehand [13]. Model-based handover is based on a mathematical model which calculates the point when a vertical handover should occur, as well as the time that mobile users would take to reach that point based on the mobile users' speed and direction. In this paper, the proposed model is a type of model-based handover. A vehicular cloud-based networks-service, delivery framework for video streaming is presented, which considers the network-slicing factor. We consider, services run on localised clouds, specifically Middlesex University cloud, depending on service demands and networks. A mathematical model is developed and a common scenario is also considered between the road-side units (RSUs) for enabling a 5G system with end-to-end (E2E) network slicing.

The rest of the paper is organised as follows: Section II presents the related studies for network slicing on cloud-based vehicular environments. The proposed cloud-based vehicular system is given in Section III. Section IV introduces the analytical model and a solution approach. Section V presents numerical results. Finally, Section VI provides the conclusion and future work.

2. RELATED WORK

Network slicing organizes computing and communication resources of a physical

infrastructure to enable flexible support of diverse use case realizations. Each physical network is sliced into multiple virtual networks with network slicing to optimize a specific application. A network slice can have its own network architecture and provision in terms of traffic flow and operation. Network slicing techniques in 5G are expected to simplify the creation and to allow mobile users to handover to another network or request that the service is migrated to a cloud system that is closer to the user's location without an interruption while streaming a video [14]. In order to support the network slicing of a wide variety of wireless networks and enable the communication in such heterogeneous networks, Y-Comm framework has introduced as a generic network architecture. It provides the required functionalities to accommodate different wireless technologies as well as supports the mobile services and advanced handovers [15].

2.1. The Y-Comm framework

The Y-Comm framework is a communication architecture to support vertical handover for multi-nodes mobile users in heterogeneous environments. The architecture has two frameworks. The first one is the peripheral framework which deals with operations on the mobile terminal. The second one is the core framework which deals with functions in the core network to support different peripheral networks. A brief explanation of Y-Comm is given in this paper however, a more detailed explanation can be found in [16]. The Y-Comm reference model is shown in Figure 1.

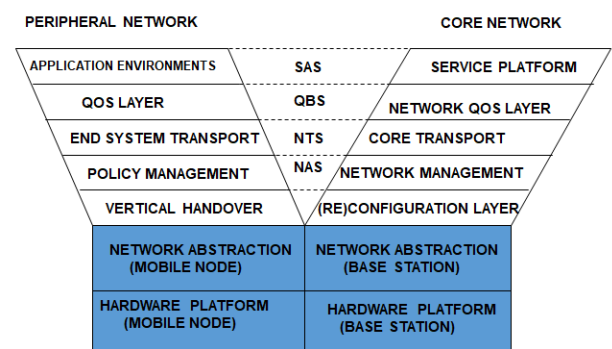


Figure 1. The Y-Comm framework: The reference model [16]

In [6] and [8] a detailed classification of handover and Y-Comm architecture are presented, respectively. However, it is necessary to know how Y-Comm architecture support advance handover in order to perform seamless handover using a mathematical model approach. In Y-Comm, the handover information is managed by the network management layer (NML) in the core framework as shown in Figure 1. The mobile user polls the NML to get information from all available networks as well as the QoS characteristics. The policy management layer (PML) determines where and when a handover should occur based on the obtained information with the direction and speed of the mobile user as well as the QoS. The PML calculates TBVH and the estimated network dwell time (NDT). This information is communicated to the vertical handover layer (VHL) which immediately requests resources to do a handover.

2.2. A Service-oriented framework for mobile services

However, in order to provide a complete set of mechanisms to enable mobile services, it is necessary to develop a new service architecture that allows services to be managed, copied or migrated to support mobile users [15]. This means that services can be managed, copied or migrated to support mobile users. The system provides algorithms that incorporate traffic management as well as the QoS requirements of the flow. This new framework was proposed in [7] as shown in Figure 2 which has six layers are briefly described below:

- The service management layer: This layer manages the service that is being provided. It specifies the functions of the service, registers and obtains a unique service ID. It also specifies the minimum resources required by networking and cloud infrastructure needed to run the service in terms of computing resources, network QoS requirements and storage needs.
- The service subscription layer: This layer handles the functions required for global clients to use the service. It, therefore, allows clients to subscribe to services. It provides the user with a unique client ID, a given service level agreement

(SLA) and sets up accounting and payment mechanisms.

- The service delivery layer: The layer is in charge of delivering the service to a given client. It first maps the SLA to a given QoS and ensures that the selected server as well as its networks can meet the required QoS. The service also receives notifications and triggers about handovers and based on these notifications, this layer may replicate or migrate the service closer to the user.
- The service migration layer: The layer handles the replication or migration of services to different cloud platforms to facilitate a good Quality of Experience (QoE) for the mobile user. Migration is done at the request of the service delivery layer.
- The service connection layer: This layer is responsible for managing the connection between a client and the service. It reports changes in transport or network parameters such as bandwidth or delay to the service delivery layer.
- The network abstraction layer: This layer allows the service to interface to different types of networks depending on network architecture and addressing. This layer maps into IP networking with TCP/IP. In more advanced systems such as Y-Comm, this functionality is spread between the QoS and transport layers in core and peripheral frameworks.

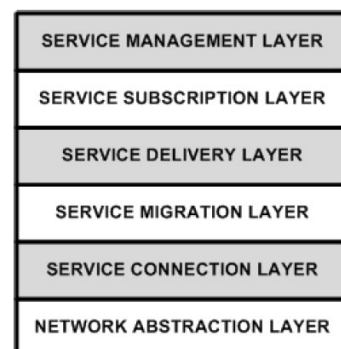


Figure 2. A Service-oriented framework for mobile services

Service migration has been proposed for many environments but has been increased in cloud environments which support virtualization. This is possible due to the virtual machine paradigm which allows entire virtual machines to be migrated [15]. Virtual machine migration can be expensive as the entire virtual machine has to be

migrated. The emergence of container technology, such as Docker, in which containers housing several services can be migrated, is gaining in prominence [1]. Unikernels in which the operating system is bounded and customized to run a single main application is the next emerging specimen in this genre and should make server migration a simpler mechanism from the management viewpoint [1]. However, these efforts assume that the communications architecture does not help facilitate server migration which makes these mechanisms difficult to use in wide area networks (WANs) [15].

The proposed model is a powerful framework that can be used to allow services to migrate from one cloud to another. In order to decide to shift/migrate a service, the works in [6] and [7] have shown how the proposed system should work. In addition, the time taken to shift/migrate the service must be compared with the amount of time the user will be in the region. Thus, the mobility model of the user must be considered. However, in [6] and [7] a simple Markov modeling technique is applied which; there is no generic analytical model as well as a solution approach that can be used for service management. In addition, according to our knowledge, nobody has considered a real vehicular testbed platform to obtain realistic QoS results for such cases. In [17] an analytical modelling approach for large-scale cloud networks is presented from a performance point of view to configure the data center parameters. In addition, analytical models are also presented in [18] for cloud computing. In [19] a hierarchical model is presented for performance analysis of large-scale cloud systems. However, all analytical models presented in the literature have not considered mobile service delivery process with service management framework. In addition, the real-time based scenarios have not been considered. However, in this study all factors and issues mentioned here are considered to get best performance measurements. In this paper, the mobile users decide if a service either migrate (move) to another RSU (network) or shift closer to the mobile user by comparison with network response time obtained based on the slicing factor. The service migration is the

network handover between mobile user and RSU while a service moves. In other words, when the mobile users move through different networks, the QoS information of RSUs passes to the mobile users. Thus, the mobile users initiate connections to the new cloud network where another instance of the service is running. In the proposed management framework mobile users have to decide to migrate the service by checking the response time (latency) of the networks (e.g., RSU). If the RSU response time is already high, the mobile user can ask for a cloud network to shift the service closer to the mobile user.

3. THE PROPOSED CLOUD-BASED VEHICULAR NETWORKS AND ANALYTICAL MODEL

In this section performance optimization of cloud-based service management for the vehicular network is investigated together with network slicing concept considering real scenarios and parameters using the Middlesex (MDX) University VANET testbed [23,24]. In order to optimize the use of the network for each mobile device, a queuing model is presented. This is achieved by E2E network slicing when the service localization and migration are desirable. The MDX vehicular testbed has seven RSUs and located as shown in Figure 3 [24]. However, in this paper, examples of different scenarios of three RSUs (namely RSU 1, RSU 2, and RSU 3) are considered for the proposed analytical model to show the effectiveness of the proposed analytical model and solution considering the common scenarios.

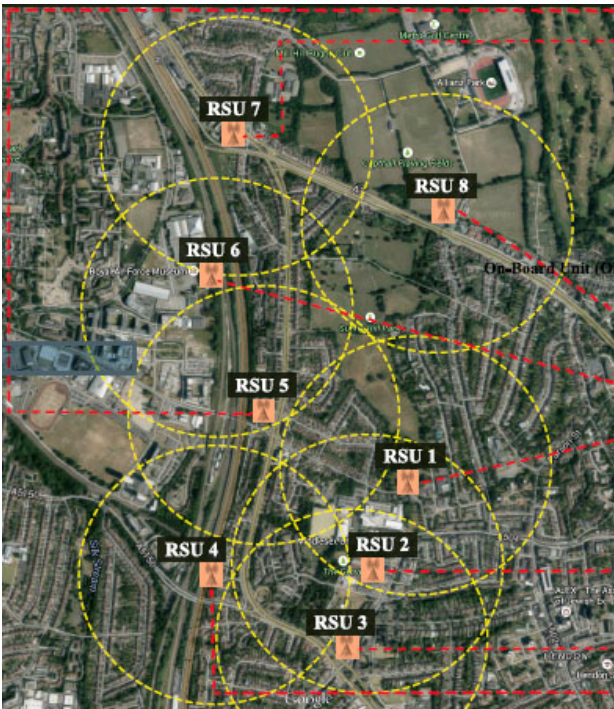


Figure 3. MDX VANET testbed, Hendon, London [23,24]

3.1. The system model

In this section, the proposed system is described as analytical modelling approach to evaluate the QoS of vehicular network considering network slicing using Markov chain. The map of the proposed model considered in this paper is shown in Figure 3. The active mobile user's trajectory is taken from RSU 1 to RSU 3. A mobile user will be connected RSU 1, RSU 2 and RSU 3 during its journey, respectively.

The inputs of the proposed system depend on the mathematical model such as average speed of the users ($E[v]$), speed of mobile users (v), radius of the RSUs (R), service time (T_s), arrival rates (λ) and network dwell times (T_{dwell}). The requests arrive to each RSU in a Poisson stream at rate λ . The required service time, T_s is independent and is distributed exponentially with rate of μ for each RSU. In addition, network dwell times, T_{dwell} for each RSU are assumed to be exponentially distributed with a mean rate of μ_{dwell} for each network. In other words, $\mu_{dwellR,i}$ and $\mu_{dwellL,i}$ are different handover rates from one RSU to another and are defined as the mobility rates in this paper. An important reactive network slicing parameter is defined and denoted as α . By considering α , the mobile user will get a share of

the maximum capacity of the network to which the mobile user is currently attached. In other words, the system will allow the mobile user to calculate the available network resources to use. Furthermore, by using the obtained measurements the mobile users will evaluate other options. The service rate is defined as the perceived rate of service responses that the mobile users are receiving. Hence, the service rates are the factor of α for each RSU. It is assumed that the service rate should be high enough so that it satisfies the request rate from the mobile users. However, the requests queue up by the network and consequently, the response time increases when this condition is not satisfied. Table 1 shows the radius, service and dwell time of all RSUs considered in this paper. Hence, the service rates, as well as the mobility rates can be calculated for proposed analytical modelling considering the real scenario. In addition, the arrival rate λ , the reactive slicing factor of the RSUs α , the average speed of the user ($E[v]=50$ Mph) are assumed for analytical tractability. In addition, the overlapping distance between RSU 1 & RSU 2 and RSU 2 & RSU 3 is 173m and 828m, respectively.

Table 1. Radius, service and dwell time of the RSUs

Networks	Radius (R)	Service Time (T_s)	Network Dwell Time (T_{dwell})
RSU 1	974m	39.57s	43.57s
RSU 2	1390m	58.19s	62.19s
RSU 3	1140m	47.00s	51.00s

The focus here is the multimedia application such as music, streaming video etc. that user wants to use during its movement. Hence, the challenge here is to use network slicing by handover techniques and server localisation to maintain a reasonable QoS in a real VANET testbed. The different formation of RSUs and mobility along a mobile user path represented as a multi-dimensional model as shown in Figure 4. Figure 4 shows the transitions of the system considered. The transitions of the system are described by i and j , specifying the RSUs configuration and number of requests in the RSUs, respectively. Thus, the steady-state probabilities of the system are described as $P_{i,j}$ s where i and j represent the

number of RSUs and number of requests in the RSUs, respectively. In Figure 4, there are three RSU configurations ($i=1,2,3$) at the x-axis. C is the capacity of each RSU. C is the buffer capacity of each RSU. Thus, C is taken as 100 requests in the paper.

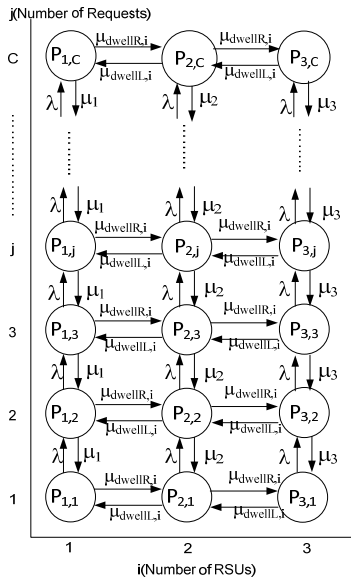


Figure 4. State transit rate diagram of the proposed model

The downward transitions indicate that the requests are being served with service rate μ_i ($i=1,2,3$) which depends on the number of RSUs. On the other hand, upward transitions take place because of the new request for applications with rate λ to the system. Please note that the proposed system is an analytical model for the mobile user. The proposed model does not analyse the given network which will have many users unlike the previous analytical models [12,17–19,21,22]. Hence, the service rate of the network as seen by the user will vary according to how many other users are using the network. The service rate of networks is defined as the comprehend rate of service responses that the mobile users receive. Thus, the service rates are the factor of α . The lateral transitions indicate mobility scenarios between RSUs. As the mobile users move between RSUs, each column shows mobile user’s performance at the given RSU. The mobility of the users between columns is expressed as a mobility rate in the model. $\mu_{dwellR,i}$ and $\mu_{dwellL,i}$, where ($i=1,2,3$), represent different rates for the mobile user to leave the RSUs by moving to the right-hand side and left-hand side, respectively. In this paper, $\mu_{dwellR,i}$ and $\mu_{dwellL,i}$

rates are obtained from the MDX VANET testbed. Each set of the vertical column represents different PoA to the same RSU. Hence, this gives a fine-grained approach to obtain QoS measurements of a single network from different PoA. In other words, for a realistic scenario, the proposed model takes into account the characteristics of each RSU where each vertical column represents different RSU.

4. MARKOV MODEL OF PROPOSED SYSTEM

In this section, the proposed analytical model is adapted and generalized for such systems. The solution of the proposed model can be used for any type of network and configuration for performance analysis.

4.1. Solution of the proposed system

In this study, state probabilities, $P_{i,j}$ s are obtained by using the solution of the successive over relaxation (SOR) method. Thus, the well-known system of balance equations is obtained in order to solve the system. Figure 4 indicates the state transit rate diagram of the proposed system however, when the system generalized we can increase the number of RSUs and request as we like. For instance, N is the maximum number of RSUs and C is the maximum number of requests that can be allocated in the system. MATLAB package is used for the solution of the proposed model. The system state probabilities are obtained with linear equations of the form of $Ax=B$ as shown in equation 1 when the balance equations are derived properly.

$$\begin{bmatrix} A_{0,0} & A_{0,1} & \dots & A_{0,N+C} \\ A_{1,0} & A_{1,1} & \dots & A_{1,N+C} \\ \vdots & \vdots & \vdots & \vdots \\ A_{N+C,0} & A_{N+C,1} & \dots & A_{N+C,N+C} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N+C} \end{bmatrix} = \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_{N+C} \end{bmatrix} \quad (1)$$

In the proposed system, A is of size $N \times C$. x is a column vector of unknown state probabilities (P_i) where $i=0,1,\dots,N \times C$. B consists of the scalars in the balance equations. A is symmetric, real and positive definite matrix. Cholesky factorization is well-known direct method for solving such linear system. It can be computed by form of Gaussian elimination. Hence, the Cholesky

method decomposes matrix A in the form of $A=L^T L$ where L is upper triangular matrix and L^T its transpose. L can be chosen so that its diagonal elements are strictly positive and it is unique. Then, a substitution performed in the equation of $Ax=B$. Hence, the resulting equation is $L^T Lx=B$ which can be solved for x . Let $Lx=Y$ and the equation can be written as $L^T Y=B$. Therefore, first solve $L^T Y=B$ for Y and then solve $Lx=Y$ for x which will give the final solution. If the Cholesky factorization fails, an asymmetric, indefinite factorization is performed. Further details can be found in [25] on the Cholesky factoring. B_i vector is denoted, where $B=\{0,0,\dots,0,1\}$. Then, the solution of the system is as follows: $A_i P_i=B_i$ where $i=0,1,\dots,N \times C$. When the model is in steady state, transition rates used to obtain the matrix A considering the balance equations are in an equilibrium.

$i=0$ and $j=0$;

$$P_{i,j} = \frac{\mu_{\text{dwellL},i+1} P_{i+1,j} + \mu_i P_{i,j+1}}{\lambda + \mu_{\text{dwellR},i}} \quad (2)$$

$i=0$ and $0 < j < C$;

$$P_{i,j} = \frac{\mu_{\text{dwellL},i+1} P_{i+1,j} + \mu_i P_{i,j+1} + \lambda P_{i-1,j}}{\lambda + \mu_{\text{dwellR},i} + \mu_i} \quad (3)$$

$i=0$ and $j=C$;

$$P_{i,j} = \frac{\mu_{\text{dwellL},j+1} P_{i+1,j} + \lambda P_{i-1,j}}{\mu_{\text{dwellR},i} + \mu_i} \quad (4)$$

$0 \leq i < N$ and $j=0$;

$$P_{i,j} = \frac{\mu_i P_{i,j+1} + \mu_{\text{dwellL},i+1} P_{i+1,j} + \mu_{\text{dwellR},i-1} P_{i-1,j}}{\lambda + \mu_{\text{dwellR},i} + \mu_{\text{dwellL},i}} \quad (5)$$

$0 \leq i < N$ and $1 \leq j < C$;

$$P_{i,j} = \frac{\mu_i P_{i,j+1} + \mu_{\text{dwellL},i+1} P_{i+1,j} + \lambda P_{i,j-1} + \mu_{\text{dwellR},i-1} P_{i-1,j}}{\lambda + \mu_{\text{dwellR},i} + \mu_i + \mu_{\text{dwellL},i}} \quad (6)$$

$0 \leq i < N$ and $j=C$;

$$P_{i,j} = \frac{\mu_{\text{dwellL},i+1} P_{i+1,j} + \lambda P_{i,j-1} + \mu_{\text{dwellR},i-1} P_{i-1,j}}{\mu_{\text{dwellR},i} + \mu_i + \mu_{\text{dwellL},i}} \quad (7)$$

$i=N$ and $j=0$;

$$P_{i,j} = \frac{\mu_N P_{i,j+1} + \mu_{\text{dwellR},N-1} P_{N-1,j}}{\lambda + \mu_{\text{dwellL},i}} \quad (8)$$

$i=N$ and $1 \leq j < C$;

$$P_{i,j} = \frac{\mu_N P_{i,j+1} + \lambda P_{i,j-1} + \mu_{\text{dwellR},N-1} P_{N-1,j}}{\lambda + \mu_N + \mu_{\text{dwellL},i}} \quad (9)$$

$i=N$ and $j=C$;

$$P_{i,j} = \frac{\lambda P_{N,L-1} + \mu_{\text{dwellR},N-1} P_{N-1,C}}{\mu_N + \mu_{\text{dwellL},N}} \quad (10)$$

When all the steady-state probabilities, $P_{i,j}$ s, are obtained, various QoS measurements could be easily computed. Thus, the mean queue length (MQL_i), throughput (γ_i) and mean response time (MRT_i) of all RSUs are computed using equations 11, 12, and 13, respectively.

$$MQL_i = \sum_{i=0}^N i \sum_{j=0}^C P_{i,j} \quad (11)$$

$$\gamma_i = \sum_{i=0}^{N_i} \mu_i \sum_{j=0}^C P_{i,j} \quad (12)$$

$$MRT_i = \frac{MQL_i}{\gamma_i} \quad (13)$$

The proposed network slicing framework heavily depends on measuring the mean response time and mean queue length of RSUs. Hence, the decision of RSU selection can be done by the mobile users obtaining the performance measurements. The proposed service management evaluates the various QoS measurements of RSUs and prepares a list of available as well as suitable networks that how much of the network slice of mobile users need.

5. SCENARIO BASED APPLICATION AND NUMERICAL RESULTS

In this section, numerical results are presented for the cloud-based vehicular network for the mobile service delivery framework considering network slicing by calculating mean response time and mean queue length of RSUs when a user is mobile while streaming a video.

In this paper, the wireless footprint method is also assumed in order to obtain mobility pattern information for the mobile user as described in [6]. The user mobility pattern information can be provided based on a user's past mobility patterns. Thus, this method can be used to estimate the next network that the user may join to obtain best QoS for the mobile users. The requirements and modelling issues are considered for video application. In addition, in terms of the QoS and the time the user spends in these networks may vary because of different RSU sizes due to the heterogeneous environment considered. However, the proposed model can be adapted

easily to different types of application as well as the traffic based on the network specifications. Hence, in order to stream a video without interruption in a mobile environment, the analysis done in [20] is followed to obtain optimal latencies for the service decision making process. In [20] authors show that if delay, $D < 40\text{ms}$, the mobile user can carry on using the same RSU without interruption. If $40\text{ms} \leq D \leq 80\text{ms}$, the mobile user should handover to better RSU in order to stream video. On the other hand, if $80\text{ms} < D$, the mobile user can appeal to switch the service closer to the user to maintain QoS. Otherwise, the mobile user will lose the connection.

5.1. Real time scenarios of video streaming

This section includes two different scenarios of a mobile user may be commonly encountered by streaming a video in the real world. The results of these case studies can assist in understanding the proposed model and multimedia content quality adaptation. Furthermore, these results prove analytically that the model is functioning as expected. The system parameters used are mainly taken from MDX VANET testbed as well as the relevant literature [6,7,17–24]. To demonstrate how this model behaves in different scenarios, we study a case where the RSUs have different service rates and keep changing according to the intensity of mobile user in the system. α is the network slicing factor. In other words, it is the probability of service rate changes in all RSUs. Hence, $\mu_i\alpha$ gives service rate changes in all RSUs while streaming a video. All scenarios express a case where a mobile user may be handed-over to any available RSUs or requested from the RSUs to shift the service closer to the mobile user in order to stream a video without interruption. In the considered scenarios, the activity of all these networks is very much dependent on what is happening during the day.

5.1.1. Scenario I

In this scenario, we consider a mobile user moving from RSU 1 to RSU 3. The RSU 1 and RSU 3 have slightly equal service rates as $\mu_1=0.0253$ reqs/sec and $\mu_3=0.0212$ reqs/sec,

respectively. The RSU 2 is located between RSU 1 and RSU 3. The service rate of RSU 2 is $\mu_2=0.017$ reqs/sec as given in Table 2. The RSUs based request rate, λ is 0.01 reqs/sec. The results will let us to predict service decision making process across various RSUs through a user's path.

Table 2. Service and mobility rates of RSUs for scenario I

Networks	Service Rate (reqs/secs)	$\mu_{\text{dwell}R,i}$ (reqs/secs)	$\mu_{\text{dwell}L,i}$ (reqs/secs)
RSU 1	0.0253	0.023	0.023
RSU 2	0.017	0.016	0.016
RSU 3	0.0212	0.019	0.019

As we see from the results in Figures 5 and 6, the mobile user requests are queued upon the RSU 2 which has an insufficient service rate when compared to other networks. This leads to increase in mean response time and mean queue length accordingly. Consequently, there will be performance degradation as the network will be busy if the mobile user joins the RSU 2. On the other hand, RSU 1 had better performance results since the reactive service rate ($\mu_i\alpha$) is higher than the other RSUs.

Finally, the RSU 2 had higher latency and therefore, the performance is insufficient for the mobile user.

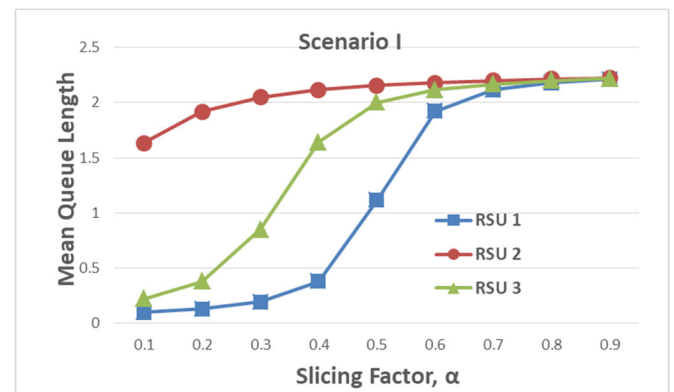


Figure 5. Mean queue length results as a function of slicing factor for scenario I

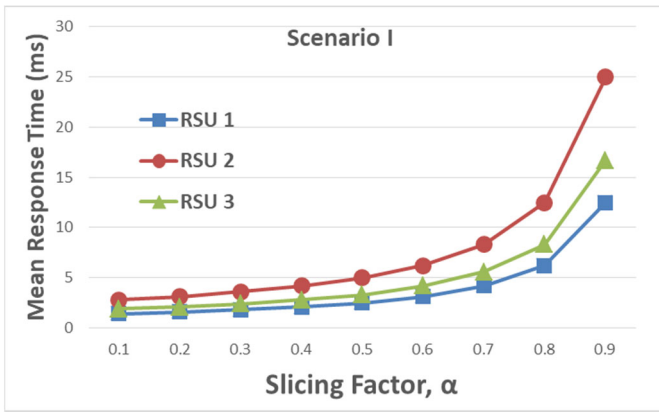


Figure 6. Mean response time results as a function of slicing factor for scenario I

In addition, when the system is significantly busy $\alpha=0.9$, the mobile user can be requested from other RSUs to migrate their services to closer. For example, it is assumed that a mobile user is out of the coverage area of RSU 1 and RSU 3. It is connected to RSU 2 and also it moves out of the coverage area of RSU 2 towards RSU 3. Figures 5 and 6 show the results from this case and we see that the performance of RSU 2 is insufficient to stream a video. The RSU 3 is also not capable of maintaining a high QoS after some threshold point (e.g. $\alpha > 0.8$). In order to achieve efficient utilisation of resources, the mobile user can request to RSU 3 to shift the service closer to it and temporarily stream music from there or shift the service to the RSU 1 without joining to RSU 3 and stream video from there.

5.1.2. Scenario II

In this case, we consider a heavy traffic in the system based on the service rates provided by the networks and the mobility rates of the user as shown in Table 3.

Table 3. Service and mobility rates of RSUs for scenario II

Networks	Service Rate (req/secs)	$\mu_{dwellR,i}$ (req/secs)	$\mu_{dwellL,i}$ (req/secs)
RSU 1	0.067	0.023	0.023
RSU 2	0.034	0.016	0.016
RSU 3	0.05	0.019	0.019

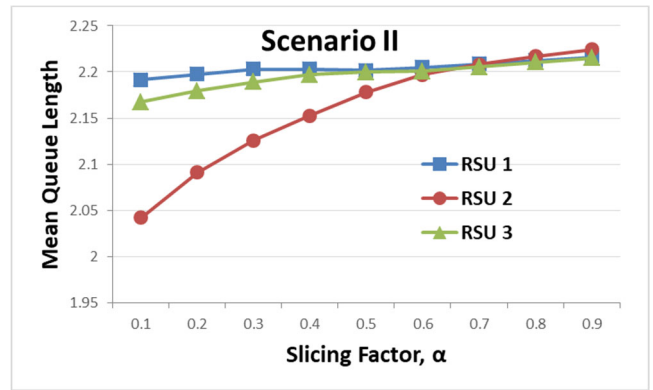


Figure 7. Mean queue length results as a function of slicing factor for scenario II

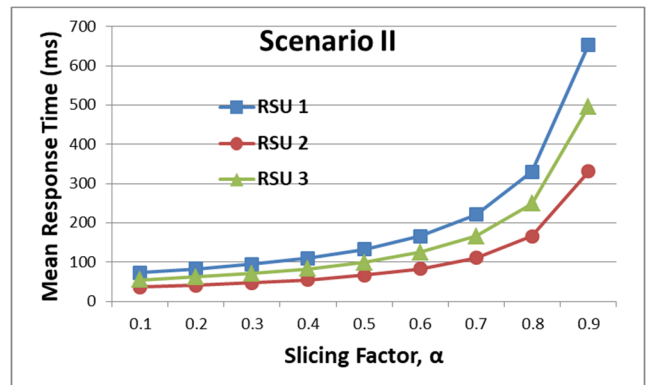


Figure 8. Mean response time results as a function of slicing factor for scenario II

Figures 7 and 8 also show the mean queue length and the mean response time results as a function of α for scenario II, respectively. As clearly seen from the figures, the performance results of the RSU 2 are better than the performance results of the RSU 1 and RSU 3. Thus, the mobile user can be handed-over to RSU 3 even though the serving capability of RSU 3 is not sufficient to stream a video without interruption when $0.1 < \alpha < 0.3$. However, a mobile user should find an alternative network to be handed-over in order to stream a video without interruption after this. In addition, the mobile user has to request a service migration from the available RSUs in order to stream a video when $\alpha \geq 0.4$ for the RSU 3. Table 4 shows the decision making results to find the sequence of RSUs that the mobile users will attach. It will either handover to available networks or request to service localization.

Table 4. Mean response times of RSUs for scenario II

α	RSU 1	RSU 2	RSU 3
0.1	73.9s	37s	55.5s
0.2	83.1s	41.6s	62.4s
0.3	95s	47.6s	71.4s
0.4	110.7s	55.5s	83.2s
0.5	132.8s	66.6s	99.9s
0.6	165.9s	83.2s	124.8s
0.7	220.8s	110.9s	166.3s
0.8	330.1s	166.3s	249.1s
0.9	653.8s	331.7s	496.3s

Hence, for such cases, we may consider the mobile user to start streaming the video from the RSU 1 as it firstly hits the coverage area of the RSU 1. When the mobile user reaches the coverage area of RSU 2, after some time, the mobile user might have to handover the service to available RSUs when $0.1 < \alpha < 0.6$ as shown in Table 4. However, the mobile user can be handed-over to the other RSUs to carry on its service without interruption when $\alpha \leq 0.5$ for RSU 2.

Table 5. Comparison of mean response time results in the absence of the proposed approach (PM: Proposed model and APM: Absence of the proposed model)

α	RSU 1		RSU 2		RSU 3	
	PM	APM	PM	APM	PM	APM
0.1	73.9s	76.4s	37s	38s	55.5s	56.8s
0.2	81.1s	90.2s	41.6s	43s	62.4s	63.3s
0.3	95s	98.7s	47.6s	50.6s	71.4s	75.2s
0.4	110.7s	120.8s	55.5s	60.2s	83.2s	89.7s
0.5	132.8s	143.8s	66.6s	80.8s	99.9s	110.9s
0.6	165.9s	200.1s	83.2s	90.3s	124.8s	140.1s
0.7	220.8s	300.6s	110.9s	150.4s	166.3s	183.2s
0.8	330.1s	440.2s	166.3s	223.5s	249.1s	320.4s
0.9	653.8s	700.2s	331.7s	452.7s	496.3s	520.8s

In order to show the effectiveness of the proposed model the performance results presented in Table 4 are also considered and presented in Table 5 in the absence of the proposed approach. It is clearly evident from the Table 5 that the proposed model increases the system performance significantly. The mean response time results show that all of the RSUs become quickly busy when the proposed model is not considered. However, the response times of each RSUs decreases (in other words, each RSUs respond much quicker) when the proposed model is applied. Thus, the QoS of each RSUs increase in terms of response time when the proposed model is considered.

Both scenarios demonstrate how various cases may be constructed considering the proposed model. It is clearly seen from the results that network slicing factor α , affect the system performance significantly. Hence, the proposed model can be used to analyze the overall performance of each scenario. This will lead the mobile users to choose an existing network to be served without interruption. In other words, the mobile users will decide to select the available RSU with better performance. Thus, this model can be applied as a network slicing framework in vehicular environments to improve QoS.

6. CONCLUSION

This paper proposed an analytical modelling approach and QoS provisioning for E2E network slicing service management framework in vehicular environments. The proposed model requests services directly from the available RSUs rather than offered services by the physical resources. Thus, this will simplify the process for end users and also gives opportunities to direct mobile user requests to the most appropriate RSUs or to run a service closer to the mobile user location. The proposed model calculates the increase in delay as a user moves while streaming a video. It offers the perspective of considering the response time of the each RSU in the system before the user requests a service. The mobile user can make a decision by analysing the RSU. It may either handover the service to available RSUs or requests to migrate the service to closer

location without an interruption while streaming a video. Overall, the results were convincing and show that the proposed approach can be useful in real networks to maintain QoS for the mobile user in vehicular networks.

ACKNOWLEDGMENTS

The author would like to thank Dr. Gelnford Mapp, Dr. Fragliskos and Mr. Vishnu V. Paranthaman for their guidance and helpful suggestions. Dr. Glenford Mapp has a funding from Department of Transport (DfT) and he is the head of Middlesex University VANET research team. Dr. Yönal Kirsal is a visiting fellow (international collaborator) of the MDX VANET research team.

REFERENCES

- [1] Q. Li, G. Wu, A. Papathanassiou, L. Wei, "End-to end network slicing in 5G wireless communication systems", *In Proc. of ESTI workshop on future radio technologies: Air interfaces*, Jan. 27-28, 2016.
- [2] P. K. Agyapong, M. I., "Design Considerations for a 5G Network Architecture", *IEEE Communications Magazine*, vol. 52, no. 11, 2014, pp. 65-75.
- [3] H. Zhang, N. Liu, X. Chu, K. Long, A. Aghvami, V. C. M. Leung, "Network Slicing Based 5G and Future Mobile Networks: Mobility Resource Management and Challenges", *Communications Magazine IEEE*, vol.55, 2017, pp. 138-145.
- [4] Ericsson white paper, 5G systems, <http://www.ericsson.com/res/docs/whitepapers/whatis-a-5g-system.pdf>, Jan. 2015. Accessed: 2015-05-29.
- [5] Nokia, Dynamic end-to-end network slicing for 5G, White paper, Finland, 2016
- [6] F. Sardis, G. Mapp, J. Loo, M. Aiash and A. Vinel, "On the Investigation of Cloud-Based Mobile Media Environments with Service-Populating and QoS-Aware Mechanisms," *IEEE Transactions on Multimedia*, vol. 15, no. 4, pp. 769-777, June 2013.
- [7] F. Sardis, "Exploring Traffic and QoS Management mechanisms to support mobile cloud computing using service localization in heterogeneous environments", School of Science and Technology, Middlesex University, August 2014, PhD Thesis.
- [8] G. Mapp, F. Katsriku, M. Aiash, N. Chinnam, R. Lopes, E. Moreira, R. P. Vanni, and M. Augusto, "Exploiting Location and Contextual Information to Develop a Comprehensive Framework for Proactive Handover in Heterogeneous Environments" *Journal of Computer Networks and Communications*, pp. 1-17, 2012.
- [9] G. Mapp, F. Sardis and J. Crowcroft, "Developing an implementation framework for the Future Internet using the Y-Comm architecture SDN and NFV", *IEEE NetSoft Conference and Workshops (NetSoft)*, Seoul, pp. 43-47, 2016.
- [10] X., Dionysis, "Handover decision for small cells: Algorithms, lessons learned and simulation study", *Computer Networks*, 100, 2016, pp. 64-74.
- [11] F. Shaikh, Intelligent proactive handover and QoS management using TBVH in heterogeneous networks [Ph.D. thesis], School of Engineering and Information Sciences, Middlesex University, January 2010.
- [12] Y. Kirsal, "Analytical Modelling of a New Handover Algorithm for Improve Allocation of Resources in Highly Mobile Environments", *International Journal of Computers Communications and Control*, 2016, 11, 6, pp. 755-770.
- [13] D. Cottingham, I. Wassell, and R. Harle, "Performance of IEEE 802.11a in vehicular contexts", *In Proceedings of IEEE 65th Vehicular Technology Conference*, 2007, pp. 854-858.
- [14] NGMN Alliance, "5G White Paper", February 2015.

- [15] A. Ghosh, V. V. Paranthaman, G. Mapp, O. Gemikonakli, J. Loo, "Enabling seamless V2I communications: toward developing cooperative automotive applications in VANET systems", *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 80-86, 2015.
- [16] G. Mapp, F. Shaik, D. Cottingham, J. Crowcroft, J. Baliosian, "Y-Comm: a global architecture for heterogeneous networking", In Proceedings of the 3rd international conference on Wireless internet (WICON '07). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, Article 22, 5 pages.
- [17] D. Bruneo, "A stochastic model to investigate data center performance and QoS in iaas cloud computing systems", *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, 2014, pp. 560-569.
- [18] J. Vilaplana, F. Solsona, I. Teixid, J. Mateo, F. Abella, and J. Rius, "A queuing theory model for cloud computing", *The Journal of Supercomputing*, vol. 69, no. 1, 2014, pp. 492–507, 2014.
- [19] R. Ghosh, F. Longo, V. K. Naik, and K. S. Trivedi, "Modeling and performance analysis of large scale iaas clouds", *Future Generation Computer Systems*, vol. 29, no. 5, 2013, pp. 1216-1234.
- [20] P. Vidales, J. Baliosian, J. Serrat, G. Mapp, F. Stajano and A. Hopper, "Autonomic system for mobility support in 4G networks", *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 12, pp. 2288-2304, Dec. 2005.
- [21] Y. Kirsal, E. Ever, A. Kocyigit, O. Gemikonakli, G. Mapp, "Modeling and analysis of vertical handover in highly mobile environments", *The Journal of Supercomputing*, 71, pp. 4352-4380, 2015.
- [22] R. Guillaume, A. G. Andres, A. Ben, "LTE Advanced and Next Generation Wireless Networks Channel Modeling and Propagation", John Wiley and Sons Ltd. 2013.
- [23] Building an Intelligent Transport Information Platform for Smart Cities, Report, <http://www.vanet.mdx.ac.uk/>, 2016. Accessed: 2017-09-29.
- [24] Building a Connected Vehicle Testbed to study the development and deployment of C-ITS in the UK, <http://www.its-ukreview.org/building-a-connected-vehicle-testbed-to-study-the-development-and-deployment-of-c-its-in-the-uk/>, 2016, Accessed : 2017-09-30.
- [25] N. J. Higham Functions of Matrices: Theory and Computation. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2008, ISBN 978-0- 898716-46-7.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Assesment of Power Quality Disturbances For Grid Integration of PV Power Plants

Authors: Gökay Bayrak, Alper Yılmaz

Recieved: 2018-03-27 17:25:14

Revised: 2018-07-12 10:09:47

Accepted: 2018-08-01 13:44:10

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 35-42

How to cite

Gökay Bayrak, Alper Yılmaz; (2019), Assesment of Power Quality Disturbances For Grid Integration of PV Power Plants. Sakarya University Journal of Science, 23(1), 35-42, DOI: 10.16984/saufenbilder.407323

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/407323>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Assesment of Power Quality Disturbances For Grid Integration of PV Power Plants

Gökay BAYRAK*¹, Alper YILMAZ

ABSTRACT

Power Quality problems, which have become an important consumer issue in recent years, are defined as changes in voltage, current, or frequency in the power system. Among the factors affecting energy quality in grid-connected PV systems are island mode operation, current and voltage harmonics, transients, flicker, interruption, DC offset, notches, frequency changes, voltage sag/swell, voltage imbalances in the system and power factor. Several transmission and distribution losses consist of both the consumers and the generators sides because of the power quality problems. The integration of PV power plants to the main grid will cause several power quality problems so a reliable operation of the grid with PV power plants is a significant issue for a distributed generation. Thus, the first step in preparing a reliable algorithm for detecting power quality events occurring in the current grid is to model a power system in which power quality impairments can be analyzed. In this study, the power quality disturbances that occur in the low-voltage grid that is fed through both the main grid and the grid-connected PV system are modeled and investigated. Developed electric power distribution model includes simulation of voltage sags caused by the three-phase fault, transformer energization and asynchronous motor switching, voltage swells caused by the three-phase fault, transients due to large capacitor bank switching, harmonics and notches caused by the load connected via the power converter. Examination of the power quality disturbances with simulation clearly revealed the resulting waveforms, the response of the electrical power system to the fault conditions. Another advantage of the realized study is that the developed model can be used to measure the performance of the PV connected distributed generation system in fault detection and classification studies.

Keywords: Distributed Generation, Power Quality Disturbances, Grid-Connected PV Systems

1. INTRODUCTION

The ever-increasing involvement of new consumers in the electricity grid, changes in user needs, electricity market expectations, and extreme dependence on today's society's advanced technological services are triggering more reliable and more flexible electrical infrastructure work. For this reason, integration

of renewable energy sources into the existing electricity network and the identification and prevention of power quality (PQ) problems have become a priority issue [1].

PQ parameters are a set of limits that allow an equipment to operate as intended without significant performance and life expectancy loss [2]. Any PQ problem that causes voltage, current and frequency deviations can lead to failure of consumer equipment or improper operation of

* Corresponding Author

¹ Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, 16330, Bursa, TURKEY, gokay.bayrak@btu.edu.tr

² Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, 16330, Bursa, TURKEY, alper.yilmaz@btu.edu.tr

equipment. Among the factors for power quality are island mode operation, harmonics, transients, flicker, interruption, DC current injection, noise, notches, frequency changes, voltage sag/swell, voltage imbalances in the system and power factor.

Distributed energy sources, which are increasing in popularity each year in the electrical energy market, play an important role in overcoming the global energy crisis. The rapid decline in the production costs of photovoltaic (PV) power plants over the last few years has led to the rapid development in PV energy systems. The integration of a grid-connected PV system with the existing grid and the identification of PQ problems and the establishment of conditions for identifying these problems have become an important issue.

The first step in preparing a reliable algorithm for detecting and mitigation PQ events occurring in the current grid is to model a power system in which PQ disturbances can be analyzed [3]. Classification based fault detection methods (artificial neural networks [4], decision trees, support vector machines [5], feature extraction) and field alternating fault methods (Fourier and Wavelet transform) utilize the fault signals obtained by modeling the electrical grid [6-8].

Waveforms of power quality impairments can be obtained by mathematical methods, simulation studies and real-time data obtained through the power system. It is proposed by a mathematical method for the analysis of waveforms resulting from PQ problems [9]. A Labview-based power quality signal generator has been proposed that can reproduce PQ disturbances that overlap with actual system data in another study [10]. One of the studies carried out, detection and classification of islanding and PQ disturbances were performed in a distributed generation (DG) based hybrid power system connected to the main grid [11]. In [12], PQ disturbances were identified and classified using real data collected from the system.

In this paper, the PQ disturbances that occur in the low-voltage grid that is fed through both the main grid and the 1 MWp grid-connected PV plant are modeled in MATLAB / Simulink environment. Developed electric power

distribution model includes different types of PQ disturbances simulation such as voltage sags, voltage swells, transients, harmonics and notches.

2. POWER QUALITY DISTURBANCES IN GRID-CONNECTED PV PLANTS

Different international standards for regulating power quality have been established. For example, IEEE 1547 standards set limits for total demand distortion (TDD) and amplitude of harmonic currents delivered to the main grid by grid-connected PV systems [13, 21]. Another important standard for research is the IEEE Std. 519, which specifies voltage and current harmonic limits [14]. Apart from these, IEEE Std. 1159 offers some suggestions for monitoring the quality of power events [15]. In this study, the fundamentals of the power quality events and the restrictions that are defined for the reliable grid operation are investigated.

