DENİZ HARP OKULU
DENİZ BİLİMLERİ VE MÜHENDİSLİĞİ DERGİSİ
TURKISH NAVAL ACADEMY
JOURNAL OF NAVAL SCIENCE AND ENGINEERING

Deniz Bilimleri ve Mühendisliği Dergisi (Journal of Naval Science and Engineering), hakemli, uluslararası düzeyde, Nisan ve Kasım aylarında olmak üzere yılda iki kez yayımlanan, çok disiplinli bilim ve teknoloji dergisidir. Dergide; Bilgisayar, Makine, Gemi İnşa, Elektrik/Elektronik, Endüstri Mühendisliği ile Temel ve Sosyal Bilimler alanlarında bilimsel nitelikli akademik yazılara yer verilmektedir. Dergi, Türk Silahlı Kuvvetlerinin artan ihtiyaçlarına bilimsel katkı sağlamayı amaçlamaktadır. Dergide yer alan makaleler İngilizce olarak yayınlanmaktadır.

Journal of Naval Science and Engineering is a peer reviewed, international, inter-disciplinary journal in science and technology, which is published semi-annually in November and April. It publishes full research papers, review articles, technical notes, short communications, book reviews, letters to the editor, and extended versions of conference papers. Topics of interest include the technological and scientific aspects of the following areas: **computer science and engineering**, **electrical and electronics engineering**, **naval/mechanical engineering**, **naval architecture and marine engineering, industrial engineering and basic/social sciences**. The papers in the journal are published in English.

EDİTÖR / EDITOR-IN-CHIEF
Dr.Dz.Kd.Alb. Mustafa KARADENİZ, Deniz Harp Okulu Deniz Bilimleri ve Mühendisliği Enstitüsü Müdürü

YARDIMCI EDİTÖRLER / ASSISTANT EDITORS
Doç.Dr.Müh.Kd.Alb. Cüneyt EZGİ, Deniz Harp Okulu
Yrd.Doç.Dr.Müh.Yb. Veysel ALANKAYA, Deniz Harp Okulu

SORUMLU MÜDÜR / MANAGING EDITOR
Öğ.Kd.Bnb. Emrah YAMAKOĞLU

YAYIN KURULU / EDITORIAL BOARD

Dr.Dz.Kd.Alb.Mustafa KARADENİZ, Dz.H.O.
Doç.Dr.Müh.Kd.Alb. Hakan TOZAN, Dz.H.O.
Yrd.Doç. Dr.Müh.Kd.Alb. İbrahim GİRGİN, Dz.H.O.
Doç.Dr.Müh.Alb. Okan TOPÇU, Dz.H.O.
Yrd.Doç.Dr.Müh.Alb. Uğur ŞİMŞİR, Dz.H.O.
Yrd.Doç.Dr.Dz.Yb. Hakan Selim CANCA, Dz.H.O.
Yrd.Doç.Dr.Müh.Yb. Murat GÜNAL, Dz.H.O.
Yrd.Doç.Dr.Müh.Yb. Mehmet TÜRKER, Dz.H.O.
Yrd.Doç.Dr.Müh.Bnb. Fatih ERDEN, Dz.H.O.
Yrd.Doç.Dr.Müh.Bnb. Erinç DOBRUCALI Dz.H.O.
Doç.Dr.Müh.Kd.Alb. Feyzi AKAR, Dz.H.O.
Doç.Dr.Müh.Kd.Alb. Cüneyt EZGİ, Dz.H.O.
Dr.Müh.Kd.Alb. Mehmet Bilge Kaan ÖNAÇAN, Dz.H.O.
Yrd.Doç. Dr.Müh.Alb. Gürkan KOLDAŞ, Dz.H.O.
Dr.Müh.Alb. Erkan KIYAK, Dz.H.O.
Yrd.Doç.Dr.Müh.Yb. Veysel ALANKAYA, Dz.H.O.
Yrd.Doç.Dr.Müh.Yb. Tolga ÖNEL, Dz.H.O.
Doç.Dr.Müh.Bnb. Yusuf KAVURUCU, Dz.H.O.
Yrd.Doç.Dr.Müh.Bnb. Mümtaz KARATAŞ, Dz.H.O.
Yrd.Doç.Dr.Öğ.Bnb.Cengiz ERDÖNMEZ, Dz.H.O.

DANIŞMA KURULU / EDITORIAL ADVISORY BOARD
Prof.Dr.A.Arif ERGİN, Gebze Teknik Üniversitesi
Prof.Dr.Ahmet TOPUZ, Yıldız Teknik Üniversitesi
Prof.Dr.A.Ercan GEGEZ, Marmara Üniversitesi
Prof.Dr.Cemal ZEHİR, Yıldız Teknik Üniversitesi
Prof.Dr.Cengiz KAHRAMAN, İstanbul Teknik Üniversitesi
Prof.Dr.Drazan Kozak, JJ Strossmayer University of Osijek
Prof.Dr.Fevzi BELLİ, Universität Paderborn
Prof.Dr.Fuat İNCE, Marmara Üniversitesi
Prof.Dr.Hakan TEMELTAŞ, İstanbul Teknik Üniversitesi
Prof.Dr.Mete ŞEN, İstanbul Teknik Üniversitesi
Prof.Dr.Nurhan KAHYAOĞLU, Piri Reis Üniversitesi
Prof.Dr.Okyay KAYNAK, Boğaziçi Üniversitesi
Prof.Dr.Oya KALIPSIZ, Yıldız Teknik Üniversitesi
Prof.Dr.Ramazan KAYNAK, Gebze Teknik Üniversitesi
Prof.Dr.Serdar PİRTİNİ, Marmara Üniversitesi
Prof.Dr.Tunç EREM, Marmara Üniversitesi
Prof.Yahya KARSLIGİL, Beykent Üniversitesi
Doç.Dr. Sergej HLOCH, Technical University of Kosice
Doç.Dr.Elif KARSLIGİL, Yıldız Teknik Üniversitesi
Doç.Dr.Gökhan YAVUZ, Yıldız Teknik Üniversitesi
Yrd.Doç.Dr.Hv.Müh.Alb. Murat ERMİŞ, Hava Harp Okulu
Dr.Hv.Müh.Alb.Süleyman BAŞTÜRK, Hava Harp Okulu

Her hakkı saklıdır.
© 2016 Copyright by Naval Science and Engineering Institute

Deniz Bilimleri ve Mühendisliği Dergisi değişik bilimsel bakış açılarını okuyucularına sunmaktadır. Doğrudan veya dolaylı olarak ifade edilen görüşler kurumun resmi görüşleri olarak görülmemelidir.
Journal of Naval Science and Engineering presents a variety of scientific viewpoints. The views expressed or implied in this publication, unless otherwise noted, should not be interpreted as official positions of the Institution.

İrtibat Noktaları / Subscription and Corresponding Address:
Öğ.Bnb.Emrah YAMAKOĞLU
Deniz Harp Okulu Deniz Bilimleri ve Mühendisliği Enstitüsü
Tuzla 34942, Istanbul - Turkey
Phone: +90 (216) 395 26 30 / Fax: +90 (216) 395 26 58

DENİZ HARP OKULU
DENİZ BİLİMLERİ VE MÜHENDİSLİĞİ DERGİSİ

CİLT : 12 SAYI : 1 NİSAN 2016 ISSN:1304-2025

İÇİNDEKİLER / CONTENTS

TURKISH NAVAL ACADEMY
JOURNAL OF NAVAL SCIENCE AND ENGINEERING

VOLUME: 12 NUMBER: 1 APRIL 2016 ISSN: 1304-2025

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.1-18
Endüstri Mühendisliği/Industrial Engineering

DEVELOPING EFFECTIVE MANUFACTURING STRATEGIES FOR PRODUCT MIX DECISIONS VIA THEORY OF CONSTRAINTS: A CASE STUDY

Gülşen Akman[1]
Burcu Özcan[2]

[1]Kocaeli University, Industrial Engineering Department, Kocaeli, Turkey,
gulsen.akman@yahoo.com

[1]Kocaeli University, Industrial Engineering Department, Kocaeli, Turkey,
burcu.ozcan@kocaeli.edu.tr

Date of Receive: 19.01.2016 **Date of Acceptance: 26.03.2016**

ABSTRACT

One of the most important problems that today's companies face the optimum product mix decision problems to maximize their profit level. Theory of constraints is one of the methods that can be used to solve these problems. In contrast to the traditional methods, TOC ensures superiority by determining the product mix. In this study, different manufacturing strategies are analyzed in the way of total income and gross profit within a firm manufacturing three types of products. Five different manufacturing strategies are investigated. Finally, it is found that the best strategy is TOC strategy according to total income and gross profit.

Keywords: Manufacturing strategy, product mix decisions, theory of constraint,

ÖZ

Günümüzün en önemli problemlerinden biri, firmaların kâr seviyelerini en büyükleyecek olurlu ürün karması problemleri ile karşı karşıya kalmalarıdır. Kısıtlar teorisi, bu tür problemleri çözmede kullanılabilecek yöntemlerden biridir. Geleneksel yöntemlerin tersine, kısıtlar teorisi bu problemlerin çözümünde üstünlüğe sahiptir. Bu çalışmada farklı üretim stratejileri toplam gelir ve brut kar yoluyla üç tip ürün üreten bir firma için analiz edilmiştir. Beş farklı üretim stratejisi incelenmiştir. Sonuç olarak kısıtlar teori yaklaşımının toplam gelir ve brut kâra gore en iyi strateji olduğu bulunmuştur.

INTRODUCTION

The TOC demonstrated an implementation of common systems in the optimization of manufacturing. TOC utilizes the company's constrained actions to divert manufacturing and process development determinations. Companies adopted the TOC to increase efficiency and quality, and also decrease stock, lead time and cycle time (Kee & Schmidt, 2000).

TOC finds the most influential factors affecting performance in a negative way, manages and eliminates these factors to improve the performance of system (Goldratt & Cox, 1992). Main topic of this theory is that each firm has a constraint at least, and constraints should be eliminated (Ünal et al., 2005). TOC sustains to be actively used in industry because of its respectable potential to define throughput problems; to assistant as a guide to solve the throughput problems; and to produce considerable improvements in productivity and efficiency (Pegels & Watrous, 2005).

This study was performed in a manufacturing firm. In order to increase profitability, constraints emerging from production process were determined. Then five different strategies related which product mix should be used in manufacturing process were investigated. These strategies are (1) Produce the highest demanded product primarily. (2) Produce the highest priced product primarily. (3) Primarily produce product with the highest unit gross profit. (4) Produce equal quantity from each product. (5) Prioritize products according to TOC approach and then produce.

This paper constructed as following; first a literature survey about product mix decisions and TOC is performed.Then TOC and its fundementals are explained. Following this, case study in a company which produces three different types of keys is performed. Finally, paper is ended with results and discussion.

LITERATURE REVIEW

Simons Jr and Simpson III (1997) aimed "to advance the state of research on constraint scheduling in several ways". They presented the solution algorithm integrated by the Goldratt Institute in their production software. Finally, they link this algorithm to alternative methods in constraint scheduling. Bengtsson and Olhager (2000) modelled the selection of a product mix by integrating the TOC and the activity-based cost (ABC) with the capacity of production-related facilities. They demonstrated that management's power over labour and overhead resources determine that the TOC and ABC cause optimal product-mix decision, and they developed a general model of the product-mix decision. Draman et al (2002) used a simple Gedunken experiment to asses the differences between strategy-driven product mix decisions based on TOC accounting and traditional cost accounting. These decisions include four strategies; contraction, market share, product quality and cost leadership. They used these strategies to provide a guidance and direction to develop strategic decisions for product mix by using traditional and TOC approaches. Lea and Fredendall, (2002) examined how three types of management accounting systems (ABC, throughput accounting and traditional cost accounting) and two methods (TOC and linear programming) to determine product mix interaction in both the short term and the long term to influence the manufacturing performance of two shops (a flat, and a deep product structure) in a highly automated industry with significantly high overhead content. Sheu et al (2003) reviewed and compared the concepts of TOC and ABC. They discussed the managerial impacts of the two theories in various decisions such as pricing, make-or-buy and product mix. They found that ABC and TOC can provide management with better performance measurement of diverse production activities. Pegels and Watrous (2005) used TOC methodology in a detailed case study of the bottleneck in the manufacturing process. The bottleneck in this case was the mold-changing operations, consisting of a plastic injection process for heavy-duty truck-lighting systems components.

Mehra et al. (2005) compared TOC and traditional accounting performance measures in a process industry by using a simulation-based approach. Study of Wu et al. (2006) aimed to solve the product mix decision problem for a mixed-yield scenario involving the simultaneous production of high-yield and low-yield products. They developed a nonlinear mathematical program in order to model the decision problem. In another study, Karataş (2012) and Razi & Karatas (2016) used nonlinear and linear programs to determine the optimal mix and allocation of resources for a decision problem. Bhattacharya et al. (2008) outlined a fuzzy approach using fuzzy linear programming (FLP) with a suitably designed smooth logistic membership function (MF) for finding fuzziness patterns at disparate levels of satisfaction for TOC-based product mix decision problems. Utku et al (2011) examined the importance of the theory of constraints compared to the conventional cost accounting in making adequate product mix decisions. They performed an application in a chemistry industry to determine product mix decisions and their effect on profitability by comparing the theory of constraints variable costing method with the full costing method in respect to the throughput approach, the contribution margin approach and the unit profit approach respectively.

THEORY OF CONSTRAINTS

Theory of constraints (TOC) is a management philosophy which is developed by Dr. Eliyahu M. Goldratt in the beginning of 1980s. Its point of origin is that performance of each firm is determined by constraints and each firm has a few constraints at least (Ruhl, 1996). According to Atwater & Gagne (1997) a constraint is “everything limiting system performance”, and TOC is a management philosophy which focuses improvement by means of managing constraints in the system.

TOC concentrates on system improvement. A system is referred as a series of interdependent processes. A chain is an analogy for a system. The chain is a group of interdependent links working together toward the overall objective. The constraint is a weakest link in the chain. The performance of the oceralle chain is limited by the strength of the weakest link (Nave, 2009). TOC consists of five steps, these are indentify the constraints, exploit the constraint, subordinate non constraints, elevate the constraints, repeat the process with the new constraints. (TOC, 2008)

Basics of the TOC can be grouped into five principles (Womack & Flowers, 1999);Systems and processes are like chains, There is a constraint at least. The weakest link of the chain is called as a bottleneck. When the weakest link is strengthen, output of the system increases. Constraints can be classified according to reasons (such as market constraints, resource constraints, capacity constraints, political constraints etc.) Any improvement performed in non-constraint resources or processes do not influence system performance.

Huff (2001) defined TOC as a production management approach which determines and manages constraints during the manufacturing processes, and any constraint can be “anything that limits a system from achieving higher performance relative to its goal.” Basic point of TOC is that, contrary to traditional thinking, each constraint provides an improvement opportunity. TOC evaluates constraints positively, because constraints define performance of a system, and elimination of constraints step-by-step increases system performance (Spencer & Cox, 1995). Any internal situation in the firm can be a constraint or bottleneck. In this case, a constraint occurs internally in the firm such as a machine with limited capacity. Also a constraint can be derived from external situations such as a market limit that would restrict total sales (Huff, 2001).

The TOC reviews manufacturing processes and organizations as “chains”. The overall performance of the system is merely determined by the weakest link. TOC aims to define the weakest link (constraint) in manufacturing processes, and than strengthen this weakest link which is no longer the limiting factor in determining the strength of the chain (or process) (Pegels & Watrous, 2005). In manufacturing processes, TOC focused on the process that slows the speed of product through the system (Nave, 2009).

The TOC optimizes efficiency of the whole system. The optimal efficiency for the production system can be reached by defining and feeding constraints. The optimization outcomes are summarized as minimizing stocks, reducing cycle times, reducing manufacturing expenditures, and yielding more aggressive and accurate customer delivery times. Feeding a constraint becomes significantly more difficult when the process is highly reentrant, and the constraint tools are met many times in the process at different process rates (Rippenhagen & Krishnaswamy, 1998).

Before defining the constraint, two prerequisites should be ensured to acquire perspective in terms of analyses (Coman & Ronen, 1995); (1) Identify the system and its aim, (2) Specify how to measure the aim of the system.

A constraint in a production system can be anything limiting the system to achieve its goal or a desired performance level. TOC is a production management technique that identifies and manages constraints in the system. TOC has three principles for production: “(1) to increase throughput, (2) to decrease operating expenses, and (3) to decrease inventory” (Huff, 2001). When companies are “making money”, they are creating value-added at a faster rate than their expenditures rate. In order to calculate "making money," TOC uses three measures; (1) Throughput, (2) Operating expense, (3) Inventory. These measures are called as global operational measures (Moore and Scheinkopf, 1998).

(1)Throughput (T) is defined as the rate at which the organization generates money through sales or interest (Ricketta, 2007) .Firms earn money by selling their products. In order product/ services. Firms have to pay money to their vendors for materials/services required for producing products/services. Thus they create value-added for customers. This "value-added" is called as throughput (Moore and Scheinkopf, 1998). Throughput is computed by using equation 1;

$$\textit{Throughput (T)} = \textit{Sales} - \textit{Raw material costs} \qquad (1)$$

(2) Operating expense (OE) is defined as all of the money the organization spends in order to turn inventory into throughput. Operating expense includes all of the expenses that we typically accept "fixed." (Moore and Scheinkopf, 1998). It contains direct labor, rent and labor, plus selling, general and administrative costs (Ricketta, 2007)

(3) *Inventory (I)* defined as all money invested in things intended for (Ricketta, 2007). It is defined with a more comprehensive definition as "the money that the system spends on things it intends to turn into throughput or the money the system spends on things it intends to sell". (Moore and Scheinkopf, 1998). It contains totally variable costs such as material, plus resources used in production such as land, machines, trucks and computers (Ricketta, 2007).

The TOC metrics of *T*, *I*, and OE provide alignment between local actions and the global aim of fulfillment "more money now and in the future" (Berry and Smith 2005).

Additionally, three following measurements can be accepted as global measurements that are mentioned above. They contain Net Profit, ROI and cash flow (Mabin & Balderstone, 2003).

$$\textit{Net Profit (NP)} = \textit{Throughput (T)} - \textit{Operating Expenses (OE)} \qquad (2)$$

$$\textit{Return on Investment (ROI)} = \textit{Net Profit (NP)} / \textit{Inventory (I)} \qquad (3)$$

These financial performance measures are a distinct measure of money moving into and out of the manufacturing system in time. Thus, a preferred optimum operating condition for the firm is to focus on what influences the “total” system and not just what influences a “part” of the system. This situation is called as the global optimum (Moore & Scheinkopf, 1998).

THE CASE STUDY

The case study is performed within a Turkish manufacturing company. The company manufactures three types of keys as Product Z, Product F and Product T. Raw material costs for three products were determined according to the used weights of grams in a product and kilogram prices of raw materials. Prices and costs are in Turkish Liras (TL). Manufacturing operations for three products are composed of three processes: (1) plastic-injection, (2) metal-cutting and pressing, and (3) assembling. All operations are performed by three operators. Sales prices of products are; *3.68 TL* for Product Z; *1.68 TL* for Product F; 3.90 TL for Product T. Production stages and raw material costs are shown in Figure 2. Daily demands of products are; 1800 units for Product Z, 2200 units for Product T, 2400 units for Product F. Operating expenses are 10570.0 TL

Finding constraints of the system: In order to find constraint process/operation, firstly, the workloads of manufacturing processes are calculated. Then the workloads and capacities of the processes/operations are compared. Then processes which their workloads exceed their capacities are determined as constraints/bottlenecks. If number of constraints more than one, the process with the highest workload is determined as CCR.

Unit processing times of all products for each operation are presented in Table 1. All processing times determined as second and they are illustrated as sec. Product Z provided after stages of frame, button, cover, fixing, plug, pad, sub-mechanism, body and product. Total unit processing time for Product Z is 36.84 sec. Product F transforms a product after stages of frame, button, cover, piston, pad, sub-mechanism and product. Total processing times for Product F is 35.12 sec. Product T transforms a product after stages of frame, button, cover, fixing, plug, pad, sub-mechanism, body and product. Total unit processing time for Product T is 44.62 sec.

Table 1: Unit processing times for products

Product	Plastic injection (sec)							Metal cutting and pressing (sec)		Assemble	Total unit time per product
	Frame	Button	Cover	Fixing	Plug	Piston	Pad	Sub mechanism	Body	Product	
Z	3.22	3.43	1.96	1.78	4.03	-	1.92	6.32	5.02	9.16	36.84
	6.89	6.02	1.96	1.78	4.03	-	1.92	7.44	5.02	9.56	44.62
F	8.41	5.21	1.96	-	-	2.02	2.65	8.89	-	5.98	35.12

Required time or work load of each process is calculated by multiplying demand of each product and unit processing time of each processes as presented in Table 2. Processes of frame, button, sub mechanism and assemble are bottlenecks because their workloads exceeds their capacities. Assemble process is determined as CCR because it has the highest workload.