2.1. Voltage Fluctuation

When PV power plants enter and exit the circuit, it may cause a voltage drop or rise. These values are limited to $\pm 3.3\%$ V according to IEC 61000-3-3 standards [16]. In addition, the effective value of the voltage provided by the PV plant may deviate by $\pm 10\%$ of the nominal value at 95% of the year, according to the standards. When voltage analysis is performed, the limit values that the system can tolerate should be analyzed.

2.2. Long and Short-time Voltage Variations

Variations in the effective value of the nominal voltage that occurs over a period of more than one minute are called long-term voltage variations, while short-term changes in the nominal voltage of the grid are called short-term voltage variations. Depending on the magnitude of the voltage change, long-term voltage variations can be classified as under-voltage, over-voltage, and continuous interruption.

Voltage sag/dip is defined as a 10% to 90% reduction in effective voltage, limited to a time interval of 10 ms to 60 s on a system operating

under nominal conditions (IEEE Std. 1159). This disturbance is caused by the switching of high power motors, overloads and short circuit faults along the line.

Voltage swell is defined as a 110% to 180% increase in voltage, limited to a time interval of 10 ms to 60 s on a system operating under nominal conditions (IEEE Std. 1159). Along with being a type of malfunction that is not often seen in the electrical system, the maneuvers in the power system, such as the disconnection of large inductive loads or the actuation of large capacitor banks, result in voltage swells. Depending on the fault location and power system conditions, a fault or PQ problem may cause a voltage drop, swelling or interruption. Waveforms of voltage sag, voltage swell, and voltage interruption are shown in Fig. 1.

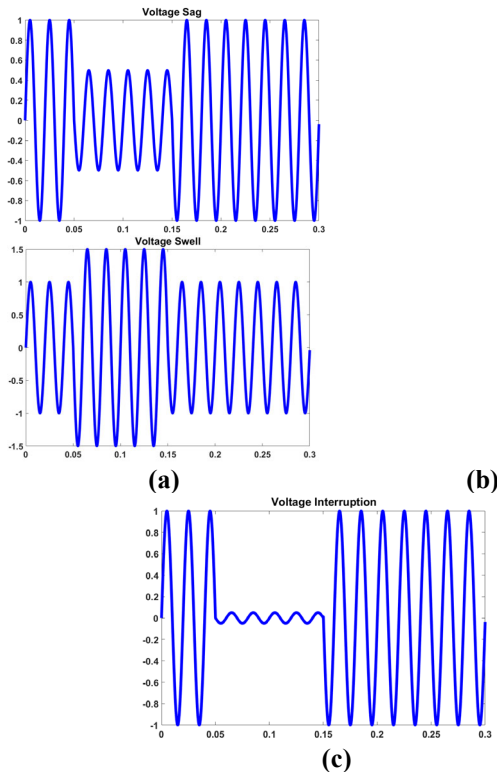


Figure 1. Signal samples of (a) Voltage Sag, (b) Voltage swell, (c) Voltage interruption

2.3. Voltage Imbalance

For the low-voltage (LV) system, the imbalance should not exceed 2% in 95% of the measurement time according to the EN 50160 standards, which are based on only the negative component of the voltage. In order to prevent

voltage imbalance, the power of the PV plant must be distributed equally.

2.4. Harmonics

The ratio of the total effective value of the harmonics outside the fundamental harmonic in the voltage and current waveform to the effective value of the fundamental harmonic is defined as the total harmonic distortion (THD).

$$THD = \frac{\sqrt{\sum_{n=2}^{\infty} V_n^2}}{V_1} \quad (1)$$

where V_n is the amplitude of the n-th harmonic components, V_1 is fundamental component.

Depending on the IEEE 519 standards, the total harmonic distortion for voltage in the power system is limited to a maximum of 8% in low voltage grids. In addition, each harmonic value is limited to be less than 5%. The most commonly used power quality standards for low voltage systems are IEC 61000-2-2, IEC 61000-3-2, and 61000-3-4 respectively. According to these standards, the THD amount must be less than 8% up to the 40th harmonic value. Fig. 2 shows harmonics caused by the universal bridge connection.

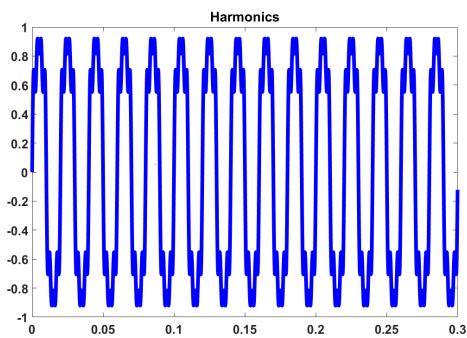


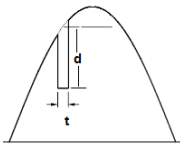
Figure 2. Harmonics caused by universal bridge connection

2.5. Notches

The notch is a voltage waveform distortion of the converter in a period of the voltage in a system operating at the grid frequency, which is proportional to the number of pulses of the converter [17]. This distortion is shorter than half a period. Since the frequency components generated by the notch effect are too high, they cannot be identified with conventional harmonic measuring devices. According to the IEEE 519-

2014 standard, the notch area and the depth-dependent limit values are as shown in Table 1.

Table 1. Notch area and depth definition (IEEE 519-2014)

System	Special Applications	General System	Systems with Converters
Notch Depth	%10	%20	%50
Notch Area (A_N)	16400	22800	36500
Notch area and depth definition (IEEE 519-2014)	 <p>% notch depth = $d/V \times 100$ $A_N = td = u \text{ sec} \times \text{volt}$</p>		

2.6. Transients

Transients that can be examined in two classes, oscillatory and impulsive, are undesirable momentary events in power systems. These events occurred in less than half a period [18]. Impulsive transients are usually sudden and unidirectional distortions caused by lightning strokes. In oscillatory transients, the resultant change in polarity of the current or voltage occurs very quickly, resulting in bi-directional distortions. Oscillatory transitions are caused by switching off/on large loads, energization of the capacitor banks or the energization of the transformers. The voltage signal waveform containing the transient is given in Fig. 3.

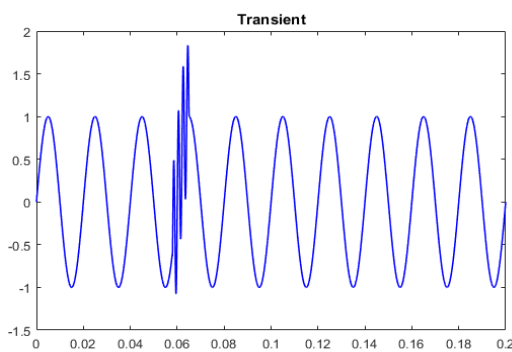


Figure 3. Transient disturbances waveform

3. DEVELOPED ELECTRICAL POWER DISTRIBUTION MODEL

This study includes simulation of voltage sags and swells, transients, harmonics, voltage fluctuations, and notches. In this paper, effective

value conversion was used to observe the effect of power quality disturbances on the voltage.

$$U_{RMS} = \sqrt{\frac{1}{M} \sum_{i=1}^M u(i)^2} \quad (2)$$

where M is window size and $u(i)$ is the voltage waveform is the i -th sample.

3.1. PV Plant Modeling

Fig 4 shows the developed grid-connected PV system [19, 20] has sub-models consisting of a PV plant, H-bridge converter, MPPT controller module, load and the grid connection modules. The 1 MWp PV power plant is connected to the DC line via the three-phase inverter and the three-phase inverter output is connected to the 10.5 kV Bus1 via the LCL filter.

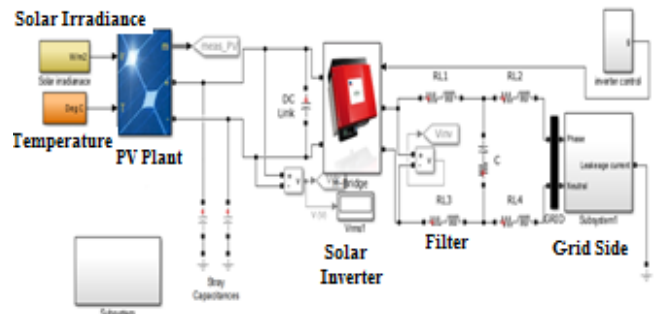


Figure 4. Developed a grid-connected PV system for residential power plants [17]

3.2. Electrical Power Distribution Model

In order to detect PQ problems, the response of the power system should be observed first in case of a fault. The simulation model was developed using MATLAB / Simulink which is shown in Fig. 5. The simulation model realized include phase-to-phase and phase-to-ground short circuit faults, starting with a large powerful asynchronous motor, transformer energizing, capacitor bank switching and system behavior under nonlinear load.

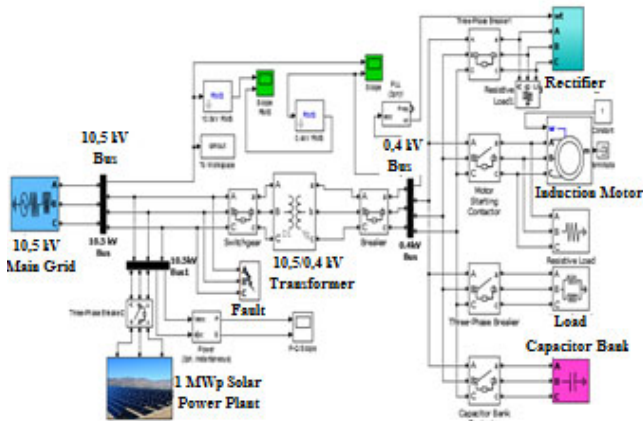


Figure 5. Matlab/Simulink model of the developed electrical power distribution system

The power distribution system model consists of 10.5 kV main grid, 10 MVA generator, 1 MWp grid-connected PV plant, a delta/star connected step down two windings transformer, inductive and resistive loads, induction motor, capacitor bank and three phase nonlinear load. The transformer of 10.5 kV/0.4 kV supplies three-phase nonlinear (20 kVA RL load) and normal (150 kVA RL load) loads, squirrel-cage 160 kW induction motor at the point of common coupling (PCC). 0.4 kV load bus also equipped with 50 kVA capacitor bank.

4. ASSESSMENT OF POWER QUALITY ISSUES

In this section, several grid conditions are investigated to indicate the power quality effects of the PV power plant to the main grid and the other electrical equipment. In three-phase systems, faults between phases and between phases and ground can be examined in two groups, symmetric and asymmetric. Fig. 6 shows the effective values of the voltages resulting from the fault occurring as a result of the short-circuiting of the three phases between 0.15 and 0.25 seconds. The resulting fault is symmetrical because it creates similar effects on all three phases.

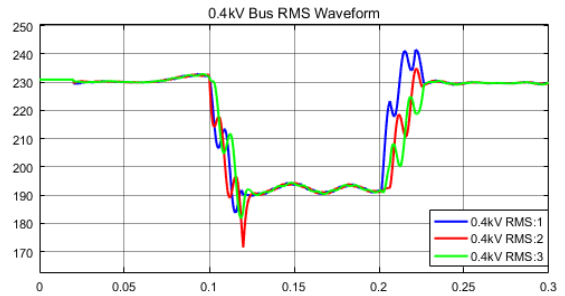


Figure 6. Effective values of voltages that occur as a result of three-phase short circuit fault

Fig. 7 shows the changes in the effective values of the voltages of the phases during faults between A and B phases between 0.1 and 0.2 seconds. The short circuit fault resistance on the model is selected as 7 ohms.

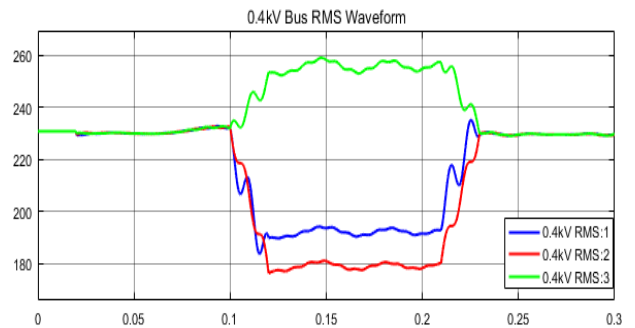


Figure 7. Effective values of phase-to-phase voltage resulting from faults

As a result of the simulation, voltage swell occurs in phase C, while voltage sag is observed in phases A and B. As seen in the Fig. 7, two-phase faults, single phase-to-ground, and two phase-to-ground faults show asymmetrical characteristics.

The breaker of the transformer was closed at 0.05 second on the model for the analysis of the voltage sag that occurred as a result of transformer energization. As a result of transformer energization, the change in instantaneous and effective values of phases A, B and C is given in Fig 8. FFT analysis used in determining the harmonic distortion is shown in Fig. 8(b). Because of the different phase angles, there is also a difference in the size of the voltage sag of each phase. The collapse of the voltage resulting from the impulse current and the saturation of the core caused a sudden drop in the system voltage and a long period of stabilization began. Simulation studies have shown that the voltage sag is larger when transformer power or line power is increased. In the FFT analysis, it

was observed that the n-th harmonic components were most affected by transformer energization of 2nd, 6th, 12th and 18th harmonics.

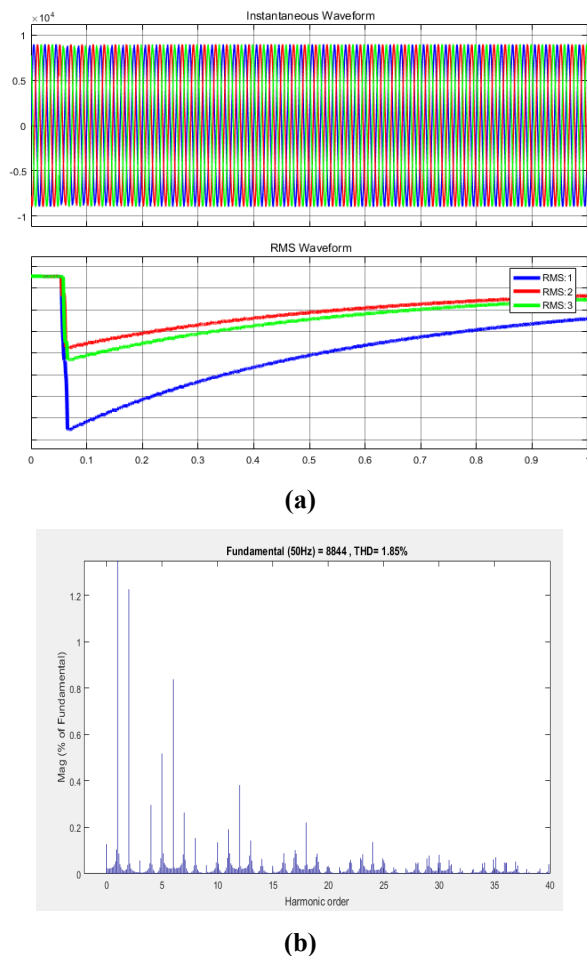


Figure 8. The transformer energization results in (a) the change in the instantaneous and effective values of phases A, B and C, (b) FFT analysis

The voltage waveforms of phases A, B, and C at the 400 V low voltage bus as a result of the asynchronous motor start-up are shown in Fig 9. Simulation duration is 0.3 seconds and 0.02nd seconds the asynchronous motor is activated by closing the circuit breaker. As a result of the voltage sag event, a voltage drop of about 15% in the voltage effective values of the 400 V bus has been observed. Simulation studies have shown that the amount of sag depends on the power of the motor. The resulting fault showed a balanced distribution unlike the transformer energization and each phase voltage showed similar behavior.

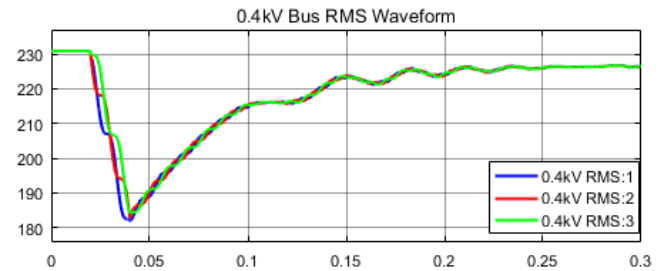


Figure 9. Voltages in the result of induction motor start-up

In order to investigate the power quality disturbance caused by the switching of the capacitor bank, 0.4 kV 50 kVA capacitor bank connected to the 400 V bar on the model was switched on with the help of a phase breaker (Fig. 4). Fig. 10 shows capacitor bank switching upon closing of the breaker at 400 V feeder line causes transient in voltage at low voltage bus at 0.03 seconds. In this short-time event, the effective value of the voltage increased about twice, especially in phase B, and the system frequency changed. After the time of the first uprising, this oscillation continued to increase and decrease in voltage and settled at the latest nominal value. Size of capacitor bank affects the voltage transient frequency. The low voltage transient frequency is caused by very large capacitor banks. Increasing the amount of load causes a higher damping factor.

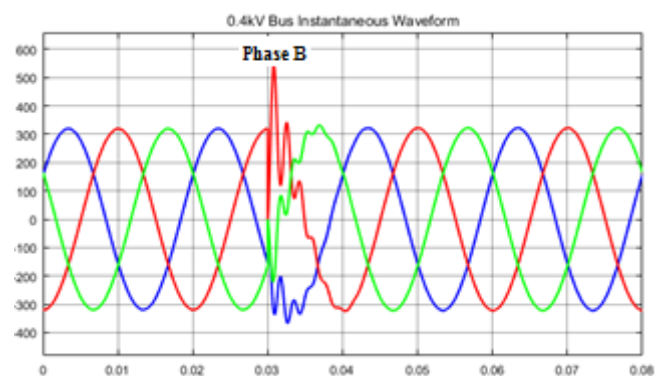


Figure 10. The voltage waveforms resulting from the switching of the capacitor bank

A nonlinear load is also fed through a switched converter to observe the notch formation. Figure 11 shows that each notch has a depth and a duration similar to those of the other. It is seen

that the harmonic values of 5th, 7th, 11th and 13th components are high in the FFT obtained as a result of notch formation. The location and depth of the voltage notch depend on the ignition angle control of the converter. The notch width depends on the amount of inductive load and increases as the number of load increases.

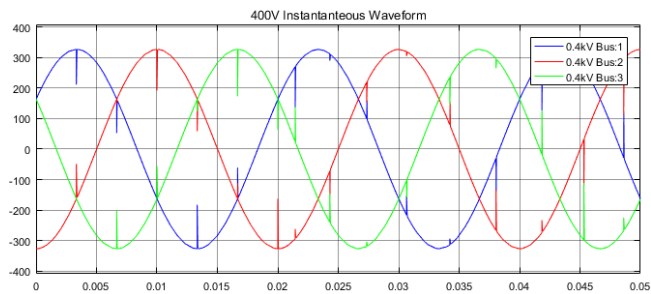


Figure 11. Voltage notch waveforms at low voltage bus

5. DISCUSSION AND CONCLUSION

In this study, the power quality (PQ) disturbances that occur in the low-voltage grid that is fed through both the main grid and the 1 MW_p grid-connected PV plant are produced by using developed electrical distribution model in MATLAB/ Simulink environment. Examination of the PQ disturbances with simulation clearly revealed the resulting waveforms, the response of the electrical power system to the PQ disturbances conditions.

The advantage of the realized study is also that the developed model can be used to measure the performance of the PV connected distributed generation system in fault detection and classification studies.

REFERENCES

- [1] Dugan, R.C., Mcgranaghan, M.F. Santoso, S. and Beaty, H.W., 2002. Electrical power systems quality, McGraw-Hill, New York.
- [2] Bollen, M.H.J. and Gu, I.Y.H., 2006, Signal processing of power quality disturbances, John Wiley&Sons, New York.
- [3] Khokhar, S., Zin, A. M., Mokhtar, A. S., & Ismail, N. A. M. (2014, October). MATLAB/Simulink based modeling and simulation of power quality disturbances. In *Energy Conversion (CENCON)*, 2014 IEEE Conference on (pp. 445-450). IEEE.
- [4] Haykin, S. 2008. “Neural networks and learning machines”, Pearson, 3rd edition.
- [5] Burges, C J.C. 1998. “A tutorial on support vector machines for pattern recognition”, *Data Mining and Knowledge Discovery*, 2, 121-167.
- [6] Silva, K.M., Souza, B.A., Brito, N.S.D. 2006. “Fault detection and classification in transmission lines based on wavelet transform and ANN”, *IEEE Trans Power Deliv*, 21, 2058–63.
- [7] Dash, P.K., Samantaray, S.R., Panda, G. 2007. “Fault classification and section identification of an advanced series-compensated transmission line using support vector machine”, *IEEE Trans Power Deliv*, 22, 67–73.
- [8] Valtierra-Rodriguez, M., de Jesus Romero-Troncoso, R., Osornio-Rios, R. A., & Garcia-Perez, A. (2014). Detection and classification of single and combined power quality disturbances using neural networks. *IEEE Transactions on Industrial Electronics*, 61(5), 2473-2482.
- [9] Rodriguez-Guerrero, M. A., Carranza-Lopez-Padilla, R., Osornio-Rios, R. A., & Romero-Troncoso, R. D. J. (2017). A novel methodology for modeling waveforms for power quality disturbance analysis. *Electric Power Systems Research*, 143, 14-24.
- [10] Simić, M., Kokolanski, Z., Denić, D., Dimcev, V., Živanović, D., & Taskovski, D. (2017). Design and evaluation of computer-based electrical power quality signal generator. *Measurement*, 107, 77-88.
- [11] Mohanty, S. R., Ray, P. K., Kishor, N., & Panigrahi, B. K. (2013). Classification of disturbances in hybrid DG system using modular PNN and SVM. *International Journal of Electrical Power & Energy Systems*, 44(1), 764-777.

- [12] Zhang, M., Li, K., & Hu, Y. (2011). A real-time classification method of power quality disturbances. *Electric power systems Research*, 81(2), 660-666.
- [13] IEEE Std. 1547, IEEE standard for interconnecting distributed resources with electric power systems, 2015.
- [14] IEEE Std 519, IEEE recommended practices and requirements for harmonic control in electrical power systems, 2014.
- [15] IEEE Std 1159, IEEE Recommended Practice for Monitoring Electric Power Quality, 2009.
- [16] Electromagnetic Compatibility- Limitation of voltage changes, voltage fluctuations and flicker in public low-voltage supply systems, equipment with rated current ≤ 16 A per phase and not subject to conditional connection, IEC 61000-3-3, 2013.
- [17] Regina Lamedica, Alberto Prudenzi, Enrico Tironi, Dario Zaninelli – Analysis of Harmonics Distortion Limits in IEC and IEEE Standards – *Electrical Power Quality and Utilisation Journal*, vol. 5, iss. 2, sept. 1999, pp. 47-52.
- [18] M. H. Bollen, *Understanding power quality problems* vol. 3: IEEE Press New York, 2000.
- [19] Bayrak. G., Cebeci. M. 2014. Grid-connected fuel cell and PV hybrid power generating system design with Matlab Simulink. *International Journal of Hydrogen Energy*. 39(1):8803-8812.
- [20] Bayrak, G., Yilmaz, A., “Design and Performance Assessment of A Grid-Connected PV System for Residential Power Generation”, 7th International 100% Renewable Energy Conference, pp: 62-67,18-20 Mayıs 2017, İstanbul, TURKEY
- [21] Bayrak, G., Kabalci, E., “Implementation of a new remote islanding detection method for wind-solar hybrid power plants”, *Renewable and Sustainable Energy Reviews*, 58:1-15, 2016.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Performance Comparison for Series and Parallel Modes of a Hybrid Electric Vehicle

Authors: Gökhan Canbolat, Halit Yaşar

Received: 2017-12-20 16:35:03

Revised: 2018-08-16 13:46:26

Accepted: 2018-10-01 14:28:07

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 43-50

How to cite

Gökhan Canbolat, Halit Yaşar; (2019), Performance Comparison for Series and Parallel Modes of a Hybrid Electric Vehicle. Sakarya University Journal of Science, 23(1), 43-50, DOI: 10.16984/saufenbilder.369149

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/369149>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Performance Comparison for Series and Parallel Modes of a Hybrid Electric Vehicle

Halit Yaşar¹, Gökhan Canbolat^{2*}

ABSTRACT

Recently, the studies in the automotive sector has concentrated on hybrid electric vehicle. The better results in fuel economy and vehicle performance values have been achieved with the hybrid electric vehicle technology. In this study, the fuel consumption and vehicle performance values are compared to the series and parallel hybrid electric modes of a conventional street sweeper by using AVL Cruise. Nowadays, there are many driving cycles used in the analysis of fuel consumption and vehicle performans. However the current driving cycles are not appropriate for the vehicles designed for a particular functions (carrying, lifting, sweeping etc.) such as street sweepers in terms of velocity and functionality. For this reason, a driving cycle is constructed by taking into consideration the daily tasks of the street sweepers with the help of AVL Cruise. The series and parallel hybrid electric configurations of the street sweeper were analyzed separately. Consequently, the demanded power (kW) from the internal combustion engine and its fuel consumption (L/h) were investiageted under the specific operating conditions in the driving cycle. The simulations show that the series hybrid mode has better fuel economy and vehicle performance.

Keywords: Hybrid Electric Vehicle, Simulation, Fuel Consumption, Vehicle Performance

1. INTRODUCTION

Since the efficiency of electric motors is higher than the internal combustion engines, many studies are being done on hybrid electric vehicles. But the current world conditions and technology have not been able to found a suitable environment to pass all-electric vehicles. Therefore, automotive companies have concentrated on hybrid electric vehicles before the all-electric vehicles. To ensure the sustainable development of the automotive industry, the research and development of electric drive systems with clean energy become a national consensus [1].

Hybrid electric vehicle technology is much more complex than conventional vehicles. Basically, the hybrid electric vehicles can be divided into series and parallel hybrid vehicles [2]. The design of the hybrid electric vehicles requires more effort and time than conventional vehicles. Simulation software packages have become indispensable in vehicle design. There are many simulation packages such as AVL Cruise, Amesim, Advisor, Adams, Msc Easy 5, Psat, etc. [3].

In the literature, there are a great number of studies Hybrid Electric Vehicle (HEV) vehicle simulations performed by using ADVISOR [2, 4-16]. In addition, AVL Cruise software developed

¹ Sakarya University, Engineering Faculty, Department of Mechanical Engineering, 54187 Esentepe-Sakarya, Turkey

² Istanbul Technical University, Institute of Science and Technology, Department of Mechanical Engineering, 34337, Gumussuyu-Istanbul, Turkey

* Corresponding Author

for this area has taken place in many automotive projects. AVL Cruise is a simulation program used in the design of vehicle and vehicle drive systems. Cruise enables modeling for vehicle design and optimization in specific applications such as fuel consumption, emission, acceleration and driving cycle tests. Previous works have shown the validity of AVL Cruise software and some works are presented here. Briggs et al. [3] studied that simulating of non automotive vehicles is highly hard according to passenger cars. A diesel-electric hybrid bus and an Internal Combustion Engine (ICE) triggered forklift is studied in the paper and concluded that the standard driving cycles are appropriate for this vehicle. Teixeira et al. [17] compared the the vehicles powered by electric motors to engine powered vehicles in terms of carbon dioxide emission and fuel economy. They presented that electric vehicles consumes the energy near four times lower than conventional vehicles at the standart driving cycle conditions. W₁ et al. [18] focused on fuel consumption reduction by changing the power systems of steering. The results demonstrate that a 1% vehicle fuel economy improvement can be achieved in a vehicle with a steering system which has electro hydraulic power compared to vehicle which has hydraulic power steering system. In addition, a 1.7% vehicle fuel economy improvement can be achieved using a full electric power steering system in a FTP-75 driving cycle. Oh et al. [19] conducted modeling of vehicles powered by gasoline and diesel ICE for testing 10 passenger cars in 7 types of driving mode. The simulations are performed by variance of driving modes and the types of vehicles, and the accuracy of the simulations results are verified by the chassis dynamometer. It is informed that the accuracy of simulations in termes of fuel economy and CO₂ emission are found compatible with test data results.

There is no HEV simulation study on street sweepers. In this article, an appropriate driving cycle was developed for Hybrid Electric Street Sweepers (HESS) by using Random Cycle Generator (RCG) of AVL Cruise. The series and parallel hybrid configurations of the HESS were simulated in this driving cycle. AVL Cruise simulations were performed to compare the

series and parallel hybrid drivetrain configurations based on fuel consumption and vehicle performance.

2. THE DRIVETRAIN SYSTEMS OF HYBRID ELECTRIC VEHICLES

HEV are defined as vehicles with more than one energy source. Hybrid electric vehicles mainly consist of an ICE which generates the power, an energy storage system and an EM.

HEVs are mainly divided into two main configurations, serial and parallel [4]. In series hybrid electric vehicles, the ICE is used in the generating electric energy and EM is used for driving the vehicle (Figure 1.). The lack of a coupling system between the EM and the ICE in series hybrid electric vehicles facilitates the design of these vehicles. ICEs are used for the generating the electrical energy that will drive the vehicle. In addition, the most important advantage of series hybrid electric vehicles is that the internal combustion engine is operated at a constant speed, resulting in low fuel consumption and high efficiency.

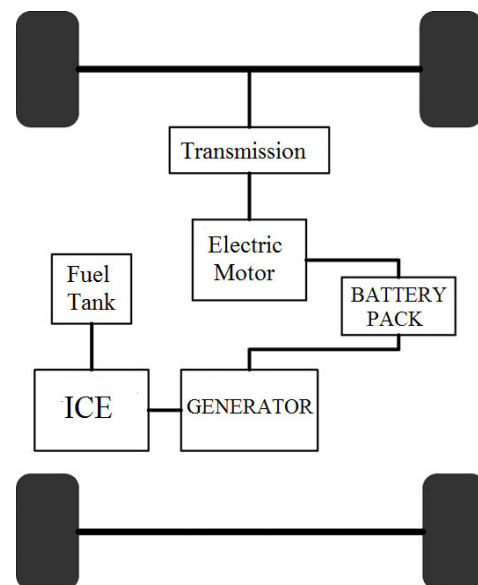


Figure 1. Series hybrid configuration

Another configuration is parallel hybrid electric vehicles (Figure 2.) In parallel hybrid configuration, the vehicle can be driven by an

ICE or EM or together depending on changing driving conditions. Parallel hybrid configuration is more complex than series hybrid configuration. In a parallel hybrid configuration, a power split unit is required to separate the ICE and EM. Depending on changing driving conditions, EM or ICE can be driven at different times.

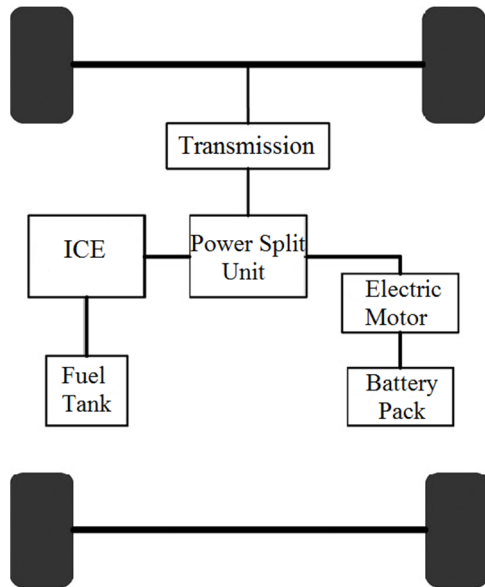


Figure 2. Parallel hybrid configuration

3. MODELING AND SIMULATION

In recent years, simulations have become a common tool used in vehicle design. In this study, the powertrain of a hybrid electric street sweeper is modeled. A driving cycle suitable for the daily operating conditions of street sweepers was created by using AVL Cruise software.

3.1. Creating an Appropriate Driving Cycle

Driving cycles are needed when examining vehicle performance, fuel consumption and exhaust emissions. Driving cycles are used to compare goals. Because they are useful to estimate the exhaust emissions, fuel economy and performance [20]. In the past a few decades, countries developed a variety of driving cycles for urban (Federal Test Procedure 72 (FTP-72), the New European Driving Cycle

(NEDC), Japan 10 15 Mode and highway driving cycles like US 06, Highway Fuel Economy Test (HWFET)) [21].

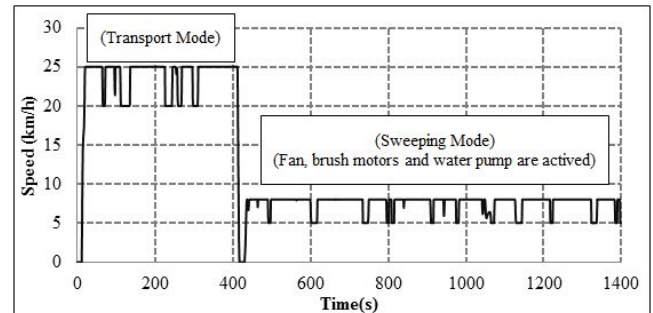


Figure 3. The vehicle velocity according to driving cycle

Figure 3. shows the driving cycle created for the hybrid electric street sweeper. This driving cycle was created by taking into account the daily tasks and operating conditions of the vehicle. The driving cycle consists of 1400 s and two driving modes (Transport and Sweeping). In the cycle, the range of 0-400 s represents the transport mode and during this mode, the speed of the vehicle is set at maximum 25 km / h. Also during the transport mode the vehicle only moves from one location to another without sweeping. In the driving cycle, the range of 400-1400 s represents the sweeping mode, during sweeping mode fan, sweeping and water pump motors are operated. During the sweeping mode, the average speed of the vehicle is set at 8 km / h, which is the most efficient sweeping velocity obtained from the field conditions.

In order to reflect the actual values, brakes and accelerations were added in driving cycle. The fluctuations seen in the cycle are the result of the added brakes and accelerations.

3.2. Hybrid Electric Street Sweeper

Hybrid electric vehicles have better fuel economy, lower emissions and energy efficiency than conventional vehicles. In recent years, many hybridization studies on heavy vehicles and electric propulsion systems have been developed for these vehicles [22].

Table 1. The basic characteristics for the hybrid electric street sweeper

Vehicle Mass (kg)	2500
Frontal Area (m^2)	$A_f=2,2$
Drag Coefficient	$C_D=0,65$
Rolling Resistant Coefficient	$\mu_r=0,01$
Vehicle-Width (mm)	1400
Vehicle-Length (mm)	3700
Vehicle-Height (mm)	1650
Wheel-Radius (mm)	320
Air Density (kg/m^3)	1,25
Maximum Speed (km/h)	25

The general characteristics of the vehicle modeled in this study are given in Table 1. In addition, the 3-D model constructed with Solidworks software is presented in Figure 4.

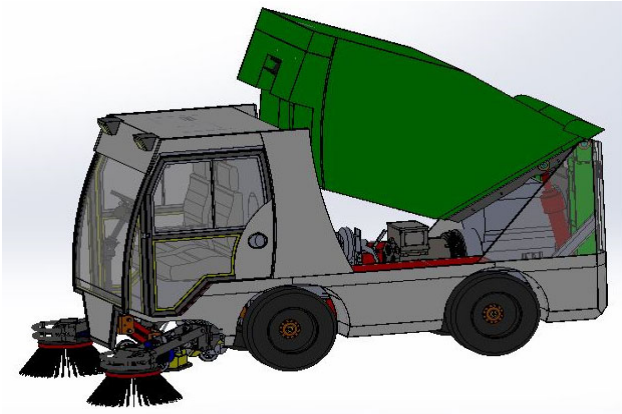


Figure 4. Hybrid electric street sweeper

In this study, a four-stroke and turbocharged diesel engine was selected for the hybrid electric street sweeper. As a result, it is predicted that the most suitable engine in terms of performance and geometry is a Mitsubishi diesel engine. The technical characteristics of the ICE were given in Table 2. The selected internal combustion engine is modeled with AVL Cruise and interpolations were made for different speeds (rpm).

Table 2. The technical specifications of internal combustion engine

Internal Combustion Engine	Mitsubishi
Type	4-stroke, cylinder in line, direct injection, turbocharged
Number of cylinders	4
Displacement volume (cm^3)	3331
Compression ratio	17:1
Bore (mm) - Stroke (mm)	94 - 120
Engine Power (kW/rpm)	36,8 / 1500
Engine Torque (Nm/rpm)	196 / 1500

Once the internal combustion engine was selected, the required step is the choice of the appropriate generator. The demanded torque values in the vehicle must be calculated correctly and the generator selection is important. In this study, a Gensan alternator was chosen by considering the vehicle's sweeping functions as well as the torque required to move the vehicle. The technical characteristics of the selected generator are given in Table 3.

Table 3. The technical specifications of generator

Generator	GENSAN
Continuous Power (kVA)	30
Standby Power kVA)	33
Speed (rpm)	1500
Frequency (hz)	50
Voltage (V)	231/400

During the sweeping mode shown in the driving cycle, generator selection is made by taking into account that the electric motor, fan motor, brush motors and water pump are activated.

Table 4. The technical specifications of the motor which drives the vehicle

AC Motor	3-Phase, 2-Pole, Asynchronous Motor
Power (kW)	15
Speed (rpm)	3000
Current (A)	28,6
Torque (Nm)	48,8
Cos (φ)	0,89
Efficiency (η)	%89
Mass (kg)	81

When selecting the electric motor; air, rolling, acceleration resistance forces acting on the vehicle are taken into account.

3.3. The Modeling of The Series and Parallel Hybrid Electric Configurations

Vehicle control algorithm for the hybrid electric street sweeper vehicle according to the driving cycle and the daily operating conditions is created. AVL Cruise models are constructed for the series and parallel hybrid modes of vehicle. Figure 5. and Figure 6. represent the series and parallel hybrid electric vehicle models. The model was created by using AVL Cruise software. Each box in the models represents one module and each module can be arranged according to the requirements of the user. The blue and red lines in the models show the mechanical and electrical connections, respectively. In addition to mechanical and electrical connections, each module is connected to each other via a communications network. Modules communicate with information by networks with each us. This control algorithm is the result of the operating conditions of the vehicle, which is also determined by the driving cycle.

It is seen in the Figure 5. that the HEV is driven only by the EM. The internal combustion engine-generator module is only used to generate the electrical energy required for the electric motors. In addition, the sweeping components are controlled according to the driving cycle.

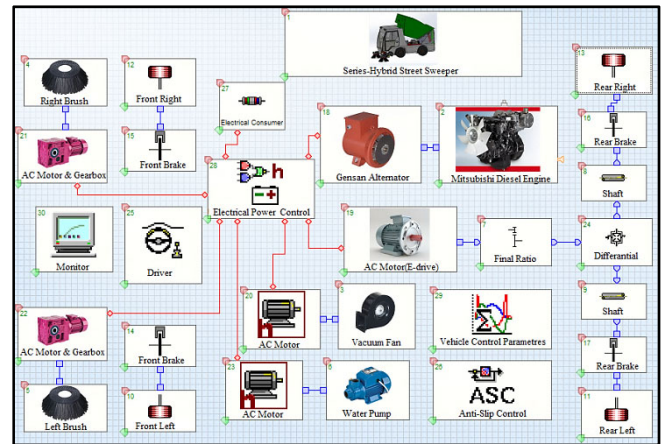


Figure 5. Series hybrid electric model by using AVL Cruise

Figure 6. shows the AVL Cruise model for the parallel hybrid electric road sweep vehicle. As it seen in the model, a parallel hybrid electric vehicle can be driven by an internal combustion engine or an electric motor with the aid of a power split unit. The vehicle can also be driven simultaneously with the ICE and the EM to meet the demanded total power according to the driving conditions.

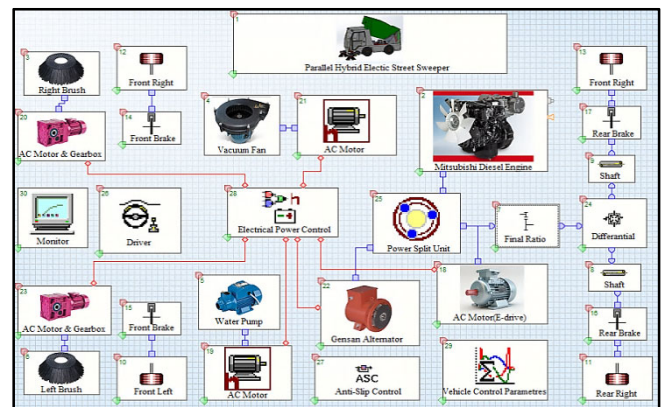


Figure 6. Parallel hybrid electric model by using AVL Cruise

In order to compare the series and parallel hybrid mode in respect of the fuel consumption and vehicle performance, the vehicle was driven only by the electric motor during sweep mode, but in transport mode the electric motor is used for series hybrid mode and the internal combustion engine is used for the parallel hybrid mode.

4. SIMULATION RESULTS

The simulations were conducted to compare series and parallel hybrid electric street sweeper in terms of fuel consumption and vehicle performance. According to the driving cycle, the sweeping components are operated through the sweeping mode and driven by the electric motors. Detailed technical informations about the used electric motors are supplied from the manufacturers and integrated into AVL Cruise software. In addition, during sweeping mode, the vehicle is powered by an EM in series and parallel hybrid modes, since the sweeping components are active and the velocity of the vehicle is very low. Therefore, the fuel consumption and vehicle performance values for the series and parallel hybrid modes will be same during the sweeping mode, so the comparisons will be detailed for the transport mode.

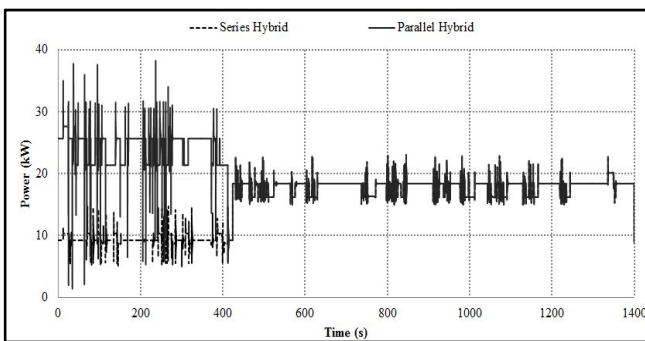


Figure 7. The demanded power according to driving cycle for series and parallel cases

Figure 7. shows the required power (kW) values for series and parallel hybrid cases. The range of 0-400 s represents the transport mode, and during this process, the average power values required for serial and parallel hybrid situations are 10 and 25 kW respectively. These results with AVL Cruise simulations show that for the series hybrid situation, 60 % less power is required. The fluctuations shown in Figure 7. are the result of accelerations and brakes added to the driving cycle and are intended to reflect the real values. The power values indicate that less power is demanded for the series hybrid mode because the vehicle is driven only by EM during the series hybrid mode. AVL Cruise simulations clearly

demonstrate that EMs have higher efficiency and torque at lower speeds than ICE.

During the sweeping mode (400-1400 s), as shown in Figure 7. sweeping components (brush, water pump and vacuum fan) are activated and the vehicle is driven by electric motor only. During this mode the internal combustion engine is only used for generating electricity with the aid of the alternator and operated at constant speed (1500 rpm). The total demanded power during the sweeping mode is 18 kW, which can rise to 23 kW according to the driving cycle conditions.

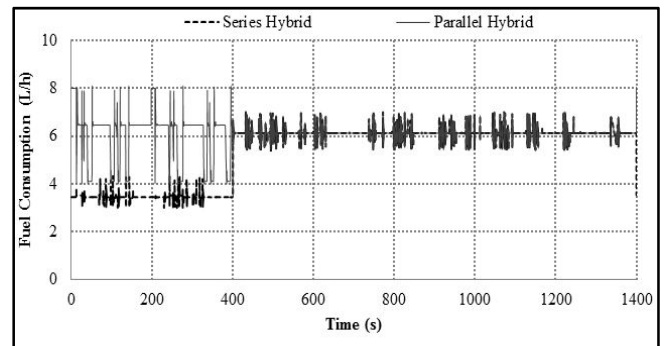


Figure 8. Fuel consumption variation for series and parallel hybrid cases according to driving cycle

Figure 8. presents the fuel consumption values of the HESS according to the driving cycle. During transport mode, the sweeping components are not activated. During this mode, average fuel consumption values for serial and parallel hybrid situations are seen as 3.8 L/h and 6.2 L/h respectively. When for the sweeping mode in the driving cycle, the sweeping components start to be driven by EMs, in addition to which the vehicle is driven only by an EM. During sweeping mode, the average fuel consumption value is 6 L / h.