Table 2: Finding CCR Bottleneck

Product	Plastic injection							Metal cutting and pressing		Assemble
	Frame	Button	Cover	Fixing	Plug	Piston	Pad	Sub mechanism	Body	Product (sec)
Z	5796	6174	3528	3204	7254	-	3456	11376	9036	16488
T	15158	13244	4312	3916	8866	-	4224	16368	11044	21032
F	20184	12504	4704	-	-	4848	6360	21336	-	14352
Work Load (sec)	41138	31922	12544	7120	16120	4848	14040	49080	20080	51872
Capacity (sec)	28800	28800	28800	28800	28800	28800	28800	28800	28800	28800
	Bottleneck	Bottleneck						Bottleneck		Bottleneck CCR

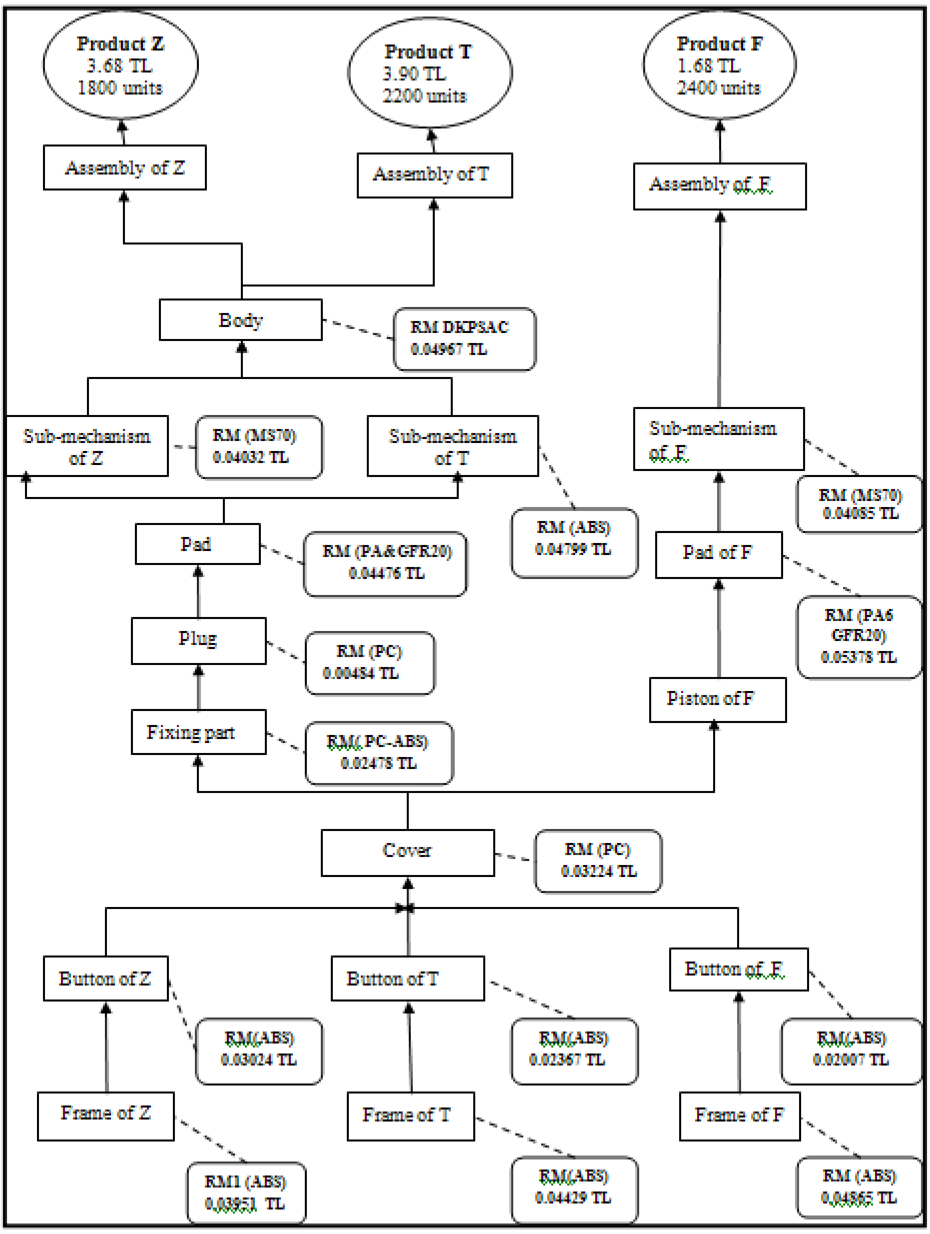


Figure 1: Process flows of Product Z, Product T and Product F

Total raw material costs are calculated according to all kinds of raw materials and their used amount for each product. Total raw material cost of Product Z is 0.26636 TL, Product T is 0.27224 TL, Product F is 0.19642 TL as shown in Table 3. Sales price of Product Z is 3.68 TL. Sales price of Product F is 1.68 TL. Sales price of Product T is 3.70 TL.

Table 3: Raw material costs for Products

	Raw material→	**ABS**	**PC**	**PC-ABS**	**PA6 GFR20**	**MS 70**	**DKP SAC**	**Unit raw material cost (TL)**
Product Z	Weight of grams	21.89	6.51	4.64	8.98	3.84	45.34	**0.26636**
	Price (TL/kg)	1.79	3.2	3	2.8	10.50	0.83	
	Item cost (TL /product)	0.09999	0.03702	0.02478	0.04476	0.04032	0.04967	
Product T	**Raw material→**	**ABS**	**PC-ABS**	**PC**	**ABS**	**PC-ABS**	**PC**	**Unit raw material cost (TL)**
	Weight of grams	21.33	4.64	6.51	8.98	4.57	45.34	**0.27224**
	Price (TL/kg)	1.79	3	3.2	2.8	10.50	0.83	
	Item cost (TL /product)	0.06798	0.02478	0.03708	0.04476	0.04799	0.04967	
Product F	**Raw material→**	**ABS**	**PC**	**Polyamide 46**	**PA 6 GFR20**	**MS 70**		**Unit raw material cost (TL)**
	Weight of grams	21.57	5.66	0.18	10.79	3.89		**0.19642**
	Price (TL/kg)	1.79	3.2	2.6	2.8	10.50		
	Item cost (TL /product)	0.06872	0.03224	0.00083	0.05378	0.04085		

APPLICATION OF DIFFERENT MANUFACTURING STRATEGIES FOR THE CASE STUDY

Companies can apply different strategies when preparing their production plans. While applying these strategies, most important considering factor is profit. Companies target maximum earnings and profit while determining their product mix. In the case study, five different strategies were investigated. These five strategies are sequenced and explained as following.

Strategy 1. Market based strategy-Produce the highest demanded product primarily.

In this strategy, manufacturing priority is given to the most demanded product. Company produces three different products; Product Z, Product F and Product T. Among these products, Product F is the most demanded product with 2400 units demand per day as presented in Table 4. In this situation all demand of Product F is produced. The rest of time is used to manufacture Product T which has second highest demand of 2200. In this time, daily production capacity of Process Assemble is 28800 sec (60x60x8). Required capacity to produce 2400 units of F is 14532 sec (2400x5.98). Rest of the capacity to produce T is 14448 sec (28800-14532). Production quantity of T is 1511 (14448/9.56). Total sales is calculated as 9926.1 TL (5894.1 + 4032.0) and, total raw material cost is calculated as 882.8 TL (Turkish Lira) (411.4 + 471.4), net income is 9043.3 TL (9926.1 - 882.8). Net profit is calculated as -1525.8 TL (9043.3 - 10570). Firm loses money.

Strategy 2. Produce the highest priced product primarily

This strategy is used by salesman while marketing products. The highest priced product is produced primarily. Product T has the highest sales price (3.9 TL). Therefore Product T is produced firstly. All demand of Product T is produced, and then respectively, Product Z with sales price of 3.78TL, Product F with sales price of 1.68 are produced as presented in Table 4. In this time, daily production capacity of assemble process is 28800 sec (60x60x8). Required capacity to produce 2200 units T is 21032 sec (2200x9.56). In rest of the capacity to produce Z is 7768 sec (28800 - 21032). Production quantity of Z is 848 (7768 / 9.16). Total sales are calculated as 11785.6 TL and, total raw material cost is calculated as 824.8 TL (225.9+598.9). Income is calculated as 10960.8 TL. Net profit is calculated as 390.8 TL (10960.8-10570.0).

Table 4: Strategies for all products

Products	STRATEGY 1				STRATEGY 2				STRATEGY 3			
	Product Z	Product T	Product F	Total	Product Z	Product T	Product F	Total	Product Z	Product T	Product F	Total
Demand	1800	2200	2400		1800	2200	2400		1800	2200	2400	
Unit Sales price(TL)	3.78	3.9	1.68		3.78	3.9	1.68		3.78	3.9	1.68	
Unit Material cost(TL)	0.26636	0.27224	0.19642		0.26636	0.27224	0.19642		0.26636	0.27224	0.19642	
Unit production time(sec) in	9.16	9.56	5.98		9.16	9.56	5.98		9.16	9.56	5.98	
Unit gross profit									3.514	3.628	1.484	
Manufacturing priority	3	2	1		2	1	3		2	1	3	
Production quantity	0	1511	2400		848	2200	0		848	2200	0	
Total production time (sec)	0	14448	14352	28800	7768	21032	0	28800	7768	21032	0	28800
Sales (TL)	0	5894.1	4032.0	9926.1	3205.6	8580	0	11785.6	3205.6	8580.0	0.0	11785.6
Raw material	0	411.4	471.4	(882.8)	225.9	598.9	0	(824.8)	225.9	598.9	0.0	(824.8)
Income(TL)	0	5482.6	3561.6	9043.3	2979.7	7981.1	0	10960.8		7981.1)	2979.7	7981.1
Operating expenses(TL)				(10570.0				(10570.0)				(10570.
Net profit(TL)				**-1525.8**				**390.8**				**390.8**

	STRATEGY 4				STRATEGY 5			
	Product Z	Product T	Product F	Total	Product Z	Product T	Product F	Total
Demand	1800	2200	2400		1800	2200	2400	
Unit Sales price(TL)	3.78	3.9	1.68		3.78	3.9	1.68	
Unit Material cost(TL)	0.26636	0.27224	0.19642		0.26636	0.27224	0.19642	
Unit production time(sec)	9.16	9.56	5.98		9.16	9.56	5.98	
Throughput					3.514	3.628	1.484	
Throughput per unit time in					0.384	0.379	0.248	
Manufacturing priority	1	1	1		1	2	3	
Production quantity	1166	1166	1166		1800	1288	0	
Total Production Time (sec)	10680.5	11146.9	6972.6	28800	16488	12312	0	28800
Sales(TL)	4407.4	4547.4	1958.9	10913.7	6804.0	5022.7	0.0	11826.7
Raw material	310.6	317.4	229.0	857.0	479.4	350.6	0.0	(830.1)
Income(TL)	4096.9	4229.9	1730.3	10056.7	6324.6	4672.1	0.0	10996.6
Operating				(10570.0				(10570.0)
Net profit(TL)				**-513.3**				**426.6**

Strategy 3. Primarily produce product with the highest unit gross profit.

In this approach, raw material costs are considered. Unit gross profits of each product are calculating by using equation 6.

$$\text{Unit Gross Profit} = \text{Unit Sales Price} - \text{Unit material cost} \qquad (6)$$

Three products are sorted according to unit gross profit values. As regard to this strategy, product with the highest unit gross profit is produced primarily. Product T is produced primarily because of that it has the highest unit gross profit of 3.62 as shown in Table 8. Firstly, all demand of Product T is produced, then, respectively, demand of Product Z with unit gross profit of 3.51, and demand of Product F with unit gross profit of 1.48. Results are same with Strategy 2 as presented in Table 4. Net profit is calculated as 390.8 TL.

Strategy 4. Produce equal quantity from each product.

In this approach, there is an assumption that balanced production is made and accordingly each product is produced equal quantity as presented in Table 4. Any sequencing is not important. According to daily capacity, 1166 (28800/ (9.16+9.56+5.98)) units from each product are produced. Total sales are calculated as 10913.7 TL. Total raw material cost is calculated as 856.5 TL. Net income is calculated as 10056.7 TL. Net profit is calculated as -513.3 TL. Firm loses money.

Strategy 5.Prioritize products according to TOC approach and then produce.

In this strategy, TOC approach is used. Unit throughput per unit time in CCR for each product is calculated by using equation 7 and equation 8 .

$$\text{Unit Throughput} = \text{Unit sales price} - \text{Unit material cost} \qquad (7)$$

$$\text{Unit throughput per unit time in CCR} = \frac{\text{Unit Throughput}}{\text{Processing time in constraint}} \qquad (8)$$

Then products are sorted according to throughput per unit time. Then Product Z with the highest throughput value (0.384) is produced primarily (Product Z), Respectively Product T (0.379), and Product F (0.248). Product Z is produced 1800 units. In this time, daily production capacity of assemble process is 28800 sec (60x60x8). Required capacity to produce 1800 units Z is 16488 sec (1800x9.16). Rest of the capacity to produce T is 12312 sec (28800-16488). Production quantity of T is 1288 (13312/9.56). Total sales is 11826.7 TL, total raw material cost is 830.1TL. Net profit is calculated as 426.6 TL.

Results of five strategies are summarized in Table 5. As seen Table 5, the best strategy is Strategy 5, TOC approach. According to these results, Application of the TOC approach provides the highest profit value of 426.6 TL when it compared with other strategies. Company should manufacture 1800 units of Z, 1288 units of T because this mix provides maximum throughput.

Table 5: Strategies and provided results

Strategy	**Sequencing**	**Z (units)**	**T (units)**	**F (units)**	**Net Profit (TL)**
Strategy1	F(2400),T(2200), Z(1800)	-	1511	2400	-1525.8
Strategy 2	T(3.90), Z(3.78), F(1.68)	848	2200	-	390.8
Strategy3	T(3.62), Z(3.51), F(1.48)	848	2200	-	390.8
Strategy 4	-	1166	1166	11666	-513.3
Strategy 5	Z (0.384), T (0.379), F (0.248)	1800	1288	-	426.6

RESULTS AND DISCUSSION

This study evaluates difference between strategy driven product mix decisions based on TOC cost approach, and traditional cost approach. This study is performed in a manufacturing company. The company produces three types of keys. In this study "assemble" process determines system's output, and it is system constraint. So system performance is related to "assemble" process. Firm developed five different strategies; Strategy 1 (Market based strategy-Produce the highest demanded product primarily); Strategy 2 (Produce the highest priced product primarily); Strategy 3

(Primarily produce product with the highest unit gross profit); Strategy 4 (Produce equal quantity from each product), and Strategy 5 (Prioritize products according to TOC approach and then produce). Five different strategies were analyzed according to revenue and gross profit. A summary of evaluation results of five strategies presented in Table 11. A review of this table leads to following conclusion. According to Strategy 1, manufacturing priority should be given to the most demanded product. In this strategy all demand which belongs to Product F is produced. It is the worst strategy for profit and revenue amongst other strategies. Strategy 2 is mostly used by salesman, in this strategy product with the highest price is produced first. Strategy 2 doesn't consider cost factor and capacity factor. Strategy 3 considers raw material costs. Strategy 3 gives importance to unit gross profit and the product with the highest unit gross profit is produced first. For both of Strategy 2 and Strategy 3, Product T is produced at first. These two strategies determine production rank of products in the same sequence. Strategy 4 insists on balanced production, and each product is produced equal quantity. If firm gives importance to product variety, it may choose this strategy. But Strategy 4 is not enough for performance targets such as higher profit and revenue,

Strategy 5 uses TOC approach. According to TOC, the constraint's capacity determines system performance. Revenue, and throughput per unit time for each product is calculated. As a result, it is the best strategy in other strategies. Obtained results from our study supports the TOC theory.

In this study, with only three products, the alternative product mix decisions are very limited and the impact of the constraint's limited capacity on alternatives can easily defined. Approach used in this study is very practical when number of different products is low. When number of different products is much larger, definition of the constraint is more difficult.

From this research, we can see how TOC based approach can affect product mix decisions. Decisions based on TOC approach produces significant effects of firm's financial performance when compared with decisions using traditional cost approaches.

TOC contains a new set of performance measures. These measures link together strategic objectives and operational capabilities. This paper uses the set of performance measures (throughput, inventory and operating expense) defined by Golddratt. Thus this paper provides support for that these measures provide a direct link between local production capabilities of the constraints and the organizational performance.

REFERENCES

Atwater, B. & Gagne, M. 1997. "The Theory of Constraints Versus Contribution Analysis for Product Mix Decisions", *Journal of Cost Management,* 11(1): 6-15.

Bengtsson, J. & Olhager, J. 2002. "Valuation of product-mix flexibility using real options", *International Journal of Production Economics*, 78 (1): 13-28.

Bengtsson, J. & Olhager, J. 2002. "The impact of the product mix on the value of flexibility", *Omega*, 30(4): 265–273.

Berry, R. a& Smith, L.B. 2005. "Conceptual foundations for The Theory of Constraints", *Human Systems Management* , 24: 83–94.

Bhattacharya, A., Vasant, P., Sarkar, B. & Mukherjee, S.K. 2008. "A fully fuzzified, intelligent theory-of-constraints product-mix decision", *International Journal of Production Research*, 46 (3):789–815.

Coman, A. & Ronen, B. 1995. "Management by Constraints: Coupling IS to Support Changes in Business Bottlenecks", *Human Systems Management,* 12: 65-70.

Draman, R.H., Lockamy III, A. & Cox III, J.F. 2002. "Constraint based accounting and its impact on organizational performance: a simulation of four common business strategies", *Integrated Manufacturing System*, 13 (4): 190-200

Goldratt, E. & Cox, J. 1992. *The Goal: A Process of Ongoing Improvement*, Great Barrington, MA, North River Press.

Huff, P. 2001. "Using Drum-Buffer-Rope Scheduling Rather Than Just-In-Time Production", *Management Accounting Quarterly*, Winter: 36-40.

Karataş, M. 2012. "Dağıtık Sualtı Sensör Ağlarının Karma Tamsayılı Doğrusal Olmayan Programlama ile Optimizasyonu" *Marmara Fen Bilimleri Dergisi*, 24(3), 77-92.

Kee, R. & Schmidt, C. 2000. "A comparative analysis of utilizing activity-based costing and the theory of constraints for making product-mix decisions", *International Journal of Production Economics*, 63 (1)*:* 1-17.

Lea, B., Lawrence, D. & Fredendall, L.D. 2002. “The impact of management accounting, product structure product mix algorithm, and planning horizon on manufacturing performance”, *International Journal of Production Economics,* 79: 279–299.

Mabin, V.J. & Balderstone, S.J. 2003. “The performance of the theory of constraints methodology”, *International Journal of Operations and Production Management*, 23(6): 568-595.

Mehra, S., Inman, R. A. & Tuite, G. 2005. “A simulation-based comparison of TOC and traditional accounting performance measures in a process industry”, *Journal of Manufacturing Technology Management*, 16(3): 328-342.

Moore, R. & Scheinkopf, L. 1998. *Theory of constraints and Lean Manufacturing: Friends or Foes?* Chesapeake Consulting, Inc, All Rights, Retrieved September 23, 2015 , from http://pm.chonnam.ac.kr/file/LeanTOC.pdf

Nave, D. 2009. “How To Compare Six Sigma, Lean and the Theory of Constraints a framework for choosing what’s best for your organization”, *Process Improvement*. Retrieved January 20, 2015, from www.lean.org/Community/Registered/ArticleDocuments/ASQStoryon QualitySigmaAndLean.pdf

Pegels, C.C. & Watrous, C. 2005. “Application of the theory of constraints to a bottleneck operation in a manufacturing plant”, *Journal of Manufacturing Technology Management*, 16(3): 302-311.

Razi, N., Karatas, M. 2016. “A multi-objective model for locating search and rescue boats” *European Journal of Operational Research*, http://dx.doi.org/10.1016/j.ejor.2016.03.026

Ricketta, J.A. 2007. *Reaching The Goal: Hor Managers Improve a Services Business Using Goldratt’s Theory of Constraints*, IBM Press. Retrieved 15 December 2015, from http://ptgmedia.pearsoncmg.com/images/9780132333122/samplechapter/0132333120_CH03.pdf

Rippenhagen, C. & Krishnaswamy, S. 1998. “Implementing the theory of constraints philosophy in highly reentrant systems”, *Proceedings of the 1998 Winter Simulation Conference*, 993-996. Retrieved 23 September 2015, from http://www.informs-cs.org/wsc98papers/134.PDF

Sheu, C., Chen, M.-H. & Kovar, S. 2003. "Integrating ABC and TOC for better manufacturing decision making", *Integrated Manufacturing System*, 14 (5): 433 - 441

Simons Jr, J.V. & Simpson III, W.P. 1997. "An exposition of multiple constraint scheduling as implemented in the goal system", *Production and Operations Management*, 6 (1): 3-22.