CONCLUSION

In this paper, simulation on series and parallel hybrid configurations of a hybrid electric street sweeper were carried out by using AVL Cruise. The powertrain of the vehicle is modeled and the control algorithm of the vehicle is presented according to driving cycle (transport and sweeping mode). The control block diagrams of

the vehicle in terms of mechanical, electrical and informational connections were constructed to reflect the real driving conditions.

Due to the fact that the current driving cycles are not appropriate for such vehicles, necessitates to create of a appropriate driving cycle for the daily working and speed conditions of the street sweepers increased. For this reason an appropriate driving cycles for street sweeper is created by considering the real operation conditions of these vehicles.

The simulation results clearly show that fuel consumption and emissions in series hybrid configuration were lower compared to the parallel hybrid case. The reason is that the ICE runs at a constant speed in series hybrid mode during the transport mode. On the other hand, the engine speed is variable in parallel hybrid mode. Therefore, the high combustion efficiency provides the lower fuel consumption.

ACKNOWLEDGMENTS

This study was performed by using AVL Cruise. We thank to AVL List GmbH for the AVL Cruise software provided us to a free annual license within in the University Partnership Program.

REFERENCES

- [1] Li X, Lu X, Han G, The Design and Simulation Research of Mazda6 Hybrid Electric Vehicle, *Electrical Review*, ISSN 0033-2097, R. 88 NR 3b/2012.
- [2] Li X, Williamson S. S, Comparative Investigation of Series and Parallel Hybrid Electric Vehicle (HEV) Efficiencies Based on Comprehensive Parametric Analysis, *IEEE*, 0-7803-9761-4/07/\$20.00, 2007.
- [3] Briggs I, Murtagh M, Kee R, Mccloug G, Douglas R, Sustainable non-automotive vehicles: The simulation challenges, *Renewable and Sustainable Energy Reviews* (2016)
- [4] Same A, Stipe A, Grossman D, Park J.W, A study on optimization of hybrid drive train using Advanced Vehicle Simulator (ADVISOR), *Journal of Power Sources* 195 (2010) 6954–6963.
- [5] Khanipour A, Ebrahimi K. M, Seale W. J, Conventional Design and Simulation of an Urban Hybrid Bus, *International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering* Vol:1, No:4, 2007.
- [6] Fu X, Wang H, Cui N, Zhang C, Energy Management Strategy Based on the Driving Cycle Model for Plugin Hybrid Electric Vehicles, *Hindawi Publishing Corporation, Abstract and Applied Analysis*, Volume 2014, Article ID 341096, 6 pages.
- [7] Zhou Y, Ou S, Lian J, Li L, Optimization of Hybrid Electric Bus Driving System's Control Strategy, *Procedia Engineering* 15 (2011) 240 – 245.
- [8] Montazeri-GH M, Mahmoodi-k M, Development a new power management strategy for power split hybrid electric vehicles, *Transportation Research Part D* 37 (2015) 79–96.
- [9] T. Markel, A. Brooker, T. Hendricks, V. Johnson, K. Kelly, B. Kramer, M. O'Keefe, S. Sprik, K. Wipke, ADVISOR: a systems analysis tool for advanced vehicle modeling, *Journal of Power Sources* 110 (2002) 255–266.
- [10] Li Q, Chen W, Li Y, Liu S, Huang J, Energy management strategy for fuel cell/battery/ultracapacitor hybrid vehicle based on fuzzy logic, *Electrical Power and Energy Systems* 43 (2012) 514–525.
- [11] Li Y, Yi P, Wang M, Investigation to Simulation of Control Strategy for Series-Parallel Hybrid Electric Vehicle, 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics.

- [12] Yu K, Yang H, Kawabe T, Tan X, Model predictive control of a power-split hybrid electric vehicle system with slope preview, *Artif Life Robotics* (2015) 20:305–314.
- [13] Markel T, Wipke K, Modeling Grid-Connected Hybrid Electric Vehicles Using ADVISOR, NREL/CP-540-30601.
- [14] Pangaribuan K. A, Purwadi A, Performance Analysis on EV Mode of the 2012 Toyota Hybrid, *Procedia Technology* 11 (2013) 1065 – 1073.
- [15] Suh B, Frank A, Chung Y.J, Lee E. Y, Chang Y.H, Han S.B, Powertrain System Optimization For A Heavy-Duty Hybrid Electric Bus, *International Journal of Automotive Technology*, Vol. 12, No. 1, pp. 131–139 (2011).
- [16] Johnson V.H, Battery performance models in ADVISOR, *Journal of Power Sources* 110 (2002) 321–329.
- [17] Teixeira ACR, Sodre JR, Simulation of the impacts on carbon dioxide emissions from replacement of a conventional Brazilian taxi fleet by electric vehicles, *Energy* (2016).
- [18] H.S Wi, Y.K. Lee, J.I. Park, J.H Lee, K.S. Park, Effects Of FTP-75 Mode Vehicle Fuel Economy Improvement Due To Types Of Power Steering System, *International Journal of Automotive Technology*, Vol. 10, No. 6, pp. 771–776 (2009).
- [19] Oh Y, Park J, Lee J, Eom M. D, Park S, Modeling effects of vehicle specifications on fuel economy based on engine fuel consumption map and vehicle dynamics, *Transportation Research Part D: Transport and Environment*, Volume 32, October 2014, Pages 287–302.
- [20] Brady J, O’Mahony M, Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas, *Applied Energy* 177 (2016) 165–178
- [21] N.T. Jeong, S.M. Yang, K. S. Kim, M. S. Wang, H.S Kim, M.W. Suh, Urban Driving Cycle For Performance Evaluation of Electric Vehicles, *International Journal of Automotive Technology*, Vol. 17, No. 1, pp. 145–151 (2016)
- [22] Lajunen A, Energy consumption and costbenefit analysis of hybrid and electric city buses, *Transportation Research Part C* 38 (2014) 1–15



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Autonomous flight performance improvement of the morphing aerial robot by aerodynamic shape redesign

Authors: Harun Çelik, Tuğrul Oktay, Metin Uzun

Received: 2017-12-05 17:51:14

Revised: 2018-09-13 11:09:12

Accepted: 2018-10-06 02:10:31

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 51-65

How to cite

Harun Çelik, Tuğrul Oktay, Metin Uzun; (2019), Autonomous flight performance improvement of the morphing aerial robot by aerodynamic shape redesign. Sakarya University Journal of Science, 23(1), 51-65, DOI: 10.16984/saufenbilder.362588

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/362588>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Autonomous flight performance improvement of a morphing aerial robot by aerodynamic shape redesign

Harun Çelik^{*1}, Tuğrul Oktay², Metin Uzun³

Abstract

In this article, autonomous flight performance of an unmanned aerial robot is advanced by benefiting aerodynamic nose and tail cone shapes redesign both experimentally and computationally. For this intention, aerodynamic performance criteria (i.e. maximum fineness) of a scaled model of autonomous aerial robot called as Zanka-II manufactured at Erciyes University Faculty of Aeronautics and Astronautics Model Aircraft Laboratory is first observed in subsonic Wind Tunnel. Results obtained in such wind tunnel are subsequently validated using computational fluid dynamic (CFD) software (i.e. Ansys). Therefore, nose and tail cone of fuselage are improved in order to improve maximum fineness of the autonomous aerial robot. Finally, a novel scaled model using optimum data is redesigned and placed in Wind Tunnel to validate Ansys results with experimental results. By using geometrical data of ultimate aerodynamically optimized aerial robot, better autonomous flight performance is achieved in both simulation environment (i.e. Matlab and Simulink) and real time flights.

Keywords: aerodynamic shape, aerial robots, autonomous performance, nose cone, tail cone

Nomenclature

α Angle of attack
 C_D Drag Coefficient
 C_L Lift Coefficient

CFD Computational Fluid Dynamic
 E_{max} Maximum fineness value
 J Cost function
 K_D Derivative gain
 K_I Internal gain

* Corresponding Author

¹ Department of Aircraft Electrics and Electronics, Erciyes University, 38038, Melikgazi, Kayseri/Turkey. E-mail: drharuncelik@gmail.com

² Department of Aeronautical Engineering, Erciyes University, Tukey.

³ Department of Airframe, Erciyes University, Tukey.

K_p	<i>Proportional gain</i>
<i>PID</i>	<i>Proportional–Integral–Derivative</i>
T_{rt}	<i>Rising time</i>
T_{st}	<i>Settling time</i>
$xwmp, xhmp$	<i>aerial morphing parameters</i>
<i>WT</i>	<i>Wind tunnel</i>
θ	<i>Pitching angle of Aerial Robot</i>
$\%OS$	<i>Percent overshoot</i>

1. INTRODUCTION

For around last five quarters, aerial robots have been largely performed for military tasks as well as in commercial operations because of the reality of that they have several eligibilities with respect to the traditional manned vehicles. Some of these eligibilities are cost-effective manufacturing and operation, not risking pilots on hazardous conditions, and elastic configuration for customer request.

Various aerial robots have been used in aerial agriculture; for instance, crop monitoring and spraying, coast guarding such as sea lane and coastline, photography, conservation such as land monitoring and pollution during civilian missions.

On the other hand, aerial robots have also been utilized for military operations. For example, they have been operated for navy such as decoying missiles via the emission of artificial signatures and shadowing enemy fleets; army such as surveillance and reconnaissance of enemy activity; and air force tasks such as radar system destruction and jamming, and airfield base security. Considerable diversity in the aerial

robots is proposed in detail by [1]. Additionally, various scientific researches on design, manufacturing and autonomous control of aerial robot have been currently investigated [2-7].

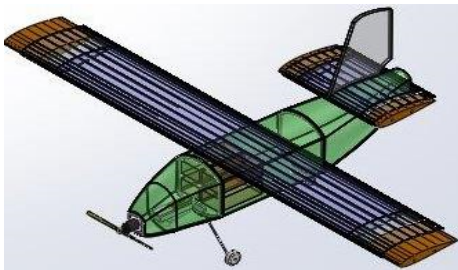
In order to obtain maximum performance for an air vehicle, several researches are conducted on aerodynamic shape improvement [8-11]. In Several tail cone shapes such as elliptical tail cone, conical tail cone and spherical tail cone are tested in order to reduce drag [9]. Also, optimal wing shapes are investigated to obtain minimum drag for an experimental unmanned aerial vehicle by using an aerodynamic shape improvement algorithm [10]. These studies show that 30% of wing drag reduction can be achieved by aerodynamic shape improvements. In another research, improvement of free-to-rotate tail fins are performed [11]. Both steady and unsteady asymmetric flows among its designs in canard fins are calculated by solving the Reynolds-Averaged Navier-Stokes solutions with Fluent [12-14]. The optimum tail fin reduced roll rate of the tail fins around 6%, and increased the normal force about 4%.

In aerodynamic shape improvement studies, this is the first journal article considering aerodynamic nose and tail cone shape improvement for performance improvement of autonomous aerial robot by simultaneous design of a load-carrying aerial robot and autopilot system. Moreover, a stochastic optimization method is first time benefited for this purpose, hence the optimum value of cost function is

found accurately and fast. In addition, although smaller values of overshoot, settling time and rise time are satisfied during tracked trajectory, aerodynamic nose and tail cone shape improvement advanced autonomous flight performance considerably [15].

2. MORPHING MECHANISM OF THE AERIAL ROBOT

Aerial robot wings yield the aerial robot to satisfy the flight conditions, but wings solely are not sufficient to satisfy the best flight performance in each condition. In order to achieve the optimal performance, mechanisms such as morphing flaps and wing tips that can change wing surface and aerial robot geometry during flight have studied by researchers and designers. After seeing accomplishment of these studies, placement of mechanical moving elements on the wing tips and remote control of wing tips are developed [16, 17]. In this study, both nose and tail cone of aerial robot are redesigned. Drawing of manufactured morphing aerial robot Zanka-II, and its side and upper view photos are illustrated in Figure 1.



(a)



(b)



(c)

Figure 1. Drawing of morphing aerial robot (morphing aspects are brown) (a), side (b) and upper view (c) photos of Zanka-II

Each one of the wing system comprises gear teeth, micro motor, bidirectional mini brushing electronic speed controller (ESC), and M3 worm gear. As seen in Figure 2 (a), micro motor is a DC motor and operating between ranges 6-9 V, creates a torque of 1.8 kg. The gear motor is placed on the shaft of the related motor (see Figure 2 (b)). This gear operates with a second gear placed on the bearing compatibly (see Figure 2 (d)). In released state of the gear mounted on bearing, the gear has M3 tooth. For turning the engine right or left by the remote control, a bistable switch is used to send signal to ESC (see Figure 2 (c)). ESC is a bidirectional and brushed circuit, and works in two sides in terms of the incoming signal directly through the

remote control switch (see Figure 2 (e)). The direction of gear rotation determines moving of motor, and hence a mechanical system that has such a gear can lead screw in-or-outward smoothly (see Figure 2(f)). This movement results to moving of wing tip parts, and therefore provides alteration of the wing area.



(a)



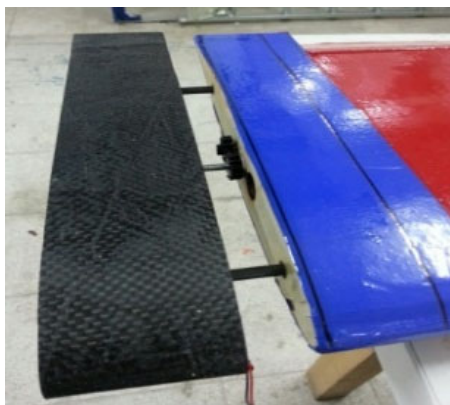
(b)



(c)



(d)



(e)



(f)

Figure 2. Morphing mechanism. DC (a) and gear motor (b) placed on shaft. Servo motor placement in the wing (c). Electronic speed controller (d). Assembly of the expanding parts of the wing (e). 3D view of the moving part of the wing (f)

3. AERODYNAMIC SHAPE IMPROVEMENT

Aerodynamic shape improvement of an air vehicle can be achieved by maximum fineness ratio that is a combination of the lift and drag forces of the air vehicle. An increase in the lift force of the air vehicle increases the controllability, hence the distance of take-off and landing decrease, and similarly the decrease in the drag force reduce fuel consumption. Maximum fineness ratio can be defined as the ratio of the lift force to the drag force in general expression. However, solely increasing the lift force or decreasing the drag force for performance improvement may not be sufficient, since the drag force is proportional to the square of the lift force as given in (1). Therefore, examining (2) can be more accurate and reliable for aerodynamic improvement.

$$C_D = C_{D_0} + KC_L^2 \quad (1)$$

$$E_{\max} = \frac{1}{2\sqrt{KC_{D_0}}} \quad (2)$$

where C_D and C_L are drag and lift coefficient. C_{D_0} depicts the drag force coefficient when the lift force is 0. $K = 1/(\pi A_R e)$ where A_R is wing aspect ratio, and e is oswald constant.

3.1. Numerical Method

Flight performance of an autonomous aerial robot improved by optimum aerodynamic nose and tail cone shape may be computationally investigated by using CFD code with commercial Ansys [20-22] software depended on finite volume technique [23]. The aerial robot has 1.06 m length and 1.60 m wingspan. For all simulations, velocity, density of air, and kinematic viscosity are defined as 16.6 m/s (about 100km/h), 1.036 kg/m³ and 1.5111x10⁻⁵ m²/s, respectively. Also, Angle of attack is selected as -4, 0, 4, 8, 12 and 16 degrees.

The turbulent viscosity was worked out through Standard k-ε Turbulence Model [24-26]. All solution variables were solved via first order upwind discretization scheme [27]. For these simulations, the convergence criterion is selected to be 10⁻⁶.

3.2. Boundary Condition

Computational domain extended to 15 wingspans (C) for upstream of the leading point and downstream of the trailing point of the aerial robot. Also 20 C was implemented from pressure outlet surface. Velocity inlet boundary condition was applied upstream and downstream with speed of 16.6 m/s. No-slip boundary condition is used at solid surfaces. Figure 3 illustrates all these properties for simulations. Velocity components are defined for each angle of attack situation. The x component of velocity is calculated by $x=16.6\cos\alpha$, and y component of velocity is defined with $y=16.6\sin\alpha$ formula, where α is the angle of attack in degree.

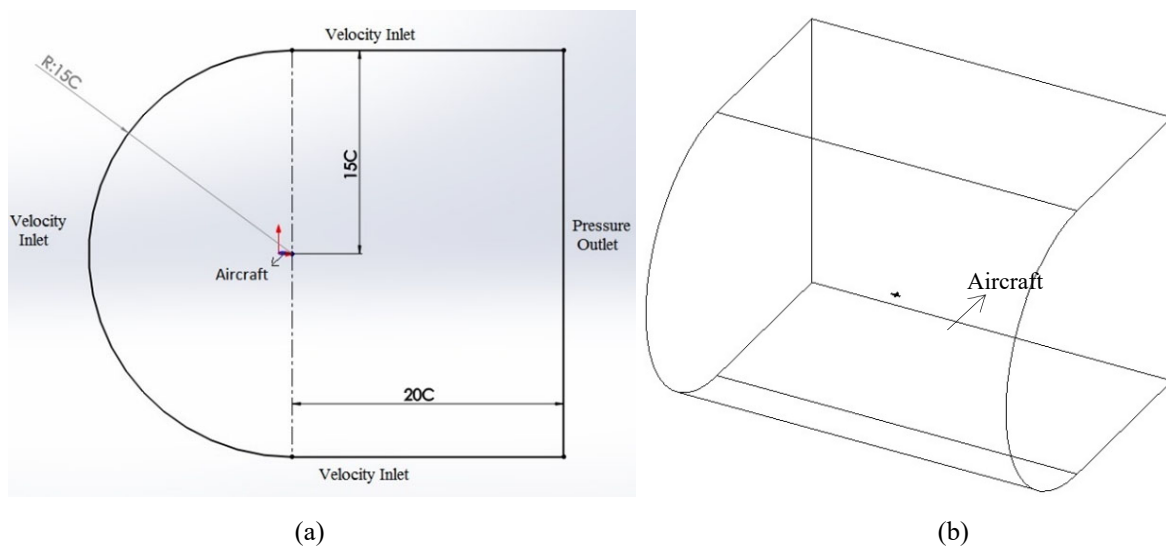


Figure 3. The dimensions (a) and boundary conditions (b) of the computational domain (nose and tail cone)

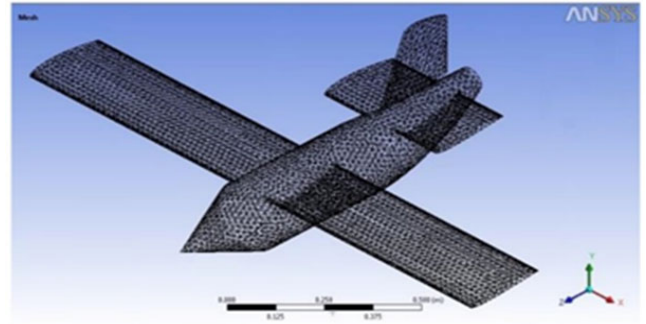
3.3. Grid Independency Study

The grid for simulations is generated by Ansys Mesh program, and nose options with mesh are shown in Figure 4. Patch conforming/sweeping method which is suitable for clean computer aided design geometries is used for meshing. For eliminating of mesh effects, optimum mesh element number should be determined. Increasing of the element number provides more accurate results, but using more element increases solving time. For this reason, grid independency study has been done with 0.1k, 0.4k, 1k, 2.4k, 6k and 10k elements where $k=10^6$. In Figure 5, lift coefficient (C_L) variation with different element numbers are given at 0 degree angle of attack. For elements number larger than 2.4k, lift coefficient values do not change considerably. So it can be decided that 2.4k elements are enough for the accurate results. Approximately all of these analyses completed between 10000 and 20000 iterations. Used computer for these analyses has 3.60 GHz CPU and 8 GB RAM. Maximum analysis time for a simulation was 2 hours. Similar to nose cone, all of these processes are reperformed for the tail cone.

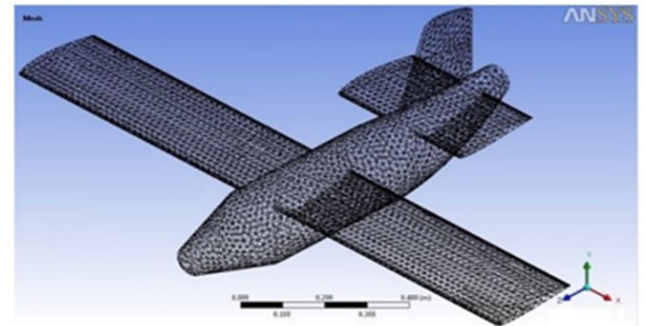
To summarize, in this section, various nose and tail cone equations have been examined to improve the autonomic performance of the aerial robot.

E_{max} values are calculated using various equations of nose and tail cone with both CFD and WT. Comparison results of CFD and WT are given in

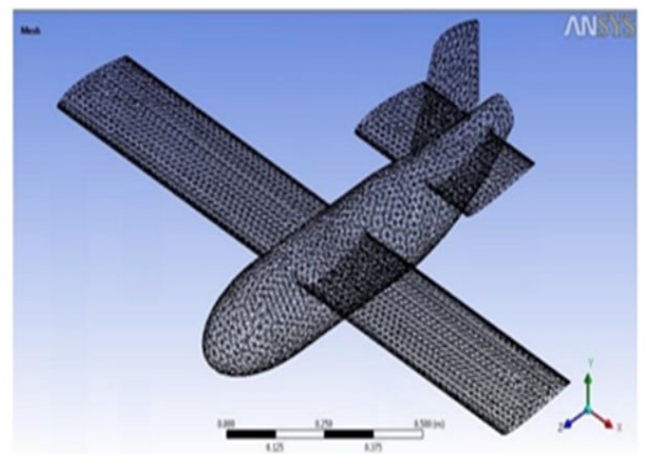
Table 1 and 2. Solid drawings of these various values performed in nose and tail cone equations are shown in Figure 4. The most appropriate equation for both nose and tail cone are obtained as conic equation with E_{max} value of 6.54 and 6.44, respectively.



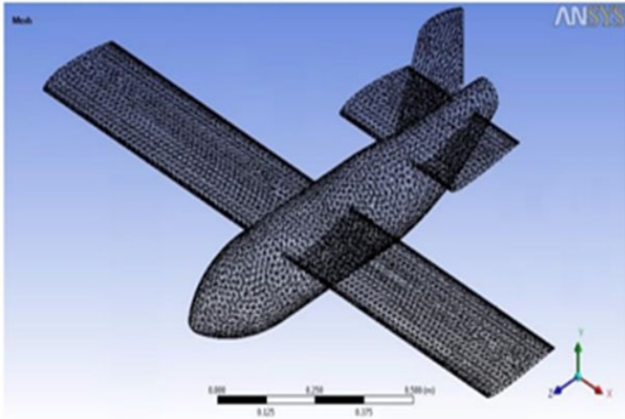
(a)



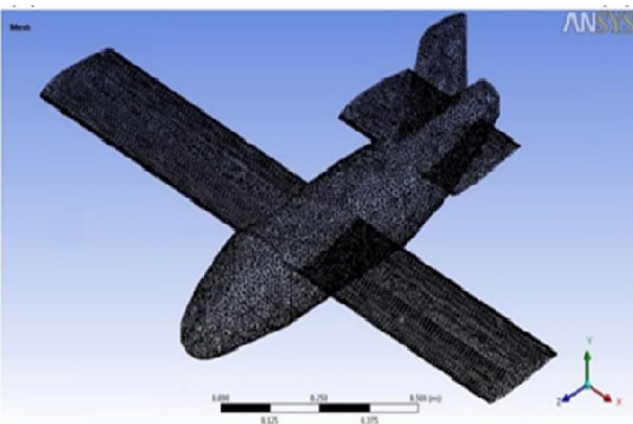
(b)



(c)



(d)



(e)

Figure 4. Mesh of nose options. Conical (a), Biconical (b), Elliptical (c), Spherically tangent ogive (d), Initial reference manufactured (e)

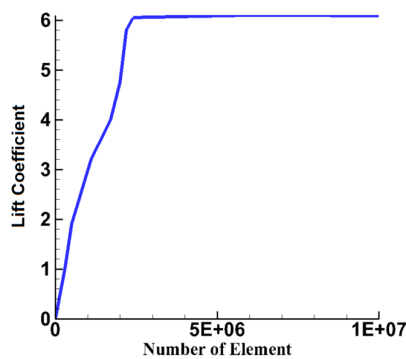
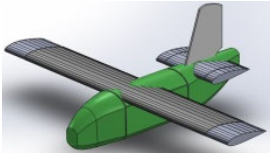
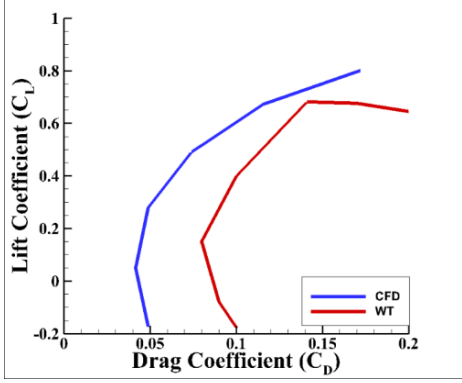
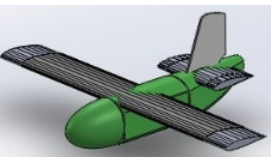
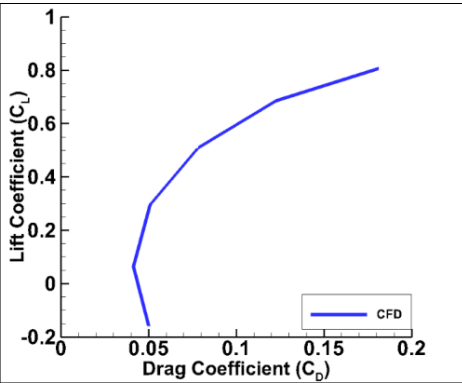
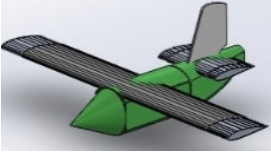
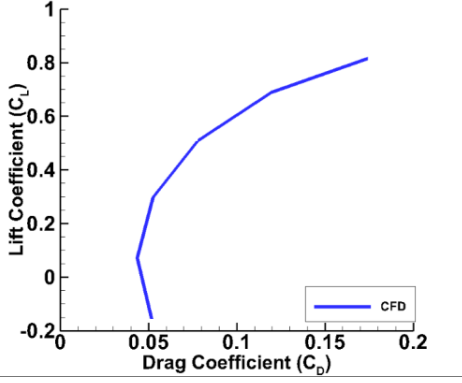
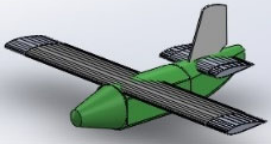
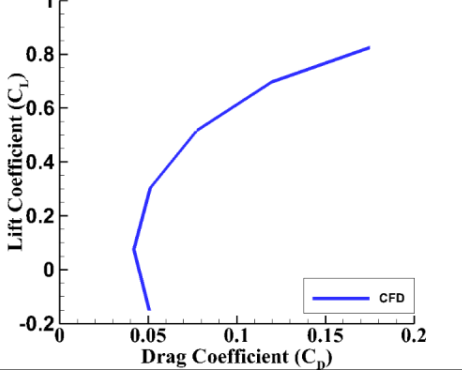


Figure 5. Grid independency result for lift coefficient (nose and tail cone)

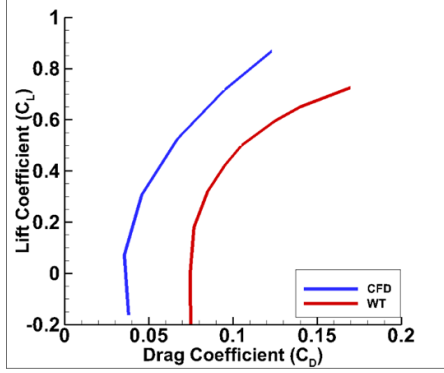
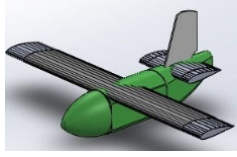
3.4. Aerodynamical Results

In this subsection, various nose and tail cone shapes such as elliptical, conical, biconical, inverse-parabolic, parabolic, spherically blunted tangent ogive, and haack series [18, 19] are investigated to maximize the value of maximum fineness (E_{max}). In Table 1 and 2, CFD drawings, CFD and Wind Tunnel (WT) results, and E_{max} values are presented for nose and tail cone shapes, respectively. From these tables, it can be seen that the maximum value of E_{max} is obtained for *spherically blunted tangent ogive* nose cone shape, and *conical* tail cone shape. The maximum values of E_{max} are around 1.4-3% larger than the one of initial reference aerial robot. This amount may seem smaller, but it has very important advantageous for performance characteristics such as maximum range, endurance and autonomous performance. It is important to state that initial cone and ultimate cone which has maximum E_{max} values (i.e. optimum geometry) are only experimented in wind tunnel. The maximum fineness value for the remainder of cones is expected computationally, so do not need to test with the aforementioned cone configurations. In Figure 6, wind tunnel experiments for initial and optimum geometry are presented.

Table 1. Improvement results for nose cone

Nose cone	CFD	C_L - C_D	E_{max}
Produced Aircraft			6.35
Elliptical			6.33
Conical			6.22
Biconical			6.36

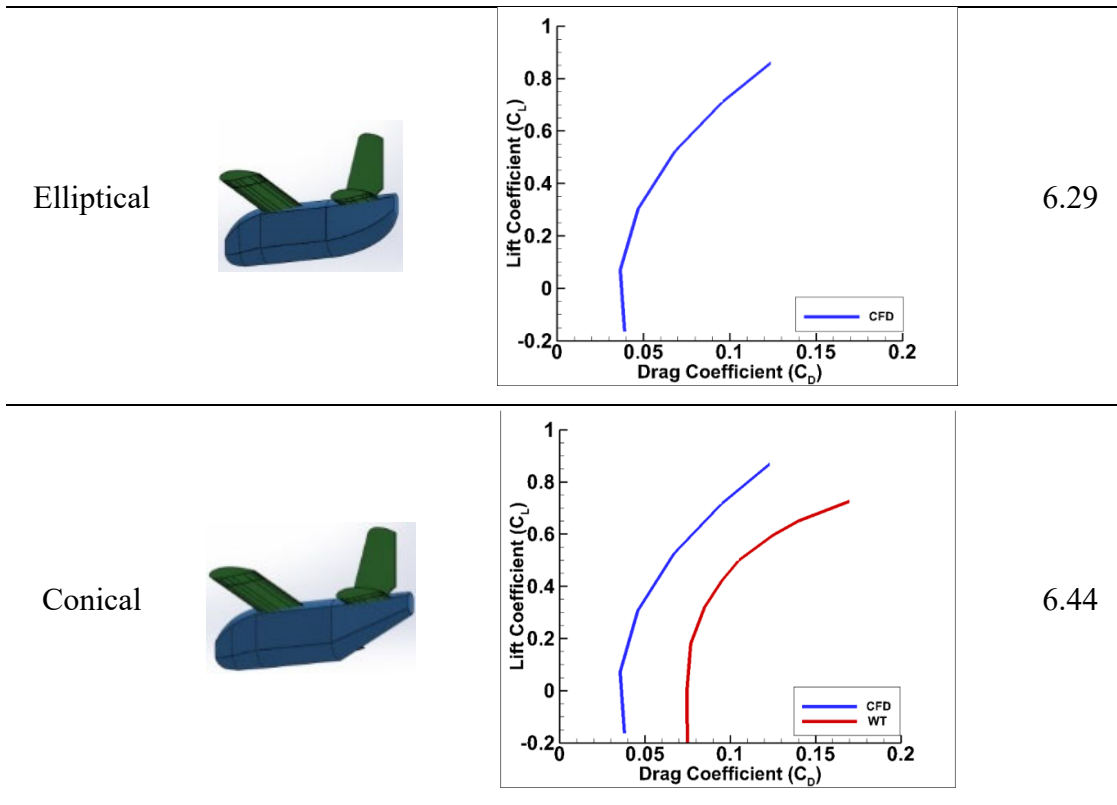
Spherically
Blunted
Tangent Ogive



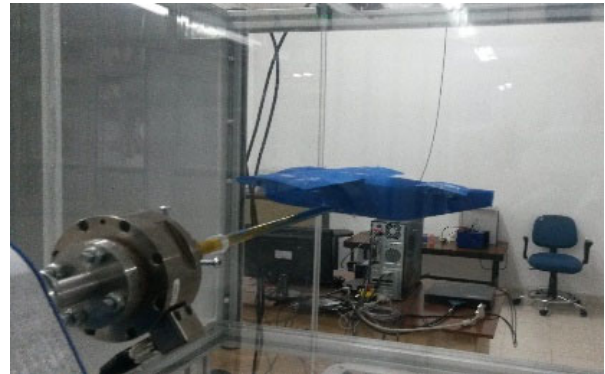
6.54

Table 2. Improvement results for tail cone

Tail cone	CFD	C_L - C_D	E_{max}
Produced Aircraft			6.35
Haack Series $C=0.3$			6.2
Inverse Parabolic			6.37



(a)



(b)

Figure 6. Wind Tunnel tests. Spherically blunted tangent ogive nose (a). Conical tail cone (b)

4. AUTOPILOT AND PERFORMANCE EVALUATION

4.1. Autopilot System

For both simulation environment (i.e. Matlab) and real-time flight study, classical PID based hierarchical autopilot system is chosen [28]. In the system structure, three layers of hierarchical PID

controller to satisfy tracking trajectory are performed as seen in

Figure 7. In this autopilot, control signals are provided by inner loop, desired pitch and yaw rates are achieved by middle loop, and the trajectory of aerial robot is tracked by outer loop. The T , h , r , ψ , θ , and ϕ indicate throttle, altitude, rudder, yaw, pitch and roll, respectively.

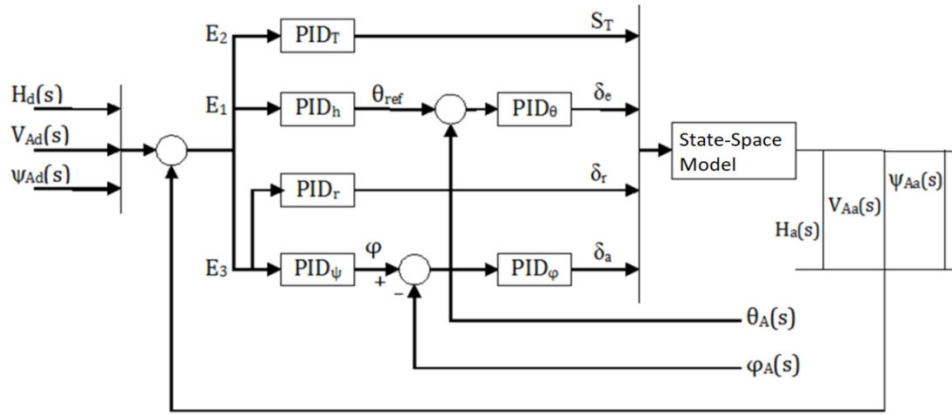


Figure 7. Hierarchical autopilot system

4.2. Autonomous Performance Evaluation

Generally, PID-based hierarchical controller system enables to control altitude, yaw angle, and velocity for tracking a definite trajectory. The proposed system includes 6 PID controllers with the 3 layers that are outer, middle, and inner loops. These PIDs have also upper and lower restrictions, and guarantee to track desired trajectory. For general situation, interested autopilot user needs to set totally 18 parameters of PID (i.e., 6 parameters for each P, I and D) to benefit from all internal PIDs. Moreover, in this study for simultaneous morphing aerial robot and autopilot design strategy, two more structural parameters (i.e. optimum extension ratios of horizontal tail and wing) are set. For high-performance trajectory tracking, a cost function consists of settling time (T_{st}), rise time (T_{rt}), and overshoot (OS) is defined as

$$J = \sum g(T_{st} - T_{st_u})^2 + g(T_{rt} - T_{rt_u})^2 + g(\%OS - \%OS_u)^2 \quad (3)$$

where subscript u denotes upper value; then optimization problem can be described as $\min\{J\}$

where J is the function of 20 terms (18 controller design parameters and 2 aerial morphing parameters) and can be given as

$$J = f \left(\begin{matrix} xwmp, xhmp, K_{P_1}, K_{I_1}, K_{D_1}, K_{P_2}, \\ K_{I_2}, K_{D_2}, \dots, K_{P_6}, K_{I_6}, K_{D_6} \end{matrix} \right) \quad (4)$$

where $xwmp$ and $xhmp$ are aerial morphing parameters, and K_P, K_I, K_D are autopilot

parameters for each PID. Terms of cost function is calculated by using simultaneous perturbation stochastic approximation (SPSA) [29, 30] as

$$\text{If } T_{st} \geq T_{st_u}, T_{st} \text{ is non-defined else} \quad (5)$$

$$T_{st} < T_{st_u}, T_{st} \text{ is its value}$$

$$\text{If } T_{rt} \geq T_{rt_u}, T_{rt} \text{ is non-defined else,} \quad (6)$$

$$T_{rt} < T_{rt_u}, T_{rt} \text{ is its value}$$

$$\text{If } \%OS \geq \%OS_u, \%OS \text{ is non-defined else} \quad (7)$$

$$\%OS < \%OS_u, \%OS \text{ is its value}$$

SPSA results of cost function J for the best nose (a) and tail (b) cone are given in Figure 8. After the simultaneous optimization process, 3% and 1.4% improvements in maximum values of E_{max} result to 2.6% and 1.1% improvement in entire autonomous performance by using nose and tail cone redesign, respectively. These improvements of E_{max} may seem small. However, the improvements are

obtained with a very small modification in aerial robot geometry, and it is very important in terms of controller performance since reducing flight energy consumption. After applying the design methodology, and when Von Karman turbulence [31] exists to model a flight condition closed to real, satisfactory tracking trajectory of the aerial robot for pitch angle is depicted in Figure 9.

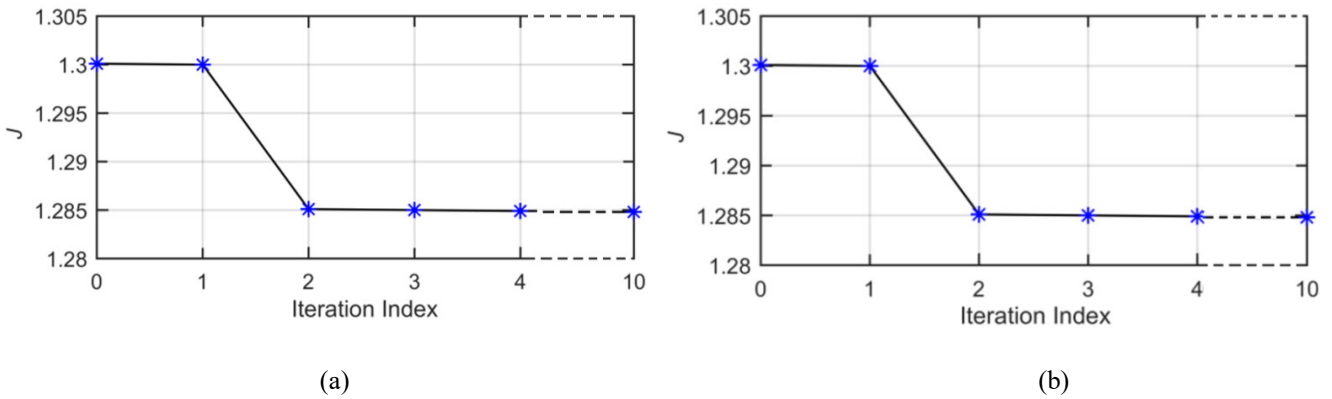


Figure 8. SPSA results of cost function J for the best nose (a) and tail (b) cone.

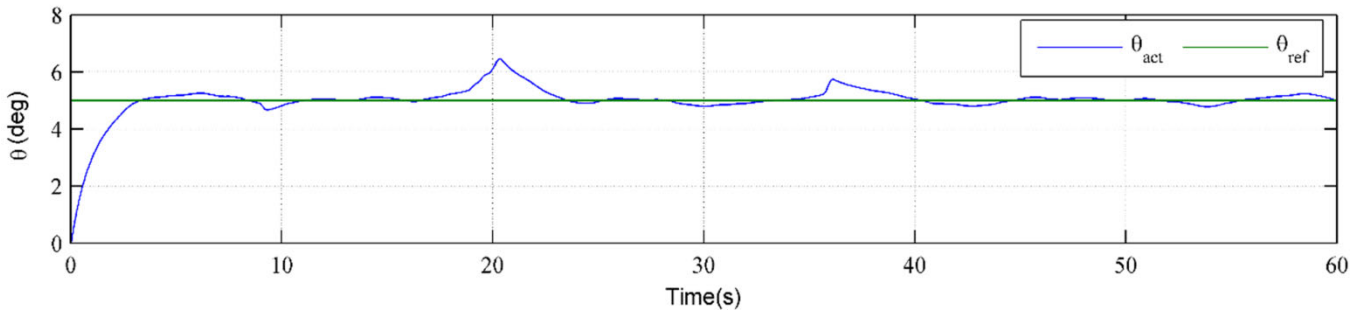


Figure 9. Trajectory tracking for resulting ultimate aerial robot

5. CONCLUSIONS

In order to improve flight performance of an autonomous aerial robot, aerodynamic nose shape optimization both experimentally and computationally were examined in this article. Aerodynamic performance criteria (i.e., maximum fineness) of a scaled model of the autonomous aerial robot manufactured at the Model Aircraft

Laboratory was first observed in subsonic Wind Tunnel. Obtained results were validated using a computational fluid dynamics software (i.e. Ansys). Nose and tail cone of fuselage were optimized in order to maximize maximum fineness of the autonomous aerial robot by using various shape equations in Ansys. After deciding the optimum shape, a new scaled model using the optimum data was then produced and placed in

Wind Tunnel in order to validate Ansys results with experimental results. By using geometrical data of ultimate aerodynamically optimized aerial robot, better autonomous flight performance was found both in simulation environment (i.e. Matlab) and real-time flight tests.

A small increase (around 3% and 1.4%, respectively) for the maximum value of fineness is obtained via aerodynamic nose and tail cone shape optimization for morphing autonomous aerial robot. Finally, a minor, but important, improvement (around 2.6% and 1.1%, respectively) in autonomous performance is achieved.

For the future, nose and tail cone can be investigated with a simultaneous design while it is studied separately in this article. Explicitly, each nose and tail cone shape is designed and then examined in computational fluid dynamic software and subsonic Wind Tunnel separately. In the future, various configurations of nose and tail cone can be examined at the same time. Thus, higher improvements can be achieved in autonomous performance.

ACKNOWLEDGMENTS

This work was supported by Research Fund of the Scientific and Technological Research Council of Turkey (TUBITAK) under Project Number: 214M282.

REFERENCES

- [1] R. Austin, *Unmanned aircraft systems*. Wiley, 2010.