Spencer, M.S. & Cox, J.F. 1995. "Master Production Scheduling Development in a Theory of Constraint Environment", *Production and Inventory Management Journal*, First Quarter: 8-15.

TOC 2008. *Theory of Constraints, Technology Roadmap for the Recreational Boat Building Sector*, Retrieved April 12, 2011, from http://www.qmisolutions.com.au/uploads/files/1_8_toc.pdf

Utku, B.D., Cengiz, E. & Ersoy, A. 2011. "Comparison of the theory of constraints with the traditional cost accounting methods in respect to product mix decisions", *Journal of Doğuş University* , 12 (2): 317-331

Ünal, E., Tanış, N. & Küçüksavaş, N. 2005. "Theory of constraints and an application in a manufacturing company", *Çukurova University Journal of Social Sciences*, 14(2): 433-448.

Womack, D. & Flowers, S. 1999. "Improving system performance : a case study in the application of the theory of constraints", *Journal of Healthcare Management*, 44(5): 397-405.

Wu, M. C., Chang, W.J. & Chiou, C.W. 2006. "Product-mix decision in a mixed-yield wafer fabrication scenario", *International Journal of Advanced Manufacturing Technology*, 29: 746-752.

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.33-50
Bilgisayar Mühendisliği/Computer Engineering

HYBRID NON-REPUDIATION PROTOCOL WITH ALL TYPES OF PAIRINGS

Yusuf KAVURUCU [1]
Ömer SEVER [2]

[1] *Ass.Prof.Computer Engineering Department,*
Turkish Naval Academy, Naval Sciences and Engineering Institute, Tuzla, Istanbul
[1] *ykavurucu@dho.edu.tr*

[2] *Cryptography Department, METU, Ankara*
[2] *severomer@gmail.com*

Date of Receive: 19.01.2016 **Date of Acceptance: 31.03.2016**

ABSTRACT

As a solution to fair exchange problem, non-repudiation protocols are being widely used over digital environment. Applications of non-repudiation protocols are spreaded over Electronic Contract Signing, Certified E-mail, electronic payment and e-commerce. In this paper we present a strong fair hybrid non-repudiation protocol which works with all types of pairings. The protocol is modeled with an on-line TTP in the first round and then works optimistic in next rounds. The protocol offers stronger security by integration of Joux tri-partite key exchange and uses certificateless ID based signature and encryption methods. All the cryptographic methods used in the protocol are based on pairing based cryptography which can be implemented on all three types of pairings.

TÜM ÇİFTLER İÇİN HİBRİD İNKAR EDİLEMEZLİK PROTOKOLÜ

ÖZ

Dürüst veri alışverişi problemine çözüm olarak, inkar edilemezlik protokolleri sayısal ortamlarda yaygın olarak kullanılmaktadır. İnkar edilemezlik uygulamaları Elektronik Sözleşme İmzalanması, Sertifakalı E-posta, Elektronik Ödeme ve Elektronik Ticarette yaygınlaşmıştır. Bu çalışmada, tüm çiftler için çalışabilecek hibrid bir inkar edilemezlik protokolü sunulmaktadır. Bu protocol, ilk turda çevrimiçi TTP ile modellenmiş ve müteakip turlarda optimistik çalışacak şekilde geliştirilmiştir. Önerilen protocol, Joux üç taraflı anahtar değişimi ile entegre edilerek daha güvenli bir model sunmakta ve sertifikasız kimlik tabanlı *imza ve şifreleme teknikleri kullanmaktadır. Bu protokolde kullanılan tüm kriptografi metodları kriptografi bilimindeki teknikler üzerine geliştirilmiş ve tüm çift kombinasyonları için kullanılabilmektedir..*

Keywords: *Crytography, Non-repudiation, Security, Digiral signature*
Anahtar Kelimeler: *Kriptografi, İnkar edilemezlik, Güvenlik, Sayısal imza*

1. INTRODUCTION

Non-repudiation protocols are used for exchange of information with evidence of non-repudiation. Applications of Non-repudiation protocols are spreaded over Electronic Contract Signing, Certified E-mail, electronic payment and e-commerce.

Although there are many different types of Non-repudiation protocols such as Certified E-mail, Contract signing, fair exchange, differing in their goals; they are related with each other and share the properties Non-repudiation and fairness in common. To show these differences with an example; when non-repudiation protocol is based on message delivery like in Certified E-mail, receiver has to provide NRR in order to get the message and obtain the NRO for that message. But when non-repudiation protocol is based on exchange of evidence of non-repudiation not the message itself like in contract signing, obtaining message content is not important but exchanging signed message/contract fairly is the main goal of the application.

2. GENERAL DESCRIPTION

2.1 Non-Repudiation Protocols

Non-repudiation is defined as a security service by which the entities involved in a communication can not deny having participated, specifically, the sender can not deny having sent a message and the receiver can not deny having received a message [1].

Non-repudiation is primarily depending on asymmetric cryptography specifically to signatures which are accepted as evidences. Regarding how used in a protocol, evidence of origin supplies Non-repudiation of Origin and evidence of receipt supplies Non-repudiation of Receipt.

Non-Repudiation Protocols can satisfy various properties in different ways like:

- Fairness: Strong, weak, light
- Non-Repudiation: NRO, NRR, NRS, NRD
- State storage: Statefull, stateless
- Timeliness: Synchronous, Asynchronous
- TTP Inclusion: In-line, On-line, Off-line, Probabilistic

These properties and non-repudiation protocols have been studied in [2,3,5,6,17].

Public key cryptography is generally based on certificates binding identities with public keys which are approved by Certificate Authorities. What is different in ID-Based Cryptography is public keys are dependant on user identities and/or identifiers. This difference brings advantages and disadvantages together as discussed in [10]. The advantages of ID-Based Cryptography are mainly achieving different encryption and signature schemes like ID-Based encryption [11], blind [12], short [13], ring [14] and verifiably encrypted [15], [22] signatures which are summarized in [4]. The disadvantage of ID-Based cryptography is if the public key is dependant only on identity of a user, key generator knows the private keys of users when generation. In this work which is an expansion of [9], we used certificateless public key cryptography described in [20].

2.2 Bilinear Pairings

Pairings in elliptic curve cryptography are functions which map a pair of elliptic curve points to an element of the multiplicative group of a finite field. Below is the simple definition of a bilinear pairing , more information on pairings like Weil or Tate pairings, divisors and curve selection can be found in [6] as a summary and in [23] in more details.

Let G_1 and G_2 be additive abelian group of order q and G_3 be multiplicative group of order q, a pairing is a function

$$e : G_1 \times G_2 \rightarrow G_3 \qquad (1)$$

e is suitable for cryptographic schemes when it is an efficiently computable bilinear pairing which satisfies the following properties:

a) e is bilinear: For all P, S $\in G_1$ and Q, T $\in G_2$ we have e(P+S,Q) = e(P,Q) e(S,Q) and e(P,Q+T) = e(P,Q) e(P,T)

b) e is non-degenerate: For all P $\in G_1$, with P $\neq$ 0 there is some Q $\in G_2$ such that e(P,Q) $\neq$ 1 and for all Q $\in G_2$, with Q $\neq$ 0 there is some P $\in G_1$ such that e(P,Q) $\neq$ 1.

Consecutive properties of bilinearity are:

- e(P,0) = e(0,Q)= 1
- $e(-P,Q) = e(P,Q)^{-1} = e(P,-Q)$
- $e([a]P,Q)= e(P,Q)^{a} = e(P,[a]Q)$ for all a $\in$ Z

As an expansion to previous work [9], here we can use all three types of pairings.

3. Protocol Definition

We present an ID-based hybrid non-repudiation protocol using the Joux tri-partite key exchange scheme. Our protocol is hybrid because in the first round of exchange TTP is on-line but in the next rounds with same entities TTP works off-line. TTP in the protocol also acts as PKG. If we had used traditional ID-Based encryption and signature methods, TTP can generate and escrow private keys of all users. But in certificateless scheme of [20] users can generate their own private keys. Also revocating a disclosed or lost private key in pure ID-Based crypto systems is difficult because you have to change the corresponding public key and so the ID of that user depends on. Using schemes of [20] TTP can not escrow keys but can revocate keys easily which is important for our non-repudiation protocol depending on pairings. All the cryptographic methods used in the protocol are based on pairing based cryptography which can be implemented on all three types of pairings.

3.1 Notation

Description of notation is as follows:

A: Sender
B: Receiver
TTP: Trusted Third Party
M_i: Message labeled i; 1 - 6
Sig_X{M}: Message M signed by agent X's private key by ID-Based Signature Scheme
(M)_k: Message M symmetrically encrypted by key k
{M}_X: Message M encrypted for agent X's public key by ID-Based Encryption Scheme
S_id: Session identifier
EOO: Evidence Of Origin
EOR: Evidence Of Receipt
EOS: Evidence Of Submission of key
EOD: Evidence Of Delivery
h(M): Hash of message M
M_id: Message identifier is equal to h(h(M),S_id)
kek_sid: Key encryption key which is equal to h(e(P,Q)^x.y.z , s_id)

3.2 Protocol Description

The protocol starts with an initialization and registration at the beginning. **Initialization:** TTP generates *setup* phase shown in Section 5 and publishes system parameters G_1,G_2, G_3, e, P, Q, P_pub, Q_pub, H_1, H_2. TTP generates s $\in$ Z*q where P_pub = [s]P, Q_pub = [s]Q and keeps s secret, TTP also generates its own public key P_pub_TTP, Q_pub_TTP and corresponding private key.

Registration: A user with identity ID registers to the TTP. First TTP sends the partial key to user ID, then user ID computes public key P_pub_ID =

[X_ID]P_pub, Q_pub_ID = [X_ID]Q_pub where [X_ID] $\in Z^*_q$ and sends to TTP over authentic channel. User ID computes his private key as shown in Section \ref{Modified}.

Execution: The sender A with public key Q_pub_A, private key d_A_1 computes [x]P and [x]Q where x $\in Z^*_q$ chosen randomly for Joux tri-partite scheme. The receiver B with public key Q_pub_B, private key d_B_1 computes [y]P and [y]Q where y $\in Z^*_q$ random element. TTP with public key Q_pub_TTP, private key d_TTP_1 computes [z]P and [z]Q where z $\in Z^*_q$ chosen randomly.

For the first time of exchange between the participants A, B and TTP round 1 procedure is executed, for next exchanges with the same participants round 2 procedure is executed.

3.2.1 Online Round

Round 1 is the online mod of the hybrid protocol. Main protocol of Round 1, shown in Figure 1 below is as follows:

Step 1

A→B: M_1 = Sig_A{A,B,TTP,S_id, h(M), [x]P, [x]Q, (A,B,TTP,{M}_B)_k}

A→TTP: M_2 = Sig_A{M_1, k_TTP}

Step 2

B→A: M_3 = Sig_B{A,B,TTP,S_id, h(M),[y]P, [y]Q, (A,B,TTP,{M}_B)_k}

B→TTP: M_4 = M_3

Step 3

TTP→B: M_5 = Sig_TTP{A,B,TTP,S_id, h(M),[z]P, [z]Q, k_{kek_sid}}

TTP→A: M_6 = M_5 + M_4

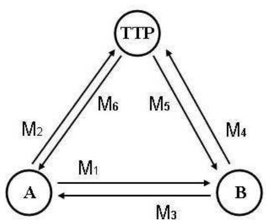


Figure 1. First Round Message Flow

Here the critical point in the protocol is the usage of signed Joux tri-partite key exchange, after the Step 3 of the Round 1, A, B and TTP has [x]P,[y]P,[z]P, [x]Q,[y]Q,[z]Q in common. This means that they can compute $e([y]P,[z]Q)^x) = e([x]P,[z]Q)^y = e([x]P,[y]Q)^z = e(P,Q)^{x.y.z}$.

The steps defined above follow previous one after some checks, as;

In Step 2 receiver B checks the identities, signature of sender A in M_1.

In Step 3 TTP checks:

First, the identities, session identifier and signature of sender A in message M_2.

Secondly, checks if the key k, which was sent in Step 1 by A is working properly. TTP decrypts the encrypted part (A,B,TTP,M_B)_k in message M_1 by the key k and checks the ID's are correct.

Thirdly, checks the identities, session identifier and signature of receiver B in message M_4 which is equal to M_3.

Finally, cross-checks the encrypted part in M_3 is same as the encrypted part in M_1.

Cancellation Sub-protocol

After Step 1 sender A can cancel the protocol by sending TTP a cancellation message. The TTP confirms the Cancellation request if the signature is valid and the request is coming from the sender of the message. The cancellation sub-protocol works as follows;

If any of these checks fail then TTP cancels the protocol. Otherwise TTP continues to Step 3, calculates the kek_sid the key encryption key which is equal to $h(e([x]P,[y]Q)^z$, s_id), encryptes the key k with kek_sid and sends the messages M_5 and M_6.

Step 1: A→B, TTP: M'_1 = Sig_A{Cancel,M_2}

Step 2: TTP→A, B: M'_2 = Sig_TTP{Cancel-Confirm,S_id,M'_1}

If A sends Cancellation request to only TTP and B sends M_3 and M_4 meanwhile, TTP gets both Cancellation request and M_4. TTP aborts the protocol in this case also. But any Cancellation request from sender after Step 3 is not accepted. Cancellation confirmation is not valid without M'_2. By this way A cannot repudiate M_1 and M_2.

After Step 1 before Step 2 receiver B can also cancel the protocol by sending TTP a Cancellation request. The TTP confirms the Cancellation request if the signature is valid and the request is coming from the receiver of the message. The Cancellation sub-protocol works as follows;

Step 1: B→A, TTP: M'_1 = Sig_B{Cancel,M_1}
Step 2: TTP→A, B: M'_2 = Sig_TTP{Cancel-Confirm,S_id,M'_1}

Dispute Resolution
After Step 2 if the receiver B did not get the key from TTP, recipient B can run Resolve sub-protocol. This is a case if the message M_3 has reached to sender, but message M_4 has not reached to TTP, because of network error or sender A blocks it as an active attack. The Resolve sub-protocol works as follows;

Step 1: B→A, TTP: M'_1 = Sig_B{Resolve,M_1,M_4}
Step 2: TTP→B: M'_2 = M_5
TTP→A: M'_3 = M_6

Before confirmation for resolve request TTP checks the same points as done in main protocol at Step 3.

3.2.2 Off-line Round
Round 2 is the off-line mod of the hybrid protocol.
After Round 1 with online TTP users can pass to Off-line TTP. A, B and TTP has [x]P,[y]P,[z]P.
A,B and TTP have previously computed $e([y]P,[z]Q)^x$), $e([x]P,[z]Q)^y$ and $e([x]P,[y]Q)^z$ respectively.
Now they can use this saved pairing for computing new kek_sid with new sid.
Main protocol of Round 2, shown in picture below is as follows:

Step 1: A→B: M_1 =

Sig_A{A,B,TTP,S_id,M_id,(M_Subj,h(M),h(M,S_id))_kek_sid},
{{A,B,TTP,S_id,{M}_kek_sid}}_TTP}

Step 2: B→A: M_2 = Sig_A{M_1, M_Subj,M_id,h(M),h(M,S_id)}

Step 3: A→B: M_3 = Sig_A{A,B,TTP,S_id,M_id,(M)_kek_sid}

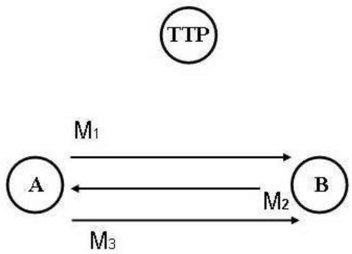


Figure 2. Second Round Message Flow

The steps defined above follow previous one after some checks, as;
In Step 2 receiver B checks the identities, signature of sender A and kek_sid is working properly by decrypting the message identifier encrypted in M_1.
In Step 3 sender A checks the identities, session identifier, signature of sender B and message subject M_Subj has been properly decrypted by B.
If any of these checks fail then TTP cancels the protocol.

Cancellation Sub-protocol
After Step 1 sender A can cancel the protocol by sending TTP a cancellation message. The TTP checks first if the signature is valid and the request is coming from the sender of the message. The TTP confirms the Cancellation request if the status of the session is not Resolved. The cancellation sub-protocol works as follows;

Step 1: A→TTP, B: M'_1 = Sig_A{Cancel,M_1}
Step 2: If (Status(S_id)==Resolved)
Step 2.a Then TTP→A: M'_2 = Sig_TTP{Cancel-Reject,S_id,M_2}
TTP→B: M'_3 = Sig_TTP{Cancel-Reject,(M)_kek_sid}
Step 2.b Else TTP→A, B: M'_2 = Sig_TTP{Cancel-Confirm,S_id,M'_1}

(Status(S_id)=Cancelled)

Dispute Resolution
After Step 2 if receiver B does not get message M_3 or the hash of the message does not match with the hash in the first message, receiver B runs Resolve sub-protocol. The Resolve sub-protocol works as follows;

Step 1: B→TTP, A: M'_1 = Sig_B{Resolve,M_1,M_2}
Step 2: If (Status(S_id)==Cancelled)
Step 2.a: Then TTP→B,A: M'_2 = Sig_TTP{Resolve-Reject,S_id, Sig_TTP{Cancel-Confirm,S_id,M'_1}}
Step 2.b: Else TTP→A,B: M'_2 = Sig_TTP{Resolve-Confirm,S_id,M_id,(M)_kek_sid,M_1,M2}
(Status(S_id)=Resolved)

4. PROTOCOL ANALYSIS

4.1 Fairness and Non-Repudiation

Proposed non-repudiation protocol satisfies fairness in both rounds. By inclusion of online TTP in first round, TTP checks the previous messages, identities, signatures and finally send complementary evidences for both sender and receiver. This achieves strong fairness at the end of the protocol as either each party gets the expexted items (NRO,NRR,Message) or none of them gets a valuable information. If the sender denies, having sent a message M, the receiver can show NRO = M_1 + M_6 and adjudicator rejects denial unless the protocol is cancelled by TTP. In case of cancellation, the sender should show a confirmed cancellation.
If the receiver denies, having received a message M, the sender can show NRR = M_3 + M_5 and adjudicator rejects denial unless the protocol is cancelled by TTP. In case of cancellation, the receiver should show a confirmed cancellation. Since *Cancellation* requests after Step 2 is not accepted, cancellation confirmation and messages M_5 and M_6 can not be present at same time.

For the second round, strong fairness is achieved by help of dispute resolution sub-protocols. Dishonest users can try to get non-repudiation evidences hindering other party to get respective evidence. As a case for dishonest sender; after Step 2, A gets successfully EOR, but can misbehave as sending a cancellation request before a resolve request. In this case since the exchange will be cancelled by TTP and confirmation of cancellation is sent to both parties. Receiever can show to adjudicator that the exchange with S_id is cancelled and EOR in his message M_2 is not valid anymore. As a case for dishonest receiver; after Step 1, B gets successfully EOO, but can misbehave as sending a resolve request to only TTP. In this case the TTP will resolve the issue only if user B sends valid EOR, and this EOR in M_2 will be forwarded to sender A also.

4.2 Timeliness

Asynchronous timeliness is achieved in the proposed protocol by means of cancellation sub-protocols without any time constraint.

4.3 TTP State

TTP works in a statefull manner as has to keeps states of protocol with respect to session identifiers. TTP also keeps securely keys for respective participating parties.

4.4 Efficiency and Comparison

The communication and computation bottleneck of the protocol is TTP for the first round. Since TTP in our protocol acts also as PKG, this situation naturally increased the burden of TTP. But this is not a necessity, PKG and TTP can be different. In that case users should get both PKG parameters and TTP pairing parameters which requires two registration. For the next rounds pairing computations on both sides seems as the reason of computational burden when compared to traditional PKI signatures and encryption.
The proposed protocol has inevitably common characteristics with previously proposed non-repudiation protocols stated in [7], [8] and [2]. It satisfies the

required properties as NRO, NRR, strong fairness and asynchronous timeliness but lacks in efficiency because of pairing computation, online TTP and statefull structure.

The advantage of using a hybrid protocol over other types (pure in-line, on-line or offline) is a kind of optimization between the security and performance. First online round embedded with Joux Tri-partite key exchange scheme enhances the security and next rounds give better performance as being off-line. Our new design does not contribute new capabilities over previous protocols at the moment but it shows that non-repudiation protocols can be built on pairing based cryptography and it is possible to extend this work by using unique properties of identity based cryptography.

4.5 Key escrow and Revocation

Generally key escrow is accepted as a positive capability for authorized third party to gain access to keys needed to decrypt encrypted data. But from view of non-repudiation key escrow property of full Identity-based cryptosystems is regarded as a negative capability. That is why we used identifier based encryption and signature schemes (Certificateless PKC) and TTP can not hold an escrow capability for the private keys of users A and B as stated in [20]. Key revocation can not be handled properly by PKG in a full Identity-based cryptosystem. But by using identifier based encryption and signature schemes (Certificateless PKC) this problem is also eliminated.

4.6 Confidentiality

Confidentiality of the message is ensured in both rounds against eavesdroppers. In the first round message is kept secret even to TTP, but in the second round message can be decrypted by TTP if the cancellation or dispute resolution sub-protocols executed. This property is inserted to improve the efficiency and generally TTP will not be joining the communication. If required this property can be changed as done in the first round.

5. CERTIFICATELESS ID-BASED SIGNATURE AND ENCRYPTION SCHEME

ID-Based signature verification and encryption schemes use publicly known variable such as identity or e-mail of a user to derive public key without any key distribution for public keys. For signing and decrypting user contacts to a Private Key Generator (PKG, CA etc.) to derive the private key which is dependant on the identity and master key of the PKG.

This scheme has some disadvantages stated in [4]

- The PKG can calculate users private keys which is a problem for confidentiality in non-rep protocols
- User has to authenticate himself to PKG
- PKG needs a secure channel to send users private key
- User has to publish PKG's public parameters

To ensure non-repudiation in our protocol we modified and used Riyami and Paterson's certificateless ID-Based encryption and signature schemes described in [20] to eliminate some of these disadvantages.

The original work of Riyami and Paterson's certificateless ID-Based encryption and signature schemes are based on only Type-I pairings. Since Type-I pairings are susceptible to recent quasi-polynomial attacks [26], [27], here we expanded their certificateless PKC to Type-II and Type-III pairings. Here we present our modification to their work.

The *setup* phase is same for both encryption and signature scheme:

Setup: Let G_1 and G_2 be additive group of prime order q and G_3 be multiplicative group of prime order q. Choose an arbitrary generator P $\in$ G_1 and Q $\in$ G_2 and a random secret master key s $\in Z_q^*$.
Set P_pub = [s]P and Q_pub = [s]Q choose cryptographic hash functions H_1 : {0,1}* $\rightarrow$ G_1 and H_3 : G_3 $\rightarrow$ {0,1}*. Public and private key pair for user ID is computed as follows:

TTP or PKG computes P_pub = [s]P, Q_pub = [s]Q and [s]H_1(ID), then send to user ID.

User ID computes P_pub_ID = [X_ID][s]P, Q_pub_ID = [X_ID][s]Q and send as public keys then computes d_ID_1= [X_ID][s]H_1(ID) as private key. Our scheme does not need to compute [s]H_2(ID) and d_ID_2= [X_ID][s]H_2(ID) and thus does not need a hash function to G_2 such that H_2 : {0,1}* → G_2 . This gives us the ability to use Type-II pairings.

5.1 Certificateless ID-Based Encryption

We adapted Riyami and Paterson [20] ID-Based Encryption Scheme to all types of pairings.

5.1.1 Encryption

- First choose a random $r \in Z_q^*$
- Message M encrypted by symmetric key k which is ciphered as C = < [r]Q, k ⊕ H_3(g_ID)^r > where g_ID = e(H_1(ID),Q_pub_ID)

5.1.2 Decryption

C=< U,V> compute k as k = V ⊕ H_3(e(d_ID_1,U))

5.1.3 Proof of Decryption

Decryption works because;

V ⊕ H_3(e(d_ID,U))

= V ⊕ H_3(e(d_ID,[r]Q))

= V ⊕ H_3(e([X_ID][s]H_1(ID),[r]Q))

= V ⊕ H_3(e(H_1(ID),Q)$^{X\_ID \cdot s \cdot r}$

= V ⊕ H_3(e(H_1(ID),[X_ID][s]Q)^r

= V ⊕ H_3(g_ID)^r

5.2 Certificateless ID-Based Signature

We also adapted Riyami and Paterson [20] ID-Based Signature Scheme to all types of pairings.

5.2.1 Signature

For signing message M user ID, chooses an arbitrary $P\_1 \in G^*\_1$ and a random $k \in Z_q^*$

First compute $r = e(P\_1,Q)^k$

$v = H(M,r)$

$u = [v]d\_ID\_1 + [k]P\_1$

The signature is the pair $<u,v> \in <G\_1, Z_q>$

5.2.2 Verification

When receiving a message M and signature $<u,v> \in <G\_1, Zq>$ verifier computes

$r = e(u,Q) \cdot e(H\_1(ID), -P\_pub\_ID\_2)^v$

Accept the signature iff $v = H(M,r)$

5.2.3 Proof of Verification

Check if $r = e(P\_1,Q)^k$:

$r = e(u,Q) \cdot e(H\_1(ID), - Q\_pub\_ID)^v$

$= e([v]d\_ID\_1+[k]P\_1,Q) \cdot e(H\_1(ID), - Q\_pub\_ID)^v$

$= e([v][X\_ID][s]H\_1(ID) +[k]P\_1,Q) \cdot e(H\_1(ID), - Q\_pub\_ID)^v$

$= e([v][X\_ID][s]H\_1(ID),Q) \cdot e([k]P\_1,Q) \cdot e(H\_1(ID), - Q\_pub\_ID)^v$

$= e(H\_1(ID),Q)^{v.X\_ID.s} \cdot e([k]P\_1,Q) \cdot e(H\_1(ID), - [X\_ID][s]Q)^v$

$= e(H\_1(ID),Q)^{v.X\_ID.s} \cdot e([k]P\_1,Q) \cdot e(H\_1(ID), Q)^{-X\_ID.s.v}$

$= e([k]P\_1,Q)$

$= e(P\_1,Q)^k$

6. CONCLUSION

We proposed a non-repudiation protocol which has new structure based on pairing based cryptography. The hybrid structure consists of two rounds described in previous sections, first round runs with an online TTP then second and next rounds run with offline TTP. Although online TTP has been regarded as a bottle-neck for security protocols, this is not a big challenge nowadays with usage of high available servers and broad band internet connection. Our main contribution here is the modification of certificateless PKC to all types of pairings. Previous works on non-repudiation protocols have used pairing based cryptography to take advantages of different properties but they also used traditional PKI for encryption and signatures. Differently our protocol is fully based on pairing based cryptography, especially certificateless ID based encryption and signature schemes which prevents some problems of pure ID-based systems.

REFERENCES

[1] NIST Glossary of Key Information Security Terms, FIPS 191.

[2] S.Kremer, O.Markowitch, J.Zhou An Intensive Survey of Non repudiation Protocols, Computer Communications 25 (2002)1606-1621, 2002.

[3] J.L.F.Gomilla, J.A.Onieva, M.Payeras Certified Electronic Mail: Propersties Revisited, Computer & Security (2009) 1-13, 2009.

[4] R.Dutta, P.Barua, P.Sarkar :Pairing Based Cryptography: A Survey, 2004.

[5] C.Calik, O.Sever, H.M.Yildirim, Z.Yuce :A Survey of Certified Electronic Mail Protocols. 4th ISC Turkey, 2010.

[6] S.Akleylek, B.B.Kirlar, O.Sever, Z.Yuce, :Pairing Based Cryptography: A Survey. 3rd ISC Turkey, 2008.

[7] C.Galdi, R.Giordano :Certified email with temporal authentication: An improved optimistic protocol. Proceedings of International Conference on Trust

and Privacy in Digital Business (TrustBus04), LNCS, vol.3184, Springer, Berlin, 2004, pp.181-190.

[8] R.Oppliger, P.Stadlin :A certified mail system (CMS) for the Internet, Comput.Commun.27 2004 1229-1235.

[9] Ö.Sever, E.Akyıldız, Hybrid Non-Repudiation Protocol with Pairing Based Cryptography, 3rd ISDFS, 9-12 May 2015

[10] M.Franklin, G.Price :A comparison between traditional Public Key Infrastructures and Identity-Based Cryptography, 2002.

[11] D.Boneh, M.Franklin :Identity Based Encryption from Weil Pairing. SIAM J.of Computing Vol.32 No.3, 2003, Extended Abstract in Crypto 2001.

[12] A.Boldyreva, :Efficient Threshold Signature, Multisignature and Blind Signature Schemes Based on the Gap-Diffie-Hellman-Group Signature Scheme. PKC 2003,LNCS 2139, pp.31-46 Springer-Verlag 2003.

[13] D.Boneh, B.Lynn, H.Shacham. :Short Signatures from the Weil Pairing. in Proceedings of Asiacrypt 2001.

[14] F.Zhang, K.Kim. :ID-Based Blind Signature and Ring Signature from Pairings. Advances in Cryptology in AsiaCrypt 2002, LNCS Vol.2510, Springer-Verlag, 2002.

[15] F.Zhang, R.Safavi-Naini, W.Susilo. :Efficient Verifiably Encrypted Signature and Partially Blind Signature from Bilinear Pairings. In Proceedings of IndoCrypt 2003, Springer-Verlag, 2003.

[16] F.Hess, :Efficient Identity Based Signature Schemes Based on Pairings, SAC 2002, LNCS 2595 \relax Springer Verlag, 2000.

[17] J.A.Onieva, J.Zhou and J.Lopez :Multi-Party Non-Repudiation: A Survey , ACM Computing Surveys, 2008.

[18] C.Galdi, R.Giordano :Certified E-mail with temporal authentication: An improved optimistic protocol, LNCS Vol.3184, 2004.

[19] A.Joux :One Round Protocol for Tripartite Diffie Hellman, LNCS Vol.1838, 2000.

[20] S.S.Al-Riyami, K.G.Paterson :Certificateless Public Key Cryptography, AsiaCrypt 2003.

[21] C.Bamboriya, S.R.Yadav :A Survey of Different Contract Signing Protocols, Ijetae V.1, I:4, January 2014.

[22] L.Chen, C.Gu :Optimistic Contract Signing Protocol Based on Hybrid Verifiably Encrypted Signature, Advances in Information Sciences and Service Sciences(AISS) V.4, N:12, July 2012.

[23] I.Blake, G.Seroussi, N.Smart :Advances in Elliptic Curves in Cryptography Number 317 in London Mathematical Society Lecture Note Series. Cambridge University Press. ISBN 0-521-60415-X, 2005.

[24] S.D.Galbraith, K.G.Paterson, N.P.Smart :Pairings for Cryptographers Elsevier 2008, Cryptology ePrint Archive, Report 2006/165.

[25] E.R.Verheul :Evidence that XTR is more secure than supersingular elliptic curve cryptosystems. in EuroCrypt 2001, 195-210.

[26] R.Barbulescu, P.Gaudry, A.Joux, E.Tomme. A Quasi-polynomial Algorithm for Discrete Logarithm in Finite Fields of Small Characteristic in EuroCrypt 2014.

[27] R.Granger, T,Kleinjung, J.Zumbragel. Breaking 128 bit Secure Binary Curves

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.19-32
Endüstri Mühendisliği/Industrial Engineering

A RICH MIN-MAX VEHICLE ROUTING PROBLEM

Ertan YAKICI

Industrial Engineering Department,
Turkish Naval Academy, Naval Sciences and Engineering Institute, Tuzla, Istanbul
eyakici@dho.edu.tr

Date of Receive: 14.01.2016 **Date of Acceptance: 28.03.2016**

ABSTRACT

We present a new variant of vehicle routing problem with a min-max objective function. The problem has different types of service demands satisfied by a heterogeneous fleet of vehicles. Unlimited service capacitated vehicles serve the demand points with multiple time windows and requirement of operation synchronization when demand is split between vehicles. A mixed integer linear programming based heuristic solution approach is proposed and a numerical study is carried out to assess the performance of the proposed method.

ÖZ

Bu çalışmada min-max amaç fonksiyonlu yeni bir çeşit araç rotalama problemi sunulmaktadır. Problemde farklı tipte araçlardan oluşan bir filo ile farklı türde hizmet talepleri karşılanmaktadır. Hizmet kapasitesi sınırsız olan araçlar, birden fazla zaman pencereli ve talebin araçlar arasında bölünerek karşılanması halinde operasyon senkronizasyonu gerektiren talep noktalarına hizmet sağlamaktadır. Problemin çözümü için tamsayılı doğrusal programlama tabanlı bir sezgisel algoritma yaklaşımı önerilmiş ve önerilen metodun persformansını değerlendirmek için sayısal bir çalışma yapılmıştır.

Keywords: *Vehicle Routing Problem; Heuristic Method; Synchronization.*
Anahtar Kelimeler: *Araç Rotalama Problemi; Sezgisel Yöntemler; Senkronizasyon.*

1. INTRODUCTION

Since the first introduction of Vehicle Routing Problem (VRP) by Dantzig and Ramser [1], a huge literature is created on VRP and its several extensions. Like different constraint environments, objective environment also may have different focuses like minimization of the maximum or average arrival time. Those objective functions is often seen in the nonprofit environments like military operations, disaster relief and humanitarian logistics (Duran et al. [2], Yakıcı and Karasakal [3], Renkli and Duran [4]). Naval operations like logistic support of warfare fleets, mine sweeping or hunting operations is the basic motivation of this research. Type of logistic support or operation, platform eligibility and other characteristics like operation speed (or economic speed), desired support or operation periods comprise several restrictions on the problem.

In the research problem presented in this work, there is a fleet that consists of vehicles having varying operational capabilities. Some vehicles are capable of conducting one of the service types, while others can operate in multiple types of service. Similarly, vehicles have different operation and transfer speeds. It is assumed that, service demands can be satisfied by one or more vehicles, demand points have multiple time windows denoting allowed periods for vehicles to arrive and all of the service demanded by each demand location should be satisfied by vehicle/vehicles arriving at the same time window. To the best of our knowledge, although there are very close problems, our problem is never introduced before in the literature.

A closely related problem is introduced in the work of Yakıcı and Karasakal [3] where various type of customer demands are satisfied by a heterogeneous fleet and split delivery is allowed. However, there are important differences between this problem and our research problem. In the problem presented in Yakıcı and Karasakal [3], the demand points are grouped in regions and the vehicles are not allowed to change their regions once they are assigned, while there is no such restriction in our problem. The other difference which makes our problem more generalized and complicated is the existence of realistic operational limitations, namely multiple time windows and operation synchronization, which are not handled in Yakıcı and Karasakal [3].

Temporal constraints are encountered in rich VRP problems frequently. However, so far, the interdependency of vehicles attracted less attention. The interdependency between the vehicles in operation synchronization is tedious to handle, especially in developing heuristic approaches. Several authors try to overcome these issue by allowing infeasible solutions, evaluating the cost approximately and using indirect search (Drexl [5]).

The closest solution technique to the one we adopt is presented in the research article of Bredström and Rönnqvist [6]. They introduce a combined VRP and scheduling problem with time windows and temporal constraints and propose an optimization based heuristic to solve real instances. In order to have a small branch and bound tree they restrict the problem significantly in number of variables and constraints, then they try to improve the best known feasible solution.

The heuristic approach we propose is based on the idea of sequential insertion of vertices and route construction by restricted model solved with a standard ILP solver.

In the following sections, the problem is introduced with its mathematical formulation, the proposed heuristic approach is described and the results of numerical experiments are reported. Finally, conclusion is given in the last section.

2. PROBLEM DEFINITION

In this section, we present the notation and the definition of our problem. The total of transfer and service time of a vehicle leaving central vertex and visiting a given sequence of demand points is referred as “travel time”. Note that since the operation is considered to finish with the satisfaction of all service demands, time elapsed in returning to central vertex is not included in the travel time. Therefore, distance from any vertex to central vertex is considered to be zero.

The following sets and parameters are used: $K = (1,...,|K|)$: set of uncapacitated vehicles. $G = (V,E)$: complete directed graph with set of vertices $V = (1,...,n)$ and set of arcs E consisting of arcs (i,j) where i and $j \in V$. $P = (p_{ij})$: distance matrix between all pairs of vertices in V. $S = (s_k)$: vector of transfer speeds of vehicles in K. $Q = (q_i)$: vector of service requirements demanded by vertices in *V*. $D = (1,...,|D|)$: set of service types. $R = (r_{kd})$: matrix of service rates (delivered service in unit time) of the vehicles in *K* for type of demands in *D*. K_d: set of uncapacitated vehicles which can serve *d* type of service. V_d: set of vertices which demand *d* type of service. *Wi*: set of time windows for vertex *i*. u_{iw} : Latest time of time window *w* for any vehicle to arrive at vertex *i*. l_{iw} : Earliest time of time window *w* for any vehicle to arrive at vertex *i*. *M*: a big number. ε: a small number.

In order to formulate the problem as an Mixed Integer Program (MIP), let $\mathbb{R}$ and $\mathcal{B}$ denote the set of real numbers and {0, 1}, respectively. The decision variables are described as follows: $x_{kij} \in \mathcal{B}$ takes value 1 if edge (*i*,*j*) belongs to the k^{th} route; 0 otherwise. $y_{kij} \in \mathbb{R}$ denotes the maximum travel time. $T_k \in \mathbb{R}$ denotes the service time of vehicle k at vertex *i*. $A_{ki} \in \mathbb{R}$ denotes the arrival time of vehicle k at vertex *i*. $tw_{kiw} \in \mathcal{B}$ takes value 1 if vehicle k arrives at vertex *i* at time window *w*. $twn_{iw} \in \mathcal{B}$ takes value 1 if vertex *i* is visited at time window *w*. Only one time window is allowed for each vertex in the context of operation synchronization. $I_{ki} \in \mathbb{R}$ is a dummy variable used for subtour elimination.

$$\min y \quad (1)$$

$$A_{ki} + T_{ki} \leq y \qquad \forall k \in K \quad (2)$$

$$\sum_{k \in K_d} \frac{r_{kd}}{q_i} T_{ki} \geq 1 \qquad \forall d \in D, \forall i \in V_d \quad (3)$$

$$\sum_{i \in V} x_{k0i} \leq 1 \qquad \forall k \in K \quad (4)$$

$$\sum_{i \in V} x_{kij} = \sum_{i \in V} x_{kji} \qquad \forall k \in K, \forall j \in V \quad (5)$$

$$I_{ki} - I_k + (M+1)x_{kij} \leq M \qquad \forall k \in K, \forall i \in V \setminus (0), \forall j \in V \setminus (0) \quad (6)$$

$$\sum_{i \in V} \frac{q_i}{r_{kd}} x_{kij} \geq T_{kj} \qquad \forall d \in D, \forall k \in K_d, \forall j \in V_d \tag{7}$$

$$A_{ki} \leq \left(1 - \sum_{i \in V} x_{kji}\right)M + u_{iw} tw_{kiw} + (1 - tw_{kiw})M$$
$$\forall k \in K, \forall i \in V \setminus (0), \forall j \in V \setminus (0), \forall w \in W_i \tag{8}$$

$$A_{ki} \geq \left(\sum_{i \in V} x_{kji} - 1\right)M + l_{iw} tw_{kiw} + (tw_{kiw} - 1)M$$
$$\forall k \in K, \forall i \in V \setminus (0), \forall j \in V \setminus (0), \forall w \in W_i \tag{9}$$

$$A_{kj} \geq A_{ki} + \frac{p_{ij}}{s_k} x_{kij} + \left(x_{kij} - 1\right)M + T_{ki}$$
$$\forall d \in D, \forall k \in K_d, \forall j \in V_d, \forall i \in V \tag{10}$$

$$\sum_{w \in W_i} tw_{kiw} \geq \sum_{j \in V} x_{kji} \qquad \forall k \in K, \forall i \in V \setminus (0) \tag{11}$$

$$twn_{iw} M \geq \sum_{v \in K} tw_{kiw} \qquad \forall i \in V \setminus (0), \forall w \in W_i \tag{12}$$

$$\sum_{w \in W_i} twn_{iw} \leq 1 \qquad \forall i \in V \setminus (0) \tag{13}$$

The objective function (1) minimizes the maximum travel time. Constraint (2) enforces the objective function value to be equal or greater than the largest travel time. Constraint set (3) ensures that all of the demands are satisfied. Constraints (4-6) ensure that each vehicle is assigned to at most only one route and each route starts and ends at central vertex without multiple visits to vertices. Constraint (7) connects the decision variables related to routing and servicing by ensuring that the vehicles should visit the vertices where they serve. Constraints (8) and (9) enforce the arrivals to occur within time windows. Constraint (10) enforce each arrival satisfies the elapsed time by previous travel. Constraint (11) connects the decision variables related to routing and time windows. Constraint (12) and (13) serve to ensure the operation synchronization by enforcing each vertex can be visited by the vehicles in only one of its time windows.