- [2] Y. Ding, Y. C. Liu and F. B. Hsiao, "The application of extended Kalman filtering to autonomous formation flight of small UAV system," *Aircraft Engineering and Aerospace Technology*, vol. 1, no. 2, pp. 154-186, 2013.
- [3] A. Drak, M. Hejase, M. ElShorbagy, A. Wahyudie and H. Noura, "Autonomous Formation Flight Algorithm and Platform for Quadrotor UAVs," *International Journal of Robotics and Mechatronics*, vol. 1, no. 4, pp. 124-132, 2014.
- [4] L. De Filippis, G. Guglieri and F. B. Quagliotti, "A novel approach for trajectory tracking of UAVs," *Aircraft Engineering and Aerospace Technology: An International Journal*, vol. 86, no. 3, pp. 198 – 206, 2014.
- [5] T. Oktay, M. Uzun, H. Celik and M. Konar, "PID Based Hierarchical Autonomous System Performance Maximization of a Hybrid Unmanned Aerial Vehicle (HUAV)", *Anadolu University Journal of Science and Technology – A Applied Sciences and Engineering*, vol. 18, no. 3, pp. 554-562, 2017.
- [6] H. Celik, T. Oktay and I. Turkmen, "Model Predictive Control and Robustness Test of the Unmanned Aerial Vehicle (Zanka-I) in Various Turbulence", *Journal of Aeronautics and Space Technologies*, vol. 9, no. 1, pp. 31-42, 2016.

- [7] T. Oktay, and S. Coban, “Simultaneous Longitudinal and Lateral Flight Control Systems Design for Both Passive and Active Morphing UAVs,” *Elektronika ir Elektrotechnika*, vol. 23, no. 5, pp. 15-20, 2017.
- [8] Z. Xuetao, *UAV Design and Manufacture*, BS Thesis, 2010.
- [9] Z. Lyu, K. W. Kenway and J. R. Martins, “Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark,” *AIAA Journal*, vol. 53, no. 4, pp. 968-985, 2015.
- [10] P. Gamboa, J. Vale, F. J. P. Lau and A. Suleman, “Optimization of a morphing wing based on coupled aerodynamic and structural constraints,” *AIAA journal*, vol. 47, no. 9, 2009.
- [11] E. Feyzioglu, “Roll characteristics and shape optimization of the free to-rotate tail-fins on a canard-controlled missile,” MS Thesis, Middle East Technical University, 2014.
- [12] R. P. Liem, J. R. Martins and G. K. Kenway, “Expected drag minimization for aerodynamic design optimization based on aircraft operational data,” *Aerospace Science and Technology*, vol. 63, pp. 344-362, 2017.
- [13] J. E. Hicken and D. W. Zingg, “Induced-Drag Minimization of Nonplanar Geometries Based on the Euler Equations,” *AIAA journal*, vol. 48, no. 11, 2010.
- [14] A. Khalid and P. Kumar, “Aerodynamic Optimization of Box Wing – A Case Study,” *International Journal of Aviation, Aeronautics, and Aerospace*, vol. 1, no. 4, 2014.
- [15] T. Oktay, M. Uzun, I. Yılmaz and M. Konar, “Aerodynamic nose shape optimization for performance maximization of morphing autonomous aerial robot,” *International Conference on Engineering and Natural Science*, Sarajevo, Bosnia and Herzegovina, 24-28 May 2016.
- [16] S. Barbarino, F. Gandhi and S. Webster, “Design of Extendable Chord Sections for Morphing Helicopter Rotor Blades,” *Journal of Intelligent Material Systems and Structures*, vol. 22, no. 9, pp. 891–905, 2011.
- [17] T. Yue and L. Wang, “Longitudinal Linear Parameter Varying Modeling and Simulation of Morphing Aircraft”, *AIAA Journal of Aircraft*, vol. 50, no. 6, pp. 1673-1681, 2013.
- [18] P. Neittaanmäki, T. Rossi, S. Korotov, E. Oñate, J. Périaux and D. Knörzer, “Overview on drag reduction technologies for civil transport aircraft,” European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS), 24-28, July 2004.
- [19] M. K. Chan, “Supersonic Aircraft Optimization for Minimizing Drag and Sonic Boom,” PhD Thesis, 2003.

- [20] S. Fu and L. Wang, "Modelling the flow transition in supersonic boundary layer with a new $k-\omega-\gamma$ transition/turbulence model," *7th International Symposium on Engineering Turbulence Modelling and Measurements-ETMM7*, Limassol, Cyprus, 4–6 June, 2008.
- [21] F. Menter, "Two-equation eddy viscosity turbulence models for engineering applications," *AIAA Journal*, vol. 32, pp. 1598–1605, 1994.
- [22] R. E. Mayle and A. Schulz, "The path to predicting bypass transition," *ASME J. Turbomach*, vol. 119, pp. 405–411, 1997.
- [23] D. Choudhury, "Introduction to the renormalization group method and turbulence modeling," *Fluent Inc. Technical Memorandum*, TM-107, 1993.
- [24] D. D. Sanders, W. F. O'Brien, R. Sondergaard, M. D. Polanka and D. C. Rabe, "Predicting Separation and Transitional Flow in Turbine Blades at Low Reynolds Numbers-Part I: Development of Prediction Methodology," *J Turbomach*, vol. 133, pp. 1-10, 2010.
- [25] P. Catalano and R. Tognaccini, "Turbulence modeling for low Reynolds number flows," *AIAA Journal*, vol. 48, pp. 1673-1685, 2010.
- [26] F. R. Menter, R. B. Langtry, S.R. Likki, Y. B. Suzen, P. G. Huang and S. Völker, "A correlation based transition model using local variables: part I-model formulation," *Proceedings of ASME Turbo Expo 2004*, Vienna, Austria, pp. 57–67, 2004.
- [27] T. Misaka and S. Obayashi, "A correlation-based transition models to flows around wings," *AIAA Paper 2006–918*, 2006.
- [28] H. Chao, Y. Cao and Y. Q. Chen, *Autopilots for Small Fixed-Wing Unmanned Aerial Vehicles: A Survey. IEEE International Conference on Mechatronics and Automation*, Harbin, China, 2007.
- [29] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE transactions on automatic control*, vol. 37, no. 3, pp. 332-341, 1992.
- [30] T. Oktay, H. Çelik and M. Uzun, "A novel learning algorithm to estimate the optimum fuselage drag coefficient," *Sakarya University Journal of Science*, vol. 21, no. 1, pp. 63-68, 2017.
- [31] U.S. Military Handbook MIL-HDBK-1797, 19 December 1997.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Passive RFID Uplink and Downlink Link Budget and Comparison of ASK and BPSK Backscatter Modulations

Authors: Kazım Evecan

Received: 2018-06-18 16:22:41

Revised: 2018-09-21 14:27:17

Accepted: 2018-10-11 11:28:03

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 66-75

How to cite

Kazım Evecan; (2019), Passive RFID Uplink and Downlink Link Budget and Comparison of ASK and BPSK Backscatter Modulations. Sakarya University Journal of Science, 23(1), 66-75, DOI: 10.16984/saufenbilder.434517

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/434517>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Passive RFID the Forward and Backward Link Budgets and Comparison of ASK and BPSK Backscatter Modulations

Kazım Evcan*¹

ABSTRACT

Reflected signal levels in passive RFID require separate TX and RX antennas to overcome the reader threshold. Using latest improvements, it is shown in this work that both the forward and backward links restrict communication distance given with equations depending upon tag power dissipation, antennas in use, type of modulation and operating environment and parameters to make both distances equal are given. Complete link budget for the forward and backward links in a RFID system are given with constituent parts in detail. After that, to further elaborate on the link budgets, reflection coefficient and power equations are obtained from a RFID front-end model with circuit theory and shown on a passive RFID system. This study shows modulation index m and antenna gain G_t for BPSK and ASK modulations to make two distances equal for the environment to be within, found from two way link equations. Also, ASK and BPSK modulations are compared on communication distance for given average backscatter difference power, minimum SNR to targeted BER on the reader by the equations obtained on the circuit model and MATLAB simulations.

Keywords: RFID, RFID link budget, communication distance, ASK and BPSK backscatter Modulation, reflection coefficient, transmitted power, reflected power.

1. INTRODUCTION

RFID technology has been employed in many promising applications with low cost in various fields, such as access control, biomedical implants, identification, tracking, logistics, sensor networks, security, fast payment system, loss prevention and shopping malls. This technology has a growing potential in massive deployment and retail stores with distant RFID tags working in UHF and microwave bands.

RFID technology is different from conventional transceiver communication systems. In passive RFID systems, reader sends electromagnetic waves to energize tag and to inform key information. After tag obtained enough power to work, tag to reader information is supplied by reflecting incoming electromagnetic waves between two impedance states backscatter modulation. During this operation, tag

switches input impedance between two different states namely matched and short circuit conditions.

Since low amount of power extracted from the field, all the circuit blocks must have low power feature. At UHF and microwave frequencies, tags work from long distance in obstructed and noisy environment and that is why this feature is very essential. On the other hand, the amount of reflected power from this weak incoming field between two impedance states must reach reader and have to be above reader sensitivity or threshold to establish communication channel. As it is understood, in RFID system we have one of two limitations namely reader to tag uplink distance and tag to reader downlink distance.

Communication distance for the forward and backward links are shown with equations and distances for incoming and reflected fields are given in latest studies. In [1], [2] and [3], backscattered power using

¹ Kütahya Dumlupınar Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü, Kütahya, kazim.evcan@dpu.edu.tr

RCS (radar cross section) is studied with classical radar equation and range equations and tag threshold are given by using RCS. Passive backscatter power for UHF band with RCS equations is shown on RFID front end model for different load conditions and measured in anechoic chamber with network analyzer for short circuit, open circuit and chip loaded conditions in [1]. From these measurements, working distance is found. Communication distance and complete link budget using RCS are shown with typical sample values in [2] and [3]. The forward and backward link budgets are given with example parameters in detail for monostatic and bistatic configurations and different environments in [4] and in reference to the latest RFIDs in the literature and their testing systems at the end in [5]. Passive RFID uplink and downlink design considerations in block level and backscatter modulation schemes with their performance comparisons are presented in [6]. Detailed RF input power and backscattered power equations for BPSK, ASK and OOK are given in [7]. In these studies, nothing is told about making two distances equal in presence of specific operation environment.

In this study, tag power dissipation, antenna gain, reader threshold considering operation environment and modulation parameters included uplink and downlink distances are shown for UHF RFID in detail and parameters to make two distance equal are given. Two way link budgets for UHF RFID with reference to power consuming hardware parts and loss mechanisms is presented in detail in second section. It is shown that reflected signal levels forces to choose bistatic configuration for distant operation. In third section, power equations and reflection coefficient are shown on a RFID front-end model. Backscattered difference power between two states found in section four is used to find minimum SNR and BER for targeted reliability and this SNR value can be used to find reader sensitivity in section five. In last section, with the equations rather than RCS, modulation parameters included uplink and downlink range are found for both ASK and BPSK modulations. It is shown with examples that range is limited with either uplink or downlink depending upon tag power dissipation, antennas in use, reader sensitivity level and modulation type. In addition, modulation index m and antenna gain G_t for BPSK and ASK modulations are obtained to make two distances equal.

2. LINK BUDGET AND CONSTITUENT PARTS

Each parameter must be utilized effectively or optimized in communication link to maximize communication range. For uplink communication, incident power on the tag antenna must be enough to perform two functions, data extraction and tag energizing, for a communication channel to exist. Incident power is given in (1) for a tag located d from reader, P_t transmitted output power from reader, G_r reader transmit antenna gain with polarization match, G_t tag antenna gain with polarization match and L path loss [4]. For lossy environments and free space, path loss d away from reader is given with (2) and (3), n path loss exponent ($n=4-6$ in obstructed buildings, $n=2-3$ in obstructed factories, etc.), d_0 reference path loss measurement distance and λ wavelength [8].

$$P_{inc} = \frac{P_t G_t G_r}{L} \quad (1)$$

$$L(dB) = L(d_0)(dB) + 10n \log_{10}\left(\frac{d}{d_0}\right) \quad (2)$$

$$L = \left(\frac{4\pi d}{\lambda}\right)^2 = \left(\frac{4\pi d}{c/f}\right)^2 \quad (3)$$

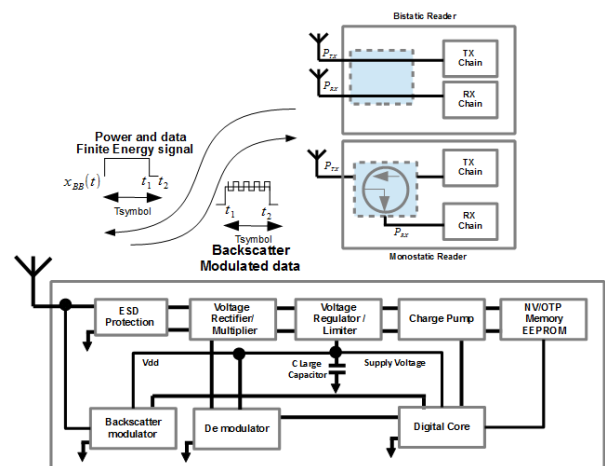


Figure 1. Link budget for a passive RFID system with the same path loss for uplink and downlink [5]

Reader EIRP (equivalent isotropically radiated power) or ERP (effective radiated power) given with $P_t G_t$ product are limited by authorities, see table 1. When we look at equation (1), only parameters we can adjust in the wireless link are G_t tag antenna gain and L path loss. G_t tag antenna gain is traded with antenna size, getting bigger with decreasing frequency restricting RFID label size and path loss

increases with frequency. Therefore, for a given size of receive antenna, the environment that RFID tag operated affects your choice of wireless link frequency for maximum power transfer. For operating frequency band, 868 MHz or 915 MHz depending upon countries are employed mostly.

Table 1. Half wavelength dipole antenna sizes and allowed maximum output power for RFID applications for different frequency bands [9]

Hw. Dipole	868 MHz EU	902-928 MHz N. A. and J.	2.45 GHz for EU
EIRP= 1.64ERP	500 mW ERP, 0.5-2W with power control	4W EIRP	500 mW EIRP, 0.5-4 W indoor with power control
Antenna size	17.3 cm	12.4 cm	6.1 cm
Antenna gain	1.64	1.64	1.64

Power obtained on RFID antenna must be utilized effectively in circuits because very low amount of power reaches RFID tag in obstructed and noisy environment. Induced voltage on an antenna is rectified or multiplied to supply the power to demodulator, backscatter modulator and digital part with peripherals on passive RFID, see figure 2. Voltage multiplier or rectifier power conversation efficiency (PCE) η also put constraints on the range because substrate couplings, parasitics, threshold voltage drop and reverse leakage current of diodes reduces efficiency. Self-synchronous and differential rectifier is proposed to effectively remove these effects at critical communication distance [10], [11] and [12], see figure 3. The best result is achieved in [11] with %66 PCE.

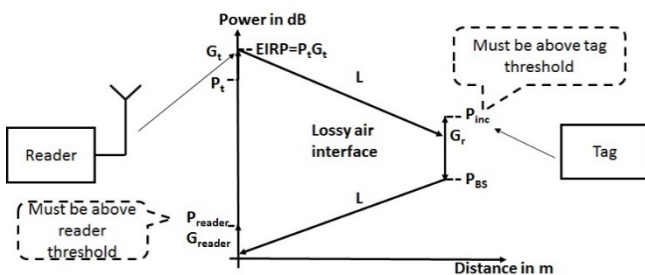


Figure 2. Passive RFID with subcircuits and monostatic and bistatic zero-IF reader

In subsequent stages of rectifier or multiplier, RFID digital core dissipates the most amount of power at one time restricting tag distance because activity of

demodulator, backscatter modulator and standard functionalities on digital core are distributed over time. By using ultra low power design techniques, a RFID processor at 0.33 V supply voltage and 1 MHz input clock with 80 nW power consumption has been designed [13].

In addition to that, when the induced power changes on the RFID antenna, non-linear effects manifest themselves and change RFID input impedance. Therefore, to have long communication distance, RFID input impedance must be matched to antenna impedance at critical communication distance in which RFID just obtains the power to run [10], called threshold.

Using the latest results in the literature %66 PCE and 80nW RFID processor P_{tag} , equal to 242.5 nW incident power on the antenna, for the range equation given in (4) for free space, it is found that range is less than or equal to 64.8 m at 868 MHz for half wavelength dipole tag antenna of gain 1.64 and 500 mW ERP in perfect matching conditions assumed. If power is increased to 2W ERP, distance becomes 129.7 m.

$$d \leq \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP} G_r}{P_{inc}}} = \sqrt{\frac{P_{EIRP} G_r}{\frac{2P_{tag}}{\eta}}}$$

(4)

On the other hand, downlink communication is restricted by the backscattered power in either one state during reflection. When it is assumed that in perfect matching condition half of incident power is reflected in ASK backscatter modulation, 121.3 nW power are reflected during high impedance states. -102 dBm power is obtained on reader antenna for half wavelength dipole antenna with 1.64 gain (G_t reader receive antenna gain), using equation (5). If power is increased to 2W ERP, -108.5 dBm power is obtained on reader P_{RX} .

$$P_{reader} = P_{RX} = \frac{P_{ref} G_r G_t}{\left(\frac{4\pi d}{\lambda}\right)^2}$$

(5)

-108.5 dBm signal is around the thermal noise level and falls below reader threshold for monostatic configuration around -80 dBm. Therefore, bi-static configuration, separate TX and RX antenna, with high antenna gain, narrow band filter and low noise feature must be employed to overcome noise floor limitation (SNR) [4], shown in figure2. For the amount of power more than 2W ERP, distant tag can

be energized but downlink signal would be masked by thermal noise in this case. The same result for P_{RX} signal level can be found by log distance path loss model with reduced distance.

3. REFLECTION COEFFICIENT AND POWER EQUATIONS ON AN RFID FRONT-END MODEL

In figure 3, passive RFID front-end model is shown and power induced on antenna, antenna impedance and RFID chip input impedance are represented by a power source with V_0 open circuit voltage, Z_{ant} and Z_{load} respectively. RFID input impedance is capacitive due to parasitic capacitances of transistors in rectifier circuit set by semiconductor technology in use and When the induced power changes on the RFID antenna, non-linear effects manifest themselves changing RFID input impedance. That is why a matched inductive antenna to RFID input impedance is employed at tag threshold. 1 and 2 indices for Z_{load} is used to show different states of input impedance during backscatter modulation.

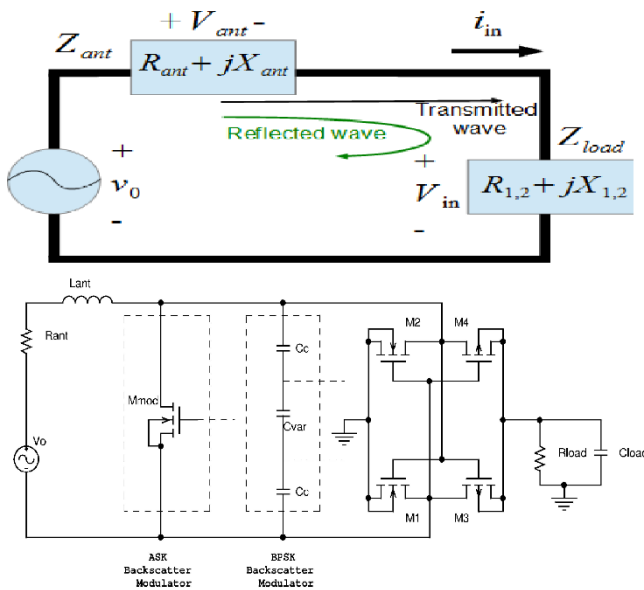


Figure 3. Passive RFID antenna model, rectifier and load @868 MHz

$$Z_{ant} = R_{ant} + jX_{ant}; Z_{load} = R_{1,2} + jX_{1,2}$$

$$\hat{v}_{in} = \frac{v_0(R_{1,2} + jX_{1,2})}{R_{1,2} + R_{ant} + j(X_{1,2} + X_{ant})}$$

$$\hat{i}_{in} = \frac{v_0}{R_{1,2} + R_{ant} + j(X_{1,2} + X_{ant})}$$

$$\rho = \frac{Z_{1,2} - Z_{ant}}{Z_{1,2} + Z_{ant}}$$

$$|\rho_{1,2}| = \sqrt{\frac{(R_{1,2} - R_{ant})^2 + (X_{1,2} + X_{ant})^2}{(R_{1,2} + R_{ant})^2 + (X_{1,2} + X_{ant})^2}}$$

$$\tau_{1,2} = 1 - |\rho_{1,2}|^2 = \frac{4R_{1,2}R_{ant}}{(R_{1,2} + R_{ant})^2 + (X_{1,2} + X_{ant})^2}$$

$$P_{aval} = \frac{P_{inc}}{2} = \frac{v_0^2}{8R_{ant}}$$

$$\begin{aligned} P_{RFIn1,2} &= \frac{1}{2} R(\hat{v}_{in}^* \hat{i}_{in}) \\ &= \frac{1}{2} R \left(\frac{v_0^2 (R_{1,2} - jX_{1,2})}{|R_{1,2} + R_{ant} + j(X_{1,2} + X_{ant})|^2} \right) \\ &= \frac{1}{2} \left(\frac{v_0^2 R_{1,2}}{(R_{1,2} + R_{ant})^2 + (X_{1,2} + X_{ant})^2} \right) \\ &= \frac{v_0^2}{8R_{ant}} \left(\frac{4R_{1,2}R_{ant}}{(R_{1,2} + R_{ant})^2 + (X_{1,2} + X_{ant})^2} \right) \\ &= P_{aval} \left(1 - \left| \frac{Z_{1,2} - Z_{ant}^*}{Z_{1,2} + Z_{ant}} \right|^2 \right) = P_{aval} (1 - |\rho_{1,2}|^2) \end{aligned}$$

$$P_{RFIn1,2} = P_{aval} \cdot (1 - |\rho_{1,2}|^2) = P_{aval} \cdot \tau_{1,2} \quad (6)$$

\hat{v}_{in} : Peak input voltage

\hat{i}_{in} : Peak input current

v_0 : Source peak voltage

$R_{1,2}$: State 1 and 2 resistance

$X_{1,2}$: State 1 and 2 reactance

R_{ant} : Antenna radiation resistance

X_{ant} : Antenna radiation reactance

$\rho_{1,2}$: State 1 and 2 reflection coefficient

$\tau_{1,2}$: State 1 and 2 transmission coefficient

P_{aval} : Incident power

$P_{RFIn1,2}$: State 1 and 2 transmitted power to RFID

$P_{aval} |\rho_{1,2}|^2$: Reflected power

Transmitted power to RFID from incoming RF field before backscatter modulation is given in (5) [7]. As we remember from transmission line theory, impedance matching, equal to zero reflection coefficient, is required to achieve maximum power transfer. In our case, conjugate match load ($Z_{ant} = Z_{load}^*$) corresponds to zero reflection coefficient and half of the power induced on antenna is transferred to load ($P_{inc}/2$).

4. BACKSCATTER MODULATION POWER EQUATIONS

There are two common backscatter modulations ASK and BPSK to be considered. On the reader side, average difference power radiated from antenna resistance between two states given in (8) is considered for data and the absolute power given in (7) must be above reader threshold at least one of two state. During backscatter communication in state 1 and 2, power transmitted to or obtained by tag is given in (9). (8) restricts communication range.

$$P_{BSdif} = \frac{|p_2 i_2 - p_1 i_1|^2}{2} R_{ant}$$

$$i_{1,2} = \frac{v_0}{Z_{ant} + Z_{1,2}} = \frac{v_0}{2R_{ant}} (1 - \rho_{1,2})$$

$$P_{DIS1,2} = \frac{1}{2} R(\hat{v}_{ant}^* \hat{i}_{in}) = \frac{1}{2} \left(\frac{v_0^2 R_{ant}}{(R_{1,2} + R_{ant})^2 + (X_{1,2} + X_{ant})^2} \right) = P_{aval} \left(\frac{\tau_{1,2} R_{ant}}{R_{1,2}} \right) \quad (7)$$

$$P_{BSdif} = P_{aval} |p_2(1 - \rho_2) - p_1(1 - \rho_1)|^2 \quad (8)$$

$$P_{RFinavg} = P_{aval} (p_1(1 - |\rho_1|^2) + p_2(1 - |\rho_2|^2)) \quad (9)$$

$P_{DIS1,2}$: State 1 and 2 dissipated power on antenna

P_{BSdif} : Average dissipated power on antenna between state 1 and 2

$P_{RFinavg}$: Average input power

$p_{1,2}$: Duty cycle for state 1 and 2,

$$0 < p_{1,2} < 1 \text{ and } p_1 + p_2 = 1$$

Matched load, open circuit and short circuit for ASK backscatter is easily grasped on the model in figure 3. For BPSK backscatter, real part of reflection coefficient is zero and imaginary part equal in amount but opposite in sign, $\pm m$. Modulation is achieved by changing RFID input capacitance equal amount and

so changing phase of voltage signal on antenna radiation resistance in equal degree. Base band signal from tag to reader is proportional to this phase signal ϕ which must be detectable on the reader side. Resistance and reactance values of RFID input impedance for BPSK equal amount of mismatched condition are found below.

$$\rho_{1,2} = \pm jm \quad 0 < m < 1 \quad (10)$$

$$\phi_{1,2} = \tan^{-1}(\pm m) \text{ Voltage phase over } R_{ant} \quad (11)$$

$$Z_{1,2} = Z_{ant} + \frac{2R_{ant}}{1 - \rho_{1,2}} = -jX_{ant} + \frac{1 + \rho_{1,2}}{1 - \rho_{1,2}} R_{ant} \quad (12)$$

$$X_{1,2} = -X_{ant} \pm \frac{2mR_{ant}}{1 + m^2} \quad (13)$$

$$R_{1,2} = \frac{(1 - m^2)}{(1 + m^2)} R_{ant} \quad (14)$$

In table 2, induced RFID chip and backscatter power equations for different modulations are presented. RFID-A and RFID-B can be run from longer distance than RFID-C increasing duty cycle up to 0.9 or 0.95 because RFID input power and average difference backscatter power have increasing trends in ASK backscatter in contrast to conflicting trends in BPSK modulation. OC case in RFID-A brings more average difference backscatter power on the reader side than SC case in RFID-B, as shown in [1]. However, in standards such as EPC Class1 Gen2 (GS1), each symbol has equal amount of time for high or low state, forcing %50 duty cycle and state duration for ASK and BPSK backscatter modulation types. In practice it is hard to create OC case. Actually, there is always reflection due to finite mismatches.

Table 2. Backscatter modulation comparisons [6]

	RFID-A	RFID-B	RFID-C
Type of Modulation	ASK backscatter	ASK backscatter	BPSK backscatter
State 1	$\rho_1 = 0$ $p_1 = DC$ Matched load	$\rho_1 = 0$ $p_1 = DC$ Matched load	$\rho_1 = -jm$ $p_1 = 0.5$ Equal mismatch
State 2	$\rho_2 = 1$ $p_2 = 1 - DC$ Open circuit (OC)	$\rho_2 = -1$ $p_2 = 1 - DC$ Short circuit (SC)	$\rho_2 = jm$ $p_2 = 0.5$ Equal mismatch
Average difference backscatter power	$P_{BSdif} = DC^2 P_{aval}$	$P_{BSdif} = (3DC - 2)^2 P_{aval}$	$P_{BSdif} = m^2 P_{aval}$
Average input power	$P_{RFInavg} = DC P_{aval}$	$P_{RFInavg} = DC P_{aval}$	$P_{RFInavg} = 1 - m^2 P_{aval}$
State 1 backscatter power	P_{aval}	P_{aval}	$P_{BS1} = P_{aval}(1 + m^2)$
State 2 backscatter power	0	$P_{BS1} = P_{aval} \left(\frac{4R_{ant}^2}{R_{ant}^2 + X_{ant}^2} \right)$	$P_{BS1,2} = P_{aval}(1 + m^2)$

5. REALIABILITY OF BPSK AND ASK MODULATIONS AND MINIMUM SNR

Minimum signal to noise ratio for a targeted BER is required to be known for reader sensitivity. Signal from tag to reader passes through multipath environment getting attenuated by reflected signals from near-by objects, called multipath fading for small amplitudes and shadowing for large amplitudes. In addition to that, channel noise must be considered to see effects of real environment challenges. On signal side, backscattered difference power is relevant for reader, given in equation (8).

To see effect of scattered signals and noise on the signal power, proper channel models are employed depending upon application environment. Weak ASK and BPSK backscatter modulated signals are assumed to pass through AWGN channel (Additive White Gaussian Noise), Rayleigh fading channel, useful model when there is no line of sight or there are many reflectors, and Rician fading channels to see their performances, used when the line of sight component

is much stronger than others figure 4. Bit error probability is given in (15), (16) and (17) for AWGN, Rayleigh and Rician channel models, E_b energy per bit and N_0 noise. When you think about situation that the same place located tags with BPSK modulation index 0.5, 0.41 change in reactance is equal to resistance, 0.32 and ASK %50 duty cycle has -3 dB, -4.7 dB, -7 dB and -3 dB lower SNR ratio with respect to a tag BPSK modulation index 0.707 figure 4. They are compared to exhibit SNR and BER performances considering differences by [simulations in theory and 10 million normal distributed samples](#) in figure 4 [14]. k is ratio of dominant LOS component power to scattered components power in Rician channel model.

$$P_b = \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{E_b}{N_0}} \right) \tag{15}$$

$$P_b = \frac{1}{2} \left(1 - \sqrt{\frac{\frac{E_b}{N_0}}{\frac{E_b}{N_0} + 1}} \right) \tag{16}$$

$$P_b = \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{k \frac{E_b}{N_0}}{k + \frac{E_b}{N_0}}} \right) \tag{17}$$

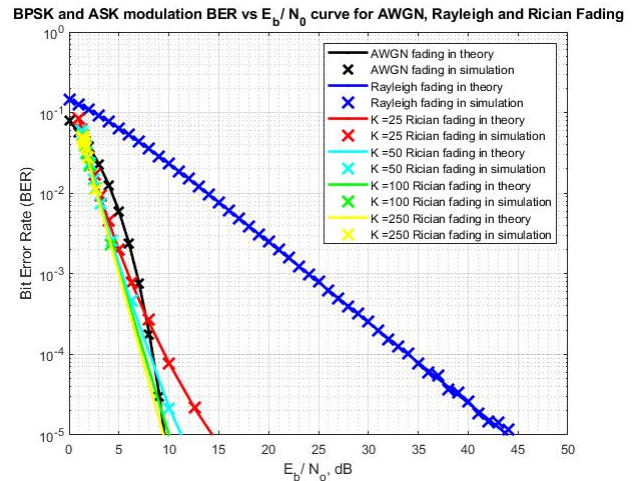


Figure 4. ASK and BPSK backscatter modulation BER curve for AWGN, Rayleigh and Rician channel models

BER value for BPSK $m=0.707$ tag is always lower than other ASK and BPSK tags and %50 ASK and BPSK $m=0.5$ are the same although input power to tag in BPSK tag is %50 higher than ASK tag. Effect of noise for low modulation indexes is higher since

signal power is low. There is always needed much higher signal to noise ratio in multiple reflector environment.

Practical value of 10^{-4} and 10^{-5} BER are used general. For example, 10^{-4} and 10^{-5} BER result in 0.05 and 0.005 read error rate for 64 byte EPC (Electronic Product Code) and similarly, 0.01 and 0.001 for 12 byte EPC. For 10^{-5} BER value, minimum SNR levels are around 9.5 dB, 11 dB and 44 dB corresponding to AWGN, Rician for $k=25$ and Rayleigh fading channels.

6. MAXIMUM RANGE

During backscatter modulation, tag can be run out of energy due to low state of symbol. For ASK case, there are two situations limiting the uplink distance. First, it is considered that tag has enough energy on the capacitor C_{load} during low level of symbol with short duration for fast enough communication not to run out of energy as in sensor tags, see figure 3. In this case, we are limited with equation (18). On the other hand, equation (19) limits uplink distance for tags with small C_{load} and low state of symbol with long duration for low speed communications. For BPSK case, equation (4) can be rewritten as in (20). Equations are given for %50 state durations.

$$r \leq \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP}G_r}{P_{inc}}} = \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP}G_r\eta(1-|\rho_1|^2)}{2P_{tag}}} \quad (18)$$

$$r \leq \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP}G_r}{P_{inc}}} = \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP}G_r\eta((1-|\rho_1|^2)+(1-|\rho_2|^2))}{4P_{tag}}} \quad (19)$$

$$r \leq \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP}G_r}{P_{inc}}} = \frac{\lambda}{4\pi} \sqrt{\frac{P_{EIRP}G_r\eta(1-m^2)}{2P_{tag}}} \quad (20)$$

For downlink communication, working distance can be rewritten for ASK and BPSK modulations in similar case to above (21), (22) and (23) assuming that tag obtains only enough power to run, G_r reader receive antenna gain. For ASK and BPSK modulations, backscatter difference power and so $P_{BS1,2}$ for either one of states must be above reader receive threshold, P_{reader_th} represents this level. Reader sensitivity can be found by $P_{reader_th} = -174 \text{ dBm/Hz} + \text{NF} + 10\log(\text{BW}) + \text{SNR}_{\min}$, NF noise figure of system, BW bandwidth and SNR_{\min} minimum signal to noise ratio of the system, sum of channel

SNR, RF chain SNR and analog baseband SNR. First three terms represents integrated noise of system called noise floor.

$$d \leq \frac{\lambda}{4\pi} \sqrt{\frac{P_{BSdif}G_rG_t}{P_{reader_th}}}$$

$$d \leq \frac{\lambda}{4\pi} \sqrt{\frac{0.25 P_{tag}G_rG_t}{\eta P_{reader_th}(1-|\rho_1|^2)}} \quad (21)$$

$$d \leq \frac{\lambda}{4\pi} \sqrt{\frac{0.5 P_{tag}G_rG_t}{\eta P_{reader_th}((1-|\rho_1|^2)+(1-|\rho_2|^2))}} \quad (22)$$

$$d \leq \frac{\lambda}{4\pi} \sqrt{\frac{P_{tag}m^2G_rG_t}{\eta P_{reader_th}(1-m^2)}} \quad (23)$$

As stated before maximum communication distance is restricted by either uplink or downlink communication, found by $\min(r,d)$ for two modulations. Using the latest results in the literature %66 PCE η , 80nW RFID processor P_{tag} , half wavelength dipole tag antenna gain 1.64 for given size, reader receive antenna gain 6.56 (8.14 dB), 2 W ERP, 0 state 1 reflection coefficient ρ_1 , -1 state 2 reflection coefficient ρ_2 and -101 dBm reader threshold P_{reader_th} with GS1 640 kHz, 5.5 dB noise figure and 9.5 dB more than thermal noise at 868 MHz in perfect matching conditions assumed, results are listed in table 3.

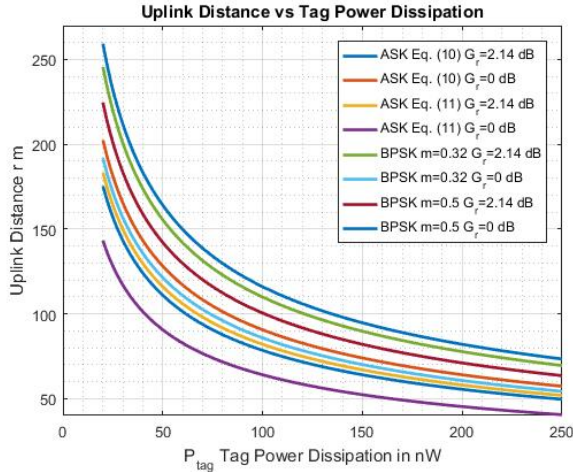


Figure 5. Uplink distance for ASK and BPSK modulations for given parameters

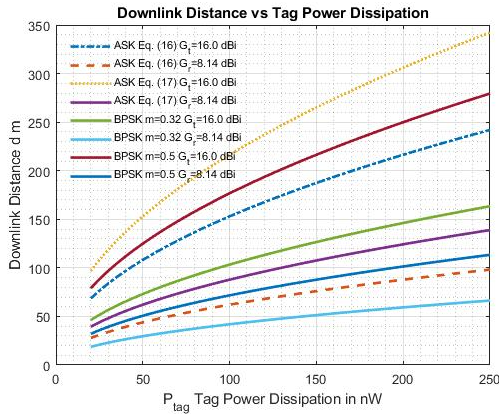


Figure 6. Downlink distance for ASK and BPSK Modulation for given parameters

During design time, rectifier efficiency η , tag power dissipation P_{tag} , tag antenna gain G_r for a given RFID tag label size, reader threshold P_{reader_th} and transmit power P_{ERP} limited by authorities are known and for given parameters above, only, reader receive antenna gain G_t and m modulation index for ASK and BPSK modulations can be changed. This point can be utilized to find antenna gain G_t and m modulation index to make d equal to r , given in (24), (25) and (26). If system is uplink limited, the distance difference brings more signal power on reader side so more SNR and reliable operation. When downlink limits distance, equations (24), (25) and (26) must be considered. We face this situation now and in further improvement of P_{tag} and η and higher P_{reader_th} requirement. The same results can be found by log distance path loss model.

$$G_t = \frac{P_{ERP} P_{reader_th} \eta^2 (1 - |\rho_1|^2)^2}{0.5 P_{tag}^2} \quad (24)$$

$$G_t = \frac{P_{ERP} P_{reader_th} \eta^2 ((1 - |\rho_1|^2) + (1 - |\rho_2|^2))^2}{2 P_{tag}^2} \quad (25)$$

$$A = P_{reader_th} P_{ERP} \eta^2 B = 2 P_{tag}^2 G_t$$

$$m = \sqrt{\frac{+2A + B \pm \sqrt{4AB + B^2}}{2A}} \quad (26)$$

Table 3. Maximum free space range for $\eta=0.66$, $P_{tag}=80nW$, $G_r=1.64$, $G_t=6.56$, $P_{reader_th}=-101$ dBm, $P_{ERP}=2W$ @868MHz

Modulation	Uplink Distance	Downlink Distance	Maximum Range
ASK %50 SC and matched load	129.7 m for Eq. (18)	55.6 m for Eq. (21)	55.6 m
	91.7 m for Eq. (19)	78.6 m for Eq. (22)	78.6 m
BPSK $m=\pm 0.707$ ($\varphi=\pm 35.3^\circ$)	91.7 m	111.2 m	91.71 m
BPSK r and d equal $m=\pm 0.66$ ($\varphi=\pm 33.4^\circ$)	97.5 m	97.7 m	97.5 m
BPSK $m=\pm 0.5$ ($\varphi=\pm 26.6^\circ$)	112.3 m	64.2 m	64.2 m
BPSK $m=\pm 0.41$ ($\varphi=\pm 22.3^\circ$)	118.2 m	50.1 m	50.1 m
BPSK $m=\pm 0.32$ ($\varphi=\pm 17.8^\circ$)	122.9 m	37.0 m	37.0 m
BPSK $m=\pm 0.22$ ($\varphi=\pm 12.4^\circ$)	126.5 m	25.4 m	25.4 m

After RFID tag design, changes can be made on G_t reader receive antenna gain and reader threshold P_{reader_th} (actually there is no change) to make two distances equal. If modulation index m is bigger than $m=0.66$, tag is limited by uplink distance. Modulation index value close to 1 cannot be used because tag cannot be energized (20). Low gain G_t antenna can be used for cost reduction or P_{reader_th} reader threshold requirement can be relaxed to higher values with low BER values while keeping the modulation index and so uplink distance the same (23). Actually, uplink limited system has more reliable operation. In ASK case, there is similar situation. On the other hand,

below $m=0.66$ value, signal will not be fetched for the distances more than maximum downlink distance. Higher gain receive antenna can bring additional degree of freedom to increase downlink distance as much as possible or similarly $P_{\text{reader_th}}$ reader threshold level can be reduced for this purpose at the cost of more BER while keeping the modulation index m and so uplink distance the same. However, it is needed to be considered that baseband signals with low modulation index m is hard to detect on reader side and effect of noise increases. That is why for the environment being operated the lowest level modulation index m must be decided first. Then, antenna gain G_t and $P_{\text{reader_th}}$ can be adjusted to make two distances equal, a kind of trade-off (20) and (23). Further reduction of P_{tag} and improvement of η has also bright future in terms of increasing distance, modulation index m and so increasing SNR on reader (23). In ASK case, situations are similar. In addition, if downlink limited distance, antenna gain could be increased to 35.7 (15.5 dBi) and 8.8 (9.5 dBi) to make two distances equal (24) and (25).

7. RESULTS

In this study, uplink and downlink distances are shown for UHF RFID in detail and parameters to make two distance equal are given. Also, it is shown that reflected signal levels require separate TX and RX antenna, bistatic configuration. Latest works show that BPSK tags are proper choice for working distance and reliability.

Complete link budget in a RFID system is given for different environments. Power consuming parts effecting this range in a passive RFID hardware are presented with power reduction techniques to have longer range because distant operation is desired in harsh and noisy working environments. It is shown on the range equations by using the results from literature that both uplink and downlink communication distance must be considered during the design. Hence, in the future we will see next generation RFID tags and readers working from longer range in challenging environments due to low power nature of RFID tags, especially in supply chain management and mid-range passive wireless sensor network (WSN) environment and lower detection threshold of readers.

Parasitics included simulations in TSMC 0.18 μm RF technology show that differential rectifier efficiency

in figure 3 is around %66-68 over wide range of input power up to more than -37 dBm with fine tunings. Latest works on RFID tag processor achieved 80 nW in similar technology node [13]. With these points and reasonable parameters given in table 3, two state difference electromagnetic power reaching RFID reader is around integrated or thermal noise of systems. This result directs to separate TX and RX antenna bistatic configuration for RFID reader.

ASK tag with large capacitor at the rectifier output and matched load have longer uplink distance than other ASK tags and BPSK tags when you look at equations (18), (19), (20) and table 3. On the other hand, Downlink limits due to noise floor of reader. For this purpose, high gain reader receive antenna must be employed to increase range considering cost containment. Similarly, higher modulation index BPSK modulation has higher downlink distance.

In the presence of current improvements, BPSK tags are proper choice in noisy environments and have higher working distance. High modulation index BPSK modulation has more reliable operation with low BER proper to use harsh environments for safe information and to decrease error. None the less, situation reverse to ASK in terms of working distance and reliability if high antenna gain G_t around 16 dBi is used. The same results and conclusions can be found by log distance path loss model except running distance.

Simulations on different communication channel models show that noise requires wide range of modulation index tags depending upon operating environment. One tag with programmable mismatch in BPSK case like capacitor banks looks more logical.

Some other opportunities are seen. First, when you look at equation (8) and (9), there is a potential to use available power in BPSK modulation for higher speed and bandwidth efficient M-ary modulation formats such as QAM. In table 3, the same situation is seen by looking at decreasing angle $\arctan(\pm m)$ and decreasing magnitude (or distance) $(1+m^2)P_{\text{aval}}$ going down to table. Realization can be done employing more number of varactors on BPSK modulator over positive x-plane at cost of the distance penalization. Advantage of M-ary modulation formats is reduction of the amount of processing time on massive deployments. Second, for portable devices, operation distance is defined and less than maximum distance, low cost reader hardware with small power amplifier

and fast operation can be used to dissipate less energy, to have long battery duration, to have less air traffic and to create less electromagnetic interference.