Note that still we can reduce the size of feasible space by demand type cuts (14) which prevents the vehicles from visiting the demand points that they cannot serve.

$$\sum_{k\in(K\setminus K_d),i\in V} x_{kij} = 0 \qquad \forall d \in D, \forall j \in V_d \qquad (14)$$

3. HEURISTIC SOLUTION APPROACH

The heuristic approach is based on the idea of restriction of the MIP model. Instead of solving the problem as a whole, we solve the problem in steps. At each step, a group of vertices are added to the model until all of the vertices are covered. The restricted problem is given a certain period of time for each step. The traversed arcs of the best solution found is carried to the next step. Each vehicle can break at most a limited number of arcs from its own route where some (or all) of the carried arcs are allowed to be broken in order to visit the vertices added. Another restriction is applied once in a while after a certain number of steps are taken in between. This restriction fixes all of the previously traversed arcs and service that are already processed in the visited vertices. Therefore, the problem restarts again with the remaining vertices. We call these steps "fixing steps". After a fixing step, a vehicle starts from the vertex it visits just before returning to the central vertex and from the time it finishes its service in that last visited vertex. In order to apply the restrictions, the following constraints are added to the model.

$$x_{kij} \geq f_{kij} - e_{kij} \qquad \forall k \in K, \forall i,j \in V \qquad (15)$$

$$\sum_{i,j\in V} e_{kij} \leq 1 \qquad \forall k \in K \qquad (16)$$

$$e_{kij} \leq a_{kij} \qquad \forall k \in K, \forall i,j \in V \qquad (17)$$

$$A_{k0} \leq A_{k\ level} \qquad \forall k \in K \qquad (18)$$

The binary parameters f_{kij}, a_{kij} and nonnegative parameter $A_{k\ level}$ respectively denote whether the arc is traversed at the previous step (= 1), whether it is allowed to be broken (= 1) at the current step and the time when vehicle k completes its job just before returning to central vertex in the last fixing step. The binary decision variable e_{kij} takes positive value if the arc (i,j) is broken by vehicle k. The last constraint (18) forces vehicle k to start at the time given by $A_{k\ level}$ after last fixing step. Note that, distance parameter in the model has to take an additional index indicating the vehicle it belongs to, and distances from new starting vertices of vehicles to all of the vertices not covered yet, p_{k0j}, have to be recalculated at fixing steps.

We define two criteria to choose vertices to be added. Given the worst scenario, the vehicles do not break any arc and continue from the latest visited vertex just before returning to the central vertex, one of them is the minimum required time to begin to serve newly added vertex, $min_{RT}(i,k)$, and the other is the maximum time that can be spent before making an attempt (leave the current vertex) to visit newly added vertex, $max_{TS}(i,k)$, where $i \in V$, $k \in K_i$ and K_i is the set of eligible vehicle set for vertex i. These values are affected by both of the time windows of the vertex to be added and the time required by vehicle to reach to that vertex. For a vertex to take priority in addition to the model, it is better to have less value from both of these criteria. Note that each of the criteria can be defined for each of the combinations made by a vertex and a vehicle that are eligible to visit that vertex. Therefore, while considering all of those combinations, it is also useful to consider the combinations that give the extreme values: the minimum value coming from the first criterion and the maximum value coming from the second one, namely, $\min_{k \in K_i} minRT(i,k)$ and $\max_{k \in K_i} maxTS(i,k)$. We define a simple function that determines the relative priorities of the vertices to be added to the model. There are four components in the function $f(i,K_i)$: the summation of values for all eligible vehicles (for first and second criteria) and the extreme values which are defined in the previous paragraph (for first and second criteria). Those four values for each vertex are separately scaled between 0 and 1. Therefore, each vertex has four values between 0 and 1. The summation of those four values, equally or differently weighted, defines the priority of vertex.

In the selection of the arcs that are allowed to be broken, the number of consecutive steps that an arc remain in the solutions is taken as the basic parameter. If this parameter is set to “0”, all traversed arcs are allowed to be broken in the next step. As the parameter increases, the number of arcs allowed to be broken decreases. In all cases, the last arc of all vehicles which is returning to central vertex is kept breakable. We accept that it is hard to give a specific comment on the best combination for the remaining parameters which are the number of added vertices at each step, the number of steps between fixing steps and the time allowed for solving the restricted problem. These parameters can be determined with respect to the problem size and the time given to have a feasible solution.

Since we solve the problem in a stepwise manner and establish a priority value for the vertices for being added to the model, it is possible to have infeasibility in subproblems because of time window restrictions of the vertices. In order to handle this case, we add penalty component to the objective function. In order the penalty mechanism to work, a new decision variable is introduced into the model. The objective function (1) and the constraint (3) is changed as indicated in the expressions (19) and (20).

$$\min\left(y + \sum_{i \in (V \setminus 0)} penalty_i M\right) \tag{19}$$

$$\sum_{k \in K_d} \frac{r_{kd}}{q_i} T_{ki} \geq 1 - penalty_i \qquad \forall d \in D, \forall i \in V_d \tag{20}$$

The new binary decision variable $penalty_i$ takes positive value if adding vertex *i* is infeasible or it takes much time than allowed to find a solution including that vertex at the current step. If there is a penalty, the algorithm turns back to the last fixing step to repair the solution to have it become feasible. In repairing, fixed solution is released in order to have the penalized vertices be introduced to the model, where all traversed arcs are kept fixed but allowing that limited number of arcs from each route can be broken just as it is applied in constructing steps. The formal ***algorithm of construction heuristic*** is given below. Since repairing procedure is straightforward, we do not denote repairing steps separately in the algorithm.

• Step 1. Set heuristic parameters and construction model (objective function 19 subject to constraints 2, 4-13, 15-18, 20), calculate vertex priority values and create an ordered list of vertices L w.r.t. addition-to-model priority. Go to Step 2.

• Step 2. If *L* is not empty, add first n_{step} vertices (or add all vertices if there are less than n_{step}) from *L* to model and update *L*, excluding the vertices added to model, go to Step 3. Otherwise stop.

• Step 3. Solve model in given time period,

> – Step 3.1. If this is not a fixing step and all vertices are served without penalty, set the values f_{kij} to 1 for all traversed arcs $(i,j) \in E$ and $k \in K$ where x_{kij} is equal to 1. Set a_{kij} to 1 w.r.t. given setting in order to determine the arcs that can be broken, go to Step 2.
>
> – Step 3.2. If this is a fixing step and all vertices are served without penalty, go to Step 4.
>
> – Step 3.3. If all vertices cannot be served, turn back to the last fixing step and apply repair procedure. If solution cannot be repaired go to previous fixing steps to apply repair procedure until the penalized vertices are added to the model. Update *L*, go to Step 2.

• Step 4. Exclude all added vertices from model. For all excluded vertices; keep the traversed arcs, served time windows and serving duration in the vertices. Update the distances from central vertex to other vertices: For all vehicles $k \in K$, set p_{k0j} to p_{kcj} where *c* indicates the last visited vertex of *k*. Set the current vertices to "0", and $A_{k\ level}$ to $A_{kc} + T_{kc}$ for all vehicles $k \in K$. Go to Step 2.

The solution constructed as it is explained above may have still room to be improved. One way is improvement within each route and the other is combining more than one routes to have an opportunity to improve. Before combining the routes, we choose to check if there is an improvement room within each route applying a simplified version of the model with considered vehicle and vertices it has served. Then, we expand this check by adding two routes, one from the set of longest routes and one from the set of shortest routes, to the model. In both improvement phases, the traversed arcs are all allowed to be broken, which means the exclusion of the constraints (15-17) from the model. Also allowed time window of each vertex that are served by other vehicles (the vehicles belong to the routes that are not considered at the current improvement step) are fixed by adding constraint (21) into the improvement model.

$$twn_{iw} \geq twn_{iw\ level} \qquad \forall i \in V \setminus (0), \forall w \in W_i \qquad (21)$$

The binary parameter $twn_{iw\ level}$ is set to positive value, if the vertex i is served in time window w by any other vehicle that is not considered at the current improvement step. While choosing the longest and shortest route, the combination, giving the greatest number of vertices in longest route that can be served by the vehicle of shortest route, is given priority. While the set of longest routes consist of only the routes that have the maximum time length, the set of shortest routes are populated starting from the route(s) having shortest time length and adding one by one until the set satisfies that the vehicles included are eligible to serve all vertices served by the set of longest routes. To complete the improvement phase within reasonable time period, each improvement iteration is given a certain period of time and after allowing a certain number of steps (or time), when there is a deterioration or no-improvement in the objective function, the algorithm is stopped. The steps of the procedure applied for the ***improvement phase of the heuristic approach*** is given below.

• Step 1. Set improvement model (objective function 1 subject to constraints 2-13, 21). Solve separately in given time periods for each vehicle $k \in K$. Update the solution. Go to Step 2.

• Step 2. Update the set of longest routes R_l and set of shortest routes R_s such that R_l covers all r_l longest routes having equal values, and R_s covers the shortest r_s routes of which vehicles can serve all of the vertices covered in R_l. Go to Step 3.

• Step 3. If the stopping criterion does not hold, solve improvement model with the combination of one route each from R_l and R_s, such that the number of vertices in the longest route that can be served by the vehicle of the short route be maximized (if there is a tie, choose the combination with shorter route, if the tie is not broken, break randomly). Update the solution, and go to Step 2. Otherwise stop.

4. EXPERIMENTS

Since there is no test instances that can serve as benchmark either as a lower bound or optimal solution to our problem, we choose our test problems from the test problems given in Yakıcı and Karasakal [3]. We also applied the same vehicle parameters as they are applied in Yakıcı and Karasakal [3]. While keeping all of the data of test problems given in Yakıcı and Karasakal [3] as they are, since it is required in our problem, we add parameters for two time windows for each vertex. Three different time window pairs, (0-15, 20-35; 5-20, 25-40; 10-25, 30-45), are determined and one of these pairs is assigned to each one of vertices except central vertex. In time window assignment to vertices, we follow the order of given numbers to vertices and the order of time windows given in the previous sentence.

Performance of the proposed method is compared to exact solution method provided by CPLEX 12.6.2.0 ILP solver. In addition, since even the small instances cannot be solved by exact method, in order to have another benchmark, we relaxed the time windows such that the first time window of each node has no upper bound. The lower bounds and best solutions achieved in twelve hours are reported along with the heuristic methods performance. In application of heuristic method, at most two minutes (and 10% relative gap criterion, if it is reached earlier) is given for solving the restricted MIP at each step and maximum ten minutes is given in improvement phase. Except 31-node instance, all instances used up ten-minute improvement period. Improvement phase is only applied after the completion of construction and stepwise approach is not used in this phase. In the construction phase, five nodes are added to the model at each step using equal weight priority function values and at each step only one arc from each route is allowed to be broken. When not all arcs, but only older arcs (the arcs which is traversed in the same route consequently at more than one in the last steps) are defined as breakable, the last arc returning to central vertex is always allowed to be broken.

Experiments are conducted with a personal computer with 2.6 GHz CPU and 8 GB RAM. The result of the experiments is reported in Table 1, where instance identity (ID) is given in the first column. In instance ID, the first section before the letter “k” denotes the test problem group and number of vertices as denoted in Yakıcı and Karasakal [3]. Vehicle quantities from each type are indicated after the letter “k”. Next four columns are related to the solutions of MIP and relaxed MIP. MIP refers to the model defined by objective and constraints (1-13), while Relaxed MIP refers to the same model where only first time windows are applied without upper bound. For MIP and Relaxed MIP, the lower bound and the incumbent solution from exact solution method is reported after twelve-hour period is elapsed. In the columns under the “Heuristic” title, solution for the problem with original time windows (not relaxed), time elapsed in the construction phase of solution process (in minutes and seconds) and parameter settings are given. In the “Settings” column, first field (before slash) denotes the number of steps that an arc should remain in the solution to be breakable and second field denotes the required number of vertices to be covered to apply fixing steps. As presented in this column, we start fixing when the step with 15 vertices is solved. Then, at every increment of 10 more vertices in the problem, fixing step is applied.

Table 1. Experiment Results

Instance ID	MIP		Relaxed MIP		Heuristic		
	LB	Sol'n	LB	Sol'n	Sol'n	Time	Setting
E031-k1/1/2/4	2.33	-	6.49	31.53	35.28	7'23"	0/15-25
E031-k1/1/2/4	2.33	-	6.49	31.53	36.53	6'57"	1/15-25
P-n51-k1/1/2/4	2.29	-	16.99	-	54.28	10'30"	0/15-25-…-45
P-n51-k1/1/2/4	2.29	-	16.99	-	58.10	10'22"	1/15-25-…-45
E-n76-k1/1/2/4	2.18	-	13.74	-	89.05	15'04"	0/15-25-…-65
E-n76-k1/1/2/4	2.18	-	13.74	-	91.04	14'15"	1/15-25-…-65

Observing Table 1 yields that we have rather poor lower bounds and no solutions at all from commercial solver. When the problem is relaxed and solved by the same solver, we can find better lower bounds which serves as lower bounds to the original problem (MIP). Since the solver does not perform well in finding solutions, we believe these lower bounds are still far away from optimal solutions. Therefore, while we cannot determine the quality of heuristic solution, we can claim that the heuristic is useful in this case when there is no other algorithm that can find better solutions.

5. CONCLUSION

We introduce a rich VRP in which there is a heterogeneous fleet having varying operational capabilities and it is assumed that, service demand of a demand point can be satisfied by one or more vehicles in one of assigned multiple time windows.

An ILP formulation of the problem is given and an ILP based heuristic is developed for solving the problem. We have employed and tailored three small test problems from an earlier study [3] about a similar problem having no time window and synchronization restrictions. Even with these small instances, the commercial solver cannot find any feasible solution in twelve hours while heuristic solution is able to find solutions in less than half of an hour for the largest instance. Also lower bounds found by the exact method solver in twelve hours are very poor. Therefore, we do not have an idea about the solution quality, however we consider that the proposed method is useful in the absence of a better solution.

REFERENCES

[1] Dantzig G.B., and Ramser J,H., (1959). *The truck dispatching problem*, Management Science, vol. 6(1), pp. 80–91.
[2] Duran S., Gutierrez M.A., and Keskinocak P., (2011). *Pre-positioning of emergency items for care international*, Interfaces, vol. 41(3), pp. 223–237.
[3] Yakıcı E., and Karasakal, O., (2013). *A minmax vehicle routing problem with split delivery and heterogeneous demand*, Optimization Letters, vol. 7(7), pp. 1611–1625.
[4] Renkli Ç., and Duran S., (2015). *Pre-positioning disaster response facilities and relief items,* Human and Ecological Risk Assessment: An International Journal, vol. 21(5), pp. 1169–1185,.
[5] Drexl M., (2012). *Synchronization in vehicle routing - a survey of VRPs with multiple synchronization constraints,* Transportation Science, vol. 46(3), pp. 297–316.
[6] Bredström D., and Rönnqvist M., (2008). *Combined vehicle routing and scheduling with temporal precedence and synchronization constraints,* European Journal of Operational Research, vol. 191(1), pp. 19–31.

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.51-79
Bilgisayar Mühendisliği/Computer Engineering

UNSUPERVISED FEATURE LEARNING FOR MID-LEVEL DATA REPRESENTATION

Emrah ERGÜL[1], Mehmet KARAYEL[2], Oğuzhan TİMUŞ[3], Erkan KIYAK[4]

[1]*Dr.,Turkish Navy Inventory Control Center, Golcuk, Kocaeli, emergul13@yahoo.com*
[2]*Turkish Navy Inventory Control Center, Golcuk, Kocaeli, m.karayel.se@gmail.com*
[3]*Dr., Turkish Navy Inventory Control Center, Golcuk, Kocaeli, oguzhantimus@gmail.com.tr*
[4]*Dr., Turkish Naval Academy, Tuzla, İstanbul, ekiyak@dho.edu.tr*

Date of Receive: 20.11.2015 **Date of Acceptance: 16.02.2016**

ABSTRACT

Attribute based approaches are commonly used in recent years instead of low level features for image classification which is one of the most important problems in the field of computer vision. The most important advantage of attribute based approach is that learning can be performed similar to human by using attributes which makes sense for people. In this study, unsupervised attributes are developed in order to avoid human related problems in supervised attribute learning. In our proposed work, the attributes are generated as random binary and relative definitions. The process of random attribute generation simplifies the data modeling when compared to other work in the literature. In addition, a major problem which is the increasing the numbers of attributes in attribute based approaches is eliminated owing to the increasing the numbers of attributes easily. Furthermore, attributes are selected more wisely using simple applicable algorithm to improve the discriminative capacity of randomly generated attribute set for image classification. The proposed approaches are evaluated with the other similar attribute based studies comparatively in the literature based on the same data set (OSR-Open Scene Recognition). Experiments show that noteworthy performance increase is achieved.

ORTA SEVİYE VERİ TEMSİLİNDE DENETİMSİZ NİTELİK ÖĞRENİMİ

ÖZ

Bilgisayarla görme alanındaki en önemli problemlerden birisi olan imge sınıflandırma için öznitelik tabanlı klasik yaklaşımların yanı sıra nitelik tabanlı yaklaşımlar son yıllarda sıklıkla kullanılmaya başlanmıştır. Nitelik tabanlı yaklaşımların en önemli avantajı, insanlar için anlam ifade eden niteliklerin kullanılması vasıtasıyla insanoğluna benzer bir öğrenme yapılabilmesidir. Bu çalışmada, denetimli nitelik öğrenme sürecinde insan faktörü sebebiyle oluşabilecek sorunlardan kaçınmak amacıyla denetimsiz yaklaşım geliştirilmiştir. Denetimsiz yaklaşımımızda niteliklerin ikili ve göreceli olarak rastgele üretilmesi sayesinde nitelik öğrenme süreci, literatürdeki diğer denetimli ve denetimsiz yaklaşımlara göre daha kolay hale gelmiştir. Ayrıca, nitelik sayısının basit bir şekilde artırılması ile nitelik tabanlı yaklaşımlarda büyük bir problem olan nitelik sayısının artırılması basitleştirilmiştir. Rastgele üretilen nitelik kümesinin imge sınıflandırma için ayırt etme kapasitesini artırmak maksadıyla, rastgele üretilen nitelikler arasından en iyileri kolay uygulanabilir bir algoritma sayesinde seçilmiştir. Çalışmada önerilen yaklaşımlar literatürdeki diğer benzer nitelik tabanlı çalışmalarla aynı veri kümesi (OSR-Açık Alan Tanıma - Open Scene Recognition) üzerinden ve farklı sınıflandırıcılar kullanılarak test edilmiştir. Yapılan deneylerde denetimsiz öğrenilen göreceli niteliklerin dikkate değer bir performans artışı sağladığı görülmüştür.

Anahtar Kelimeler: Göreceli Nitelikler; Denetimsiz Nitelik Çıkartımı; Nitelik Seçimi; Görsel Tanıma.

Keywords: Relative Attributes; Unsupervised Feature Extraction; Attribute Selection; Visual Recognition.

1. INTRODUCTION

Attributes constitute intermediate layer data representation between the low-level image features (i.e. color/edge histograms, bag of visual words, quantized pixel values, GIST, SIFT, Fourier/Laplace/Hough/Wavelet transforms etc.) and the top level categories. Because attributes are common properties of the object categories, intermediate representations can be achieved by using classes in combinations with respect to the shared attributes, and this leads to generating new discriminative spaces for visual recognition.

Visual attributes are important for understanding object appearance and can be used for describing objects. In detail, visual attributes include color, modal, textural, functional, structural, and conceptual or any kind of semantic properties of objects. In addition to visual or semantic distinction, the representation of attribute is also varied as binary or relative. The presence or absence of an attribute in binary and the strength of an attribute in relative become important in attribute representation. One may think binary correlations (i.e. existence or absence of an attribute in a class) would be sufficient while the others claim real-valued ranking scores are essential to measure the attribute strength among categories [6,15].

Attributes can be learned by supervised or unsupervised manners. Supervised methods are firstly proposed in the literature and then unsupervised approaches become more popular. In supervised attribute learning, images are labeled with attributes by human effort. Hence, many difficulties occur. These difficulties can be summarized as; more general and intuitive attributes are determined instead of discriminative attributes which are indeed appropriate for classification purposes. In addition, some discriminative attributes may be overlooked or could not be expressed by words. Furthermore, erroneous attribute tagging can be performed. Finally, the process of attribute extraction become exhaustive and it takes a long time in large datasets that may contain many attributes [16]. In addition to above mentioned difficulties; attribute labeling of datasets in supervised methods needs a great deal of human laboring and budget. Moreover, extracting attributes by searching the related images on the internet as in [9] seems to be a clever idea for cost reduction, but the discovered attributes can be irrelevant with the image categories.

Since attributes are commonly shared amongst different top level categories, one of the major advantages of attributes is that fewer training examples are required to train an attribute and a classifier established on the basis of attributes. Consequently, the main idea is to learn attributes at the intermediate level for separating visual categories efficiently in attribute learning. However, the uppermost main target is to discriminate classes and it is not to learn some attributes perfectly.

In this work, we aim the image classification with the visual attributes which are used as the new feature space at mid-level. This kind of representation is achieved in an unsupervised way such that binary and relative attributes are learned by random binary predicates or class based relative orderings. Additionally, we select some of randomly generated orderings distinctively by implementing Kendall Tau metric which computes the distance between two sequences.

The contribution of this work is two folded. Firstly, we get unsupervised data representation at a new mid-level feature space with binary/relative attributes. The class based attributes are generated randomly, and binary SVM scores are used out of binary attributes while ordering scores are handled for relative ones. So the new feature space is assumablely expanded and established more discriminatively. On the other hand, we train the basis vectors of the new feature space with a very limited number of training instances. Secondly, we also select some of randomly generated attributes with a distance based algorithm where more discriminative sequences are picked. We also try three classifier (kNN, decision tree and SVM) for accuracy performance comparisons.

In Section 2, the development of attribute notion and attribute based approaches are explained in mixed form. In Section 3, our proposed algorithms based on random binary and relative attributes are introduced while in Section

4 experimental results are detailed. Finally, the experimental results are concluded.

2. RELATED WORK

The literature of attribute-based computer vision problems can be generally summarized in the types and extraction methods of attributes, applications and datasets on which they are implemented, and performance criteria in the experiments, as shown in Figure 1.

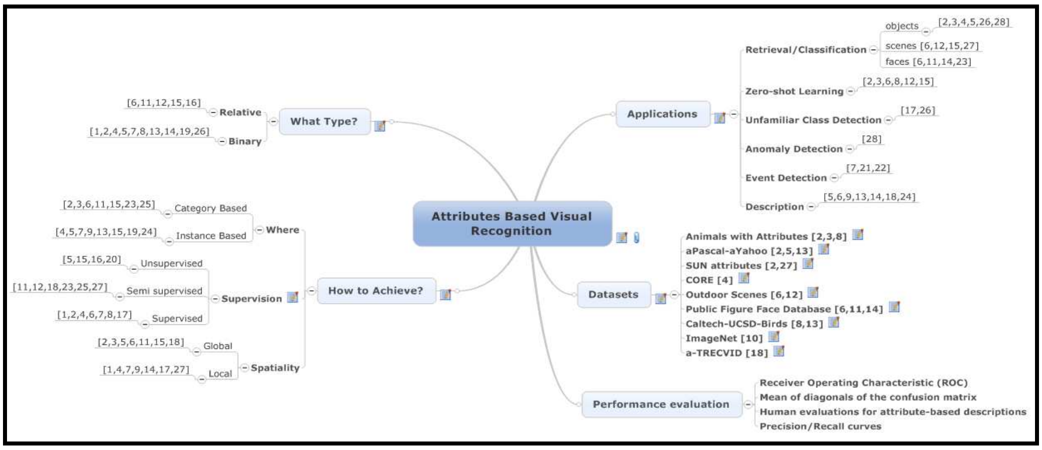


Figure 1: Summary of the literature works on attribute-based computer vision problems.

Ferrari and Zisserman propose a probabilistic generative model which infers whether an image contains a learned binary attribute and determines which regions over an object image the attributes may cover [1] in a weakly supervised manner. They use simple attribute like shape, color or texture; and two adjacent segments produce a complex attribute like spotted, striped, checked etc. Lampert et. al. study object recognition for categories which are not seen during training at all [2, 3].

Farhadi et. al. describe objects by the spatial arrangement of their appearance and part based attributes, and they benefit from the interactions between attributes [4]. In addition, Farhadi et. al. [5] embed object recognition problem into describing objects and mainly focus on attribute learning on the basis of semantic (i.e. nameable) attributes like object parts and discriminative attributes which are achieved by splitting the visual feature space into different regions by comparing some classes randomly in a binary fashion. The discriminative intuition is that some attributes cannot be nameable although they may be very useful for discrimination.

The relative attributes are first introduced by Parikh and Grauman in [6] with the assumption that semantically more enriched data descriptions and discrimination will be achieved if we use relative class memberships on attributes, instead of binary relations. SVM-like algorithm is implemented as a ranking function in which not only maximizing the margins between class boundaries but also ordering of classes over attributes space is aimed by using Newton's method. They use predetermined nameable attributes and class orderings are given on these attributes in a supervised manner. The main restriction of the algorithm is that equality acceptance in category ordering according to an attribute. In the notion of equality acceptance, human can not differentiate two image belong to different categories. On the basis of the supervised relative attributes [8], many studies [33,34] whose aim is to learn more robust and precise relative attributes have emerged in the literature. In [33], it is claimed that relative attribute learning method is insufficient in indistinguishable image pairs which the human can't sort or differentiate two image belong to different categories on the basis of an attribute. Namely, it is aimed to sort the image pairs which are assumed as equal situation in [6]. For this purpose, Bayesian local learning is proposed in [18]. In addition, instead of creating new method for generating relative attributes, Verma et al [34] improve the performance of the basic relative attribute method using patch-based features instead of global (GIST and Color Histogram) which are used in [6]. Verma et al claimed that their representation capture local shape in an image comparison to global features.

Simple solution to a multi-attribute query is to train a classifier for each attribute independently and combine their scores in retrieval. But some attribute conjunctions may be more useful since such combinations can be learned more easily and they discriminate visual data more. At this point, it is critical to determine which combinations of attributes should be trained without trying all combinations intensively. Rastegari et. al. [13] focus on learning more discriminative attributes by merging some of them, instead of learning each attribute individually. Kumar et. al. [14] open an interesting discussion about attributes in that similarity of faces with respect to other specific people as references may help for achieving more discriminative attributes for face verification, called 'similes'. With such visual traits, for example, a face might be described as having a forehead like Barack Obama's and eyes like Jennifer Lopez's.

It is often intractable for a human to predefine and label all the attributes in large datasets explicitly. Furthermore, some attributes may be more valuable for recognition although we can't name them. Ma et. al. [15] implement an algorithm to learn class-level relative attributes in an unsupervised manner, unlike [6] where relations between pairs of classes on attributes are already given relatively.

Instead of using pre-determined binary/relative attribute labels or class orderings on attributes, Karayel and Arica [16] follow the similar way of unsupervised attribute learning like in [5, 14, 15]. But in here, binary and relative attributes are produced completely randomly where classes are separated into positive and negative sides for binary attributes, while class ranks are selected for each relative attributes. Wang and Mori [17] purpose to model objects discriminatively for classification as the dependencies among attributes are captured using an undirected graphical model built from a training set. The main distinction from other works is that they unify object class and attribute predictions in a joint framework since classes and their attributes are closely related concepts.

So far, we mainly build category-attribute correlation matrices or dependency matrix among attributes for object recognition. Assuming that we have a limited number of nameable attributes which are pre-determined, such matrices would not be sufficient for large-scale computer vision problems. Yu et. al. [18] add another intermediate layer for multi-attribute based image retrieval which corresponds to a large pool (i.e. 6000) of weak attributes. Weak attributes are comprised of automatic classifier scores or other mid-level representations that can be easily acquired with little or no human labor. Chen et. al. focus on learning a regression model which introduces a cumulative attribute representation [19]. In details, each attribute is not only discriminative but also cumulative such that all other attribute values depend on their relative positions in a scalar value.

Human efforts involved in the class-attribute relationship designing are costly to obtain, subjective for evaluation and not scalable to large-scale datasets. Given images with category labels, Yu et.al. [20] formalize a category-attribute co-occurrence matrix for cross-category generalization. This is different from randomly generating ‘category splits’ in those geometric properties of category separability and attribute learnability are used.Chen et. al. [23] build facial classifiers which are based on appearance similarity of people with the same birth name. Another work of human description by visual attributes is proposed by Sadovnik et. al. [24]. The task is to describe a person in a group that distinguishes her from the others. The description will contain as minimum number of visual attributes as possible while it is maximizing the likelihood that a listener will correctly guess which person description refers.

3. UNSUPERVISED FEATURE LEARNING SCHEME

The overall flow chart of the visual recognition in this work can be basically split into three stages: Unsupervised data representation via attributes at mid-level, category based domain modeling, and evaluation of the classification performance. Given the dataset $X = \{x^{(i)} \mid i=1,2,3,\ldots,N\}$; where N is the number of train instances and $x^{(i)} \varepsilon R^d$ represents the low-level feature vector, we first divide it into three non-overlapping subsets randomly. Train set, $X_{train} = \{x^{(j)}, y^{(j)}\}_{j=1}^{K}$; where $X_{train} \subset X$, $y^{(j)} = \{1,2,3,\ldots,C\}$, and C is the number of classes in X_{train}, hold for the class label, is used in both unsupervised attribute learning and category modeling. Test set, $X_{test} \subset X$, is utilized in classifier evaluation while the free parameters of classifiers (i.e. kNN, SVM and Decision Tree) are optimized in a grid search method on a validation set, $X_{validation}$. Also note that $X_{validation}$ is achieved by inserting small amount of white noise to the samples in the dataset.

As mentioned earlier, we learn the classifier discriminants in a new feature space as the mid-level data representation, instead of simply using low level features. So we learn class based binary and relative attribute models in an unsupervised manner which will be detailed in subsections A and B, respectively. The binary attributes define output of the binary SVMs where the scores are computed by dot (i.e. scalar) products of the input samples, $x^{(j)}$, and the learned weight vectors, $w \in R^d$, of the binary SVMs. We call it Score Related Attribute (SRA) space. On the other hand, the relative attributes are modeled in Newton algorithm of [6]. Although this resembles of SVM method very much, the input signal would be the difference of related feature vectors, and the comparative condition determines the positive and negative sides, instead of tagging binary instance labels.

After we define binary and relative attributes in an unsupervised manner and learn them on the train set, X_{train}, we then model our classifiers; kNN, SVM and C4.5 decision tree still on the very limited set of train set. $X_{validation}$ is used to optimize the parameters of classifiers at hand whereas we compute the accuracy performance on the X_{test}, eventually. The flow chart of the proposed work is depicted in Fig. 2.

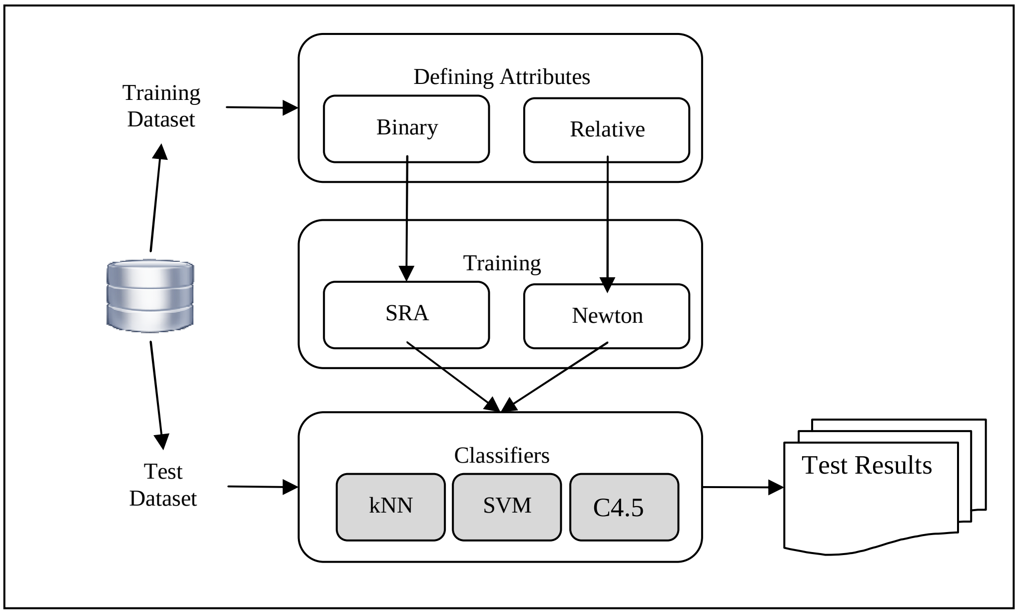


Figure 2: Schematic flow chart of the methodology

A. SRA Representation with Random Binary Attributes

We start with binary attributes for unsupervised feature learning. As the name explains itself, a binary attribute refers to whether it exists or not in the visual data. As speaking of class based attributes, we generalize them throughout each class specifically. For example in the statement “Attribute a_m exists in Class A, but Class B does not have it”, we hypothesize that all instances in a class contains or does not contain the mentioned attribute, a_m, at all. Although it seems to be more convenient if the attributes are assigned per instance individually, the literature work [6, 15] claim that class based attribute definitions result in consistency for learning the attribute models at the mid-level. Additionally, instance based labeling would consume much more effort while this process also hinders the unsupervised learning of data representation.

In detail, we define class based binary attributes randomly for the unsupervised learning in mid-level feature representation with some constraints to the random binary sequence generation. First, all ones or all zeros for an attribute $a_m = \{0,1\}^C$, is discarded from the list because they do not help us discriminating the visual classes in the attribute space. Next, we include random binary sequences into the consideration only if they have at least two positions different from each pattern which has been added to the list of attribute definitions already. Note that the number of positions which is set for discrimination is strongly related to the number of classes in the train set, X_{train}, since the length of each attribute sequence equals to it, C. Moreover, we explicitly limit the number of random sequences out of all combinations, 2^C. Nevertheless, binary definitions are produced as many as possible, and we select some of them for training, randomly. After we finish the random definitions, this process results in M class based binary attributes $A = \{a_m \mid m = 1,2,3,\ldots, M\}$, and we carry on the next step of modeling the attributes in binary SVMs.

SVM is a powerful tool for supervised learning which separates the feature space linearly into two categories: positive and negative. It tries to maximize the margin between positive and negative sides [35, 36]. The margin in SVM represents the gap between support vectors of both sides which are data samples acceptably close in limits to the opposite ones. The main idea is to find the optimum hyperplane that achieves the total minimum distance between the support vectors and the hyperplane. After training the SVM, a new sample is classified simply as either positive or negative by the result of the dot product with the optimized weight vector. In fact, SVM is a linear discriminant function and it obviously does not handle nonlinearly separable datasets with a satisfactory accuracy. Kernel functions (Gaussian, polynomial, chi-square, histogram intersection etc.) are introduced in the literature to establish nonlinear SVM classifiers and they achieve a justifiable popularity with SVM. Actually, SVM still preserves its linearity in this perspective but the input data are transferred into a new feature space via nonlinear kernel functions beforehand. Hence, we get more complex feature spaces with higher dimensionality instead of complicating the discriminant function itself.

In the second part of the introduced architecture, we propose binary attribute learning concept that is based on two-class SVM topology. Given the input data $X_{train} = \{x^{(j)},y^{(j)}\}$ and their class based binary attribute assignments $a_m = \{0,1\}^C$; input signals for the SVM are firstly achieved by Gaussian Kernel Function (GKF), $K(x^{(j)},X_{train})$:

$$K(x^{(j)},X_{train}) = e\frac{-\left\|x^{(j)} - X_{train}\right\|^2}{2\sigma^2} \; ; \; x^{(j)} \in X_{train}$$

(1)

where σ is the scale parameter that factors the neighborhood. So each data sample is now represented by its GKF responses to the data samples (i.e. landmarks) in X_{train}. Note that the input dimensionality now equals to the number of instances, K. Since SVM is a supervised learning algorithm, the sample-attribute assignments obtained in the previous class based attribute definition, $a_m : y^{(j)} \rightarrow \{0,1\}$, are now used as the data labels for supervision, instead of $y^{(j)}$ itself. The unconstrained objective function of the SVM is:

$$y^{(j)}(w^T x^{(j)} + b) \; > \; 1 \; ; \forall j$$

(2)

$$J_{G_X,y}(W,b) = P\frac{1}{K}\sum_{j=1}^{K} max(1\text{-}w^T G_{x^{(j)}} y^{(j)},0)^2 + \frac{1}{2}\sum\left\|W\right\|^2 ; y^{(j)} \in \{-1,1\}$$

(3)

where P is the trade-off constant, penalizing data points which violate the margin requirements. G_x represents GKF output vector of (1) for each sample, $x^{(j)}$, that is the new input signal to the SVM. W is the matrix, which embeds the parameter vectors, including biases, b. They are assumed to be orthogonal to the hyperplanes that separate both sides (i.e. binary assignments) and initialized randomly. As aforementioned, SVM simply separates the space into two parts. So the desired output signal for each input, $y^{(j)}$, is achieved by assigning 1 for the classes which have the attribute, a_m, and -1 for the rest. The stochastic gradient descent algorithm is then employed as:

$$h_W(G_x) = W^T G_x \tag{4}$$

$$\delta w = \frac{\delta J(W,b)}{\delta h_W} = -2Py\ max(1-h_W(G_x)y,0) \tag{5}$$

$$\Delta w = \alpha \left(\frac{1}{N} \sum (G_x^T \delta w) + w \right) \tag{6}$$

$$w_{new} = w_{old} - \Delta w \tag{7}$$

where δw is the back propagated derivation of the error signal per data sample, Δw is the average weight correction that includes L_2 regularization without bias terms. Also note that α is the learning rate and h_w is the hypothesis function of the SVM. Once the SVM is set up, we optimize the weight parameters iteratively. The hyperplane is updated with the max-margin objective function to separate the samples of each side based on the static sample-attribute assignments. The iteration is terminated when the saddle point is reached.

After we find optimum parameters of the SVM, w^T, the data sample, $x^{(j)}$, can simply be conveyed to the new feature space by its related binary attribute score with Score Related Attribute (SRA):

$$SRA(x^{(j)}; w_{a_m}) = w_{a_m}^T x^{(j)} \tag{8}$$

where w_{a_m} is the weight vector of SVM which corresponds to the binary attribute definition a_m, including the bias term, b. So we train an independent SVM for each random binary attribute with the given train set, X_{train}, and the visual data is now in a new M (i.e. number of binary attributes) dimensional feature space by their SVM scores. Eventually, we implement our classifier algorithms on the train set X_{train}, where the feature vectors are now represented in the mid-level, instead of their original space R^d, where M << d.

B. Rank Based Representation with Relative Attributes

The relative attribute definitions are first introduced in [6] and they have attracted much attention so far [15, 16, 30]. Unlike binary attributes, they infer the relative strength of an attribute on the visual data, instead of exposing the existence (or non-existence). As it can be seen in the statement of "Class A has attribute a_m more than Class B, but less than Class C," the class based relative attributes order the visual categories on the basis vectors of a new feature space by comparative constraints; i.e. more/less than. They have obvious advantages over the binary definitions in those: 1. More input data are fed into the attribute learning models because the input data are now the pairwise comparisons of the samples. Assuming that each class has K examples and we have C categories in the training data set, X_{train}, then the number of input data will be $C\binom{C}{2}K^2$, instead of KC. So we assume that more training data would increase the accuracy performance in learning the attribute models. 2. Since we randomly define the relative attributes by ordering the classes in each attribute basis, the total number of possibly generated ordering patterns equals to the permutation of the number of classes, C. Thus, one can produce many random ordering sequences more than the binary predicates, and more discriminative patterns may be selected among them.

Given a class based ordering, $a_m = \{c^{(1)} > c^{(2)} > c^{(3)} > \ldots > c^{(i)}\}$; $c^{(i)} \in C$, which relates every category to each other with a less/more condition, we use the Newton method of [6] for a relative attribute as:

$$r_m(x^{(j)}) = w_m^T x^{(j)} \quad (9)$$

$$\forall (i,j) \in O_m : w_m^T x_{C_a}^{(i)} > w_m^T x_{C_b}^{(j)} ;\ i \in c_a ,\ j \in c_b ,\ c_a > c_b \quad (10)$$

$w_m^T(x^{(i)} - x^{(j)}) \geq 1\text{-}\gamma_{ij} \;\; ; \; \forall(i,j) \in O_m \, , \gamma_{ij} \geq 0$

(11)

$argmin_{w_m}(\frac{1}{2}\left\|w_m^T\right\|_{L_2}^2 + T\sum\gamma_{ij}^2$

(12)

where r_m is the reel ranking score of the training instance, $x^{(j)}$, on the attribute basis, a_m, $w_m \in R^d$ is the parameter vector of the relative attribute model, O_m is the set which consists of pairwise data instances holding for the more/less conditions. When we look into (11) closer, the equation is very similar to that of the SVM. But the input signal is now the difference of pairwise feature vectors from the set, O_m, not the low-level feature vectors itself. So the optimum solution would then order the classes on the weight vector, w_m, by minimizing the cost function of (12); where T is the constant that regulates the balance between weight decreasing and the non-negative slack variables, γ_{ij}. This results in maximizing the margin between classes in the order definition, a_m.

Once we optimize the free parameters, w_m, the attribute strength is computed as in the binary attribute score. Hence, we convey the original input data $x^{(j)}$ into a mid-level feature space by M (i.e. number of generated relative attributes) dimensional ranking scores, M << d. The next step is to answer how one may generate class orderings for relative attribute modeling which will be detailed in the subsections below.