REFERENCES

- [1] P. Nikitin and K. V. S. Rao, "Theory and measurement of backscattering from RFID tags," *IEEE Antennas Propag. Mag.*, vol. 48, no. 6, pp. 212–218, Dec. 2006.
- [2] S. S. Srikant and R. P. Mahapatra, "Read Range of UHF Passive RFID," *Int. J. Comput. Theory Eng.*, vol. 2, no. 3, pp. 323–325, 2010.
- [3] Y. Gao, Z. Zhang, H. Lu, and H. Wang, "Analysis and calculation of read distance in passive backscatter RFID systems," *LISS 2012 - Proc. 2nd Int. Conf. Logist. Informatics Serv. Sci.*, vol. 2, no. 1, pp. 905–912, 2013.
- [4] J. D. Griffin and G. D. Durgin, "Complete Link Budgets for Backscatter-Radio and RFID Systems," *IEEE Antennas Propag. Mag.*, vol. 51, no. 2, pp. 11–25, Apr. 2009.
- [5] P. Nikitin, K. V. S. Rao, S. Lam, and W. Ave, "UHF RFID TAG CHARACTERIZATION : OVERVIEW AND STATE-OF-THE-ART," no. 2, pp. 2–7.
- [6] F. a. Hussien, D. Z. Turker, R. Srinivasan, M. S. Mobarak, F. P. Cortes, and E. Sanchez-Sinencio, "Design considerations and tradeoffs for passive RFID tags," vol. 5837, pp. 559–570, Jun. 2005.
- [7] U. Karthaus and M. Fischer, "Fully Integrated Passive UHF RFID Transponder IC With 16 . 7- W Minimum RF Input Power," vol. 38, no. 10, pp. 1602–1608, 2003.
- [8] M. Viswanathan and S. Edition, "SIMULATION OF DIGITAL COMMUNICATION SYSTEMS USING MATLAB."
- [9] P. Note, "ERC Recommendation 70-03," no. February, 2014.
- [10] M. H. Ouda *et al.*, "for Implantable Intraocular Pressure Monitoring," pp. 1–8.
- [11] K. Kotani, A. Sasaki, T. Ito, and S. Member, "High-Efficiency Differential-Drive CMOS Rectifier," vol. 44, no. 11, pp. 3011–3018, 2009.
- [12] R. Vaddi, R. P. Agarwal, S. Dasgupta, and T. T. Kim, "Design and Analysis of Double-Gate MOSFETs for Ultra-Low Power Radio Frequency Identification (RFID): Device and Circuit Co-Design," *J. Low Power Electron. Appl.*, vol. 1, no. 3, pp. 277–302, Jul. 2011.
- [13] W. Shi, C. S. Choy, J. Guo, C. F. Chan, K. N. Leung, and K. P. Pun, "A 90nm RFID tag's baseband processor with novel PIE decoder and uplink clock generator," *Midwest Symp. Circuits Syst.*, pp. 644–647, 2010.
- [14] DSPLOG, "Communication Channels." [Online]. Available: <http://www.dsplog.com>.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Numerical and experimental approach of various sectioned new concept of the crash-boxes to determine the reliability and crashworthiness of the vehicles during frontal impacts

Authors: İbrahim Kutay Yılmazçoban, Ömer Adanur, Ahmad Bakhtiyar, Aslı Ergün

Received: 2018-09-14 16:40:19

Revised: 2018-09-26 16:51:53

Accepted: 2018-10-11 11:29:05

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 76-84

How to cite

İbrahim Kutay Yılmazçoban, Ömer Adanur, Ahmad Bakhtiyar, Aslı Ergün; (2019), Numerical and experimental approach of various sectioned new concept of the crash-boxes to determine the reliability and crashworthiness of the vehicles during frontal impacts. Sakarya University Journal of Science, 23(1), 76-84, DOI: 10.16984/saufenbilder.460078

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/460078>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Numerical and experimental approach of various sectioned new concept of the crash-boxes to determine the reliability and crashworthiness of the vehicles during frontal impacts

Ibrahim Kutay Yilmazcoban^{*1}, Omer Adanur¹, Ahmad Bakhtiyar¹, Asli Ergun¹

ABSTRACT

Overview on vehicle collisions shows us a perspective to evaluate the impact behavior of the vehicles and occupant safety. To decrease the deformations of the vehicle body and occupant injuries, many products have been developed already. The mentioned safety measures can be defined via active and passive security systems. This study is about, passive safety systems used naming “the crash-box” for vehicles during accidents, are investigated experimentally to reduce the shock effects of the frontal impacts. For crash-boxes square, circular or other standard closed sections are preferred regularly due to ease of the production. For this research, open section crash-boxes are studied. Up to the manufacturing restrictions one of the designs in four, is preferred. The new concept of crash-box is, assembling the specimens in front of the chassis and just behind front bumper, instead of using it like regular crash-boxes. Cross-section type of w shape, steel, sheet-metal crash-boxes are manufactured as a shock absorber unit. Height of 2.88m drop test setup was used for the experiments. Considering the limitations of the test setup, one eighth of a real accident scenario for a 1200kg vehicle can be processed with this experimental system. The impact tests are handled with 150kg drop weight and one crash-box unit as an accepted collision scale. To understand the most appropriate thickness for the absorbers is decided between 2mm, 1.5mm, 1mm and 0.8mm thick samples. Finally, 1mm thick St37 w shape cross-sectional sheet metal crash-box is strong enough to absorb the impact energy of the frontal collisions.

Keywords: Vehicle Reliability, Impact, Shock Energy, Crash-box, Occupant Safety

* Corresponding Author: kyilmaz@sakarya.edu.tr

¹ Sakarya University, Faculty of Engineering, Department of Mechanical Engineering, Sakarya/Turkey

INTRODUCTION

Due to the advancement of vehicle technologies with the comfort and speed, accidents have also increased concerning the human factors. During collisions even sophisticated safety procedures, injuries and sometimes fatalities are inevitable. One of the most important problem of vehicles collision is frontal impact for decades [1-8]. Thus, manufacturers and establishments (Euro NCAP etc.) have been trying to facilitate rules and create different type of prevention systems.

Recent studies indicate about features; inside the chassis and under the hood staying in front of the driver like protection bars to decrease the collision effects and prevent the passengers. This safety system defined as a crash-box [9-11]. Crash box is a system converting the kinetic energy caused by the collision, via deforming itself in plastic region of the material and absorbing the impact energy and shock waves of the accident and is expected to be collapsed with absorbing crash energy prior to the other body parts so that the damage of the main cabin frame is minimized, and passengers have saved their lives. Crash-box or thin walled safety structure is designed for absorbing approximately 50% of kinetic energy of vehicles during frontal impact collision [12-14]. Widely known cross section of energy absorbers is mainly utilized as a rectangular/square shape. Previous studies indicated to understand impact behavior of rectangular/shape under static or dynamic axial loading. The most significant design parameters

for the crash-box can be described as energy absorption efficiency and light weight [15,16]. Shock absorbing capability advances gradually by using pre-notched crash-boxes; otherwise in non-notched samples observed deformation is materialized in an uncontrolled way. Yet, because of the additional manufacturing processes for the notching, it is increasing the costs of the production. Considering the effective, simple and economic way, non-notched samples have been preferred for this study.

Instead of regular shapes known as rectangular or circular cross sections [17,18], different geometries like W folded crash-boxes are selected for this research. W shaped design could make many fold ways [19-21], but up to the manufacturing ability and limitations, one of the designs are accepted.

The manufactured specimens could not be carried out in the form of actual vehicle crash tests due to real collisions, laboratory conditions and material availability. Instead, the tests were accomplished using the drop test setup. The data obtained from the current test setup were analyzed and evaluated by the experimental approach. Considering the deformations of the different thicknesses of crash-boxes, 1mm thick specimen gain enough amount of the shock energy of the impact scenarios.

MATERIAL AND TESTING

Experimental study of the impact procedures is handled via drop test module at laboratory

facilities in Sakarya University, Turkey. While the real accidents are very difficult to simulate in the laboratory conditions, some parts of the collisions are partially imitable to actualize the daily life incidents. In this perspective frontal crash case will be evaluated with a drop test approximation.

2.1. Experimental Setup

The drop test module (Figure 1) is able hold 150 kg and drop it from 2.88m height. Because of the sample height is 300mm, it can reach a crash-box contact maximum speed of 24.6024km/h measured by the speed sensor.

The impact velocities were measured by the velocity sensor just before the contact of the crash-box and steel plate. To understand the absorbing capability, four distinct specimen thicknesses are selected considering previous studies for different cross sections [22]. Deciding which thickness is the most appropriate for the absorbers, 2mm, 1.5mm, 1mm and 0.8mm thick sheet metal boxes crashed respectively via drop test experiments.

Up to the calculations, the experiments represents eight crash-box mounted in front of a 1200kg vehicle, 4 by 4 symmetrically in a 24.5km/h impact scenario.

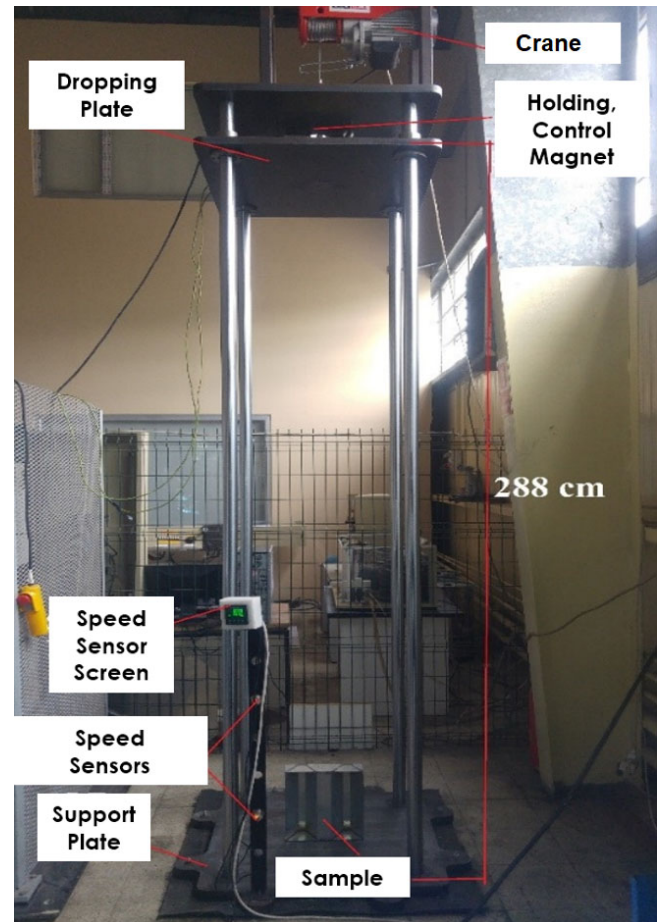


Figure 1. Drop Test Setup

2.2. Material Properties

When creating a description for the safety precautions of the accidents, companies usually try to facilitate economical methods to reduce the deformations and injuries of the occupants. Sometimes this way is enough, nevertheless unfortunately not enough for all. Protective systems should be developed and the system should not much expensive for the mass production. Regarding the automobile industry St-37 mild steel sheet-metal material was applied to the W shaped cross-sectional designs just because the St-37 material is financially suitable for the manufacturers.

Just because the regular shapes had been studied before for decades, different profiles were designed to demonstrate the behavior of the W shaped folded crash boxes. The designs can be seen in Figure 2. With the limitations of the manufacturer, the most convenient design 4 was selected as a crash-box during the frontal impact scenario of the vehicle.

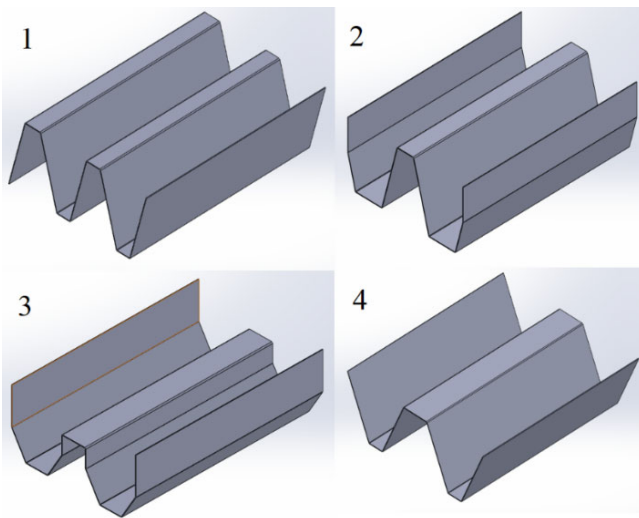


Figure 2. The designs of the W shaped crash boxes

The dimensions and perimeter of the columns are designed up to the recent studies and appropriate space imagined between the chassis and the frontal bumper of the daily used vehicles. When the height is 304mm, width and depth are 260mm and 90mm respectively. The height and the perimeter of the samples are considered by the effects of the collisions during frontal impacts of the vehicles to absorb enough amount of the shock energy. And also the thicknesses are decided up to the availability of the sheet metal

market as 0.8mm, 1mm, 1.5mm and 2mm. The dimensions of the design 4 is given in Figure 3.

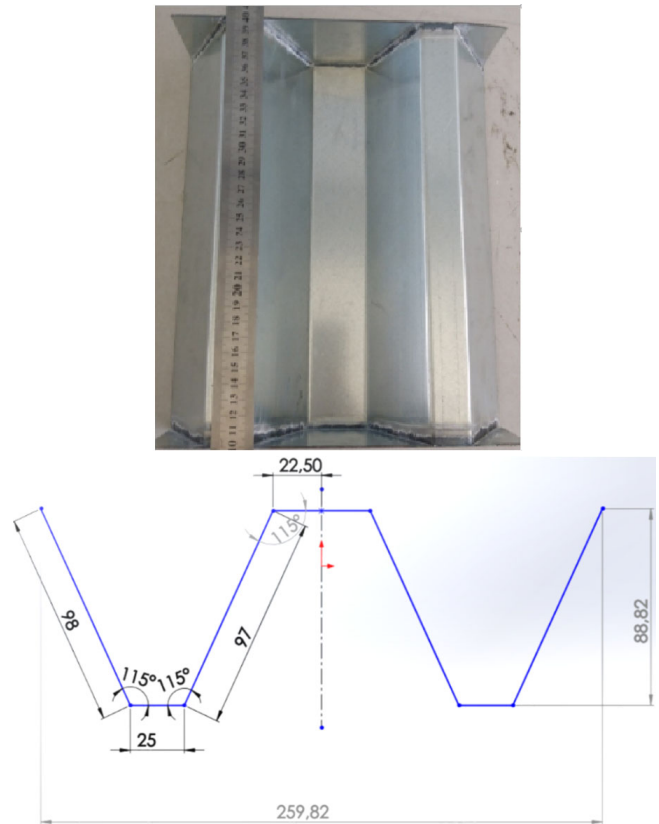


Figure 3. Dimensions of the W cross section

2.3. Experiments

The main idea in the impact shock absorbers like crash boxes is to deform enough to absorb sufficient amount of destructive energy. But if the specimens deform all the way in a large deformation scale, there will be no space to fold on itself and protect the vehicle anymore. Thus, initially the optimum thickness should be evaluated between very large deformations and not enough deformable behavior of the crash-boxes. During the tests the folding mechanism usually starts from the bottom side, opposite of the initial contact region as a reaction attitude.

2mm thick specimen: Initially a little more buckling seen at the left side. Folding deformations are determined above from the right side of the column is 69mm, left side is 79mm and the center is 75mm. After the impact test the average height became 279mm. This sample is so rigid and is not able to absorb harmful shock energy. The crash-box will transfer the shock wave directly to the vehicle. The impact deformation of the 2mm thick specimen can be seen on Figure 4.



Figure 4. The impact deformation of the 2mm thick specimen

1.5mm thick specimen: As seen in the Figure 5, buckling at left side is more than the right. After first fold reached to the bottom, second fold at right is 33mm above from the bottom, and for the left is 29mm above. Defined plastic deformation at right column is 14mm, at left 16mm and at middle 15mm. Right column decreased to 247mm, and left decreased to 243mm. 1.5mm thick plate sample crash-box is not able to absorb enough energy because of rigidity.



Figure 5. Drop test result for the 1.5mm thick specimen

0.8mm thick specimen: After the impact tests for the 0.8mm thick specimens, large deformations were determined. Opposite of the 2mm and 1.5mm thick samples, 0.8mm thick crash-boxes were not able to absorb enough energy considering that the all 0.8mm thick specimens deformed in an extra-large behavior. Two other specimens are behaved identically.



Figure 6. Drop test result for the 0.8mm thick specimen

Results up to this section shows that 2mm and 1.5mm thick samples are acted rigidly and not sufficient folding mechanism is detected. On the

other hand, all of the three 0.8mm thick samples present extra-large deformation and this sample is not capable to withstand the impact forces.

According to the results to gain optimum thickness for the crash-box should be between 1.5mm and 0.8mm. Subsequently, 1mm thick easily obtainable St37 sheet metal is decided for the next experiment.

1mm thick specimen: The samples 0.5mm thinner than the previous one, thence the deformation result and mechanism is different. The increase of number of the folding mechanism represents an acceptable folding process for an efficient impact absorbing capability, besides the crash-box is not indicating large-deformation which is favorable. To depict the correlation between the numbering of the samples and specimen numbering in the Figure 7 subsequently the trial sample (W01_I1_S01_T01: specimen 1) which is not given, the specimen numbering starts from the number 2. After the first trial impact test, the second impact test (Figure 7-2) represents the gradually and acceptable folding deformation for absorbing enough energy. The results for the 2nd specimen shows that it absorbed the shock waves and energy during the deformation between 304mm to 199mm. The drop test for the 3rd sample is carried out, and the result of the test is depicted in Figure 7-3. The outcome for the 3rd specimen demonstrate that absorbing the impact energy during deformation, is starting from 304mm to approximately 200mm. Impact test for

the fourth specimen (Figure 7-4), the height of the crash-box decreased to, between 189mm to 184mm. The second folding is not at the bottom like the 2nd and 3rd; it is on the top region of the crash-box. When the 5th specimen is examined in Figure 7-5, two folds were observed just above the lower surface. Just below the upper surface, folding starts in the two columns, but the plastic deformation of buckling in the right column is greater than in the left column. The result of 5th specimen was found to be 209mm on the right side from 300mm and 211mm on the left side. For the 6th and 7th (See Figure 7-6 and 7) specimens, the folding behaviors are close to 4th one. One fold at the bottom and the other fold at the top surface of the crash-box. Results of 6th specimen shows that the height decreased from 300mm to 187mm on the right side and from 300mm to 185mm on the left side. And the 7th specimen's height decreased from 300mm to 181mm on the right side and from 300mm to 182mm on the left side.



Figure 7. Drop test results of the 1mm thick specimen
When the all experiments are considered, the results at Table 1 should be investigated.

Table 1. Experimental results of the drop tests

Specimens	Velocity (m/s)	Absorbed Energy (J)	Mean Deformation (mm)	Deformation
W01-I2-S01-T01 (2mm thick)	6,815	3072	25	Not Enough (Rigid)
W01-I1.5-S01-T01 (1.5mm thick)	6,821	3119	57	Not Enough (Rigid)
W01-I1-S01-T01 (1mm thick-trial sample)	6,79	3164	98	
W01-I1-S01-T02 (1mm thick)	6,827	3202	101	
W01-I1-S01-T03 (1mm thick)	6,809	3189	99	
W01-I1-S01-T04 (1mm thick)	6,811	3204	111	Acceptable
W01-I1-S01-T05 (1mm thick)	6,814	3173	91	
W01-I1-S01-T06 (1mm thick)	6,834	3209	115	
W01-I1-S01-T07 (1mm thick)	6,825	3226	118	
W01-I0,8-S01-T01 (0.8mm thick)	Error	Error	245	
W01-I0,8-S01-T01 (0.8mm thick)	6,741	3343	256	Large-Def.
W01-I0,8-S01-T01 (0.8mm thick)	6,88	3447	230	

After comparing the results, 1mm thick St37 W shaped crash-box absorbers exhibit better results to eliminate the shock effects of the frontal collision scenario. For the further studies, 1.2mm thick samples will be added to the W shaped cross sections expanding with regular sections thickness of 1mm to compare the different shapes acts. After deciding the best shape and thickness the height optimization will take place for a superior crash-box design. To improve the experiments, accelerometer is going to be included while the numerical approximation of

explicit dynamics procedures will be conducted with LS-DYNA software.

CONCLUSION

According to the experimental study of the frontal impact simulations via drop test setup is admissible to describe the direct collisions as expected. Although crash-box developments spread on spacious studies, there are still sufficient amount of undiscovered areas exist. To saturate some part of not defined areas of the impact absorbing field, this study has accomplished using not regular shapes.

Instead of using standard cross-sections, w cross-section open form crash-boxes are preferred to see the mechanical behavior of the new concept. And also for the further studies, the W shaped cross sections will be investigated via comparing with the standard cross sections.

Thickness optimization of the crash-box has been processed by the experiments. The impact absorbing capability of the 2mm and 1.5mm thick crash boxes is not appropriate because of the rigid and not enough deformable behavior.

When 0.8mm thick boxes are investigated the mechanical strength was found so weak to absorb enough impact energy, considering all samples have large-deformations.

According to the results to gain optimum thickness for the crash-box should be between 1.5mm and 0.8mm. Hereby, 1mm thick easily obtainable St37 sheet metal is capable to absorb

enough impact energy by folding two times on itself while the rigid ones just fold once and the soft samples fold all the way.

Finally, repeated 1mm thick specimen experiments, verify the absorbing ability during many drop tests. And also the results for the new design of crash-box concept are promising.

REFERENCES

- [1] S. Boria, J. Obradovic, and G. Belingardi, "Experimental and numerical investigations of the impact behaviour of composite frontal crash structures," *Composites Part B: Engineering*, vol. 79, pp. 20–27, 2015.
- [2] K. Solabannawar, S.I. Bekinal, "To simulate and study the behaviour of vehicle front structure through tox joints in frontal collision," *International Research Journal of Engineering and Technology (IRJET)*, vol. 03, no. 7, pp. 1702-1707, 2016.
- [3] A. Ghadianlou, S.B. Abdullah, and A. Agarwal, "Crashworthiness design of vehicle side door beams under low-speed pole side impacts," *Thin-Walled Structures*, vol. 67, pp. 25-33, 2013.
- [4] N. Tanlak, "Cross-sectional shape optimization of thin-walled columns enduring oblique impact loads," *Thin-Walled Structures*, vol. 107, pp. 65-72, 2016.
- [5] N.N. Hussain, S.P. Regalla and Y.V.D. Rao, "Low velocity Impact Characterization of Glass Fibre Reinforced Plastics for Application of Crash Box," *Materialstoday: Proceedings*, vol. 4 (2 PartA), pp. 3252-3262 2011.
- [6] J.A.C. Ambrósio, "Contact and impact models for vehicle crashworthiness simulation," *International Journal of Crashworthiness* 8, vol. 8, no. 1, pp. 73-86, 2003.
- [7] Y. Nakazawa, K. Tamura, M. Yoshida, K. Takagi and M. Kano, "Development of Crash-box for passenger car with high capability for energy absorption," *CIMNE 2005 VIII International Conference on Computational Plasticity*, 2005.
- [8] G. Belingardi, A.T. Beyene, E.G. Koricho and B. Martorana, "Alternative lightweight materials and component manufacturing technologies for vehicle frontal bumper beam," *Composite Structures*, vol. 112, pp. 1-10, 2014.
- [9] C. Zhou, Y. Zhou and B. Wang, "Crashworthiness design for trapezoid origami crash boxes," *Thin-Walled Structures*, vol. 117, pp. 257-267, 2017.
- [10] G. Belingardi, A.T. Beyene, E.G. Koricho and B. Martorana, "Crashworthiness of integrated crash-box and bumper beam made by die-forming composite," *16th European Conf. on Composite Materials*, 2014.
- [11] H. Chul, K. Dong, K. Shin, J.J. Lee and J.B. Kwon, "Crashworthiness of aluminum/CFRP square hollow section

- beam under axial impact loading for crash box application,” *Composite Structures*, vol. 112, pp. 1-10 2014.
- [12] G. Suna, T. Panga, C. Xua, G. Zhenga and J. Songc, “Energy absorption mechanics for variable thickness thin-walled structures,” *Thin-Walled Structures*, vol. 118, no. 1, pp. 214-218, 2017.
- [13] F. Tarlochan, F. Samer, A.M.S. Hamouda, S. Ramesh and K. Khalid, “Design of thin wall structures for energy absorption applications: Enhancement of crashworthiness due to axial and oblique impact forces,” *Thin-Walled Structures*, vol. 71, pp. 7-17, 2013.
- [14] A.A.A. Alghamdi, “Collapsible impact energy absorbers: An overview,” *Thin-Walled Structures*, vol. 39, pp. 189-213, 2001.
- [15] T. Wierzbicki, W. Abramowicz, “Collapsible impact energy absorbers: An overview,” *Thin-Walled Structures*, vol. 50, pp. 727-734, 1983.
- [16] S.A. Meguid, Y.P., “. Hou, Crush behaviour of foam-filled thin-walled conical frusta: analytical, numerical and experimental studies,” *Acta Mechanical*, vol. 227, pp. 3391–3406, 2016.
- [17] Y. Zhou, F. Lan and J. Chen, “Crashworthiness research on S-shaped front rails made of steel–aluminium hybrid materials,” *Thin-Walled Structures*, vol. 49, pp. 291-297 2011.
- [18] N. Tanlak, F.O. Sonmez and M. Senaltun, “Shape optimization of bumper beams under high-velocity impact loads,” *Engineering Structures*, vol. 95, pp. 49-60, 2015.
- [19] H. Ozera, Y. Canb and M. Yazıcıa, “Investigation of the Crash Boxes Light Weighting with Syntactic Foams by the Finite Element Analysis,” *Acta Phys. Pol. A*, vol. 132, pp. 734-737, 2017.
- [20] U. Ozsaraca, S. Isika, F. Varolb, “M.E. Unata, C. Ozdemira and S. Aslanlara, Investigation of Tensile Properties of Aluminium 6082-T6 Alloys Joined by Cold Metal Transfer Method by Using Different Working Time,” *Acta Phys. Pol. A*, vol. 132, pp. 705-707, 2017.
- [21] M. Urbanek and D. Blaszkiewicz, “FEM Based Improvement of CAD for Non-Conventional Railway Track,” *Acta Phys. Pol. A*, vol. 128, pp. 241-242, 2015.
- [22] I. K. Yilmazcoban, Omer Adanur, “Experimental Crash-Box Optimizations for The Frontal Impact Scenario of a Vehicle to Decrease The Shock Effects of Collisions,” *ICCESEN 2017 4th International Conference on Computational and Experimental Science and Engineering*, pp. 387, 2017.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: A Modeling Study on Surface Roughness of Spinneret Mold Sections Machined By WEDM

Authors: Erdoğan Kanca, Volkan Cem Taşkın, Ali Günen

Received: 2018-05-30 00:37:31

Revised: 2018-10-11 12:50:50

Accepted: 2018-10-23 09:27:12

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 85-93

How to cite

Erdoğan Kanca, Volkan Cem Taşkın, Ali Günen; (2019), A Modeling Study on Surface Roughness of Spinneret Mold Sections Machined By WEDM. Sakarya University Journal of Science, 23(1), 85-93, DOI: 10.16984/saufenbilder.428457

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/428457>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

A Modeling Study on Surface Roughness of Spinneret Mold Sections Machined By WEDM

Erdoğan Kanca^{*1} Volkan C. Taşkın¹, Ali Günen²

ABSTRACT

The Wire Electrical Discharge Machining (Wire EDM) is widely used in the cutting of Bulk Continuous Filament (BCF) spinneret molds. Because of the low surface roughness of the spinneret molds obtained by the Wire EDM method, it ensures that the polypropylene material has a steady flow, volume and cross-sectional shape. Since the yarn extruded from the sections on the BCF acquire a number of physical and visual properties surface roughness of these molds have a great of importance. In this context, a new model was developed to predict the surface quality of spinneret mold sections, AISI 431 martensitic stainless steel, machined using Wire EDM by Analysis of variance (ANOVA). Machining parameters such as voltage, current, pulse time and wire feed rate were used as independent input variables and surface roughness was used as dependent output parameters. Contribution of input variables into the output variable determined by means of analysis of variance. Developed mathematical model estimations have been found to be in good agreement with the measured ones. The parameters with the most effect on surface roughness are listed as voltage, current, pulse, and feed, respectively. It is predicted that the steel used in spinneret mold can be machined more economically and practically by using the empirical formula obtained from this study.

Keywords: Wire EDM, Surface Quality, ANOVA

1. INTRODUCTION

BCF (Bulk Continuous Filament) yarns are manufactured by means of extruding polymer through spinneret molds. Stability of yarn flow and yarn quality are determined directly by surface quality of spinneret mold sections. Holes with different sections are machined on spinnerets by means of wire electrical discharge machining (Wire EDM).

Wire EDM is based on electric discharges between wire electrode and workpiece body so that it is a non-contact machining process.

Consequently material removal is not dependent upon material hardness [1]. Wire EDM is the best choice for most of the machining operations because of its ability to produce intricate shapes with good dimensional accuracy and surface roughness.

Performance measures of Wire EDM processes are classified as material removal rate, dimensional accuracy and surface quality. Analytical and statistical methods are used to determine useful parameters for optimal machining performances [2].

* Corresponding Author

¹ Iskenderun Technical University/Mechanical Engineering Department, Hatay, Turkey

² Iskenderun Technical University/ Metallurgy and Materials Engineering Department, Hatay, Turkey

Discharge current, pulse duration and pulse frequency are found to be Most affecting factors material removal rate by Scott et al. on the other hand they found that wire speed, wire tension and dielectric flow rate have least effect on MRR [3]. An approach for determining parameters setting based on Taguchi design method and ANOVA proposed by Liao et al. it is conclude that MRR is influenced easily by feed rate and pulse on time [4]. Similar results have been reported by Rozenek et al and Huang and Liao [5], [6]. Hsue et al presented a systematic analysis for MRR in corner cutting by Wire EDM and they formulated the MRR of geometrical cutting [7].

Lots of research tried to improve dimensional accuracy of Wire EDM by using different approaches. Firouzabadi et al. have investigated errors of small radius convex and concave successive cutting (two roughing and one finishing). They have found that roughing is the most influential stage of Wire EDM cutting and in can be better by optimization of process parameters [8]. Sanchez et al. also concluded that errors produced by previous cuts must be considered during optimization of corner radius [9]. Chen et al. have achieved to reduce corner radius error 50% by optimization of control factors [10].

Surface roughness is a very important performance parameter for Wire EDM as well as other machining processes. Durairaj et al. have used multi

objective optimization technique grey relational theory to get optimum values of gap voltage, wire feed, pulse on time and, pulse off time for machining of Stainless Steel (SS304) to get minimum surface roughness and the results have been validated with experimental results [11]. In

Table 1. Chemical composition of (wt. %) AISI 431 martensitic stainless steel used in the experimental studies

	%C	%Si	%Mn	%Cr	%Ni
Standard	0.12-0.22	1.00 (max)	1.50 (max)	15-17	1.50-2.50
Chemical Analysis	0.172	0.274	0.482	15.4	2.11

another study, Pulse-on time, pulse-off time, peak current, spark gap voltage, wire feed rate, and wire tension have been selected as input variables of optimization of surface roughness of Inconel 718. Mathematical models have been developed by using surface response methodology and surface roughness was predicted with error less than 5% [12]. As a result of a study on Wire EDM of titanium alloy pulse off time has been determined as the most significant parameter on material removal rate, surface roughness and kerf with [13].

There are a vast amount of papers on effects of cutting parameters of Wire EDM on dimensional accuracy and surface quality. But, a little of them about Wire EDM machining of martensitic stainless steel, which is material of spinnerets. In this study effects of Wire EDM cutting parameters (voltage, current, pulse time and wire feed rate) on surface roughness of spinneret material, which is 1.4057 stainless steel, will be studied.

2. MATERIALS AND METHOD

2.1. Experimental

Because of high pressure and high temperature conditions, DIN X17CrNi16-2 (1.4057) martensitic stainless steel (AISI 431) is used for spinneret mold applications. Blocks with 35x10x5 mm dimensions made from AISI 431 was prepared for Wire EDM cutting operations. Required and chemical composition of the material according to the standard and chemical analysis of the used specimen are given in Table 1. The used cutting pattern for this experiment is shown in Figure 1.

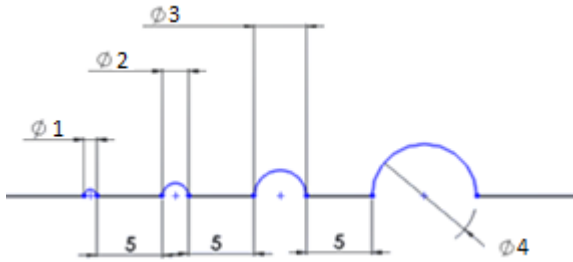


Figure 1. Cutting pattern used for experiments

During experimental cutting copper plated brass wire with 0.25 mm thickness used on SPM EZ20S EDM machine tool. Used cutting parameters during experiments are listed in Table 2. The used specimens marked by cold stamping is shown in Figure 2.

Table 2. Cutting parameters

Parameters	Used Values
Voltage (V)	38, 44, 50
Current (A)	7, 9, 11
Pulse On Time (μ s)	1, 2, 3
Feed Rate (m/min)	3, 4, 5

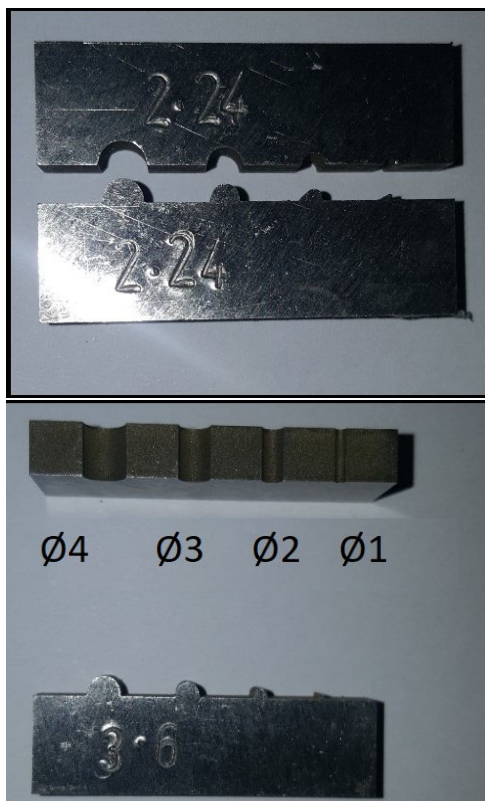
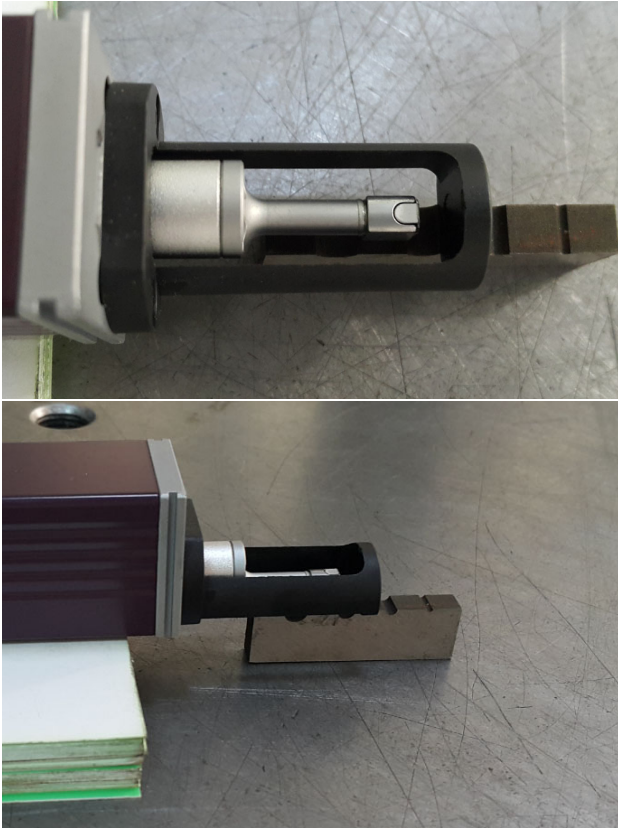


Figure 2. Marked and cut sample specimens

Surface roughness values were measured with respect to JIS 01- 0.25x5 standard as shown in Figure 3. Three measurements were applied for each experiment and mean values were recorded.

2.2. ANOVA

During ANOVA analysis the total sum of squares was calculated by adding the sum of the squares of residual random error into the sum of sum of squares of individual factors. Corresponding sum of squares of the factors were divided by associated degrees freedom to calculate mean squares of the factors. Then, null hypothesis was tested for individual factors to evaluate the effect or significance at a particular probability level of them. For this, the ratio of mean squares of factors to the mean squares of the residual error, i.e. F-statistic, was calculated and compared to the tabulated F-values related to Fisher distribution. Number of degrees of freedom of the individual factors, number of degrees of freedom of residual error and the probability level affects the F-values [14]. Degree of contribution (ρ %) of each significant factor in model was also determined according to computed value of F distribution. The ratio of F-value of each factor to the sum of computed F values give the degree of contribution of each significant factor. The $q\%$ values in Table 3. Define the degree of contribution of each independent factor to the measured dependent parameter.



The test results listed in Table 4 were analyzed to find out the variation in the surface roughness depending on the cutting parameters i.e. voltage, current, pulse and feed rate. Analysis of variances (ANOVA) have been performed by using a commercial statistical software (Design-Expert 7.0.3). Cubic regression model was determined as the best fitted among others by comparing estimations with measured values.

Figure 3. Measurement of surface roughness

Table 3. Results of ANOVA

Independent Parameters	Degree of Freedom	Sum of Squares	Mean Square	F Value	P value	$\rho\%$ (% effect on model)
<i>Model</i>	16	2.53				
<i>A-Voltage</i>	1	0.66	0.66	31.6	< 0.0001	25.8
<i>B-Current</i>	1	0.49	0.49	23.4	0.0003	19.1
<i>C-Pulse</i>	1	0.11	0.11	5.25	0.0394	4.3
<i>D-Feed</i>	1	1.57E-03	1.57E-03	0.075	0.7886	0.1
<i>AB</i>	1	0.062	0.062	2.97	0.1082	2.4
<i>AC</i>	1	0.05	0.05	2.4	0.1455	2.0
<i>AD</i>	1	0.17	0.17	8.01	0.0142	6.5
<i>BC</i>	1	0.011	0.011	0.54	0.4767	0.4
<i>BD</i>	1	9.03E-05	9.03E-05	4.31E-03	0.9486	0.0
<i>CD</i>	1	1.02E-03	1.02E-03	0.049	0.8284	0.0
<i>A²</i>	1	0.14	0.14	6.9	0.0209	5.6
<i>BCD</i>	1	0.25	0.25	11.73	0.0045	9.6
<i>A²C</i>	1	0.11	0.11	5.27	0.0389	4.3
<i>A²D</i>	1	0.016	0.016	0.75	0.4007	0.6
<i>AB²</i>	1	0.47	0.47	22.51	0.0004	18.3

Table 4. Predicted and measured surface roughness values with respect to independent variables

Voltage (V)	Current (Ω)	Pulse On Time (μ s)	Feed Rate (m/min)	Surface Roughness		% Error
				Predicted (μ m)	Measured (μ m)	
44	9	2	4	3.035	3.047	0,48
44	9	2	3	3.063	3.013	-1,65
44	9	2	5	3.007	2.957	-1.70
44	9	3	4	2.687	2.687	0.00
44	7	2	4	2.783	2.797	0.049
44	11	2	4	3.112	3.110	-0.06
38	9	2	4	3.404	3.410	0.18
38	7	1	3	2.878	2.987	3.64
38	7	1	5	2.665	2.617	-1.85
38	7	3	3	2.966	2.817	-5.30
38	7	3	5	2.290	2.370	3.38
38	11	1	3	3.282	3.123	-5.08
38	11	1	5	2.565	3.187	19.51
38	11	3	3	2.981	3.177	6.16
38	11	3	5	2.791	2.660	-4.93
50	9	2	4	2.254	2.260	0.27
50	7	1	3	2.316	2.203	-5.11
50	7	1	5	2.513	2.557	1.71
50	7	3	3	2.628	2.773	5.24
50	7	3	5	2.361	2.277	-3.70
50	11	1	3	2.970	3.130	5.11
50	11	1	5	2.662	2.567	-3.71
50	11	3	3	2.892	2.697	-7.24
50	11	3	5	3.112	3.243	-4.05

3. RESULTS AND DISCUSSIONS

A statistical analysis was performed to determine the statistically significant dependent parameters and interactions of them on roughness.

Results of ANOVA are given in Table 3. The F value in the table provides an information of the degree of contribution of the independent parameters to the measured dependent parameter (roughness). If the F is high, the contribution of the factors to that particular response is high. P

values are related with the significance of independent parameters on the dependent parameter. P values smaller than 0.05 means that related parameter is statistically significant on result.

A regression model in reduced cubic polynomial form is built as a result of ANOVA. The coefficients of terms of the model equation are listed in Table 5.

Table 5. Regression coefficients of model equation

Coefficient	Factor
-98.512	Constant
2.83989	* Voltage
17.70618	* Current
20.58278	* Pulse
-4.73774	* Feed
-0.38086	* Voltage * Current
-0.86624	* Voltage * Pulse
+0.24745	* Voltage * Feed
-0.23450	* Current * Pulse
-0.12506	* Current * Feed
-0.54944	* Pulse * Feed
-0.015138	* Voltage ²
-0.96548	* Current ²
+0.061938	* Current * Pulse * Feed
9.95E-03	* Voltage ² * Pulse
-2,62E-03	* Voltage ² * Feed
+0.021448	* Voltage * Current ²

Evaluation of regression model by comparing predictions versus measured values is given by Figure 4. and Table 4. Figure 4 depicts actual values of the experiment results versus the predicted ones. The points on the graph shows a uniform distribution in a region close to the 45° line which represents the perfect fit. In addition actual and corresponding predicted values of

surface roughness values and % error are listed in Table 4. Maximum absolute % error in this table is 19.51 % and the rest of the errors are less than 7 %.

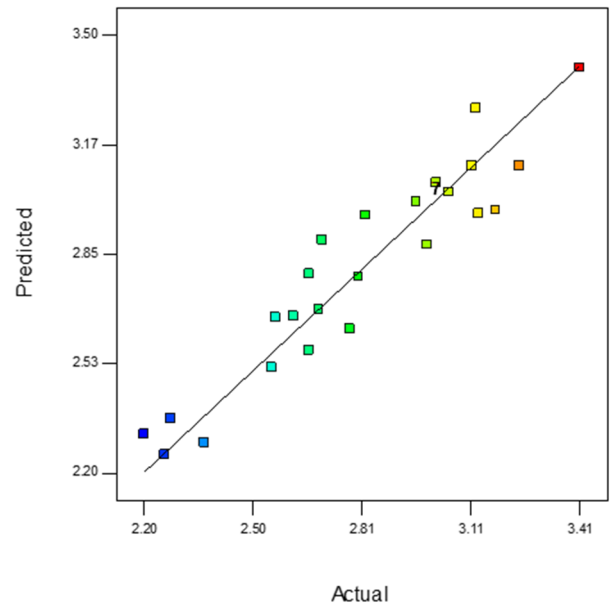


Figure 4. Evaluation of model estimations

It is seen that from the Table 3. the most significant variables are voltage, current and pulse on time. In addition interactions of factors i.e. multiplication of them, which have significant effect on the result (surface roughness value) are, voltage*pulse on time, voltage², current*pulse on time* feed, voltage²*pulse on time, voltage*current². The findings revealed by ANOVA are compatible with literature [2], [15], [16]. Developed model have four basic variables so that it can be represented by only a surface in a five dimensional hyperspace. Consequently the model depicted in three dimensional space by keeping constant two of the variables and constructing the model graph in three dimensions or contour graphs for changing values of remaining two variables. Three dimensional graphs of 6 combinations the variables are presented in Figures 5 to 10. It must be noted that the graphs are valid just for stated values of the constant factors. The graphs are varied for the changing values the constant factors.

Effects of voltage and current on surface roughness for constant values of pulse on time at

2 μ s and feed rate at 4 m/min is depicted in Figure 5. Increase in current have a positive effect up to about 10 A. After this value it decreases slightly. On the other hand surface roughness decreases with increase of voltage.