(1) Random Relative Attributes

We follow the same approach of binary attribute generation described in section 3.A. The relative definitions, $A = \{a_m \mid m = 1,2,3,\ldots, M\}$, indicate ordering the visual categories, $\{c^{(1)} > c^{(2)} > c^{(3)} > \ldots > c^{(i)}\}$; $c^{(i)} \in C$, randomly for each attribute, a_m. The class ordering expands the feature space much more than binary attributes and we have many options this time. The random class based ordering sequences are included into the consideration only if they have at least four positions different from each pattern which has been added to the list of attribute definitions already. Note that the number of different positions is twice that of binary predicates. So we produce many unique orderings (say 1,000) and select M sequences out of them randomly.

(2) Selective Relative Attributes

To make attribute definitions more discriminative, we propose a new approach for picking some orderings based on Kendall Tau (KT) correlation metric [15], instead of selecting randomly. For each pair of randomly generated attribute definition, KT is computed as:

$$KT = \frac{n_c - n_d}{C\binom{C}{2}}$$

(13)

where n_c and n_d are the number of concordant and discordant pairs between the two orderings and the denominator refers to the total number of pairs. The range of KT is then in [-1,1], and it is -1 if two orderings are complately different (1 if they are the same). Thus, we first compute a correlation matrix in which each element is the KT value of pairwise orderings of all generated ones, next the average correlation values of all definitions are sorted in the decreasing order, finally we select the top M random orderings (i.e. least correlated) among them. Thereafter, the preselected classifiers (i.e. kNN, SVM and C4.5 decision tree) are modeled on the train set, X_{train}, while optimizing their free parameters with $X_{validation}$.

4. EVALUATION OF THE PROPOSED WORK

A. Experimental Setup

We use Outdoor Scene Recognition (OSR) Dataset [6] containing 2,688 images of 8 scene categories. The distribution of images for the dataset is shown at Table I. Note that the number of samples in each class varies. OSR dataset is also utilized in [6] and [15] which are the recent studies in attribute based object recognition literature. Besides, the provided low level features (i.e. GIST) and the same train/test splits for multiple runs are used as the initial input in multi-category classification schemes. Most of the outdoor scenes in OSR dataset display large intra-class variability, meaning that object contents within a scene category are very different while inter-class variance is small especially for the natural scene categories. This issue makes the object classification problem harder when working with OSR dataset. Example images from OSR are displayed in Fig. 3, respectively.

Table I. THE DISTRIBUTION OF IMAGES FOR OSR DATASET.

OSR Dataset	Coast	Forest	Hihgway	Insidecity	Mountain	Open Country	Street	Tall Building
	360	328	260	308	374	410	292	356

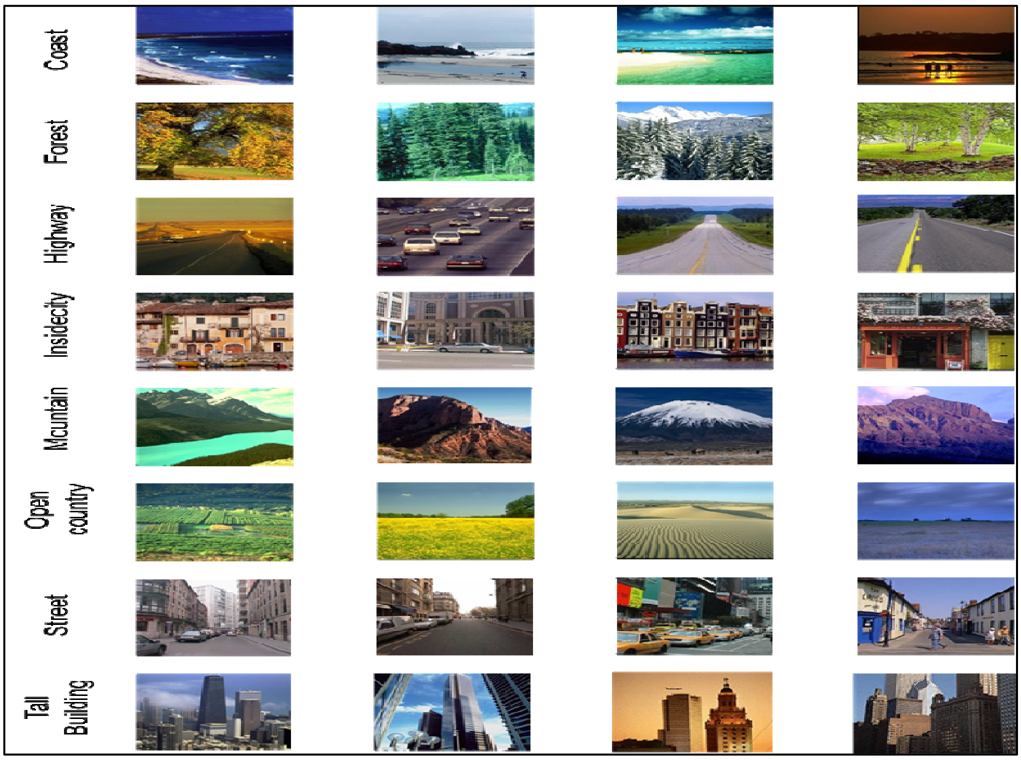


Figure 3: OSR dataset sample images.

For the training phase of both attribute models and visual classes, we randomly select 30 instances from each class as X_{train}, and the rest is used as the test set, X_{test}. Note that the X_{train} is very limited due to the mid-level attribute representation when compared to the low-level features in classification. The experiments are repeated 20 times, and the mean and standard deviation values are noted at tables for comparative results whereas the average accuracies are used in the figures. Additionally, we limit the number of both randomly generated relative and binary attributes to 28 for the sake of comparison to the other literature work.

Furthermore, we evaluate three algorithms to measure their classification accuracies in the mid-level attribute space: SVM, kNN, and C4.5 decision tree. We select these methods as they are powerful and popular discriminants on the shelf. So WEKA toolbox [37] is used to implement them while we optimize their free parameters (i.e. the regulator constant, C, for SVM; the number of nearest neighbors, k, for kNN; the pruning confidence, C, and the minimum number of samples, M, for the decision tree) on the $X_{validation}$. Additionally, we normalize the feature vectors of attribute scores as the new inputs to the classifiers by whitening process of [38] in order to achieve zero mean and unit standard deviation for each dimension.

Finally, we also use the supervised binary and relative attribute definitions which are given in [6] to promote the benefits of unsupervised (i.e. randomly generated) definitions. Fig. 4 displays the usage combinations of all attribute patterns that are utilized for the experiments, detailed in the next subsection.

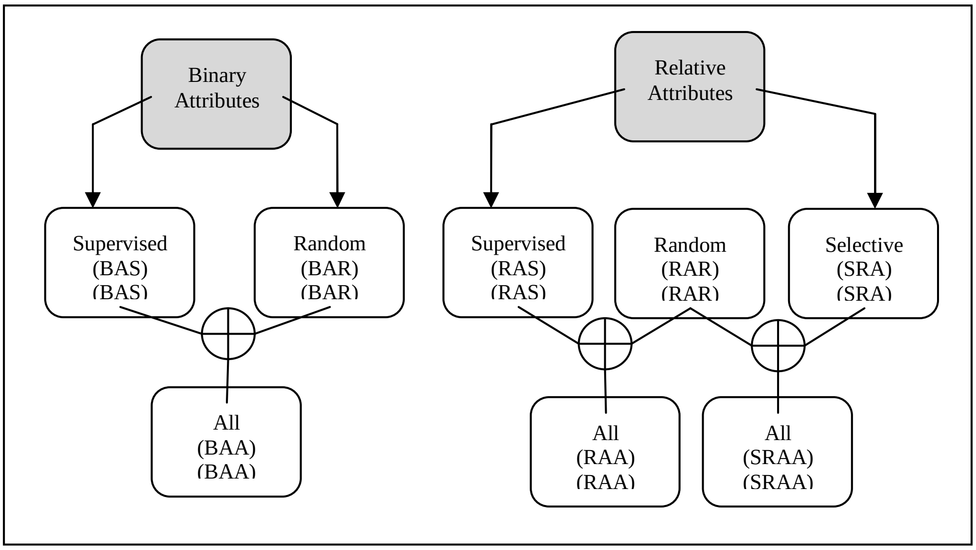


Figure 4: Attribute based comparisons scheme.

B. Classification Results

In this subsection, we first analyze the classification results in different configurations of attributes and classifiers. The experiments are repeated 20 times and the mean and standard deviations are noted for binary attributes as SRA results and relative attributes as Newton ranking scores at Table II and III, respectively. Also note that we use the 6 binary and relative attribute definitions which are already established on OSR dataset in a supervised manner [6]. Additionally, we generate/select 28 random binary and relative attributes to compare the classification results with the other literature work, although we may produce them as many as needed that is dependent with the number of visual classes at hand.

Table II. SRA RESULTS WITH BINARY ATTRIBUTES ON OSR DATASET.

	Classifier Accuracies (%)		
Attribute Type	**kNN**	**Decision Tree**	**SVM**
binary_attributes_supervised (BAS)	52.99 ± 2.58	49.07 ± 2.32	54.88 ± 3.68
binary_attributes_random (BAR)	75.27 ± 1.52	62.31 ± 3.99	74.28 ± 1.63
binary_attributes_all (BAA)	76.38 ± 1.55	64.38 ± 3.34	**76.73 ± 1.59**

For speaking of supervision, randomly generated attributes outperform the human labeled attributes at both tables considerably. This is due to the fact that we can generate more definitions randomly at no cost and this expands the mid-level feature space discriminatively which results in better accuracies. We claim that supervision may sometimes divert the learning system into a worse situation as it is subject to the human experience, and hard work of labeling. Nevertheless, we can surely append the supervised attributes into the unsupervised patterns if they exist. We achieve almost 2 % increase in the performance at both tables when they are concatenated with the unsupervised attributes. Additionally, relative attributes overcome the binary definitions about 2-3 %. We assume that the class orderings which we may produce randomly is related to the permutation of the number of categories, not power of 2, and that gives many more choices for selection. Additionally for the relative attributes as detailed in section 3.b, we run the KT algorithm to select more discriminative ordering patterns from the randomly generated pool, instead of random selection. We see that selective relative attributes indeed increase the performance more than 2 %, and this confirms our previous assumptions. On the other hand, SVM algorithm achieves better accuracies than kNN and C4.5 decision tree overall while C4.5 is the worst. Note that the kNN gets the similar, even better results than SVM although it is the simplest instance-based classifier. We assume that non-parametric learning of the kNN method benefits the attribute based feature space more than the others.

Table III. NEWTON RESULTS WITH RELATIVE ATTRIBUTES ON OSR DATASET.

	Classifier Accuracies (%)		
Attribute Type	**kNN**	**Decision Tree**	**SVM**
relative_attributes_supervised (RAS)	62.17 ± 1.02	54.54 ± 2.05	63.12 ± 1.79
relative_attributes_random (RAR)	76.78 ± 1.97	69.77 ± 2.63	76.34 ± 1.52
relative_attributes_all (RAA)	77.15 ± 1.52	73.57 ± 2.04	77.86 ± 2.13
selective_relative_attributes (SRA)	77.24 ± 1.87	70.81 ± 3.62	77.12 ± 1.96
selective_relative_attributes_all (SRAA)	78.36 ± 2.01	72.66 ± 2.91	**79.86 ± 2.52**

Next, the proposed method is compared with the similar approaches in literature on the same experimental setups, and the mean accuracy results of the multiple experiments are listed at Table IV. BINs, PCA and FLD algorithms are actually used for dimension reduction and these references are not related to the attribute learning. Nevertheless, the basis vectors (i.e. like attribute weight vectors, w) which are extracted during the implementations help representing the data in a new features space, so they are included as baselines for this reason. Besides, the other methods generate supervised/unsupervised attributes in the intermediate level for visual recognition, like the proposed work.

The results at in Table IV show that the proposed method outperforms the other approaches for about a minimum of 1 % with the selective relative attributes. In general, it is observed that the attribute-based methods achieve much better accuracies than the other baseline works. So the attributes do not only reduce the dimensionality but also do they constitute a more representative space in the mid-level. On the other side, the unsupervised attributes display increased performance when compared with the supervised ones. Additionally, the accuracy rises up even further when we combine the both types. We assume that the expanded number of unsupervised attributes with distinct class orderings establish a better representation without human laboring, leading to more effective classifiers.

Table IV. PERFORMANCE COMPARISON OF THE ALGORITHMS.

Algorithms	**# of Attributes**	**Mean Accuracy (%)**
BINs [15]	28	76.05
PCA [15]	34	71.46
FLD [15]	28	63.10
Supervised Attributes (SAT) [6]	6	72.82
Unsupervised Attributes (UAT) [15]	28	76.57
SAT+UAT [15]	34	77.88
RAS [16]	34	78.64
Our Binary ALL	34	76.73
Our Selective Relative ALL	34	**79.86**

Additionally, we evaluate the behavior of mid-level feature space by changing the number of attributes that we generate randomly in the proposed work, and the graphical results are displayed in Fig. 5. Note that we use the SVM results as it is better than kNN and C4.5, comparatively. When we take into account the results of supervised attributes (i.e. 6 binary/relative definitions) at Table II and III, the accuracy performance is almost the same with 10-12 randomly generated binary and relative attributes, respectively. After this point, we outperform the supervised attributes obviously, and it confirms that the performance is increased as we enlarge the feature space with more attributes, although we select them randomly. Moreover, the relative definitions achieves better accuracies than the binary attributes. We think the main reason is that the Newton method orders the visual categories by maximizing the sequential margins with many more pairwise inputs, and we can generate more orderings than the binary predicates. Also, the selective relative attributes gets the best performance since the KT correlation metric is used to pick the more distinctive orderings, instead of simply selecting them randomly. Another point is that we can have even better results if the supervised attributes are concatenated with the unsupervised orderings. One may use the unsupervised definitions as the supplementary feature space if the supervised attributes already exist.

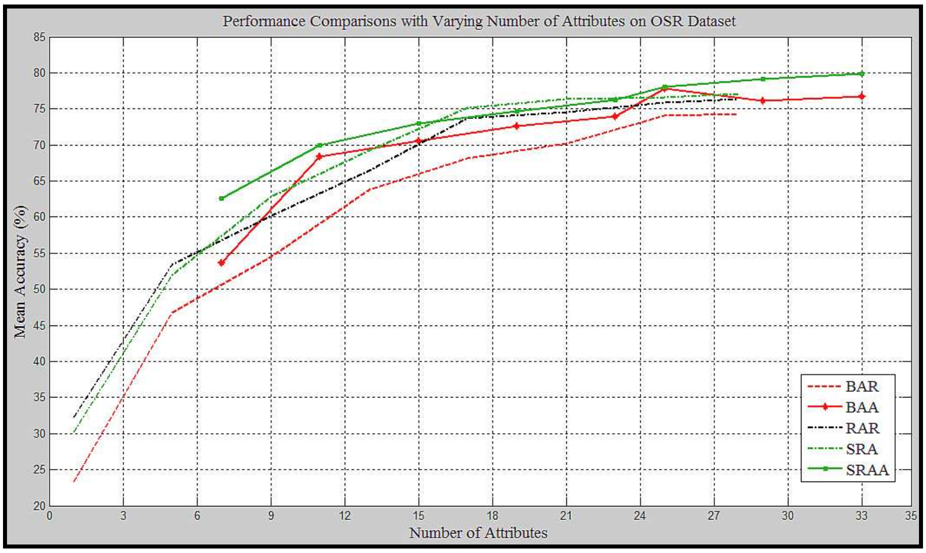


Figure 5: Performance comparisons with varying number of attributes.

Receiver Operating Characteristic (ROC) curve is frequently used in literature to evaluate the performance of classifiers. Basically, the ratio of false and true positive samples is plotted by changing thresholds in a step-wise manner. The classifier is regarded as more successful when its plot rises up earlier and sharper than the others. Eventually, we compare the performances of binary, relative and the selective attributes with their supervised and combined (i.e. supervised + unsupervised) versions on ROC curves for OSR dataset in Fig. 6. As seen, the accuracy is increased obviously when all attributes are used together, and this confirms that the unsupervised attributes add discriminative power in dimensionality. Additionally, the selective relative attributes outperform the others clearly while relative definitions are better than the binary predicates.

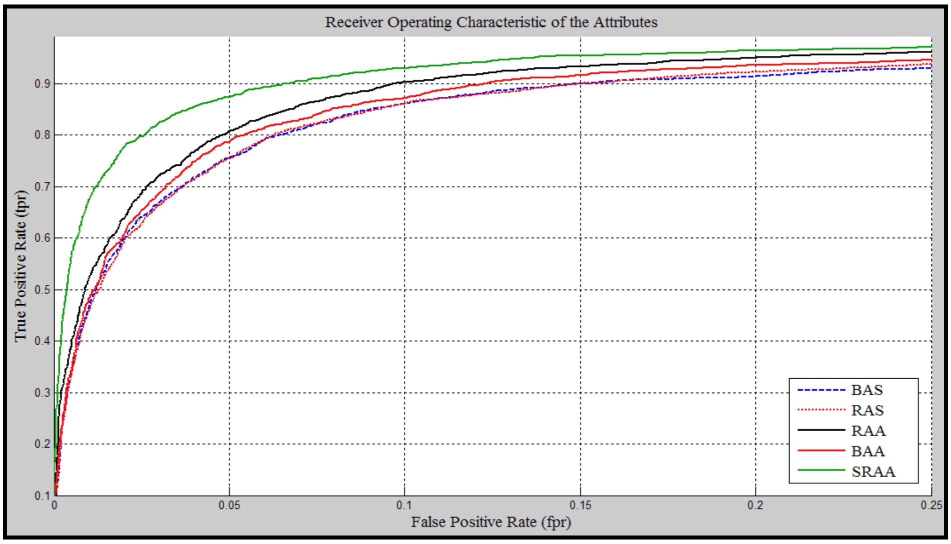


Figure 6: ROC analysis of the attribute types on OSR dataset.

5. CONCLUSION

In this work, we introduce two approaches for the mid-level visual data representations in an unsupervised manner which is based on the binary and relative attributes, respectively. Binary attributes mainly split the low-level feature space into two sides; i.e. positive and negative. Then, the SVM algorithm is established to maximize the margin, and its scores are used as the new data representation. On the other hand, the Newton method tries to maximize the gap between visual categories based on a definition which describes the relative ordering. So we first generate random attribute definitions with some limited constraints that assure to get exclusively different binary predicates and relative orderings. Thereafter, we convey the low-level feature vectors into a more discriminative attribute space by using their new representations, and the classification is carried on this new space.

In the experiments, we utilize a mid-scale visual recognition dataset, OSR, to evaluate the combinational attribute types and classifiers, namely SVM, kNN, and C4.5 decision tree. Also note that only a limited set of train data is used for learning both the attribute and classification models which benefits the mid-level data representation. The results reveal that the unsupervised attributes outperform the supervised definitions clearly although we produce them randomly without any effort. Additionally, KT correlation metric is used to pick the more discriminative orderings among randomly generated sequences, instead of simply selecting them randomly. This also boosts the accuracy performance slightly. Moreover, we have even better results if the supervised attributes are concatenated with the unsupervised orderings. We conclude that the unsupervised definitions can be used as the supplementary features if the supervised attributes already exist.

For the future work, we focus on the relative attribute selection issue since it already proves to be an important tool for the performance increase. Also, the classifier algorithms can surely be used in a combined form, called mixture of experts, to make better decisions at the end of the classification process. Lastly, an incremental learning scheme can be established on the proposed work which refers to learning the attribute space and category models simultaneously in an iterative way.

REFERENCES

[1] Ferrari V. and Zisserman A. “Learning visual attributes” Advances in Neural Information Processing Systems, Vancouver CA, December 2007.

[2] Lampert C.H., Nickisch H. and Harmeling S. “Attribute-Based classification for zero-shot visual object categorization” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 3, 2014.

[3] Lampert C. H., Nickisch H., and Harmeling S. "Learning To Detect Unseen Object Classes by Between-Class Attribute Transfer" Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2009.

[4] Farhadi A., Endres I. and Hoiem D. "Attribute-centric recognition for cross-category generalization" CVPR, 2010.

[5] Farhadi A., Endres I., Hoiem D. and Forsyth D. "Describing objects by their attributes" CVPR 2009.

[6] Parikh D. and Grauman K. "Relative attributes" Int'l Conference on Computer Vision (ICCV), 2011.

[7] Sharma G., Jurie F., and Schmid C. "Expanded parts model for human attribute and action recognition in still images" CVPR, pp. 652 – 659, 2013.

[8] Akata Z., Perronnin F., Harchaoui Z. and Schmid C. "Label-embedding for attribute-based classification" CVPR, pp. 819 – 826, 2013.

[9] Tamara L.B., Alexander C.B. and Jonathan S. "Automatic attribute discovery and characterization from noisy web data" ECCV, pp. 663-676, 2010.