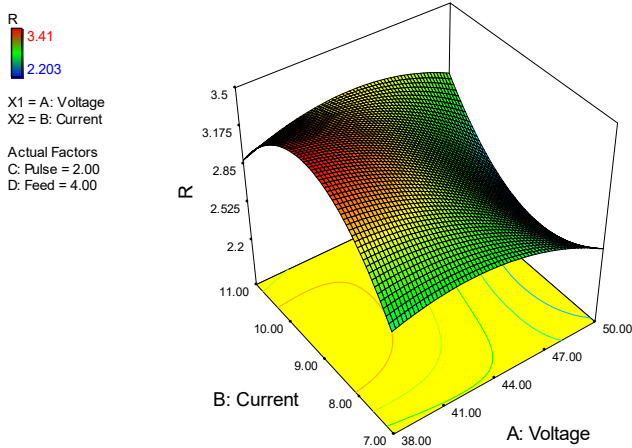


Figure 5. 3d plot of change of surface roughness with current and voltage for constant values of pulse on time =2 μ s and feed = 4 m/min.

Pulse on time have almost no effect on surface roughness for constant values of current at 9 A and feed rate at 4 m/min as seen in Figure 6. Increase in voltage causes decrease of roughness values in similar manner with Figure 5 but in this case the rise is quite steep.

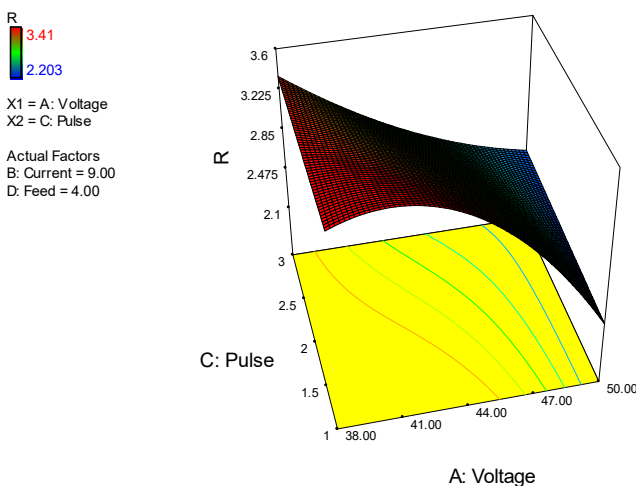


Figure 6. 3d plot of change of surface roughness with pulse on time and voltage for constant values of current = 9 A and feed = 4 m/min.

Change in surface roughness voltage and feed rate for the constant values at current value of 9 A and pulse on time of 2 μ s is shown in Figure 7. As in Figures 5 and 6. roughness have sharp rise with increase in voltage. Increase in feed rate have a negative effect on roughness for lower values of voltage although it have almost no effect on roughness for higher voltages.

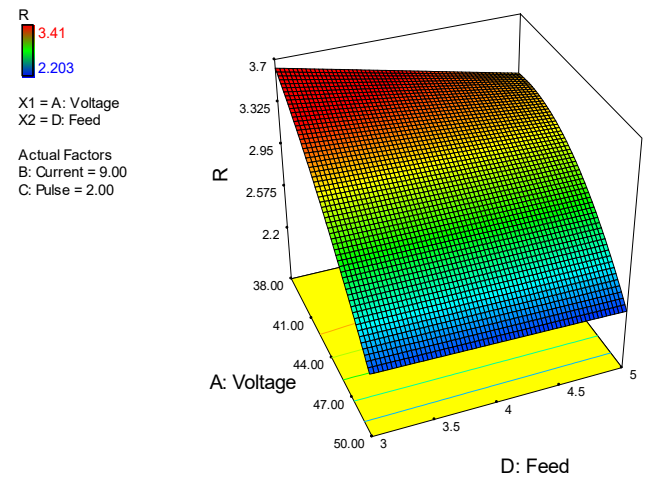


Figure 7. 3d plot of change of surface roughness with feed and voltage for constant values of current =9 A and pulse on time = 2 μ s.

Figure 8. shows change in surface roughness with respect to current and feed rate for voltage of 44 V and pulse on time of 2 μ s. As being compatible with Figure 5. roughness increases with increase in current. Similar with Figure 7. roughness values decreases slightly with increase in feed rate.

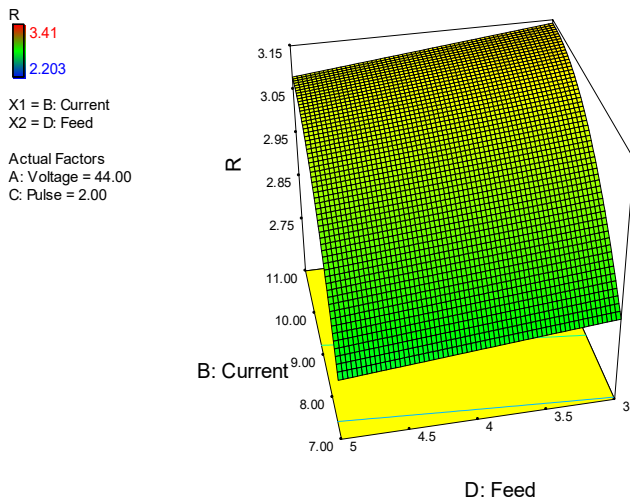


Figure 8. 3d plot of change of surface roughness with feed and current for constant values of voltage = 44 V and pulse on time = 2 μ s.

Effect of pulse on time and feed rate for constant values of voltage at 44 V and current at 9 A on surface roughness is depicted in Figure 9. Roughness increases with increase in feed rate. On the other hand roughness is increasing very slightly with decrease in feed rate.

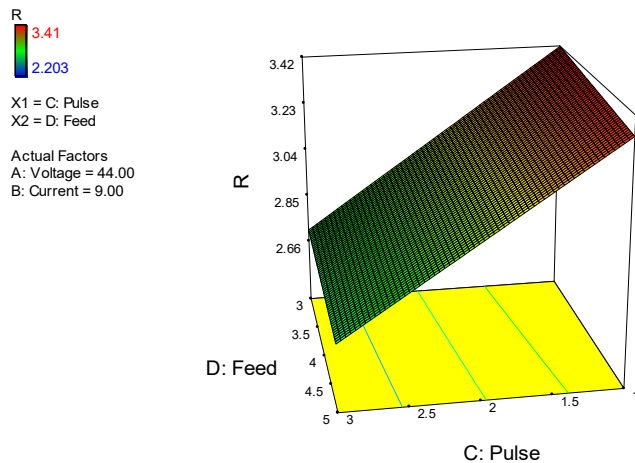


Figure 9. 3d plot of change of surface roughness with feed and pulse on time for constant values of voltage = 44 V and current = 9 A.

Change in surface roughness with respect to pulse on time and current for constant values of voltage of 44 V and feed rate of 4 m/min is shown in Figure 10. Current influences the roughness positively however pulse on time influence it negatively.

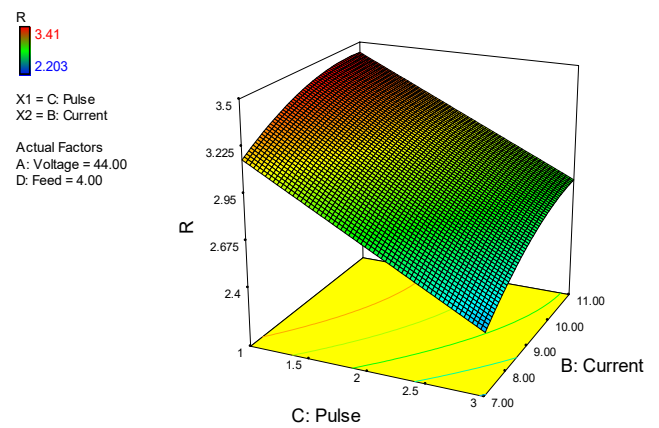


Figure 10. 3d plot of change of surface roughness with pulse on time and current on time for constant values of voltage= 44 V and feed = 4 m/min

4. CONCLUSIONS

A mathematical model to estimate surface roughness of Wire EDM cut spinneret material (AISI 431 martensitic stainless steel) have been developed during this study. Voltage, current, pulse on time and feed rate have been chosen as input variables. Statistical regression analysis have been conducted to get contribution of input variables and their products into the output parameter (surface roughness).

As a result of the study: A mathematical model in a form of cubic polynomial developed to predict surface roughness as a function of voltage, current, pulse on time and feed rate. Developed model predicts the surface roughness with maximum relative error of 19.51%. Voltage (25.8%) and current (19.1%) have been determined as most effective factors on surface roughness. Other effective parameters are listed as interaction of current pulse feed (9.6%), interaction of voltage feed (6.5%) and pulse (4.3%), parameters respectively.

REFERENCES

- [1] S. Kalpakjian and S. Schmid, Manufacturing, Engineering & Technology, 5th ed. Lebanon, Indiana, USA: Prentice Hall, 2005.
- [2] K. H. Ho, S. T. Newman, S. Rahimifard, and R. D. Allen, "State of the art in wire

- electrical discharge machining (Wire EDM),” *Int. J. Mach. Tools Manuf.*, vol. 44, pp. 1247–1259, 2004.
- [3] D. SCOTT, S. BOYINA, and K. P. RAJURKAR, “Analysis and optimization of parameter combinations in wire electrical discharge machining,” *Int. J. Prod. Res.*, vol. 29, pp. 2189–2207, Nov. 1991.
- [4] Y.-S. Liao, J. T. Huang, and H. C. Su, “A study on the machining-parameters optimization of wire electrical discharge machining,” *J. Mater. Process. Technol.*, vol. 71, pp. 487–493, Nov. 1997.
- [5] M. Rozenek, J. Kozak, L. Dąbrowski, and K. Łubkowski, “Electrical discharge machining characteristics of metal matrix composites,” *J. Mater. Process. Technol.*, vol. 109, pp. 367–370, Feb. 2001.
- [6] H. J.T and LiaoYunn-Shiuan, “Optimization of machining parameters of Wire EDM based on Grey relational and statistical analyses,” *Int. J. Prod. Res.*, vol. 41, pp. 1707–1720, May 2003.
- [7] W. J. Hsue, Y. S. Liao, and S. S. Lu, “Fundamental geometry analysis of Wire electrical discharge machining in corner cutting,” *Int. J. Mach. Tools Manuf.*, vol. 39, no. 1, pp. 651–667, 1999.
- [8] J. Abyar Firouzabadi, H., Parvizian and A. Abdullah, “Improving Accuracy of Curved Corners in Wire EDM Successive Cutting,” *Int J Adv Manuf Technol*, vol. 76, pp. 447–459, 2015.
- [9] J. A. Sanchez, J. L. Rodil, A. Herrero, L. N. L. De Lacalle, and A. Lamikiz, “On the influence of cutting speed limitation on the accuracy of Wire -EDM corner-cutting,” *Int. J. Adv. Manuf. Technol*, vol. 182, pp. 574–579, 2007.
- [10] Z. Chen, Y. Huang, Z. Zhang, and H. Li, “An analysis and optimization of the geometrical inaccuracy in Wire EDM rough corner cutting,” *Int. J. Adv. Manuf. Technol*, vol. 74, pp. 917–929, 2014.
- [11] M. Durairaj, D. Sudharsun, and N. Swamynathan, “Analysis of Process Parameters in Wire EDM with Stainless Steel Using Single Objective Taguchi Method and Multi Objective Grey Relational Grade,” *Procedia Eng.*, vol. 64, pp. 868–877, Dec. 2013.
- [12] V. Aggarwal, S. Sehijpal, and R. Garg, Parametric modeling and optimization for Wire electrical discharge machining of Inconel 718 using response surface methodology, vol. 79. 2015.
- [13] A. Ramamurthya, R. Sivaramakrishnan, and T. Muthuramalingamc, “Taguchi-Grey computation methodology for optimum multiple performance measures on machining titanium alloy in Wire EDM process,” *Indian J. Eng. Mater. Sci.*, vol. 22, pp. 181–186, Apr. 2015.
- [14] A. Rutherford, *Introducing ANOVA and ANCOVA: a GLM approach*. Sage Publications, 2001.
- [15] D. Ghodsiyeh, A. Golshan, and J. A. Shirvanhdeh, Review on current research trends in Wire electrical discharge machining (Wire EDM), vol. 6. 2013.
- [16] M. Y. A. and A. B. and M. A. Bakar, “Influence of Wire Electrical Discharge Machining (Wire EDM) process parameters on surface roughness,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 290, no. 1, p. 12019, 2018.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Secrecy Outage Probability of Modified TAS/Alamouti-STBC Schemes under Pilot Contamination Attacks

Authors: Ahmet Faruk Coşkun

Received: 2018-08-13 22:38:50

Revised: 2018-09-13 11:26:22

Accepted: 2018-10-24 17:35:44

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 94-105

How to cite

Ahmet Faruk Coşkun; (2019), Secrecy Outage Probability of Modified TAS/Alamouti-STBC Schemes under Pilot Contamination Attacks. Sakarya University Journal of Science, 23(1), 94-105, DOI: 10.16984/saufenbilder.453314

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/453314>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Secrecy Outage Probability of Modified TAS/Alamouti-STBC Schemes under Pilot Contamination Attacks

Ahmet Faruk COŞKUN*¹

ABSTRACT

Modified transmit antenna selection (M-TAS)/Alamouti orthogonal space-time block coding (STBC) schemes have been shown to achieve superior error performance together with a reduced-rate feedback channel in the presence of feedback imperfections when compared to the conventional TAS/Alamouti-STBC schemes. By shifting the focus of the investigation, this paper answers the query on whether the modified schemes provide enhancements in also the secrecy outage probability (SOP) performance of multi-antenna schemes in wiretap channels. The exact expressions of the SOP for the M-TAS/Alamouti-STBC schemes in Rayleigh fading channels have been derived and validated via Monte Carlo simulations. Additionally, the deteriorating effect of active eavesdropping has been demonstrated by simulating the pilot contamination attacks (PCAs) on the feedback channel. The extensive investigation and comparisons to the conventional schemes have shown that M-TAS/Alamouti-STBC schemes employed at the transmission end of the legitimate link provide considerable enhancements in the secrecy rates of wireless communications systems in the presence of PCAs.

Keywords: Physical layer security, TAS, Alamouti-STBC, pilot contamination attack

1. INTRODUCTION

Due to the broadcasting nature of wireless communications systems, the information content of the legitimate transmitter - receiver link might be intercepted by eavesdroppers in the same network. Accordingly, communications security is becoming a more challenging design issue to be resolved especially as the state-of-the-art and forthcoming wireless communications standards encourage to successfully operate in heterogenous networks each serving massive number of user terminals. Traditional measures that promise secure communications mainly

employ cryptographic protocols to be implemented in the network layer [1]. Nevertheless, the involved secret-key distribution and management processes might be unaffordable and extremely fragile to attacks [2]. Hence, by exploiting the characteristics of wireless channels, physical layer security concept is becoming more prominent for the purpose of providing secure data transmission between the transmitter and legitimate user(s) in wiretap channels.

The pioneering researches that have focused on communications security by physical-layer means have examined the single-input single-

* Corresponding Author

¹ The Scientific and Technological Research Council of Turkey, Kocaeli, Turkey. E-mail: ahmet.coskun@tubitak.gov.tr

output (SISO) wiretap scenario constituted by transmitter, receiver and eavesdropper terminals with single transceiver [1]-[5]. Afterwards, as the multi-antenna transmitter and/or receiver designs have attracted utmost attention for the sake of their spectrally-efficient and fading-resistant characteristics in rich multi-path scattering environments, enhanced-security physical layer designs consolidated by multi-input multi-output schemes have come up [5]-[20]. The investigations prosecuted within these studies have examined the communications security concept of several well-known transmit and/or receive diversity schemes from the perspective of secrecy outage probability (SOP) in the presence of single- or multi-antenna eavesdroppers. SOP which might be briefly defined as the probability that the secrecy capacity is less than a specific transmission rate, constitutes a useful metric for the researchers to assess the communications secrecy. The average SOP achieved by a single-antenna transmitter and a maximal-ratio-combining (MRC) receiver at the legitimate user in the presence of an eavesdropper with also MRC-receiver has been examined in [8] for Rayleigh-distributed channel amplitudes. The study in [9] has focused on the wiretap scenario constituted by single-antenna transmitter and receiver ends in the legitimate link and a multi-antenna eavesdropper that is considered to employ MRC or selection combining (SC) in Rician fading environments. In addition to receive-diversity-based schemes employed at the legitimate user and eavesdropper users, the usage of transmit-diversity based schemes consisting of or including transmit beamforming (TBF), space-time block coding (STBC) and transmit antenna selection (TAS) have also attracted the researchers' attention. The communications security has been shown in [10]-[13] to be enhanced efficiently with the help of TBF. Despite being probably the most efficient means of transmit diversity scheme that enables secure communications, TBF suffers from its dependency on the precise channel state information (CSI) of the legitimate and the wiretap link at the transmitter. The high feedback

burden and computational complexity faced by TBF schemes especially in case of increased number of transmit radio-frequency (RF) chains might be seen as unaffordable from the implementation perspective.

As an efficient way to achieve enhanced transmit diversity orders and reliable communications with reduced feedback requirements and hardware complexity, TAS-based transmission strategies have been of interest [14]-[20]. The simplest TAS scheme that switches the single transmit antenna maximizing the received SNR at the legitimate user has been investigated for MRC receivers [15]-[18] and generalized SC receivers [19] at the legitimate user and eavesdropper users in Rayleigh [14], [16]-[19] and non-identically distributed Nakagami- m [15] fading environments. The examination in [16] has introduced the average SOP performance of TAS/MRC scheme in the presence of an MRC-enhanced eavesdropper by also taking the practical conditions of time-delayed feedback (TDF) and binary symmetric channel-based feedback errors (FEs) into account. The authors of [17] have also concentrated on the same scheme in TDF conditions, and have proposed a modified TAS/MRC scheme that promises to achieve enhanced communications security in the presence of feedback delays. Relying on the practical assumption that the CSI of eavesdropper channels is difficult to be available at the transmitter especially if the eavesdropper is not a registered user of the same network or is a dedicated wiretapper, the antenna selection criterion employed within the studies [15]-[17], [19] and [20] is based on maximizing the instantaneous combined SNR at the legitimate receiver.

The conventional form of the combined TAS/STBC scheme has only been examined in [20] for Rayleigh fading environments from the communications security perspective. Here, under the assumptions of delayless and error-free feedback between the legitimate user and transmitter terminals, the authors have provided an extensive analysis on the average secrecy

performances achieved in the passive wiretapping scenario where the transmitter employs combined TAS with dual-branch Alamouti-STBC, and the legitimate user and eavesdropper users employ MRC. By providing comparisons between the SOP performances of single TAS and the conventional² TAS/Alamouti-STBC (C-TAS/Alamouti-STBC) schemes, [20] has demonstrated the superiority of the combined scheme when compared to the single-branch TAS.

Being motivated by the lack of an investigation on the effects of active eavesdropping (e.g., PCAs) on the secrecy performance of C-TAS/Alamouti-STBC schemes and the potential enhancements offered by the modified TAS/Alamouti-STBC (M-TAS/Alamouti-STBC) schemes under practical imperfections (as shown in [21] for imperfect feedback case), this paper focuses on the SOP performances of both schemes in the presence of PCAs. By focusing on the practical scenario where the transmitter does not have any CSI of the eavesdropper's channel, this paper analyzes the usage of the conventional and M-TAS/Alamouti-STBC schemes at the transmitter end of a legitimate wireless link in the presence of PCAs in flat Rayleigh fading channels, and makes the following specific contributions:

- The exact SOP expression for M-TAS/Alamouti-STBC schemes,
- the exhibition of the advantages such as the average SNR required to achieve a specific level of communications secrecy that are obtained by employing M-TAS/Alamouti-STBC schemes instead of the conventional ones,
- effects of active eavesdroppers that perform pilot contamination attack (PCA) [5] to the

feedback channel in the legitimate (i.e., main) link,

- a useful perspective for the design of multi-antenna diversity schemes that are more robust to active eavesdropping techniques, and that provide enhancements in the average communications secrecy.

2. SYSTEM MODEL AND SNR STATISTICS

This paper focuses on the average secrecy performances of the multi-antenna diversity schemes that are constructed by performing Alamouti-STBC signaling combined with TAS at the transmission sessions and MRC scheme at the reception sessions. As sketched in Figure 1-(a), the investigated diversity schemes employ conventional and modified versions of the combined TAS/Alamouti-STBC with two active RF chains among n_A total antennas of the transmitter (namely Alice). And at the receive ends of both the legitimate and the eavesdropping users (namely Bob and Eve, respectively), optimal receive-combining (i.e., MRC) is employed through n_B and n_E antennas, respectively.

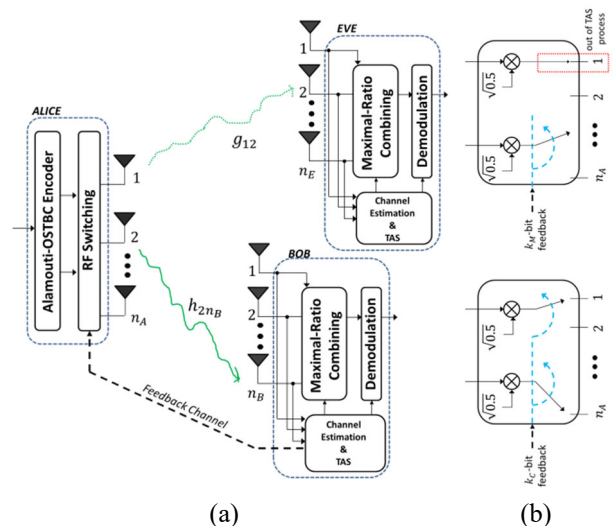


Figure 1 (a) Schematic representation of the TAS/Alamouti-STBC techniques in the presence of an

² Here, the term conventional is used to distinguish the straightforward but optimal antenna-subset-

selecting scheme from the modified versions that have been previously introduced in [21].

eavesdropper (b) Detailed view of RF switching block for the conventional (at bottom) and modified (at top) schemes

For clarity, we first recall the system descriptions of both the C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC schemes. As shown in Figure 1-(a), both schemes occupy a feedback channel that is dedicated to convey the antenna subset index to Alice and to assist the transmission process through the best transmit-receive branches between Bob and Alice.

2.1 Conventional TAS/Alamouti-STBC

Since modified versions of combined TAS/STBC scheme have been proposed and analyzed in [21] in the presence of feedback imperfections, the scheme with the most straightforward approach to perform transmission over an optimal subset has been labeled as conventional. Note that, the antenna subset selection strategy has been shown to provide the optimal solution with the highest total channel gain in ideal feedback and antenna switching conditions, whereas it is not the case in practical conditions as shown in [21].

Without regarding the feedback burden of high-complexity brute-force subset selection and its complications on a practical feedback link, the C-TAS/Alamouti-STBC scheme performs dual-branch STBC transmission after selecting the antenna subset that consists of the transmit antennas with the ordinal numbers $o_{C,1}$ and $o_{C,2}$ among n_A total transmit antennas as seen in Figure 1-(a). Hence, the ordered antenna indices $(o_{C,1}, o_{C,2})$, where $o_{C,j} \in \{1, 2, \dots, n_A\}$, $j = 1, 2$, comprise only a subset among $K_C = \binom{n_A}{2}$ total combinations, and denote the orders of the transmit antennas selected (perfectly or erroneously). It would be clear that in the case of ideal subset selection (i.e., no PCAs or other imperfections), the ordinal numbers would be $o_{C,1} = 1$ and $o_{C,2} = 2$. Besides, this type of subset selection is shown to result in a feedback-bit load of $k_C = \lceil \log_2 K_C \rceil$ in [22] since the feedback information should only convey the subset index corresponding to two transmit antenna indices.

Here, $\lceil \cdot \rceil$ denotes the smallest integer that is greater than or equal to its argument.

2.2 Modified TAS/Alamouti-STBC

Modified TAS/STBC (M-TAS/STBC) schemes are constructed by modifying the antenna selection/switching strategy at the transmitter ends of communication links, and are proposed to reduce the average feedback-bit requirement and to enhance the average error probability in the presence of practical feedback impairments. They are shown to efficiently reduce the feedback requirement of the C-TAS/STBC and to provide average SNR savings when compared to C-TAS/STBC schemes in the presence of FEs [22].

As seen in Figure 1-(b), the modified scheme with the single antenna selection allocates the first transmit antenna for transmission regardless from the antenna selection process, and switches the second transmit antenna by determining the best antenna with the highest total channel gain among the remaining $n_A - 1$ ones. Since the size of antenna set that is subject to the single-selection is shrunked to $K_M = n_A - 1$ (when compared to the conventional case), the feedback-bit load required to switch the selected antenna is reduced to $k_M = \lceil \log_2 K_M \rceil$ per each Alamouti-STBC transmission period.

The Alamouti-STBC mapper at Alice inserts the digitally-modulated message symbols x_1 and x_2 into the first row of the Alamouti transmission matrix that had been denoted by \mathcal{G}_2 and defined as

$$\mathcal{G}_2 = \begin{pmatrix} x_1 & -x_2^* \\ x_2 & x_{21}^* \end{pmatrix} \quad (1)$$

as also provided in Eqs. (3) and (4) of [21]. Each row of the matrix consists of the symbols that are transmitted from each antenna while columns correspond to transmission time slots. Hence, the symbols x_1 and $-x_2^*$ will be conveyed through the first transmit antenna (i.e., dedicated regardless from the TAS process). After the selection of the single transmit antenna that had been determined to maximize the output SNR at the legitimate

receiver, the signal transmission of the second row of the Alamouti STBC codeword matrix is directed to the selected (single) transmit antenna.

2.3 SNR Statistics of M-TAS/STBC Schemes

The envelopes of the complex Gaussian-distributed fading coefficients $h_{j,i}$, $i = 1, 2, \dots, n_B$, $j = 1, 2, \dots, n_A$ defined through the transmitter and receiver antennas of the legitimate user, and $g_{j,i}$, $i = 1, 2, \dots, n_E$ defined through the transmitter and receiver antennas of the eavesdropper are assumed to be i.i.d. Rayleigh fading variables with the squared mean of unit magnitude $\Omega = E[|h_{j,i}|^2] = E[|g_{j,i}|^2] = 1$, $\forall i, j$. Here, $E[\cdot]$ denotes the expectation operator. Under the assumption that the channel estimation processes are perfectly done at Bob and Eve, the combined instantaneous SNR at each receiver output of Bob and Eve through the j^{th} transmit antenna would be obtained as: $\gamma_{B,j} = \bar{\gamma}_B \sum_{i=1}^{n_B} |h_{j,i}|^2$ and $\gamma_{E,j} = \bar{\gamma}_E \sum_{i=1}^{n_E} |g_{j,i}|^2$ where the subscripts B and E stand for the main link (i.e., the link between Alice and Bob) and the wiretap link (i.e., the link between Alice and Eve), respectively. Besides, $\bar{\gamma}_{B/E} = \frac{E_s}{N_{o,B/E}}$ are the average branch SNRs defined in terms of the average energy per symbol (E_s) and the one-sided power spectral densities of the additive white Gaussian noise (AWGN) ($N_{o,B}$ and $N_{o,E}$) at each receiver antenna of Bob and Eve, respectively. The instantaneous SNR $\gamma_{B,j}$ and $\gamma_{E,j}$ will follow Gamma distribution with the PDF and CDF given as $f_{L,j}(x) = \frac{x^{n_L-1} e^{-x/\bar{\gamma}_L}}{\bar{\gamma}_L^{n_L} \Gamma(n_L)}$ and $F_{L,j}(x) = \frac{\psi(n_L, x/\bar{\gamma}_L)}{\Gamma(n_L)} = 1 - e^{-x/\bar{\gamma}_L} \sum_{n=0}^{n_L-1} \frac{[x/\bar{\gamma}_L]^n}{\Gamma(n+1)}$, $x \geq 0$, respectively where $\Gamma(\cdot)$ and $\psi(\cdot)$ denote the Gamma and the incomplete Gamma function, respectively, and the label L becomes B and E for the main and wiretap links.

2.3.1 SNR Statistics of the Legitimate Link

For multi-antenna diversity systems employing M-TAS/Alamouti-STBC scheme, the first transmit antenna is employed regardlessly as

seen at the top of Figure 1-(b), and the single transmit antenna which maximizes the received instantaneous SNR at Bob is selected among the remaining $K_M = n_A - 1$ antennas by comparing the instantaneous SNRs $\gamma_{B,j}$, $j = 1, 2, \dots, n_A$: $\gamma_{B,max} = \max_{j=1,2,\dots,n_A} \{\gamma_{B,j}\}$. After switching the m^{th} , $m \in \{1, 2, \dots, K_M\}$ best transmit antenna as a result of possible erroneous decoding, the resulted instantaneous post-processing SNR at the receiver output of the legitimate user will be

$$\gamma_{B,M}^{(m)} = \frac{1}{2} (\gamma_{B,(m)} + \gamma_{B,1}). \quad (2)$$

Here $\gamma_{B,(m)}$ denotes the m^{th} maximum instantaneous SNR among K_M transmit branch SNRs and $\gamma_{B,1}$ denotes the instantaneous combined SNR of the first transmit branch used regardlessly from the antenna selection process. As seen from Eq. (2), the total transmit power is equally distributed among two active transmit antennas. The statistics of the instantaneous combined SNR in Eq. (2) could be examined by evaluating the moment-generating function (MGF)

$$\mathcal{M}_{\gamma_{B,M}^{(m)}}(s; m) = E \left[e^{-\frac{s}{2} \gamma_{B,(m)}} \right] E \left[e^{-\frac{s}{2} \gamma_{B,1}} \right]. \quad (3)$$

Since the PDF expression related to the order statistics $\gamma_{B,(m)}$ has been defined in [21] as

$$f_{\gamma_{B,(m)}}(x; m) = \frac{\binom{n_A-2}{m-1} f_B(x) [F_B(x)]^{n_A-m-1}}{(n_A-1)^{-1} [1-F_B(x)]^{1-m}} \quad (4)$$

using the Laplace transform pair in [25, Eq. (2.2.1-2)], $\mathcal{M}_{\gamma_{B,(m)}}(s; m) = E \left[e^{-\frac{s}{2} \gamma_{B,(m)}} \right]$ is expressed as

$$\mathcal{M}_{\gamma_{B,(m)}}(s; m) = b_0 \sum_{n=0}^{n_A-m-1} b_1 \sum_{r=0}^{(n+m-1)(n_B-1)} \times \frac{\mu_{r,n+m-1,n_B}}{2^{r+n_B}} \frac{\Gamma(r+n_B)}{[s + 2(n+m)/\bar{\gamma}_B]^{r+n_B}} \quad (5)$$

Where $\mu_{r,n+m-1,n_B}$ is the multinomial coefficient that has been defined in [23, Eq. (0.314)], $b_0 = \frac{n_A-1}{\bar{\gamma}_B^{n_B} \Gamma(n_B)} \binom{n_A-2}{m-1}$, $b_1 = (-1)^n \binom{n_A-m-1}{n}$. By substituting Eq. (5) and the MGF corresponding

to the single-diversity-branch SNR $E \left[e^{-\frac{s}{2}\gamma_{B,1}} \right] = \left[1 + \frac{s\bar{\gamma}_B}{2} \right]^{-n_B}$ into Eq. (3), and applying inverse Laplace transform, the overall PDF of the combined instantaneous SNR for the M-TAS/STBC scheme could be obtained as:

$$f_{\gamma_{B,M}}(x; m) = b_0 \sum_n b_1 \sum_r b_2 w(x). \quad (6)$$

Here, the parameter is $b_2 = \frac{\Gamma(r+n_B)\mu_{r,n+m-1,n_B}}{\bar{\gamma}_B^{n_B} 2^{r+2n_B}}$ and the x -dependent function $w(x)$ is obtained by after partial fraction decomposition (PFD) prescribed in [23, Eq. (2.102)] and the inverse Laplace transform:

$$w(x) = \begin{cases} \frac{x^{v_1+v_2-1} e^{-x\kappa_1}}{\Gamma(v_1+v_2)}, & n=0, m=1 \\ \sum_{j=1}^2 \sum_{i=1}^{v_j} \frac{e_{ij} x^{i-1} e^{-x\kappa_j}}{\Gamma(i)}, & \text{otherwise.} \end{cases} \quad (7)$$

Here $v_j = n_B + r(j-1)$, $\kappa_1 = 2/\bar{\gamma}_B$, $\kappa_2 = \kappa_1(n+m)$ and e_{ij} denotes the PFD coefficients.

2.3.2 SNR Statistics of the Wiretap Link

The eavesdropper user, assumed to have perfect knowledge of the channel coefficients $g_{j,i}$, $j = 1, 2, \dots, n_A$, $i = 1, 2, \dots, n_E$, intercepts the information conveyed in the legitimate link (between Alice and Bob) by employing n_E -branch MRC. Since the wiretap channels can only make use of dual transmit branches that is accordingly switched due to the channel quality of the legitimate link, Eve could achieve no additional diversity gain provided by TAS but the STBC combining gain with the asymptotic diversity order of $2n_E$. The combined instantaneous SNR at the receiver output of Eve could be defined as

$$\gamma_E = \frac{1}{2} (\gamma_{E,t} + \gamma_{E,j \neq t}), t, j \in \{1, 2, \dots, n_A\} \quad (8)$$

as given in [18] and [20]. With the help of independence between branch SNRs, the MGF of Eq. (8) might be easily obtained as $\mathcal{M}_{\gamma_E}(s) = \left[1 + \frac{s\bar{\gamma}_E}{2} \right]^{-2n_E}$. By using the Laplace transform

pair given in [25, Eq. (2.1.2-71)], the PDF of the combined SNR at the receiver output of Eve could be expressed as

$$f_{\gamma_E}(x) = \frac{x^{2n_E-1} e^{-2x/\bar{\gamma}_E}}{[\bar{\gamma}_E/2]^{2n_E} \Gamma(2n_E)}, x \geq 0. \quad (9)$$

As inferred from the instantaneous combined SNR expression in Eq. (8) and its PDF in Eq. (9), the SNR statistics would have no dependency on TAS process, which later will be mentioned in Section V to provide advantages by means of communications secrecy.

3. FEEDBACK CHANNEL

For C-TAS/Alamouti-STBC technique, the feedback information between Bob and Alice conveys the single index corresponding to the antenna index vector that consists of the selected antenna indices. Since each combination of the antenna indices would correspond to a combination of ordered antenna indices (i.e., $\mathbf{o}_{C,m} = (o_{C,1}, o_{C,2})$), the index $m \in \{1, 2, \dots, K_C\}$, will also represent the index of the vector that consists of ordinals. For a transmitter configuration of $n_A = 3$, there would only be $K_C = \binom{3}{2} = 3$ total combinations of the ordered antenna index vector (i.e., $(o_{C,1}, o_{C,2}) = (1, 2), (1, 3)$ and $(2, 3)$), and the corresponding index of an ideally or erroneously activated antenna subset would be $m = 1, 2, 3$, respectively. The ordered index vector corresponding to the ideal antenna selection/switching case would be $\mathbf{o}_{C,1} = (1, 2)$ that is meant to use the first and second best transmit antennas for two-branch STBC. However, due to the FEs, the feedback message might be decoded to other antenna index vectors $\mathbf{o}_{C,m}$, $m \neq 1$: $\mathbf{o}_{C,2} = (1, 3)$ or $\mathbf{o}_{C,3} = (2, 3)$. For different values of n_A , the antenna index combinations and the corresponding index values (i.e., m) might be easily associated and employed via a look-up table-like mechanism. Whereas for M-TAS/Alamouti-STBC technique, the feedback information conveys simply the single index corresponding to the selected antenna: $m \in \{1, 2, \dots, K_M\}$.

The indices related to the selected transmit antenna subsets are fed back from Bob to Alice over an open (non-secure) and low-rate feedback link that is considered to have fading characteristics from a more realistic viewpoint [22]. Here, note that Eve might access the antenna indices determined by Bob, she will not be able to attain any diversity gains since she has no information about the main link's CSI and even though she had, the indices are tailored due to the main link. However, in the case of active eavesdropping, Eve might contaminate the feedback channel from Bob to Alice in order to manipulate the antenna switching procedure through her favor.

The k_c - and k_M -bit feedback informations are sent to Alice followed by \mathcal{L}_c - and \mathcal{L}_M -phase shift keying (PSK) modulation where $\mathcal{L}_c = 2^{k_c}$ and $\mathcal{L}_M = 2^{k_M}$. Due to degrading effects of the feedback channel with fading characteristics and the PCAs on the feedback channel, the index information might be decoded erroneously. For C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC cases, the antenna subset with an ideally or erroneously determined index vector $\mathbf{o}_{C,m}$ and the index m will be activated with the probabilities $p_m \triangleq \mathbf{P}\{\mathbf{o}_{C,m}\}$ and $p_m \triangleq \mathbf{P}\{m\}$ while the other antennas are kept silent. The a priori probabilities are defined as $p_m = \sum_{j=1}^{K_{C/M}} p(s_j) \mathbf{P}(s_j \rightarrow s_m)$, where s_j denotes the baseband symbol belonging to $\mathcal{L}_{C/M}$ -PSK modulation, $p(s_j)$ represents the occurrence probability of the symbol s_j and $\mathbf{P}(s_j \rightarrow s_m)$ is the pairwise error probability (PEP) related to the j^{th} and m^{th} symbols corresponding to the j^{th} and m^{th} best transmit antennas (or subsets). The feedback message, delivering $k_{C/M}$ bits in order to represent $K_{C/M}$ transmit antenna combinations, can be erroneously decoded to one of the $\mathcal{L}_{C/M} \geq K_{C/M}$ index values because of feedback imperfections and PCAs. This might cause the proper transmit antenna (subset) indices $1, 2, \dots, K_{C/M}$ to be de-mapped to any of the improper transmit antenna (subset) indices $K_{C/M} + 1, K_{C/M} + 2, \dots, \mathcal{L}_{C/M}$ (which can be called

feedback failure (FF)). Thus, assuming that each transmit antenna subset is selected equally-likely in the presence of FF as in [22], the a priori probabilities would be increased by the factor $\frac{1}{K_{C/M}} \sum_{j=K_{C/M}+1}^{\mathcal{L}_{C/M}} p_j$ yielding

$$p'_m = p_m + \frac{1}{K_{C/M}} \sum_{j=K_{C/M}+1}^{\mathcal{L}_{C/M}} p_j. \quad (10)$$

Hence, it would be possible to cover also the case of $\mathcal{L}_{C/M} > K_{C/M}$ rather than the limited case of $K_{C/M} = 2^b$, $b \in \mathbb{Z}^+$.

4. THE EFFECTS OF ACTIVE EAVESDROPPING

This section describes the effects of active eavesdropping on the average SOP values of TAS-aided diversity schemes (e.g., conventional and modified TAS/STBC, single TAS/MRC). In order to investigate the degradations on the communications secrecy, the eavesdropper is assumed to act not only to scrounge the information of the legitimate link but also to contaminate the TAS and antenna switching processes at Alice in order to manipulate the information flow from Alice in her favor. For this purpose, the active eavesdropping scheme based on pilot contamination attack (PCA) introduced in [5] has been involved. As could be found in Section 3 of [5], PCA might be applied by eavesdroppers in both in full-duplex and half-duplex communications modes. Hence, eavesdroppers would be able to contaminate the training and reverse training phases performed between Alice and Bob. Here, by assuming that the communications between Alice and Bob are half-duplex and Eve has synchronized to the timelines of the legitimate link, the training phase (i.e., run via the feedback transmission) between Alice and Bob could be misled by the intruding signal replica of Eve. Firstly, both Bob and Eve determine their optimal transmit antenna subsets individually. Then, during the feedback transmission phase (i.e., accomplished for the purpose of enhancing the communications

security between Bob and Alice via TAS), both Bob and Eve transmit their modulated symbols to the feedback decoder at Alice. Hence, because of the high-rate coherence between the feedback signals conveyed to Alice, the baseband symbols received at Alice could be expressed as in [5, Eq. (4)]:

$$y_{A,p} = \sqrt{P_{B,p}}h_{B,A}s_{m_B} + \sqrt{P_{E,p}}h_{E,A}s_{m_E} + \eta_A \quad (11)$$

Here, $\sqrt{P_{B,p}}$ and $\sqrt{P_{E,p}}$ are the average power values related to the signal transmissions over the feedback links Bob→Alice and Eve→Alice, respectively. Similarly, $h_{B,A}$ and $h_{E,A}$ are the complex Gaussian-distributed channel gains of both links, s_{m_B} and s_{m_E} denote the unit-magnitude (i.e., $|s_{m_B}| = |s_{m_E}| = 1$) modulated PSK symbols those correspond to the transmit antenna subset indices m_B and m_E determined by Bob and Eve, respectively, and η_A is the complex AWGN with zero mean and variance of $N_{o,A}$. Since the average SNR of the legitimate feedback link is defined as $\bar{\gamma}_{FC}^{(L)} \triangleq P_{B,p}/N_{o,A}$, the average SNR of the eavesdropping link would be similarly obtained as $\bar{\gamma}_{FC}^{(E)} \triangleq P_{E,p}/N_{o,A}$. Assuming that the channel estimation at Alice is perfect, Alice would attempt to decode the received PSK symbol in order to determine the selected transmit antenna subset index and to enhance the communications security. After defining the power ratio $\Gamma_{FC} \triangleq P_{E,p}/P_{B,p}$, it is intuitively obvious to state that the increasing values of Γ_{FC} would cause Eve to contaminate and dominate the feedback information. In the case of Eve's relatively increased feedback transmit power, Alice would decide on the transmit antenna indices in the favor of Eve.

5. SECRECY PROBABILITY PERFORMANCES

After having the SNR statistics of both the main link and the wiretap link derived, we have focused on the examination of M-TAS/STBC schemes from the communications security perspective. In order to examine the SOP performances of multi-antenna diversity schemes

that employ M-TAS/Alamouti-STBC techniques at the transmitter, we have focused on deriving the exact expressions for the PDF, CDF and average SOP. With the help of the derived performance metrics, it would be possible to examine the variation of the average SOP of the modified schemes due to the several system configurations and practical imperfections such as PCAs.

SOP is the probability that the secrecy capacity C_s is less than a specific transmission rate \mathcal{R}_0 (bits/channel-use) where C_s is expressed in [17, Eq. (6)] as

$$C_s = \begin{cases} \log_2 \left(\frac{1 + \gamma_B}{1 + \gamma_E} \right), & \gamma_B > \gamma_E, \\ 0, & \gamma_B \leq \gamma_E. \end{cases} \quad (12)$$

The SOP that is formulated as $P_{SOP}(\mathcal{R}_0) = \Pr(C_s < \mathcal{R}_0)$ has been shown in [17, Eq. (14)] to be evaluated as

$$P_{SOP}(\mathcal{R}_0; m) = 1 - \int_{y=0}^{\infty} \int_{x=\beta(y)}^{\infty} f_{\gamma_{B,M}}(x; m) f_{\gamma_E}(y) dx dy \quad (13)$$

where $\beta(y) = 2^{\mathcal{R}_0}y + 2^{\mathcal{R}_0} - 1$. Substituting the PDF expressions $f_{\gamma_{B,M}}(x; m)$ and $f_{\gamma_E}(y)$ given in Eq. (6) and Eq. (9) into Eq. (13) would result in

$$P_{SOP}(\mathcal{R}_0; m) = 1 - b_0 w_0 \sum_n b_1 \sum_r b_2 W(\mathcal{R}_0). \quad (14)$$

Here, $w_0 \triangleq [\bar{\gamma}_E/2]^{-2n_E}/\Gamma(2n_E)$ and $W(\mathcal{R}_0)$ is derived for the case of $n = 0, m = 1$ as

$$W^{(n=0, m=1)}(\mathcal{R}_0) = e^{-\kappa_1(2^{\mathcal{R}_0}-1)} \sum_{k=0}^{v_1+v_2-1} \frac{\kappa_1^{k-v_1-v_2}}{\Gamma(k+1)} \times \sum_{q=0}^k \binom{k}{q} \frac{2^{q\mathcal{R}_0} [2^{\mathcal{R}_0} - 1]^{k-q} \Gamma(q + 2n_E)}{[2^{\mathcal{R}_0}\kappa_1 + (1/\bar{\gamma}_E)]^{q+v_1+v_2}} \quad (15)$$

and for the case of $n \neq 0, m \neq 1$ as

$$W^{(n \neq 0, m \neq 1)}(\mathcal{R}_0) = \sum_{j=1}^2 \sum_{i=1}^{v_j} \sum_{k=0}^{v_j-1} \frac{e_{ij} e^{-\kappa_j(2^{\mathcal{R}_0}-1)} \kappa_j^{k-v_j}}{\Gamma(k+1)}$$

$$\times \sum_{q=0}^k \binom{k}{q} \frac{2^{q\mathcal{R}_0} [2^{\mathcal{R}_0} - 1]^{k-q} \Gamma(q + 2n_E)}{[2^{\mathcal{R}_0} \kappa_j + (1/\bar{\gamma}_E)]^{q+v_j}}. \quad (16)$$

Consequently, Eqs. (14)-(16) constitute the exact average SOP performance of the M-TAS/Alamouti-STBC scheme. Note that, the average SOP performance $P_{SOP}(\mathcal{R}_0; m)$ is only valid in the case of (ideally or erroneously) activating the m^{th} combination of transmit antenna subset $m \in \{1, 2, \dots, K_M\}$. Hence, the overall average SOP performances of the modified schemes would be obtained by simply weighting by the a priori probabilities p'_m defined by Eq. (10) and superposing the above-mentioned ordinal-dependent performance metrics over the entire possible transmit antenna indices and the subset combinations:

$$P_{SOP}(\mathcal{R}_0) = \sum_{m=1}^{K_M} p'_m P_{SOP}(\mathcal{R}_0; m). \quad (17)$$

6. NUMERICAL RESULTS

This section presents numerical results consisting of exact SOP performance results together with Monte Carlo simulations of diversity schemes that employ C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC under active eavesdropping effects. In order to examine the practical performances of these schemes and to highlight the robustness of the modified schemes, several performance results are provided in Figures 2 and 3 consisting of comparisons for different system and channel configurations. In both figures, the numerical results related to C-TAS/Alamouti-STBC, M-TAS/Alamouti-STBC and TAS/MRC schemes are depicted by black, blue and green curves respectively. Besides, by comparing the secrecy performances of M-TAS/STBC schemes to those of the conventional ones under same system and channel conditions, the enhancements gathered by the modified schemes have been exhibited. For this purpose, the SOP values of the C-TAS/STBC schemes under passive eavesdropping conditions (perfect feedback w/o PCAs) are evaluated by using the

theoretical expressions provided by the pioneering work in [20]. For benchmarking purposes, the average SOP performance results of single TAS/MRC scheme have been included in Figures 2 and 3. The SOP performances of both the conventional schemes in the presence of PCAs and all cases of single TAS/MRC scheme are evaluated via Monte Carlo simulations. Besides, the simple decoding procedure for the received feedback signal defined in Eq. (11) has been performed via Monte Carlo simulations that would result in the a priori probabilities given in Eq. (10).

With the aim of examining the variation in the average SOP performance of TAS/STBC and single TAS/MRC schemes in the presence of PCAs, we have focused on the scenario with the parameters $n_A = 3, n_B = 2, n_E = 1, \mathcal{R}_0 = 1$ bit/s/Hertz, $\bar{\gamma}_E = 5$ dB, $\bar{\gamma}_{FC}^{(L)} = 0$ dB. In Figure 2, the effects of active eavesdropping based on PCA are demonstrated in the presence of a noisy feedback channel (i.e., $N_{0,A} \neq 0$).

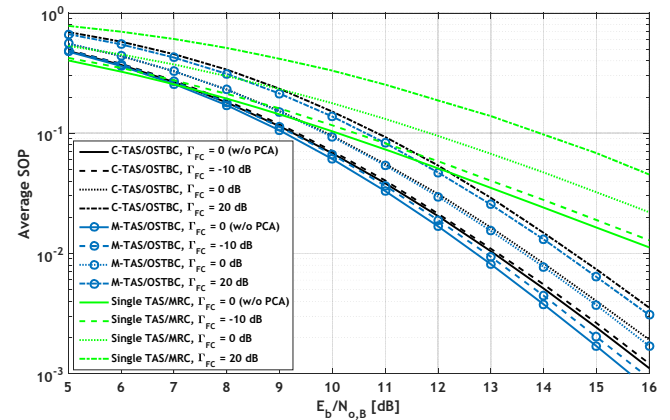


Figure 2 Average SOP of TAS/STBC and single TAS/MRC schemes for $n_A = 3, n_B = 2, n_E = 1, \mathcal{R}_0 = 1$ bit/s/Hertz, $\bar{\gamma}_E = 5$ dB, $\bar{\gamma}_{FC}^{(L)} = 0$ dB and under the effects of PCA ($\Gamma_{FC,dB} \in \{-\infty, -10, 0, 20\}$)

In Figure 2, the C-TAS/Alamouti-STBC, M-TAS/Alamouti-STBC and single TAS/MRC schemes are seen to face degradations with the increasing values of Γ_{FC} . When compared to the single TAS/MRC scheme which severely suffers from vanishing transmit diversity order (i.e., unity as the non-transmit-diversity SISO/SIMO

schemes) in the presence of practical imperfections such as PCAs, C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC schemes keep maintaining at least the transmit diversity order achievement of twice the SISO/SIMO schemes. Besides, the modified schemes are seen to achieve average SNR gains of 0.35 dB, 0.25 dB, 0.1 dB and 0.2 dB for $\Gamma_{FC,dB} \in \{-\infty, -10, 0, 20\}$, respectively at an average SOP of 10^{-2} when compared to the conventional schemes. This clearly shows that the modified schemes are more robust against PCAs when compared to the conventional (and also to single TAS/MRC) schemes.

Another examination on SOP performances of single TAS/MRC, C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC schemes has been made by considering a noise-free feedback channel (i.e., $N_{o,A} = 0$) and demonstrated in Figure 3 for $n_A = 4$, $n_B = 2$, $n_E = 2$, $\mathcal{R}_0 = 1$ bit/s/Hertz, $\bar{\gamma}_E = 5$ dB and $\Gamma_{FC,dB} \in \{-\infty, -10, 0, 20\}$.

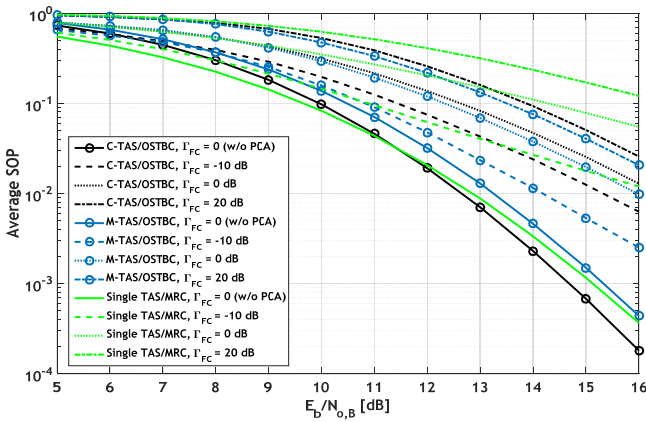


Figure 3 Average SOP of the TAS/STBC and single TAS/MRC schemes for $n_A = 4$, $n_B = 2$, $n_E = 2$, $\mathcal{R}_0 = 1$ bit/s/Hertz, $\bar{\gamma}_E = 5$ dB, $\bar{\gamma}_{FC}^{(L)} \rightarrow \infty$ and under the effects of PCA ($\Gamma_{FC,dB} \in \{-\infty, -10, 0, 20\}$)

As seen from the SOP curves given in Figure 3, for the ideal case of $\bar{\gamma}_{FC}^{(L)} \rightarrow \infty$ and $\Gamma_{FC,dB} \rightarrow -\infty$ (i.e., perfect feedback w/o PCAs), single TAS/MRC scheme outperforms both the C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC schemes in lower SNR (i.e., < 10 dB) region. For higher SNR values, the C-TAS/Alamouti-STBC scheme achieves the best SOP performance among three

schemes. However, as the dominance of the Eve’s contamination signal becomes evident (by the increasing values of Γ_{FC}), the M-TAS/Alamouti-STBC schemes are seen to provide additional SNR savings when compared to both single TAS/MRC and C-TAS/Alamouti-STBC schemes. Similar to the cases in Figure 2, employing modified schemes instead of the conventional ones provides average SNR savings of 0.6 dB, 0.3 dB and 0.37 dB, respectively for $\Gamma_{FC,dB} = -10, 0, 20$ at an average SOP of 10^{-1} .

7. CONCLUSIONS

We have examined the advantages obtained by the usage of the M-TAS/Alamouti-STBC schemes at the transmitter of a closed-loop diversity scheme from the perspective of physical layer security. The analytical derivations and the numerical results presented within this paper have exhibited the robustness of the modified scheme also to the active eavesdropping techniques such as PCAs when compared to the conventional scheme. Considering fading channel characteristics at both the legitimate/wiretap links and the feedback links, statistics related to the output SNR and the performance metrics have been presented in the presence of PCAs. SOP curves have shown that, in the presence of PCAs both schemes maintain the asymptotic diversity order that is provided by the pure Alamouti-STBC transmission and the optimal-diversity reception provided by the MRC. Analytical and numerical performance results clearly point out the potential of secrecy performance degradations that might be faced by both the C-TAS/Alamouti-STBC and M-TAS/Alamouti-STBC schemes due to the feedback imperfections caused by PCAs, and the superiority of M-TAS/Alamouti-STBC scheme to C-TAS/Alamouti-STBC scheme in the presence of PCAs.

Consequently, the considerable average SNR gain achieved against the conventional scheme in the presence of PCAs, provide the modified scheme to maintain significant importance in

real-world wireless communications system design from also the perspective of communications secrecy.

The authors think that further researches would be worth to be carried out on proposing novel antenna selection strategies and optimizing them by paying regard to probable imperfections in channel estimation and feedback processes (e.g., channel estimation errors, feedback delay and errors, and active eavesdropping).

REFERENCES

- [1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Technol. J.*, vol. 28, pp. 656-715, Oct. 1949.
- [2] B. Schneier, "Cryptographic design vulnerabilities," *Comput.*, vol. 31, no. 9, pp. 29-33, Sep. 1998.
- [3] A. Wyner, "The wire-tap channel," *Bell Syst. Technol. J.*, vol. 54, no. 8, pp. 1355-1387, Oct. 1975.
- [4] S. K. Leung-Yan-Cheong, and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451-456, July 1978.
- [5] X. Zhou, B. Maham and A. Hjørungnes, "Pilot Contamination for Active Eavesdropping," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 3, pp. 903-907, March 2012.
- [6] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas - part I: the MISOME wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088-3104, July 2010.
- [7] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas - part II: the MIMOME wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515-5532, Nov. 2010.
- [8] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961-4972, Aug. 2011.
- [9] F. He, H. Man, and W. Wang, "Maximal ratio diversity combining enhanced security," *IEEE Commun. Lett.*, vol. 15, no. 5, pp. 509-511, May 2011.
- [10] V. U. Prabhu and M. R. D. Rodrigues, "On wireless channels with M-antenna eavesdroppers: characterization of the outage probability and ϵ -outage secrecy capacity," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 853-860, Sep. 2011.
- [11] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351-361, Jan. 2011.
- [12] S. A. A. Fakoorian and A. L. Swindlehurst, "MIMO interference channel with confidential messages: achievable secrecy rates and precoder design," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 640-649, Sep. 2011.
- [13] X. Jiang, C. Zhong, X. Chen, T. Q. Duong, T. A. Tsiftsis, and Z. Zhang, "Secrecy performance of wirelessly powered wiretap channels," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3858-3871, September 2016.
- [14] Y. Zhang, Y. Ko, R. Woods, and A. Marshall, "Defining spatial secrecy outage probability for exposure region-based beamforming," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 2, pp. 900-912, February 2017.
- [15] H. Alves, M. De C. Tomé, P. H. J. Nardelli, C. H. M. De Lima, and M. Latva-Aho, "Enhanced transmit antenna selection scheme for secure throughput maximization without CSI at the transmitter," *IEEE Access*, vol. 4, pp. 4861-4873, 2016.
- [16] N. Yang, P. L. Yeoh, M. Elkashlan, R. Schober, and I. B. Collings, "Transmit antenna selection for security enhancement in MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 144-154, Jan. 2013.
- [17] J. Xiong, Y. Tang, D. Ma, P. Xiao, and K.-K. Wong, "Secrecy performance analysis

- for TAS-MRC system with imperfect feedback," *IEEE Trans. Inf. Foren. Sec.*, vol. 10, no. 8, pp. 1617-1619, August 2015.
- [18] J. Hu, Y. Cai, N. Yang, and W. Yang, "A new secure transmission scheme with outdated antenna selection," *IEEE Trans. Inf. Foren. Sec.*, vol. 10, no. 11, pp. 2435-2446, November 2015.
- [19] J. Zhu, Y. Zou, G. Wang, Y.-D. Yao, and G. K. Karagiannidis, "On secrecy performance of antenna-selection-aided MIMO systems against eavesdropping," *IEEE Trans. Vehic. Tech.*, vol. 65, no. 1, pp. 214-225, Jan. 2016.
- [20] N. Yang, P. L. Yeoh, M. ElKashlan, R. Schober, and J. Yuan, "MIMO wiretap channels: Secure transmission using transmit antenna selection and receive generalized selection combining," *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1754-1757, Sept. 2013.
- [21] S. Yan, N. Yang, R. Malaney, J. Yuan, "Transmit antenna selection with Alamouti coding and power allocation in MIMO wiretap channels," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 3, pp. 1656-1667, March 2014.
- [22] A. F. Coşkun, O. Kucur, "Feedback-rate efficient transmit antenna selection/Alamouti scheme with robust error performance in the presence of feedback errors," *IEEE Comm. Lett.*, vol. 17, no. 5, pp. 908-911, May 2013.
- [23] I. Gradshteyn, I. Ryzhik, *Tables of Integrals, Series and Products*, Academic Press: San Diego CA, 1994.
- [24] H. A. David and H. N. Nagaraja, *Order Statistics, 3rd ed.* Hoboken, NJ: Wiley, 2003.
- [25] P. Prudnikov, Y. A. Brychkov, O. I. Marichev, *Integrals and Series, Vol. 5*, Gordon and Breach Science Publishers: New York, 1986.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Performance comparison and analysis of Linux block I/O schedulers on SSD

Authors: Yunus Ozen, Abdullah Yildirim

Received: 2018-11-01 15:03:41

Revised: 2018-11-11 23:53:28

Accepted: 2018-12-13 09:51:56

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 106-112

How to cite

Yunus Ozen, Abdullah Yildirim; (2019), Performance comparison and analysis of Linux block I/O schedulers on SSD. Sakarya University Journal of Science, 23(1), 106-112, DOI: 10.16984/saufenbilder.477446

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/477446>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Performance comparison and analysis of Linux block I/O schedulers on SSD

Yunus Ozen^{*1}, Abdullah Yildirim¹

ABSTRACT

A computer system's one of the slowest operation is disk seek operation. Sending out read and write requests to the block devices such as disks as soon as the request arrives results in poor performance. After performing sorting and merging operations, the operating system kernel issues block I/O requests to a disk for improving the overall system performance. The kernel subsystem to perform scheduling the block I/O requests is named as the I/O scheduler. This paper introduces performance comparison and detailed analyses of Deadline, CFQ, Noop and BFQ block I/O schedulers that are contained in the Linux 4.1x kernel. The tests have been carried out on an SSD block device that is common in hardware combinations of both personal and professional use-case scenarios. The performance of the schedulers has been evaluated in terms of throughput. Each scheduler has advantages in different use-case scenarios and provides better throughput in a suitable environment.

Keywords: Block I/O Scheduler, Deadline, Noop, CFQ, BFQ.

1. INTRODUCTION

Block devices such as hard drives or flash memories run in a random access fashion to write or read fixed-size pieces of data. That data is named as a block. Whenever a piece of data is requested for a block device, the read/write head seeks from a position to another position. This seeking operation is a slow operation. Since block devices are performance-sensitive the kernel has a dedicated sub-system called block I/O layer to optimize seeking operations. The main motivation of the Linux kernel version 2.5 development was to optimize the block I / O layer. The bio struct proposed with version 2.5 in addition to the bufferhead struct is still an essential part of the modern Linux kernel [1]. The

bio struct is the basic container for block I/O requests in the Linux kernel. It stores the active block I/O operations as a list of segments. A segment represents a bunch of buffers that are contiguous in memory. The bio struct provides flexibility to perform multiple block I/O operations with its segments based approach.

Block devices have request queues to schedule pending read or write requests. The kernel subsystem to perform scheduling the block I/O requests is called the I/O scheduler. Sending requests to the block device immediately ends up in poor performance. The scheduler organizes the request order in the queue and dispatching time to the block device. The main objective is reducing seeks to improve overall throughput of

* Corresponding Author: yunus@yunus.gen.tr

¹ Yalova University, Computer Engineering, Yalova, Turkey

the system. The kernel issues the requests to the device after performing some merging and sorting operations in this purpose. Merging is the bundling of two or more requests into a single request. Merging the requests reduces overhead while decreases the seek operations. The whole request queue is kept sorted according to sector positions. The purpose in sorting is minimizing the number of seeks by keeping the disk head moving into the same direction [2].

Linus Elevator was the first I/O scheduler in the Linux. It performs both merging and sorting operations to optimize the number of seeking. The request is added to the tail of the request queue, if a suitable location is not found for a request to merge [3]. If an existing request is older than a threshold, the new request is added to the tail of the queue. That is not efficient but prevents several requests to be starved. This improves latency but causes to request starvation.

Several schedulers such as Deadline, CFQ, and Noop are introduced after version 2.6 to overcome this starvation problem. The main motivation of those earlier schedulers was reducing the number of seek operation on rotational magnetic block devices such as hard drives [1].

Rotational magnetic block devices have been replaced by solid state drives (SSDs) recently in areas ranging from smart devices to large data center implementations. SSDs have some advantages over traditional HDDs in terms of throughput, reliability, and energy consumption. They are free from the latency of the seeking time of rotational magnetic disks. However, existing I/O hardware and schedulers have been designed and optimized for rotational magnetic disk specifications [4].

The SSDs are getting research interest for their potential to make appropriate optimizations with a new motivation [5].

The literature already has some recent studies that modify queue structures of existing state of the art schedulers and exploit the internal parallelism of SSDs. FlashFQ [6] analyzes request size to estimate the response time and uses the start-time fair queueing to provide fairness among concurrent tasks on SSDs. Gao et

al. [7] proposed a scheduler called PIQ for minimizing the access conflicts among the I/O requests in one batch. External mergesort is a common sorting algorithm to sort large amounts of data. FMsrt focuses on enhancing the merge phase of external mergesort for SSDs [8]. Several studies have been proposed to take advantage of the internal parallelism of SSDs to improve performance [9]. Mao et al. designed a new I/O scheduler called Amphibian by utilizing internal parallelism of SSDs. Amphibian performs size-based request ordering to prioritize requests with small sizes [10]. Chen et al. performed experiments to show the results of optimization based on internal parallelism for the performance improvement [11]. Guo et al. [12] proposed a scheduler called SBIOS considering full use of read internal parallelism and avoiding the block cross penalty. Most of the studies in the literature focus on existing complex schedulers for adapting them to hardware opportunities of SSDs. Revisiting state of the art schedulers in terms of throughput and highlighting the potential for a simple scheduler is needed.

The focus of our study is to make a comparison between the state of the art I/O schedulers in the Linux kernel in terms of throughput.

The rest of the paper is organized as follows. The schedulers Deadline, CFQ, Noop, and BFQ block I/O are presented in Section 2. Benchmark setup, experimental platform details, performance metrics, and preferred workloads are described in Section 3. Performance evaluation is presented in Section 4. We finally conclude this paper in Section 5.

2. I/O SCHEDULERS

The I/O scheduler merges and sorts the pending block I/O requests into the request queues and sends them to the system. This section briefly describes the Deadline, CFQ, Noop, and BFQ block I/O schedulers that are compared in terms of throughput and analyzed in this paper. Deadline, CFQ, Noop and BFQ block I/O schedulers are chosen, because they are contained in most of the Linux distributions with the 4.1x kernel.

2.1. The Deadline I/O Scheduler

The Deadline scheduler is one of the earliest schedulers that focus on the starvation problem of the Linux Elevator. Linus Elevator targets merging I/O requests to a specific portion of the disk and this causes starvation of the requests to another portion of the disk. The read requests are generally dependent on each other because of data locality. The scheduler merges them to minimize the seek operation and this causes starvation. There is a tradeoff between minimizing seeks and preventing starvation. The Deadline scheduler contains a request queue that is sorted sectorwise on disk and merged like Linus Elevator. The Deadline scheduler inserts the request into another queue according to the type of request. Write requests are inserted into a write FIFO queue and read requests are inserted into a read FIFO queue. Deadline scheduler maintains a balance to make these operations fair with its multi-queue structure. It gives smaller expiration value to the read requests than write requests to prevent write requests starving read requests. The simplified diagram of deadline scheduler is shown in Figure 1.

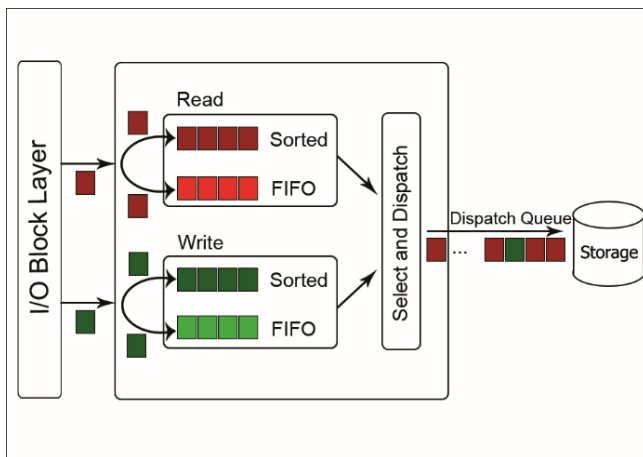


Figure 1. Deadline scheduler.

2.2. The Complete Fair Queuing I/O Scheduler (CFQ)

The CFQ scheduler organizes incoming I/O requests to the queues based on the processes. The newly submitted I/O request is combined with neighboring requests and insertion sorted sectorwise in every queue. The CFQ scheduler

differs from other schedulers with its per-process queues. The round-robin structure lets a number of requests to be dispatched before continuing on to the next one. Each process gains an equal slice of disk bandwidth and this algorithm provides fairness at a per-process fashion. The simplified diagram of CFQ scheduler is shown in Figure 2.

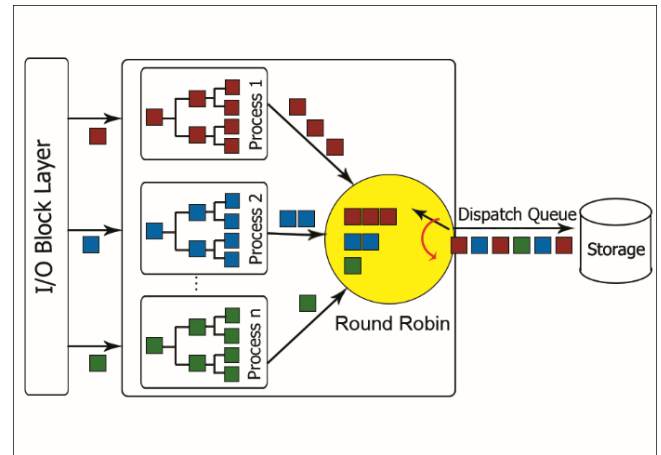


Figure 2. CFQ scheduler

2.3. The Noop I/O Scheduler

The Noop scheduler does not sort requests before inserting to the queue. It merges the new request to the adjacent request and maintains a single request queue in a near-FIFO order. It is said to be the first I/O scheduler that targets the block devices such as flash memories that run in a completely random-access fashion. The simplified diagram of Noop scheduler is shown in Figure 3.

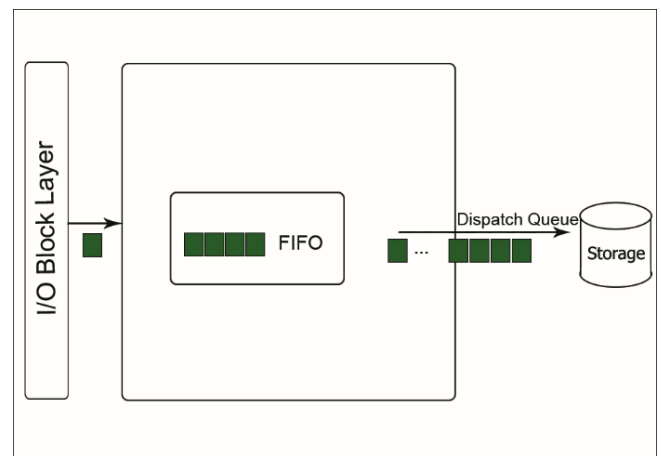


Figure 3. Noop scheduler.

2.4. The Budget Fair Queueing I/O Scheduler (BFQ)

The BFQ scheduler is an equal-share disk scheduling algorithm. It is based on CFQ that is default I/O scheduler in several Linux distributions. BFQ converts time intervals based round-robin to the number of sectors based round-robin. It assigns a sector budget to each request instead of a time slice. The simplified diagram of BFQ scheduler is shown in Figure 4.

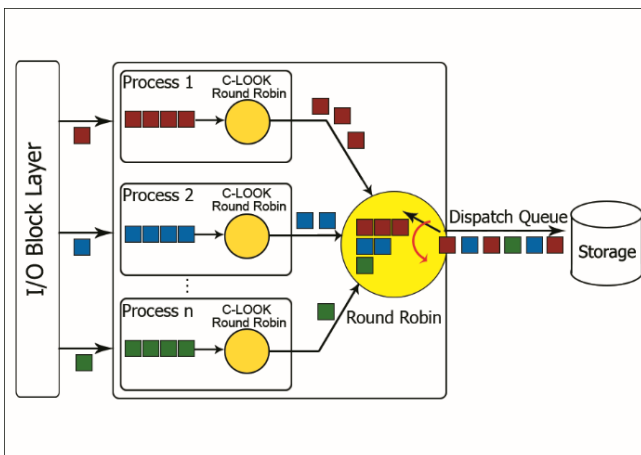


Figure 4. BFQ scheduler.

3. EXPERIMENTAL EVALUATION

The benchmark setup and the experimental platform with the different file sizes used for the performance analysis of the Deadline, CFQ, Noop, and BFQ schedulers are presented in this section. The results provided are read, reread, write and rewrite results of different schedulers as already explained in Section 2. For our analysis, we used IOzone to perform benchmarking of selected I/O schedulers.

3.1. Benchmark Setup

The preferred benchmark tool for the study presented in this paper is the IOzone benchmarking tool [13]. It generates workloads for several file operations. The IOzone provides a framework to run different scenarios to measure the performance of the system. It can measure the performance of file operations with different file sizes and different sized chunks of the file at a time. These chunks are particular

spots within a file to read or write in one try. The size of these chunks affects the I/O performance. The experiments have been executed for each of the schedulers on the selected platform with constant 64 KB sized chunks and varying file sizes from 64 KB to 500 MB to measure the throughput performance against varying file sizes. The maximum throughput obtained for each of the file operations has been reported.

3.2. Experimental Platform

The experiments have been carried out on 4 cores 1.90 GHz Intel i7 351U processor system, with 4GB main memory, 256KB L2 cache, 4MB L3 cache running Manjaro Distribution (Linux 4.19.0-3). The SSD was a 240GB Sandisk U100 SATA 600. An EXT4 file-system has been used on the drive. The computer has been rebooted before each experiment to remove cache related effects.

3.3. Performance Metrics

The throughput of disks is finite and comparatively small while reading and writing. It causes bottlenecks that block I/O schedulers intend to improve overall system performance by changing the throughput performance. Total disk throughput (in KB/s) has been used as the metric to show the performance of the schedulers in the benchmark experiments. I/O intensive process execution time has been used to measure disk throughput. Any other processes have been killed except the daemons before benchmark execution. In these experiments, larger throughput (smallest execution time) means better scheduling for that file operations.

3.4. Workloads

The write, rewrite, read, and reread workloads have been used for the experiments using IOzone benchmark tool.

The read test is for measuring the reading performance of an existing file. The reread test is for measuring the reading performance of a file that was recently read. In this case, the performance tends to be higher as the data is

cached by the OS as it is recently accessed. The write test is for measuring the writing performance of a new file to the disk. Whenever a new file is written, the metadata is written on the block device in addition to the data itself. The rewrite performance becomes higher than the performance of writing a file because of this overhead. The rewrite test is for measuring the writing performance of a file that already exists on the disk. Whenever an existing file is written, the required effort is lower as the metadata is not written again.

4. ANALYSIS AND RESULTS

This section shows a comparative performance analysis using the workloads described in Section 3 for the Linux I/O schedulers Deadline, CFQ, Noop, and BFQ. The aim of the analysis is to understand how different schedulers perform under different workloads in terms of throughput.

For multiprocess throughput evaluation, we let IOzone run 3 processes for the initial write, rewrite, read, reread tests. These tests have been carried out 10 times and the results have been analyzed through the average of these tests.

Figure 5 shows the initial write test results. Noop gives the best performance with its SSD-ready structure. CFQ scheduler gives equal chance to every process and it has better write performance comparing to Deadline and BFQ. The schedulers without process priority have worse throughput results. Noop has an average 10% better write performance than its closest competitor CFQ.

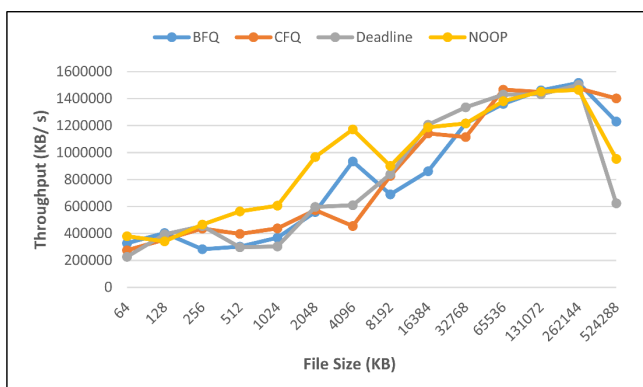


Figure 5. Write test results.

Figure 6 presents the rewrite test results. The throughput results have nearly the same ratio

with the write test results. The write tests write the data and also the metadata for the file, but the rewrite test writes only the data to disk. All schedulers have higher rewriting performance than writing performance. The average rewrite performance of all schedulers is 38% better than the write performance. Noop has an average 4% better rewrite performance than its closest competitor CFQ.

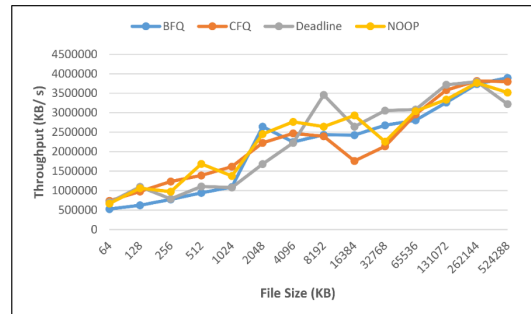


Figure 6. Rewrite test results.

The read and reread tests are shown in Figure 7 and Figure 8. Both read and reread throughput results are better than write and rewrite results for all schedulers. Noop has the best results. Deadline scheduler has better results than CFQ and BFQ in both read and reread tests because it prioritizes reads more than writes. Noop has an average 7.1% better read and 10.2% better reread performance than its closest competitor Deadline. Deadline has an average 2.9% and 6.7% better read performance than CFQ and BFQ respectively. It also has an average 1.5% and 7.9% better reread performance than CFQ and BFQ respectively.

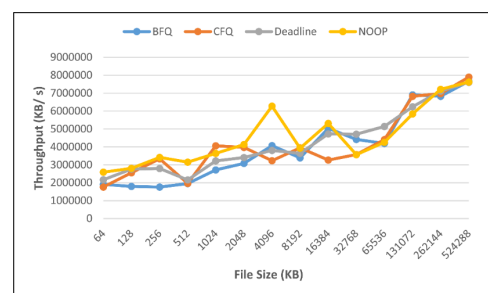


Figure 7. Read test results.

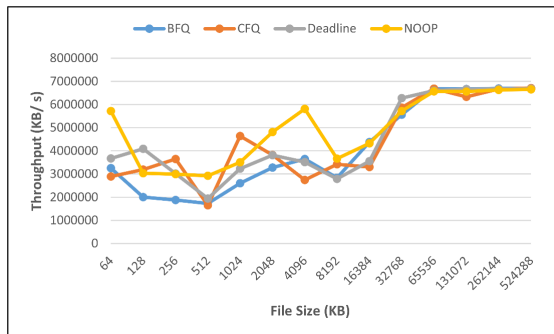


Figure 8. Reread test results.

The average of all tests is shown in Figure 9. CFQ scheduler provides an equal chance to each process in its round-robin structure. This makes it not suitable for environments that might need to prioritize request types for processes. Deadline scheduler is suitable for read-intensive works. Its default timeout values prioritize reads more than writes. These values are configurable according to the features of the work. Noop scheduler is optimized for systems that do not need a specific I/O scheduler. It has a modest structure with its single FIFO queue. It is suitable for the environments where the operating system is in a hypervisor. The underlying host operating system does scheduling itself in hypervisors or cloud environments that the operating systems inside virtual boxes do not need complex I/O schedulers. The BFQ scheduler is optimized for interactive tasks of personal-use scenarios instead of server scenarios. It focuses on delivering the lowest latency rather than reaching higher throughput. The operating system distributions focusing on personal usage do not perform the heavy read or write operations generally and lower latency is the prioritized to throughput.

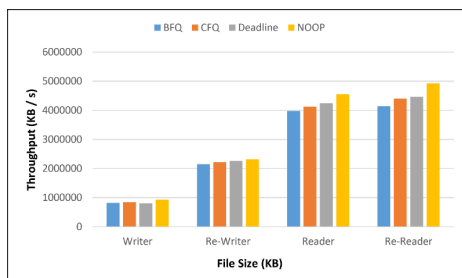


Figure 9. Average of all test results.

5. CONCLUSION

Block devices maintain request queues and approaches to schedule pending read or write requests. The kernel I/O scheduler subsystem is responsible for request optimization. The performance of Deadline, CFQ, Noop, and BFQ block I/O schedulers that are included in the latest Linux 4.1x kernel are compared in terms of throughput this paper. The tests have been carried out on an SSD block devices that are common in ranging from small handheld devices to large-scale data center configurations. According to the test results, each scheduler has different advantages over others. CFQ scheduler is suitable for the systems that require balanced I/O access and do not need process prioritization. Deadline scheduler has better performance on read-intensive works. Noop is for the systems on the cloud or hypervisors. BFQ performs better on interactive use-case scenarios. Noop is the simplest scheduler and it is considered to have the potential for optimized new implementations targeting SSD block devices.

REFERENCES

- [1] R. Love, "The Block I/O Layer," in *Linux Kernel Development*, Crawfordsville, Indiana, Addison-Wesley, 2010, pp. 290-304.
- [2] F. Chen, R. Lee, and X. Zhang. "Essential roles of exploiting internal parallelism of flash memory based solid state drives in high-speed data processing," *In HPCA'11*, San Francisco, CA, 2011.
- [3] J. Fusco, "The I/O Scheduler," in *The Linux programmer's toolbox*, Upper Saddle River, NJ, Pearson Education, 2007, pp. 282-284.
- [4] J. Kim, J. Kim, P. Park, J. Kim and J. Kim, "SSD Performance Modeling Using Bottleneck Analysis," in *IEEE Computer Architecture Letters*, vol. 17, no. 1, pp. 80-83, 1 Jan.-June 2018.
- [5] S. Mittal and J. S. Vetter, "A Survey of Software Techniques for Using Non-Volatile Memories for Storage and Main Memory Systems," in *IEEE Transactions*

- on Parallel and Distributed Systems*, vol. 27, no. 5, pp. 1537-1550, 1 May 2016.
- [6] K. Shen and S. Park, "FlashFQ: A fair queueing I/O scheduler for flash-based SSDs," in *Proc. USENIX Annu. Tech. Conf. (USENIX ATC)*, San Jose, CA, USA, Jun. 2013, pp. 67–78.
- [7] C. Gao *et al.*, "Exploiting Parallelism for Access Conflict Minimization in Flash-Based Solid State Drives," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 1, pp. 168-181, Jan. 2018.
- [8] J. Lee, H. Roh and S. Park, "External Mergesort for Flash-Based Solid State Drives," in *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1518-1527, 1 May 2016.
- [9] W. Wang and T. Xie, "PCFTL: A plane-centric flash translation layer utilizing copy-back operations," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 12, pp. 3420–3432, Dec. 2015.
- [10] B. Mao and S. Wu, "Exploiting request characteristics and internal parallelism to improve SSD performance," in *Proc. 33rd IEEE Int. Conf. Comput. Design (ICCD)*, NY, USA, Oct. 2015, pp. 447–450.
- [11] F. Chen, R. Lee, and X. Zhang, "Essential roles of exploiting internal parallelism of flash memory based solid state drives in high-speed data processing," in *Proc. 17th Int. Conf. High-Perform. Comput. Archit. (HPCA)*, San Antonio, TX, USA, Feb. 2011, pp. 266–277.
- [12] J. Guo, Y. Hu and Bo Mao, "SBIOS: An SSD-based Block I/O Scheduler with improved system performance," *2015 IEEE International Conference on Networking, Architecture and Storage (NAS)*, Boston, MA, 2015, pp. 357-358.
- [13] W. D. Norcott, D. Capps, "Iozone filesystem benchmark," [Online]. Available: www.iozone.org. [Accessed 4 October 2018].



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: Optimization of Surface Roughness of AISI 1040 Stainless Steel in Milling Process Using Taguchi Method

Authors: Neslihan Özsoy, Murat Özsoy

Received: 2018-10-01 12:00:12

Revised: 2018-11-14 11:23:03

Accepted: 2018-12-13 10:05:04

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 113-120

How to cite

Neslihan Özsoy, Murat Özsoy; (2019), Optimization of Surface Roughness of AISI 1040 Stainless Steel in Milling Process Using Taguchi Method. Sakarya University Journal of Science, 23(1), 113-120, DOI: 10.16984/saufenbilder.466053

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/466053>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

Optimization of Surface Roughness of AISI 1040 Stainless Steel in Milling Process Using Taguchi Method

Neslihan Özsoy*¹, Murat Özsoy¹

ABSTRACT

Surface roughness significantly affects the work efficiency and life of machine parts interacting with each other. There are lots of parameters that affect surface roughness such as processed material, cutting tool, cutting parameters, cooling type. For this reason, optimization of the machining parameters and proper processing conditions are very important. In this study, the optimum machining conditions were determined by investigating the surface roughness of the milled AISI 1040 steel alloy depending on the feed per tooth (0.08 mm/tooth, 0.12 mm/tooth, 0.16 mm/tooth, 0.20 mm/tooth), spindle speed (2000 rpm, 3000 rpm, 4000 rpm, 5000 rpm) and the cooling type (liquid, air) parameters. The Taguchi experimental design method was used to help achieve results at acceptable levels and to save time and cost in achieving optimal results. Experiments designed with the Taguchi method were based on the L16 orthogonal array and signal / noise (S/N) ratios were used in the evaluation of the experimental results. The optimum levels of control factors for minimizing surface roughness were determined using S/N ratios. The ideal conditions of surface roughness were seen at A1B1C2 (i.e., feed per tooth=0.08 mm/tooth and 2000 rpm spindle speed, cooling type=liquid). Variance analysis (ANOVA) was also conducted in the study. According to the results of analyses, it was found that spindle speed was the most dominant parameter for surface roughness. Lastly, confirmation tests were run to check the success of the optimization.

Keywords: Taguchi method, optimization, surface roughness, ANOVA, milling of stainless steel

1. INTRODUCTION

Due to the demand for more precise and complex products in the automotive, tooling and aerospace sectors, the increase in workability studies is taking place. Especially the flexibility of high-speed CNC machine tools, CAD/CAM software and the improvements in cutting tool

technology show that the work in this area is dispersed in a very wide range.

AISI 1040 steel is widely used in the machinery and manufacturing industries and is used in a variety of machine parts, automotive parts, molds, fixtures and many other areas. Technological developments in the field of machining in recent years have led to many

* Corresponding Author: nerken@sakarya.edu.tr

¹ Sakarya University, Engineering Faculty, Mechanical Engineering Department, Sakarya, Turkey

problems affecting machining performance, although machining makes machining more efficient.

The work in this area is focused on examining the cutting parameters (cutting speed, axial depth of cut, radial depth of cut and feed rate), cutting tool geometry/material and workpiece material changes on surface roughness. In the 1960s, the Taguchi method proposed by Genichi Taguchi is widely used due to its proven success in improving industrial product quality. The relevance to the "Taguchi" method of researchers is increasing day by day, especially because of the low number of experiments, ease of application and easy evaluation of qualitative variables.

Some studies on the subject:

Tosun et al. examined the surface roughness of AA7075-T6 using conventional and air cooling methods. They used a mixture of boron oil and water as the cutting process. They selected parameters as cutting tools, spindle speed and feed rates. As a result it was determined that the surface roughness increased when the feed rate increased and the surface roughness decreased when the spindle speed increased. [1].

Kahraman studied about optimizing cutting parameters to reduce surface roughness in turning of studs manufactured from AISI 5140 steel. She chose rotational speed, feed rate and depth of cut as control factors for the surface roughness and L9 orthogonal array for experiment trials. She considered rotational speed, feed rate and depth of cut as control factors for the surface roughness, and used L9 orthogonal array for experiment trials. She found optimal surface roughness as 1.70 μm according to S/N ratio analysis[2].

Kıvıak experimented to investigate the machinability of hadfield steel with PVD TiAlN- and CVD

TiCN/Al₂O₃-coated carbide inserts in dry milling conditions. He used the Taguchi method and regression analysis in his study. He selected cutting tool, cutting speed and feed rate as machining parameters. He obtained that feed rate had been the most effective parameter according to variance analysis [3].

Günay, AISI 316L austenitic stainless steel processing of the Fc and Ra cutting parameters and tool nose radius examined the optimization. He performed the experiments using Taguchi method [4]. Cakiroglu and Acir investigated drilling performance of Al2014 materials in dry cutting. They found out that cutting speed, feed rate and cutting tool had been effective on the performance [5].

In another study Debnath et al investigated effects of cutting speed, feed rate, depth of cut and cutting fluid conditions on surface roughness and tool wear in turning process of mild steel bar. They carried out experiments according to Taguchi's L9 orthogonal array and found that feed rate and the flow rate of the cutting fluid were the most effective parameters on surface roughness. On the other hand they explained that cutting speed and depth of cut were the dominant factors influencing tool wear [6].

Manivel and Gandhinatan optimized surface roughness and tool wear in hard turning of austempered ductile iron. They designed experiments by Taguchi method and selected cutting speed, feed rate and depth of cut as parameter [7].

Pillai et al conducted an experimental study of optimization of surface roughness and machining time in milling of Al6005A alloy. They used Taguchi-Grey relational and ANOVA to investigate effects of parameters [8].

Mandal et al conducted experimental work based on L9 orthogonal array with three parameters. These

were depth of cut feed rate, cutting speed at three levels. They investigated machinability of AISI 4340 steel and found that depth of cut had maximum contribution on tool wear [9].

Vishnu et al made a study to investigate effects of type of machining conditions, cutting speed, feed rate, depth of cut, type of tool on surface roughness of EN-353 Alloy Steel in turning operation. They used L27 orthogonal array in experiments and ANOVA to deduce influence of parameters. As a result they found that type of coolant had been the most effective factor [10].

In this study, experiments were planned according to Taguchi method, surface roughness was evaluated statistically using ANOVA in the milling of AISI 1040 stainless steel.

2. MATERIALS AND METHODS

2.1. Taguchi method

Taguchi method is an experiment design and optimization method based on parameter design, system design and tolerance design. Most commonly used in statistical analysis of data collected within the scope of quality assurance systems. Taguchi's experimental design method is a very useful method to detect the optimum combination between different levels of different parameters. This technique had been popular in the world after the 1980s. The benefit of Taguchi design is that many factors can be considered together. It also searches for nominal design points that are insensitive to changes in production and user environments to improve productivity in production and reliability in the performance of a product. At the same time noise

factors can be checked. Although similar to the experimental design (DOE/design of experiment), the Taguchi design performs only combinations of balanced (orthogonal) experiments, making Taguchi design more effective than fractional factorial design. Industries can greatly reduce product development cycle time for both design and production, thereby reducing costs and increasing profits using the Taguchi method. Besides, Taguchi allows examine to the variability caused by noise factors, which are usually neglected in the traditional DOE approach [11].

In cases where a considerable amount of experimental work is required for all combinations involving each level of each parameter, it is possible to achieve a much lower number of experimental studies using the Taguchi method. The methodology developed in the Taguchi experimental design method consists of three basic concepts; system design, parameter design and tolerance design [12, 13].

2.2. Work piece material

The work piece material selected for investigation is AISI 1040 steel with the compositions as shown in table 1. The AISI 1040 steel used in the test is of the general production grade class and is widely used in the market. Since it is suitable for heat treatment, it is used in many mold sets, apparatus making, automotive sector, parts subjected to various difficulties,

transferring miller and some gears. It has a tensile strength of 600 MPa, a modulus of elasticity of 210 GPa and a 24% elongation. The dimensions of the work piece used in tests are 20X150X80 mm³.

Table 1. Chemical content of the workpiece material

C %	Si %	Mn %	P _{max} %	S _{max} %
0,40-0,50	0,25-0,35	0,60-0,90	0,04	0,05

2.3. Experimental plan and procedure

Three different parameters were chosen in the experiment design: Feed per tooth (Fpt), spindle speed (rpm), cooling type (Ct). Selected parameters and levels are specified in Table 2. The most suitable orthogonal array L₁₆ (4²x2¹) was selected to determine optimum conditions and analyse the parameters [14]. The L₁₆ mixed level orthogonal array shown in table 3.

Table 2. Control parameters and their levels

Symbol	Parameters	Level 1	Level 2	Level 3	Level 4
A	Fpt mm/tooth	0.08	0.12	0.16	0.20
B	rpm	2000	3000	4000	5000
C	Ct	Air	Liquid	-	-

All statistical analyses were performed using the Minitab 17 statistical package program at 90% confidence level.

Experimental samples were prepared at TAKSAN brand TMC 700V model CNC vertical machining center in Sakarya University Mechanical Engineering Department Laboratory.

Table 3. The L₁₆ mixed level (4²x2¹) orthogonal array

Experiment #	Factor A	Factor B	Factor C
1	1	1	1
2	2	1	1
3	3	1	2
4	4	1	2

5	1	2	1
6	2	2	1
7	3	2	2
8	4	2	2
9	1	3	2
10	2	3	2
11	3	3	1
12	4	3	1
13	1	4	2
14	2	4	2
15	3	4	1
16	4	4	1

MAHR-MARSURF PS1 model desktop roughness measurement instrument was used to measure surface roughness values and is shown in figure 1. Measurements were taken at room temperature. The instrument measures according to DIN EN ISO 3274 standards. Measurements were taken from the workpiece on the same axis at one point. The cutting length (Lc) was chosen as 0.8 mm and the exemplification length (Lt) as 5.6 mm for the measurement of the surface roughness values which occurred during machining on the workpiece.

APXT1035PDSR-MM coated cementitious carbide endmill cutter manufactured by Korloy, was used for milling in the sample AISI 1040 stainless steel at 20x150x80 mm³. The diameter of the endmill was selected as 16 mm.

3. RESULTS AND DISCUSSION

3.1. S/N ratio results

The average surface roughness (µm) measurements made and their average values are given in Table 4. Optimization was done with the help of the results obtained. In this optimization process, there are three different convenient functions known as the Taguchi loss function, also referred to as the signal to noise ratio (S/N) function. They are " smallest best, the highest (the largest) is the best, the nominal 'is the best' cases. The aim of this work was to decrease

surface roughness. Therefore “the smallest the better” quality characteristic was used as shown in equation 1.

$$\eta = S/N = -10 \log \left[\frac{1}{n} \sum_{i=1}^n y_i^2 \right] \quad (1)$$

Here y_i =the observed data at the i_{th} experiment and n = the number of the experiments [4].



Figure 1. Desktop roughness measurement instrument

Table 4. The experimental results

Experiment #	Feed per tooth (mm/tooth)	Spindle speed (rpm)	Cooling type	Surface Roughness, Ra (µm)	S/N ratio for Ra (dB)
1	0.08	2000	Air	1.121	-0.9902
2	0.12	2000	Air	1.279	-2.1374
3	0.16	2000	Liquid	1.336	-2.5161
4	0.20	2000	Liquid	1.392	-2.8697
5	0.08	3000	Air	1.313	-2.3653
6	0.12	3000	Air	1.501	-3.5291
7	0.16	3000	Liquid	1.599	-4.0770
8	0.20	3000	Liquid	1.497	-3.5059
9	0.08	4000	Liquid	2.350	-7.4232
10	0.12	4000	Liquid	3.586	-11.0928
11	0.16	4000	Air	2.628	-8.3917
12	0.20	4000	Air	3.687	-11.3341
13	0.08	5000	Liquid	2.287	-7.1853
14	0.12	5000	Liquid	1.919	-5.6615
15	0.16	5000	Air	2.367	-7.4830
16	0.20	5000	Air	2.882	-9.1924

Taguchi analysis, " the smallest better " signal noise ratios (S/N) according to the surface roughness (µm) of the values indicating the order of the effect of the values shown in table 5. The optimal levels for optimum surface roughness can be seen from this table. The graphical forms of the levels of control factors for Ra given in Table 5 are illustrated in figure 2. The best level of each factor was found by looking at the highest S / N ratio in the levels of its control factor. Accordingly, factor A (Level 1, S/N=-4.491), factor B (Level 1, S/N=-2.128) and factor C (Level 2, S/N=-5.541) were the factors that gave the best Ra value. That means an optimum Ra

value can be acquired from these levels of factors.

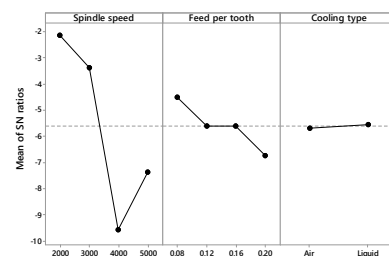


Figure 2. S/N graphs for surface roughness

Table 5. S/N response table for surface roughness

Level	A	B	C
1	-4.491	-2.128	-5.678
2	-5.605	-3.369	-5.541
3	-5.617	-9.560	-
4	-6.725	-7.381	-
Delta	2.234	7.432	0.136
Rank	2	1	3

3.2. ANOVA results

With Variance Analysis (ANOVA), the effects of which factors are effective on the process are determined statistically. In the analysis of variance, the aim is to determine the effects of the factors examined on how much they affect the output values selected to measure the quality and how the different levels cause the variability. In addition, the statistical

reliability of the results obtained is also tested. Analysis was conducted with ANOVA to determine the effects of feed per tooth, spindle speed and cooling type on surface roughness. The ANOVA results of the surface roughness are given in table 6. The analysis was performed with a 10% significance level and a 90% confidence level.

Table 6. Results of ANOVA for surface roughness

Variance source	Degree of freedom (DoF)	Sum of squares(SS)	Mean square (MS)	F ratio	Contribution rate (%)
A	3	0.7338	0.24459	1.71	7
B	3	8.1682	2.72272	19.09	80
C	1	0.0411	0.04108	0.29	0.4
Error	8	1.1412	0.14265	-	13
Total	15	10.0842	-	-	100

The importance of control factors in ANOVA is defined by comparing the F values of each control factor. The percentage value of each parameter contribution and the degree of impact on the process performance is shown in the last column of the table. The contribution of A, B and C factors to surface roughness was found to be 7%, 80% and 0.4%, respectively (Table 6). So the most dominant factor affecting the surface roughness was feed per tooth (factor B, 80%).

3.3. Confirmation tests

The last step of Taguchi method is to control optimum levels. So, verification experiments of the control factors were performed with optimum levels. Confirmation test results are shown in table 7.

Table 7. Confirmation test results

	Prediction	Experimental
Level	A1B1C2	A1B1C2
Parameters	2000 0.08 Liquid	2000 0.08 Liquid
Average surface roughness	0.952	1.03

According to the difference between actual result and prediction calculated error is %7.5. This value is within acceptable level. That means, the optimization for surface roughness was obtained in confidence interval.

4. CONCLUSION

In this work, the effects on the average surface roughness of the feed per tooth, spindle speed, cooling type parameters in the milling of the AISI 1040 Stainless Steel were evaluated using the

Taguchi test design method. The following results may be drawn.

- The best average surface roughness value was at $A_1B_1C_2$ (i.e., feed per tooth=0.08 mm/tooth, spindle speed=2000 rpm, cooling type=liquid).
- It was observed that the spindle speed was the most significant factor in the roughness change with the ratio of 80 %, and this parameter was followed by feed per tooth (7%) and cooling type (0.4%) interactions according to the order of significance.
- Measured values were within the 90% confidence level according to the confirmation experiments.

These results show that the Taguchi method is a trusty process for reducing the machining time and production costs of milling AISI 1040 stainless steel. These results can be used for future academic research as well as industrial applications.

REFERENCES

- [1] N. Tosun, C. Kuru, E. Altintas, O. E. Erdin, "Investigation of surface roughness in milling with air and conventional cooling method," J. Fac. Eng. Arch. Gazi Univ. , vol. 25, pp.141-146, 2010.
- [2] F. Kahraman, "Optimization of cutting parameters for surface roughness in turning of studs manufactured from AISI 5140 steel using the Taguchi method," Materails Testing, vol. 59, no. 1, pp. 77-80, 2017, DOI:10.3139/120.110968
- [3] T. Kivak, "Optimization of surface roughness and flank wear using the Taguchi method in milling of Hadfield steel with PVD and CVD coated inserts," Measurement vol. 50, pp. 19–28, 2014, DOI: 10.1016/j.measurement.2013.12.017
- [4] M. Günay, "Optimization with Taguchi method of cutting parameters and tool nose radius in machining of AISI 316L steel," Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 28, pp. 437-444, 2013.
- [5] R. Çakıroğlu, A. Acır, "Taguchi optimization method of tool chip interface temperature depending on the cutting parameters in drilling operations," Electronic Journal of Machine Technologies, vol.10, pp. 73-86, 2013.
- [6] S. Debnath, M. M. Reddy, Q. S. Yi, "Influence of cutting fluid conditions and cutting parameters on surface roughness and tool wear in turning process using Taguchi method", Measurement 78 (2016) 111-119. DOI:10.1016/j.measurement.2015.09.011
- [7] D. Manivel, R. Gandhinathan, "Optimization of surface roughness and tool wear in hard turning of austempered ductile iron (grade 3) using Taguchi method", Measurement 93 (2016) 108-116, DOI:10.1016/j.measurement.2016.06.055
- [8] J. U. Pillai, I.Sanghrajka, M. Shunmugavel., T. Muthuramalingam, M. Goldberg., G.Littlefair, "Optimisation of multiple response characteristics on end milling of aluminium alloy using Taguchi-Grey relational approach", Measurement 124 (2018) 291-298, DOI: 10.1016/j.measurement.2018.04.052
- [9] N. Mandal, B. Doloi, B. Mondal, R. Das, "Optimization of flank wear using Zirconia Toughened Alumina (ZTA) cutting tool: Taguchi method and regression analysis",

Measurement 44 (2011) 2149-2155,
DOI:10.1016/j.measurement.2011.07.022

- [10] A.Venkata Vishnu, M. V. Ramana, K.B.G.Tilak, “Experimental Investigations of Process Parameters Influence on Surface Roughness in Turning of EN-353 Alloy Steel under Different Machining Environments”, *Materials Today: Proceedings* 5 (2018) 4192–4200.
- [11] J. Z. Zhang, J. C. Chen, E. D. Kirby, “Surface roughness optimization in an end-milling operation using the Taguchi design method,” *Journal of Materials Processing Technology* vol. 184, pp. 233–239, 2007, DOI: 10.1016/j.jmatprotec.2006.11.029
- [12] K. Palanikumar, “Experimental investigation and optimization in drilling GFRP composites,” *Measurement*, vol. 44, pp. 2138-2148, 2011, DOI: /10.1016/j.measurement.2011.07.023
- [13] I. Asilturk, H. Akkus, “Determining the effect of cutting parameters on surface roughness in hard turning using the Taguchi method,” *Measurement*, vol. 44, pp. 1697-1704, 2011, DOI: 10.1016/j.measurement.2011.07.003
- [14] K. R. Ranjit, “A Primer on the Taguchi Method,” *Competitive Manufacturing Series*, New York, 1990.



Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |

<http://www.saujs.sakarya.edu.tr/>

Title: The influence of canola oil biodiesel on performance, combustion characteristics and exhaust emissions of a small diesel engine

Authors: Mehmet Şen

Received: 2018-10-18 17:02:29

Revised: 2018-11-18 13:02:19

Accepted: 2018-12-15 11:14:50

Article Type: Research Article

Volume: 23

Issue: 1

Month: February

Year: 2019

Pages: 121-128

How to cite

Mehmet Şen; (2019), The influence of canola oil biodiesel on performance, combustion characteristics and exhaust emissions of a small diesel engine .

Sakarya University Journal of Science, 23(1), 121-128, DOI:

10.16984/saufenbilder.472112

Access link

<http://www.saujs.sakarya.edu.tr/issue/38708/472112>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

The influence of canola oil biodiesel on performance, combustion characteristics and exhaust emissions of a small diesel engine

Mehmet Şen^{*1}

ABSTRACT

In this work, the influence of canola oil biodiesel addition to eurodiesel was evaluated on combustion, performance characteristics and exhaust emissions of a single cylinder diesel engine. In the experiments, fuel mixtures obtained by adding 10%, 20% and 50% canola oil biodiesel (named COB10, COB20 and COB50 respectively) to eurodiesel fuel (COB0) were used. The test engine was loaded at full load with electrical dynamometer and data was recorded between 1500 rpm and 3000 rpm at 500 rpm intervals. The result show that addition of biodiesel to eurodiesel reduced cylinder pressure, engine torque and BTE while increased BSFC. Ignition delay decreased slightly with the addition of biodiesel. NO_x emission and smoke density were decreased as the biodiesel content increase in the fuel blends.

Keywords: Canola oil biodiesel, performance, combustion, exhaust emissions

1. INTRODUCTION

Biodiesel is a renewable resource and a good alternative to diesel fuel. Waste oils, used cooking oils, vegetable oils and animal fats and microalgae can be used as biodiesel raw materials. HC, CO, CO₂ and PM emissions except NO_x emissions are less emitted with the use of biodiesel in diesel engines. The lubricant feature extends engine life without any additional lubricant additives. It is also more economical as it is derived from local and renewable sources, so that it is a solution for energy demand [1]. The use of 100% biodiesel in diesel engines is not recommended because of the high density and viscosity of biodiesel which causing injector and fuel line problems [2]. For these reasons, biodiesel is usually blended with diesel fuel at various proportions without any engine modifications [3].

Canola seeds are very suitable for biodiesel production with very high oil content

(approximately 40-45%) from other oily seeds [4]. Canola biodiesel is well mixed with diesel fuel because the density difference between the canola oil biodiesel and diesel fuel (canola 883.5 kg/m³, diesel fuel 831.5 kg/m³) is not as large as the viscosity difference (canola 4.32 mm²/s, diesel fuel 2.6 mm²/s) [5]. One of the advantages of biodiesel compared to diesel fuel is oxygen content. Canola biodiesel has approximately %10 oxygen content. The high oxygen content cause the reduction in the ignition delay and exhaust emissions (PM, CO and HC) and improve the combustion [6].

There are many studies in the literature using edible vegetable oil biodiesel in diesel engines. Can et al. [7] studied the effects of canola biodiesel blends in different mixing ratios at different loads in a single-cylinder diesel engine. As a result, they reported that the canola biodiesel blends caused a shorter ignition delay for all engine loads because of the earlier combustion timing. When the ratio of canola biodiesel increased, maximum in-cylinder

* Corresponding Author

¹ Bolu Abant İzzet Baysal University, Bolu Technical Sciences Vocational School, Bolu, Turkey, sen_m@ibu.edu.tr

pressures, maximum heat release rate and brake thermal efficiency (BTE) decreased while brake specific fuel consumption (BSFC) increased. In addition, it was stated that canola biodiesel mixtures caused a decrease in smoke, HC and CO emissions and increased NO_x and CO₂ emissions.

Ozsezen et al. [8] investigated the effect of injection characteristics, performance and combustion behaviors of a direct injection diesel engine when it was operated with palm oil and canola oil methyl esters under full load. The results indicated that the brake power reduced and the BSFC increased when the test engine was operated with both methyl ester blends. Also, methyl esters caused reductions in CO, HC, CO₂ and smoke emissions. However, both methyl esters caused more NO_x emissions. Zhang and Van Gerpen [9] examined the of soybean oil methyl esters blends in aturbocharged diesel engine. The results showed that the methyl ester blends caused similar combustion characteristics as diesel fuel.

Many studies were reported in the literature on engine performance and emission characteristics of various edible vegetable oil biodiesel blends. However, the studies on the analysis of combustion are limited. The main objective of this study is to evaluate the effects of canola oil biodiesel blends on combustion, performance, and emissions.

2. MATERIALS AND METHODS

The fuel mixtures obtained by adding 10%, 20% and 50% canola oil biodiesel (named COB10, COB20 and COB50 respectively) to eurodiesel fuel (COB0) were used in the experiments. Table 1 shows some basic properties of the test fuels. The experimental set-up is shown in Figure 1.

Table 1. Important features of the test fuels.

Properties	Diesel	COB10	COB20	COB50
Density (kg/m ³)	831.5	836.7	841.9	857.5
Kinematics				
Viscosity (mm ² /s)	2.40	2.60	2.79	3.36
Lower heating value (MJ/kg)	43.20	42.73	42.26	40.85
Cetane Number	58.8	58.3	57.9	56.5

The engine was connected to a 15 kW Kemsan brand DC dynamometer. The test engine was loaded at full load and data was recorded between 1500 rpm and 3000 rpm at 500 rpm intervals. Torque measurement was made by a flange system, consisting of the torque measuring unit KiTorq rotor Type 4550A and the torque evaluation unit KiTorq stator Type 4541A. In order to measure the cylinder pressure, Kistler 6052c having a piezoelectric crystal which achieves high sensitivity has been used. The Kistler 4065B compact high pressure sensor is attached on the fuel line to detect the start of injection. Kistler 2614B type optical crank angle encoder was used to calculate the crank angle and TDC data (a resolution of 720 x 0.1° crank angle degree (CAD)). Table 2 shows the specification of the Lombardini brand 15 LD 350 model diesel engine used in the experiments.

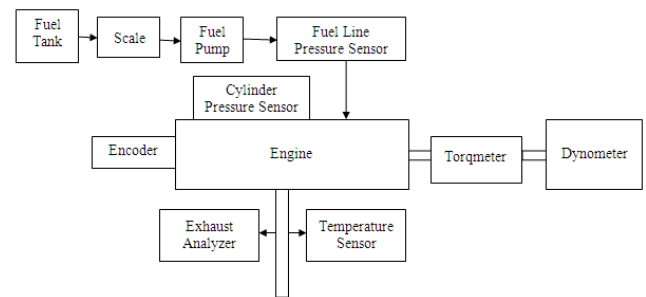


Figure 1. Schematic representation of the experimental setup

Table 2. The specification of the Lombardini 15 LD 350

Items	Specifications
Cylinder	1
Cooling	Air cooled
Maximum torque	16.6 Nm/2400 rpm
Compression ratio	20.3/1
Displacement	349 cm ³
Bore	82 mm
Stroke	66 mm
Nozzle opening pressure	207 bar
Injection pump type	QLC type

All data excluding emission values were collected with KiBox To Go data acquisition system and evaluated with KiBox Cockpit software. Mobydick 5000 was used for obtain smoke density and NO_x values. The basic feature of the Mobydick 5000 is shown in Table 3. Before collecting data for each test the engine was run for 10 minutes until it reaches operating temperature. A scale and stopwatch were used to measure fuel consumption.

Table 3. The basic feature of the exhaust analyzer

Emission	Unit	Range	Accuracy
Smoke Density	m ⁻¹	0-20	0.01
NOx	ppm	0-5000	1

The following formulas were used in the calculations:

The Engine brake power (kW):

$$P_b = \frac{M.n}{9459} \quad (1)$$

The brake specific fuel consumption (g/kWh):

$$BSFC = \frac{m_{fuel}.3600}{P_b} \quad (2)$$

The Brake Thermal Efficiency (%):

$$BTE = \frac{P_b}{m_{fuel}.LHV} \cdot 100 \quad (3)$$

The Rate of Heat Release (RoHR) (J/CAD):

$$RoRH = \frac{k}{k-1} P \frac{dV}{d\theta} + \frac{1}{1-k} V \frac{dP}{d\theta} \quad (4)$$

5% of total heat release and 90% of total heat release were taken as the start of combustion (SoC) and the end of combustion (EoC), respectively. Duration of combustion (DoC) is difference between SoC and EoC. The start of injection (SoI) is the crank angle corresponds to at which fuel line pressure reaches nozzle opening pressure (207 bar). Ignition delay (ID) is the difference between SoI and SoC.

3. RESULTS AND DISCUSSION

Figure 2 shows the torque variation of test fuels with respect to different engine speed. Unlike to gasoline engines, torque is less dependent on engine speed in diesel engines. Due to the increase in the volumetric efficiency and the decrease in the heat losses with gas leaks per stroke, the torque has increased upto 2500 rpm for all fuel. The reduction of torque at 3000 rpm depends on the decrease in volumetric efficiency and the increase of friction losses. Due to the low heating value of the biodiesel, engine torque was decreased as the amount of the biodiesel in the blend increase. Also,

the high viscosity and density of biodiesel reduce the atomization efficiency, therefore engine torque decrease [10].

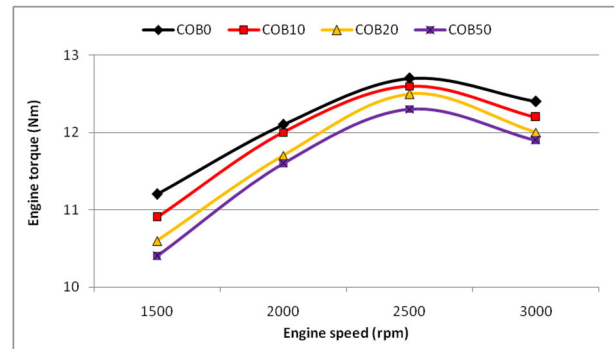


Figure 2. Engine torque graphs of the test fuels

The BSFC of test fuels are shown in Figure3. The decrease in BSFC upto 2500 rpm in all test fuels is thought to be due to increase in volumetric efficiency and decrease in gas leakage and heat loss. The most important factor affecting BSFC is heating value [11]. Low heating value of canola biodiesel increased BSFC as seen in Figure3.

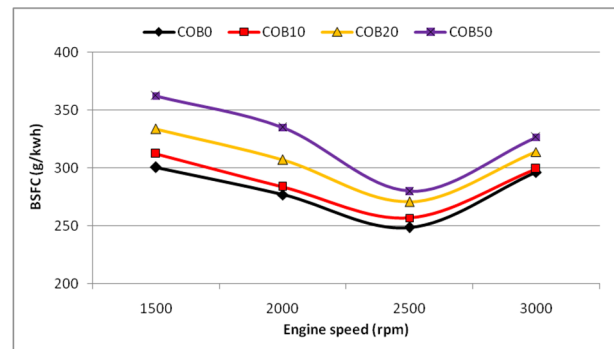


Figure 3. BSFC graphs of the test fuels

BTE is the ratio of the effective power measured at the engine output to the power to be obtained by the complete combustion of the fuel. It provides information about combustion efficiency. BTE changes of the test fuels according to different engine speed at full load are shown in Figure 4. BTE value was decreased with addition of biodiesel because biodiesel has low heating value. In spite of the high oxygen content of COB, BTE has decreased due to the high density and high viscosity of biodiesel, which worsens the fuel atomization [12]. A similar trend was observed at BTE for all test fuels at all engine speed. BTEs increased with engine speed increase, reached peak values at 2500 rpm then slightly decreased at

3000 rpm as shown in Figure 4. The BTE differences at high engine speeds were observed as smaller when compared with lower engine speeds.

Physical and chemical fuel properties such as density, viscosity, cetane number and heating value affect the combustion characteristics, engine performance and emissions. By evaluating the cylinder pressure, RoHR and ID, the combustion performance of different fuels can be compared [13].

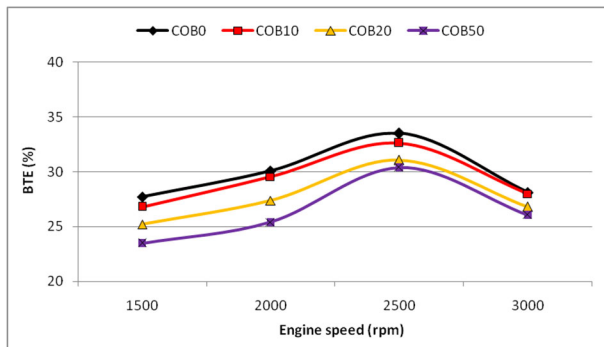


Figure 4. BTE graphs of the test fuels

The cylinder gas pressure and RoHR for diesel fuel and blends are presented in Figure 5. Because of the high heating value of diesel fuel, cylinder pressures are higher than biodiesel blends for all engine speeds as seen in Figure 5. The torque curves in Figure 2 confirm this. It is seen that the maximum pressure values of diesel fuel are higher than around 1-2 bar compared to the biodiesel blends. This is related to the higher heating value, but also to the longer ignition delay time of diesel fuel. As it can be seen more clearly in the RoHR curves, the rapid combustion of fuel accumulated in the cylinder during the ignition delay, which is called the pre-mixed combustion phase, causes the higher in-cylinder pressure and heat.

Figure 6 shows the change of SoI, SoC, and EoC of the canola oil biodiesel at four different engine revolutions. With the increase in engine speed, the premixed combustion phase decreased and the diffusion combustion duration increased because of the moving of SoC toward TDC, which caused the decrease in the maximum pressure values. In spite of similar RoHR values, the torque values at 3000 rpm are lower than those of 2500 rpm, due to increased friction losses.

In the engines with mechanical fuel pumps, the fuel line pressure increases due to the low compressibility, high density and high viscosity of biodiesel blends and thus the injector opens earlier [14]. Although the SoC for all engine speed is quite close to each other, SoC of COB50 is about 0.3-0.5 CAD earlier in comparison to the diesel fuel as seen in Figure 7.

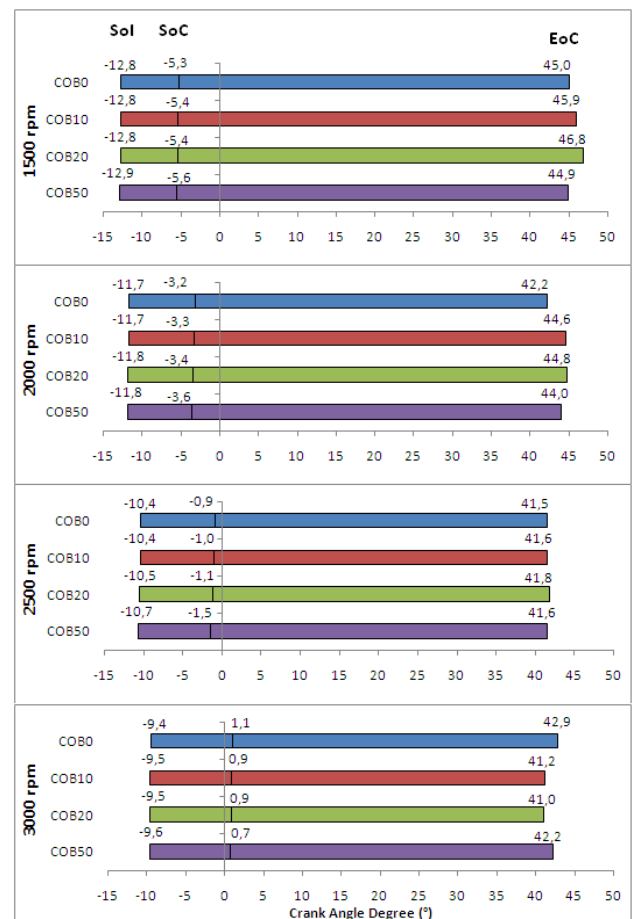


Figure 6. Change of SoI, SoC and EoC of the test fuels

The ignition delay is an important determinant for the combustion characteristics. Oxygen content, atomization of the fuel, cylinder pressure and especially cetane number influence the ignition delay. Since ID is the time between SoI and SoC, all the parameters affecting them also affect the ID. The addition of biodiesel to the eurodiesel reduced the ID slightly as shown in Figure 7. Although the cetane number of diesel is slightly higher than those of the blends, the oxygen content of biodiesel is effective in reducing the ID. In addition, the small amount of aromatic content of biodiesel is one of the factors that reduce the ID [15].

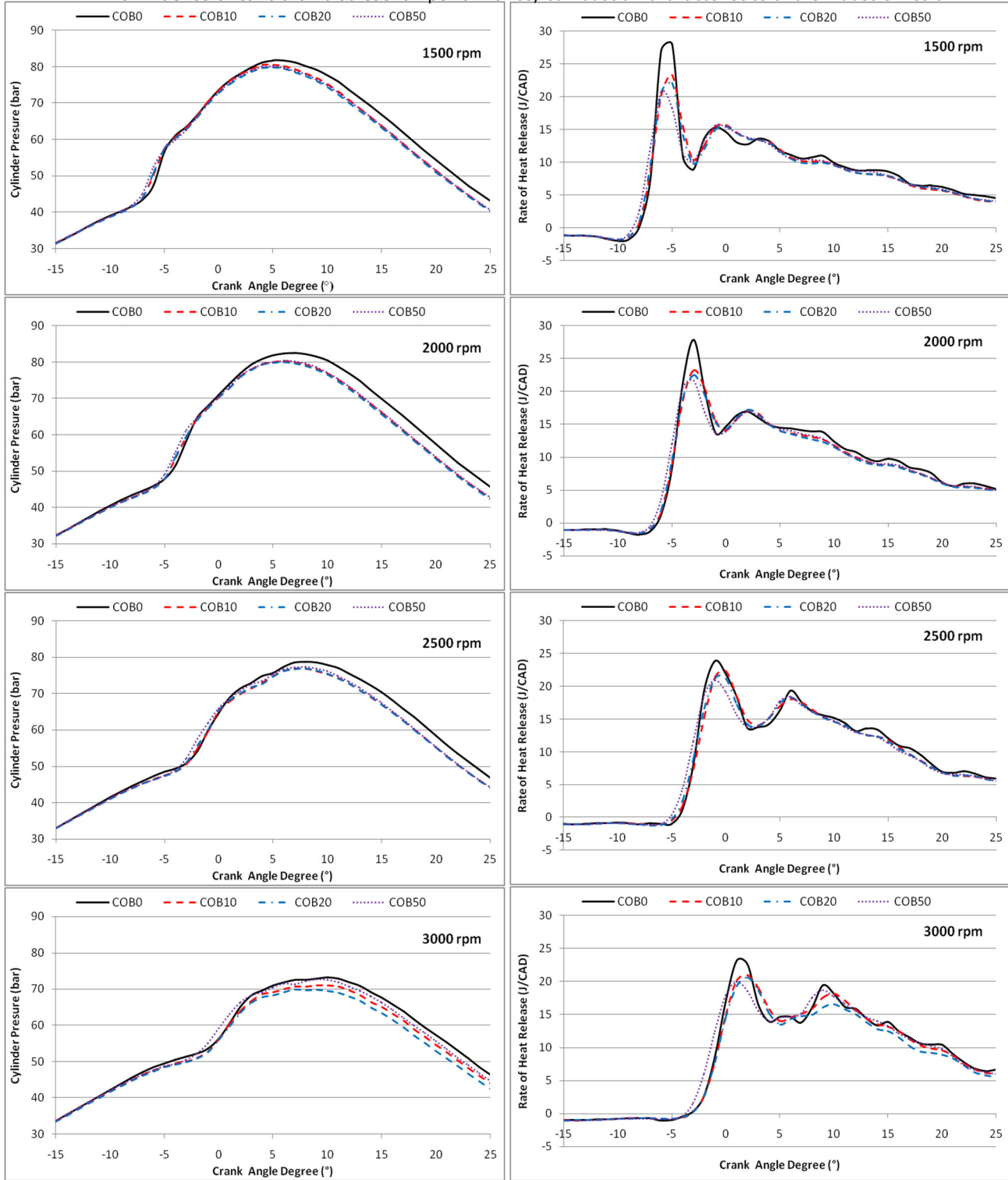


Figure 5. Cylinder pressure and RoHR graphs of the test fuels.

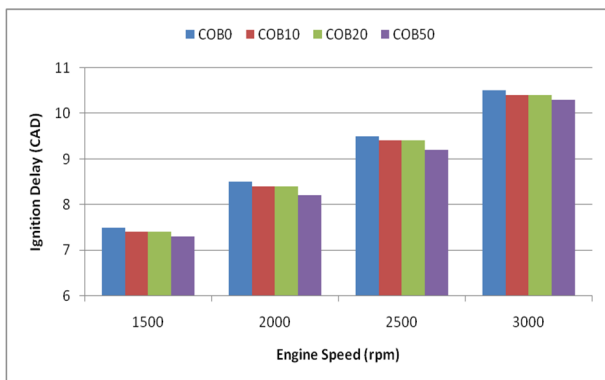


Figure 7. ID graphs of the test fuels.

Since the fuel injected into the cylinder is exposed to higher temperatures and pressures as the SoI is late, the ignition takes less time. However, with increasing engine speed, this time corresponds to a larger crank angle [16].

The change in exhaust temperature is shown in Figure 8. As the engine speed increases, frequency of combustion increases in unit time, so that the exhaust temperature increases. The increase in the diffusion combustion phase is also effective in

increasing the exhaust temperature as shown in RoHR graphics. The increase in the amount of biodiesel in the blends for each engine speed reduces the exhaust gas temperature.

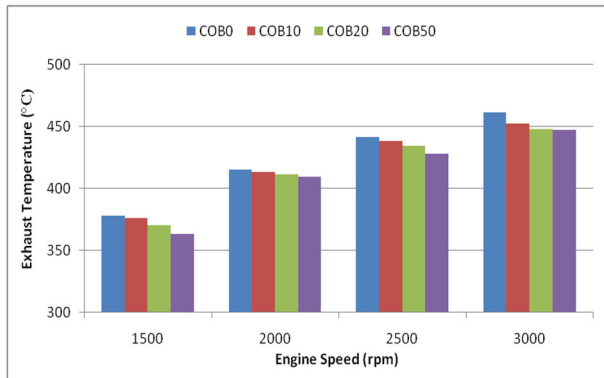


Figure 8. Exhaust temperature graphs of the test fuels

The change in NO_x emissions is shown in Figure 9. As defined in Zeldovich Mechanism, the oxygen and nitrogen in the air react with each other due to the high temperature in combustion chamber. Thus, NO_x emerge as products. The formation of NO_x depends on ignition delay, oxygen content of fuel and combustion temperature [17].

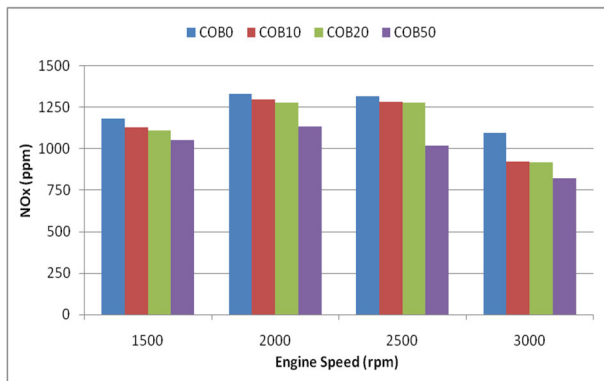


Figure 9. NO_x graphs of the test fuels

As seen in Figure 9, the concentration of NO_x was decreased as the biodiesel content increase in the fuel blends, because the low heating value of biodiesel causes in lower cylinder pressure and temperatures.

Smoke density is the coefficient of light absorption of the exhaust emission gases released by the engine. It increases with the increase of particles formed by unburned carbon atoms. Smoke density depends on the temperature of cylinder, the amount of oxygen and soot particle residence time

in the combustion chamber [18]. The change of smoke density of test fuels according to engine speed is shown in Figure 10. The addition of COB to euro diesel reduced the smoke density. Smoke density of the test fuels was lower at 1500 rpm but increased at 2000 and 2500 rpm and dropped again at 3000 rpm. As seen in Figure 10, the diesel fuel has highest smoke density at all engine speeds. COB10 comes after diesel fuel and is slightly lower than those of diesel fuel. Approximately 40% reduction was observed in COB50 according to diesel fuel for all engine speeds. The low C/H ratio and low oxygen content of canola oil biodiesel is thought to have an effect on reducing smoke density.

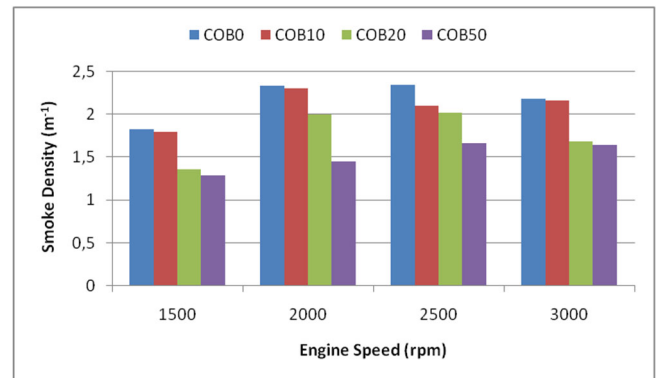


Figure 10. Smoke density graphs of the test fuels

4. CONCLUSIONS

In this study, the influence of canola oil biodiesel addition to eurodiesel was evaluated on combustion, emissions and performance characteristics of a single cylinder diesel engine at full load. In the light of the results, the following conclusions were obtained:

1. With the increase of the biodiesel in the fuel blends, engine torque decreased and BSFC increased due to the low heating value of the biodiesel.
2. Despite the high oxygen content of biodiesel, BTE decreased because of the high density and viscosity of biodiesel, which worsens fuel atomization.
3. The cylinder pressures of the diesel fuel are higher than those of the biodiesel blends for all engine speeds because of the high thermal value of diesel fuel.

4. The maximum cylinder pressure values of diesel fuel are higher than around 1-2 bar compared to the biodiesel blends.
5. The addition of COB to the eurodiesel reduced the ID slightly.
6. Because the low heating value of biodiesel causes lower cylinder pressure and temperatures, the NO_x emissions decreased with the increase the biodiesel content in the fuel blends.
7. The addition of COB to euro diesel reduced the smoke density.

REFERENCES

- [1] H.M. Mahmudul, F.Y. Hagos, R. Mamat, A.A. Adam, W.F.W. Ishak and R. Alenezi, "Production, characterization and performance of biodiesel as an alternative fuel in diesel engines—A review," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 497–509, 2017.
- [2] Y.C. Sharma, B. Singh and S.N. Upadhyay, "Advancements in Development and Characterization of Biodiesel: A Review," *Fuel*, vol 87, pp. 2355–2373, 2008.
- [3] G. Dwivedi, S. Jain and M.P. Sharma, "Diesel engine performance and emission analysis using biodiesel from various oil sources—Review," *J. Mater. Environ. Sci.*, vol. 4, pp. 434–447, 2013.
- [4] D.K. Yadava, S. Vasudev, N. Singh, T. Mohapatra, and K.V. Prabhu, "Breeding Major Oil Crops: Present Status and Future Research Needs. In Technological Innovations in Major World Oil Crops," Volume 1: Breeding; Springer: New York, NY, USA, 2012; Chapter 2; Volume XIII, 405p, ISBN 978-1-4614-0355-5.
- [5] S.K. Yoon, M.S. Kim, H.J. Kim, and N.J. Choi, "Effects of canola oil biodiesel fuel blends on combustion, performance, and emissions reduction in a common rail diesel engine," *Energies*, vol. 7, pp. 8132–8149, 2014.
- [6] D. Singh, K.S. Subramanian, M. Juneja, K. Singh and S. Singh, "Investigating the effect of fuel cetane number, oxygen content, fuel density, and engine operating variables on NO_x emissions of a heavy duty diesel engine," *Environ. Prog. Sustain. Energy*, vol. 36, pp. 214–221, 2017.
- [7] Ö. Can, E. Öztürk and H. S. Yücesu, "Combustion and exhaust emissions of canola biodiesel blends in a single cylinder DI diesel engine," *Renewable Energy*, vol. 109, pp. 73–82, 2017.
- [8] A.N. Ozsezen, and M. Canakci, "Determination of performance and combustion characteristics of a diesel engine fueled with canola and waste palm oil methyl esters," *Energy Conversion and Management*, vol. 52, no 1, pp. 108–116, 2011.
- [9] Y. Zhang and J.H. Van Gerpen, "Combustion Analysis of Esters of Soybean Oil in a Diesel Engine," *SAE paper* 960765.
- [10] A. Keskin, "Experimental investigation of the effect of cottonseed oil biodiesel eurodiesel mixtures on combustion, performance and emissions at full load," *Afyon Kocatepe University Journal of Science and Engineering*, vol. 17, pp. 797–809, 2017.
- [11] M. Şen, A.O. Emiroğlu and A. Keskin, "Production of Biodiesel from Broiler Chicken Rendering Fat and Investigation of Its Effects on Combustion, Performance, and Emissions of a Diesel Engine," *Energy & Fuels*, vol. 32, no 4, pp. 5209–5217, 2018.
- [12] M. Mohamed Musthafa, S.P. Sivapirakasam and M. Udayakumar, "Comparative studies on fly ash coated low heat rejection diesel engine on performance and emission characteristics fueled by rice bran and pongamia methyl ester and their blend with diesel," *Energy*, vol. 36, no 5, pp. 2343–2351, 2011.
- [13] A. O. Emiroğlu and M. Şen, "Combustion, performance and exhaust emission characterizations of a diesel engine operating with a ternary blend (alcohol-biodiesel-diesel fuel)," *Applied Thermal Engineering*, vol. 133, pp. 371–380, 2018.

- [14] C. Sayin, M. Gumus, and M. Canakci, "Effect of fuel injection pressure on the injection, combustion and performance characteristics of a DI diesel engine fueled with canola oil methyl esters-diesel fuel blends," *Biomass and Bioenergy*, vol. 46 pp. 435–446, 2012.
- [15] A.N. Ozsezen, M. Canakci, A. Turkcan and C. Sayin, "Performance and combustion characteristics of a DI diesel engine fueled with waste palm oil and canola oil methyl esters," *Fuel*, vol. 88, no 4, pp. 629-636, 2009.
- [16] H. An, W.M. Yang, S.K. Chou, S. K. and K.J. Chua, "Combustion and emissions characteristics of diesel engine fueled by biodiesel at partial load conditions," *Appl. Energy*, vol. 99, pp. 363–371, 2012.
- [17] A.O. Emiroğlu, A. Keskin, and M. Şen, "Experimental investigation of the effects of turkey rendering fat biodiesel on combustion, performance and exhaust emissions of a diesel engine," *Fuel*, vol. 216, pp. 266–273, 2018.
- [18] İ. Çelikten, and M.A. Arslan, "Investigation of Diesel Fuel, Rape Oil and Soybean Oil Methyl Esters Effects on a Direct Injection Diesel Engine Performance and Emissions," *Journal of The Fac. of Eng. and Arc. of Gazi University*, vol. 23, no 4, pp. 829–836, 2008.