[10] Russakovsky O. and Fei-Fei L. "Attribute learning in large-scale datasets" ECCV Workshops, pp. 1-14, 2010.

[11] Biswas A. and Parikh D. "Simultaneous active learning of classifiers & attributes via relative feedback" CVPR, 2013.

[12] Parkash A. and Parikh D. "Attributes for Classifier Feedback" European Conference on Computer Vision (ECCV), vol. 3, pp. 354-368, 2012.

[13] Rastegari M., Diba A., Parikh D., Farhadi A. "Multi-attribute queries: To merge or not to Merge" CVPR, 2013.
[14] Kumar N., Berg A.C., Belhumeur P. N., and Nayar S. K. "Attribute and smile classifiers for face verification" ICCV, 2009.

[15] Ma S., Sclaroff S. and Cinbis N.I. "Unsupervised learning of discriminative relative visual attributes" ECCV Workshop on Parts and Attributes, 2012.

[16] Karayel M. and Arica N. “Random attributes for image classification” IEEE 21th Conference on Signal Processing and Communications Applications, 2013.

[17] Wang Y. and Mori G. “A discriminative latent model of object classes and attributes” ECCV, pp. 155-168, 2010.

[18] Yu F.X., Ji R., Tsai M., Ye G. and Chang S. “Weak attributes for large-scale image retrieval” CVPR, 2012.

[19] Chen K., Gong S., Xiang T. and Loy C.C. “Cumulative attribute space for age and crowd density estimation” CVPR, pp. 2467 – 2474, 2013.

[20] Yu F.X., Cao L., Feris R.S., Smith J.R. and Chang S. “Designing category-level attributes for discriminative visual recognition” CVPR, 2013.

[21] Li W., Yu Q., Sawhney H. and Vasconcelos N. “Recognizing activities via bag of words for attribute dynamics” CVPR, pp. 2587 – 2594, 2013.

[22] Ma Z., Yang Y., Xu Z., Sebe N., Yan S. and Hauptmann A.G. “Complex event detection via multi-source video attributes” CVPR, 2013.

[23] Chen H., Gallagher A. and Girod B. “What's in a name: first names as facial attributes” CVPR, 2013.

[24] Sadovnik A., Gallagher A. and Chen T. "It's not polite to point: describing people with uncertain attributes" CVPR, 2013.

[25] Choi J., Rastegari M., Farhadi A. and Davis L.S. “Adding unlabeled samples to categories by learned attributes” CVPR, 2013.

[26] Wah C. and Belongie S. “Attribute-based detection of unfamiliar classes with humans in the loop” CVPR, pp. 779 – 786, 2013.

[27] Wang S., Joo J., Wang Y., and Zhu S.C. “Weakly supervised learning for attribute localization in outdoor scenes” CVPR, 2013.

[28] Saleh B., Farhadi A. and Elgammal A. “Object-centric anomaly detection by attribute-based reasoning,” CVPR, 2013.

[29] Bosch A., Xavier M. and Marti R. “A review: which is the best way to organize/classify images by content?” Image and Vision Computing, 2006.

[30] Ergül E., Ertürk S. and Arica N. “Unsupervised Relative Attribute Extraction” IEEE 21th Conference on Signal Processing and Communications Applications, 2013.

[31] Chang C.C. and Lin C.J. “LIBSVM : A library for support vector machines” ACM Transactions on Intelligent Systems and Technology, pp. 1-27, 2011.

[32] Shrivastava A., Singh S. and Gupta A. "Constrained semi-supervised learning using attributes and comparative attributes", ECCV, vol 3, pp. 369-383. 2012.

[33] Yu, A., and Grauman, ,K., “Just Noticeable Differences in Visual Attributes” ICCV, 2015.

[34] Verma, Y., and Jawahar, C.V., “Exploring Locally Rigid Discriminative Patches for Learning Relative Attributes” ICCV, 2015.

[35] Alpaydın E., “Support Vector Machines,” in Introduction to machine Learning, The MIT Press, London, 2004, pp. 218-225.

[36] Cortes C. and Vapnik V., "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.

[37] Waikato University, Waikato Environment for Knowledge Analysis (Weka) Versiyon 3.7.11, Waikato University, Hamilton, 2014.

[38] Coates A., Lee H. and Andrew Y. Ng. “An analysis of single-layer networks in unsupervised feature Learning,” International Conference on Artificial Intelligence and Statistics (AISTATS), 2011.

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.105-129
İşletme/Business Administration

THE EFFECTS OF PERSONALITY, USABILITY AND TECHNOLOGICAL FACTORS TO M-SHOPPING USE

Dr.Mustafa KARADENIZ[1]
Mustafa GÖZÜYUKARI[2]

[1]*Naval Science and Engineering Institute, Naval Academy, Istanbul, Turkey,*
mkaradeniz@dho.edu.tr

[2] *Naval Science and Engineering Institute, Naval Academy, Istanbul, Turkey,*
mgozuyukari@hotmail.com

Date of Receive: 26.12.2015 **Date of Acceptance: 29.03.2016**

ABSTRACT

Nowadays mobile devices such as smartphones and tablets become widespread and this led mobile shopping to be conducted anytime and anywhere and increased the attention to m-shopping. In this paper we studied the factors effecting the use of m-shopping value from Personalization, Self-Efficacy, Intimacy, Simplicity, Mobility, Connectivity perspectives. The m-shopping value that users experience during m-shopping can be divided into utilitarian value and hedonic value. The results show that personalization, self-efficacy, intimacy, simplicity, mobility, and connectivity variables have effect on m-shopping value.

In this study sample consists of 342 people above 18 years of age living in Istanbul. A public survey is used as data collecting method and a factor analysis, T-tests, an ANOVA/Welch test and a reliability analysis are performed for the acquired data by using the SPSS package program. Moreover, the model structured for the study is tested through a LISREL structural equation model.

KİŞİLİK, KULLANILABİLİRLİK VE TEKNOLOJİK FAKTÖRLERİN MOBİL ALIŞVERİŞE OLAN ETKİSİ

ÖZ

Günümüzde akıllı telefonlar ve tabletler gibi taşınabilir cihazların yaygınlaşması bu cihazlarla her zaman ve her yerde alışveriş yapılmasına imkan vermiş ve yapılan alışverişin artmasına neden olmuştur. Bu makalede mobil alışverişi etkileyen faktörler "kişiselleştirme", *“öz yeterlilik”, “mahremiyet”, “sadelik”, “mobilite” ve “bağlanılabilirlik” gibi değişkenler üzerinden incelenmiştir. Kullanıcıların mobil alışveriş sırasında tecrübe ettikleri mobil alışveriş değeri faydacı ve hazcı olmak üzere iki unsurdan oluşmaktadır. Sonuçlar “kişiselleştirme”, “öz yeterlilik”, “samimiyet”, “sadelik”, “mobilite” ve “bağlantı” gibi değişkenlerin mobil alışveriş değeri üzerinde etkileri olduğunu göstermektedir.*

Bu çalışmanın örneklemi İstanbul ilinde yaşayan 18 yaş ve üzeri 342 kişiden oluşmaktadır. Veri toplamak maksadıyla anket yöntemi kullanılmış ve elde edilen veriler SPSS istatistiksel paket programı kullanılarak faktör analizi, t-testi, Anova/Welch testleri ve güvenilirlik analizi testlerine tabi tutulmuştur. Ayrıca kurulan model LISREL yapısal eşitlik modellemesi vasıtasıyla test edilmiştir.

Keywords: *Mobile shopping, utilitarian value, hedonic value.*

Anahtar Kelimeler: *Mobil alışveriş, Faydacı değer, Hoşlanma değeri*

Literature Review

Worldwide mobile phone sales to end users totaled 417 million units in the third quarter of 2010, a 35 percent increase from the third quarter of 2009, according to Gartner, Inc. Smartphone sales grew 96 percent from the third quarter last year, and smartphones accounted for 19.3 percent of overall mobile phone sales in the third quarter of 2010 [1]. Worldwide mobile voice and data revenue will exceed one trillion dollars a year by 2014, according to Gartner, Inc. Mobile will generate revenue from a wide range of additional services such as context, advertising, application and service sales, and so on. Each of these will be a significant business worth several tens of billions of dollars per year [2].

Smartphone technology is exponentially evolving and significantly impacting consumers’ behavior, marketing and business activities, education and mobile industry. As a consequence, studying and understanding key factors that affect adoption of smartphone technology has become more important in business and marketing activities, improving product, and meeting consumers’ expectations. Besides, scholars from different fields and interests mostly agreed on the importance of smartphone technology as critical evolutions in the information technology. Smartphone technology' importance and popularity is increasing and showing more promising futures. According to Gartner Inc. as of third quarter of the year 2011, Smartphone represented 26% of mobile phone sale and that represent an increase of about 42% of the third quarter of year 2010. Also, Smartphone' sales are expected to sharply increase in the next 3 years [3]. According to an eMarketer forecast global mobile app audience is expected to pass 2 billion this year, according to September 2014 estimates from 451 research. The research firm predicted that the number of active mobile app users worldwide would rise from 1.81 billion to 2.17 billion between 2014 and 2015. By 2018, it expected this total to pass 3 billion [4]. According to IDC (International Data Center) research data, smartphone sales increased by 6.8% in 3rd quarter of 2015 comparing to the same quarter of 2014 [5].

Table 1: Top Five Smartphone Vendors, Shipments, Market Share and Year-Over-Year Growth, (Units in Millions)

Vendor	3Q15 Shipment Volumes	3Q15 Market Share	3Q14 Shipment Volumes	3Q14 Market Share	Year-Over-Year Change
Samsung	84.5	23.8%	79.6	23.9%	6.1%
Apple	48.0	13.5%	39.3	11.8%	22.2%
Huawei	26.5	7.5%	16.5	5.0%	60.9%
Lenovo	18.8	5.3%	16.9	5.1%	11.1%
Xiaomi	18.3	5.2%	17.3	5.2%	5.6%
Others	159.1	44.8%	163.0	49.0%	-2.4%
Total	355.2	100.0%	332.6	100.0%	6.8%

Source: http://www.idc.com/getdoc.jsp?containerId=prUS25988815, (06.03.2016)

Scholars have done significant effort in studying various aspects related to Smartphone technology to explore and better understand users' adoption of Smartphone technology. Smartphone technology evolves fast, and its popularity increased grasping more attention among scholars in both industry and academia. As it appears in figure 1, publication on research in subject related to adoption of Smartphone technology is increasing continuously specially in the last five years which indicates importance of understanding and studying the adoption of Smartphone technology among scholars in various field [3].

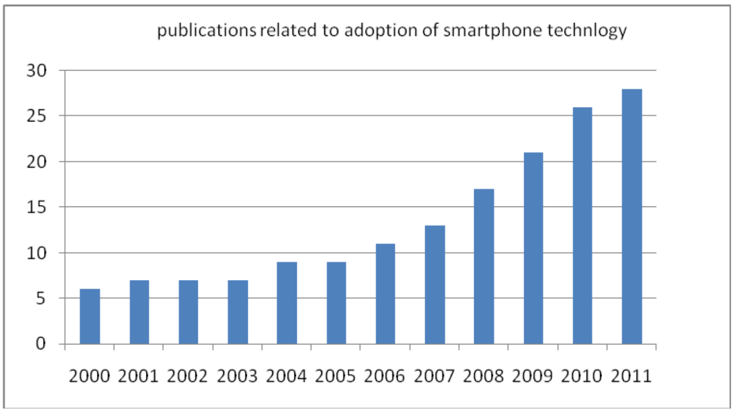


Figure 1: Publications Related to Smartphone Technology
Source: Aldhaban, 2012, p.2759

With the rapid changing atmosphere of mobile devices consumer preferences are also changing and this transforms the way of consumers' shopping experience. Technology force consumers to use their mobile devices to make purchases and to buy anything which they possibly need and want immediately from anywhere and accessed at anytime [6,7,8,9]. This new type of shopping mode, named in different ways such as; Internet shopping, e-shopping, net shopping, web-based shopping, online shopping, or mobile shopping. In this type of shopping customers are free from having to personally visit physical stores [10].

M-shopping can be defined as a popular approach for modern consumers to order or pay for goods using mobile devices [11]. Yang and Kim [12] describes m-shopping as an influential medium for connecting customers with retailers and ultimately in generating sales.

Mobile business (m-business) applications have also grown exponentially even though they have been slow to catch on implementing mobile applications for consumers. M-business applications have created tremendous business opportunities and provided benefits such as lowered operational costs, improved productivity and created fast shopping [13].

With the rapid growth of the mobile internet, shopping becomes extremely flexible in terms of space, time and channels. Mobile devices have a number of characteristics, such as ultra-portability, location sensitivity and personal nature which enable consumers to use these devices for a number of shopping activities: creating shopping lists, query, search, comparison, purchase, and post-purchase. Consumers use their mobile devices for numerous pre-purchase activities, such as finding store locations, finding promotions, consulting opening hours, making price comparisons, finding retailers of particular products, browsing for product information and product reviews, checking product availability in-store and purchasing. With m-commerce, consumers can access retailers' offers and product information anywhere and anytime while shopping becomes extremely flexible in terms of time and space. Consumers can visit a retail website via a mobile device even in a competing retail store, and even purchase at a competitor's on-line shop without leaving your brick-and-mortar store! In 2014, three-quarters of American smartphone owners believed they would be more likely to shop at a store offering services via a mobile application [14].

The popularity of smartphone usage has resulted in increased mobile shopping (m-shopping). Mobile devices especially phones facilitate the use of mobile shopping anytime and anyplace; this has heightened people's expectations and interest in this new shopping type. Unlike in the case of PC-based business, mobile devices play a critical role in m-shopping. Previous studies have showed that as a value-added service within m-commerce, m-shopping appears to be a new opportunity for increasing revenue through the use of mobile devices anytime and anywhere. This indicates a need for a better understanding of the reason behind the rapid growth of shopping that is done with smartphones [15].

However, despite the industry's conviction that the mobile Internet is the next "killer application," the reactions of actual users are quite negative in terms of usability. Their disappointing experiences with the mobile Internet result from the limitations that distinguish mobile devices from conventional desktop PCs [16]. Smaller screen sizes on mobile phones increase the cost to the user of browsing for information. In addition, a wider range of offline locations for mobile Internet usage suggests that local activities are particularly important. It is found that ranking effects are higher on mobile phones suggesting higher search costs links that appear at the top of the screen are especially likely to be clicked on mobile phones and the benefit of browsing for geographically close matches is higher on mobile phones: stores located in close proximity to a user's home are much more likely to be clicked on mobile phones. Thus, the mobile Internet is somewhat less "Internet-like": search costs are higher and distance matters more. [17].

The tendency of shopping behavior happening today is related to consumers' underlying motivations to shop. Shopping activity is initially done by the consumers with rational motives regarding with the benefits of products. Another value influencing consumers' shopping activity is emotional value known as hedonic. Moreover, consumers will take extra aspects into their consideration covering pleasure and joy aspects (hedonism) that can be gained apart from the product profits that can be enjoyed through shopping activities. Today, consumers are more recreation-oriented that accentuates pleasure, joy, and entertainment aspects when shopping [18].

Hedonism comes from the Greek word 'hedone', meaning pleasure. The central theory of hedonism is that the natural objective of human life is to attain pleasure, considered as the highest good, and to refrain pain. There are many different views of pleasure some involving a hierarchy of different pleasures. In British philosophy, the hedonistic current is linked to utilitarianism which is descibed as the greatest happiness principle [19]. Hedonic shopping value can be defined as shopping's potential entertainment and emotional worth, whereas utilitarian value reflects shopping with a work mentality [20].

Consumers are influenced by both hedonic and utilitarian shopping value when they make the decision to buy. They prefer some products to meet their utilitarian expectations and some to satisfy their hedonic desires. These two shopping motives are considered as the opposite of each other. However, consumers are influenced by both types of shopping value together when shopping for most products. Consumers’ purpose to satisfy hedonic desires and acquire utilitarian expectations may happen at the same time or different times. For example, a tooth paste provides utilitarian value by preventing caries and hedonic value with its nice taste. It means that utilitarian and hedonic reasons or motivations don’t necessarily exclude each other for consumption [21].

Utilitarian shopping is a consumer behaviour, which based on acting rationally and effectively to look for solutions to problems, realizes a specific purpose, and finally acquires the optimal value [20, 22]. The decision making processes of a consumer in utilitarian shopping go through rational processes. This approach is related to utilitarian benefit, and the consumer focuses of the functional features of a product. For hedonic shopping, it is first needed to understand hedonism which is a philosophical current. Hedonism is defined as a life style dedicated to pleasure. While it is a rare behaviour that an individual devotes oneself completely to pleasure, the search for hedonic experiences is very common. Hedonic shopping value describes the value which is acquired from the multisensory, fantasy, and affective aspects of the shopping experience [20, 23]. According to this definition, hedonic shopping value not tangible as in pragmatic shopping value, it is rather experimental and affective. Shopping is not just a boring task that needs to be completed by consumers; but an activity providing pleasure.

Hedonic shopping motivation is a person’s motive to buy something based on sensory pleasures, emotional responses and dreams. Kusuma et al. [18] mentions six hedonic shopping motivations: Gratification shopping motivation is when consumers shop to relieve stress, alleviate negative mood, and forget about present problems; Adventure shopping motivation that occurs when consumers shop for stimulation, adventure, and the feeling of being in their own world; Role shopping motivation that happens when consumers feel enjoyment from shopping for others rather than for their own selves; Value shopping motivation is when consumers assume shopping as a bargaining game, hence they seek stores that offer discounts, sales or bargains; Social shopping motivation that occurs when consumers feel enjoyment and gain a lot of information on potential product by shopping with family and friends, and view shopping as a social activity with other consumers or workers at the mall; idea shopping motivation that happens when consumers shop to keep up with the latest fashion trends and see new products and innovations.

2. Research

Research Model

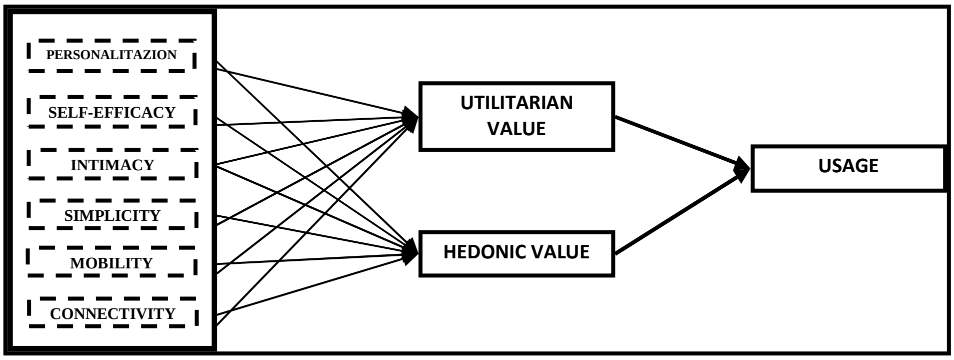


Fig. 1.1: Research Model

Source: [15]

2.1. Research Objective

The aim of this study is to investigate the factors influencing the use (Usage) of mobile shopping (m-shopping) value from Personalization, Self-Efficacy, Intimacy, Simplicity, Mobility, Connectivity perspectives. The m-shopping value that users experience during m-shopping can be divided into utilitarian value and hedonic value.

2.2. Sample Size and Sampling Technique

In this study, a total of 342 people above 18 years old chosen, and are surveyed between 16 January- 10 February. Using three indicators for each latent variable and sample size of above 200 is enough for a research [24].

2.3. Research Instrument

Research data are obtained through a two-part survey conducted to the sample given above. In the first part of the survey there are 36 statements measured on five point Likert scale (1=Strongly agree, 2=Agree, 3=Neither agree nor disagree, 4=Disagree, 5=Strongly disagree). The second part of the survey contains demographic characteristics such as gender, marital status, age, education, profession and income level.

2.4. Data Analysis

SPSS (Statistical Package for Social Sciences) statistics package has been used in the analysis of interpretation of data, while LISREL structural equation model has been chosen for testing the validity and reliability of the developed model to see if the model is valid for Turkey. Statistical analyses have been performed and survey results have been examined through SPSS program. Statistical analyses and tests used in research data analysis are as follows: Frequency Analysis, Factor Analysis, Reliability Analysis, t-Tests and ANOVA/Welch tests.

Demographic features of survey participants have been tested by frequency analysis and then the service quality dimensions tested by factor analysis. Following the factor analysis, reliability of dependent and independent variables have been tested by Cronbach's Alfa method. The result of that analysis shows that answers given to survey questions have had a high rate of internal consistency.

In order to test the significance and reliability of research model, a second-order confirmatory factor analysis has been performed, following which goodness of fit statistics, t value and standardized solution results have been examined. Consequently, the model has been found to be significant and reliable, along with being tested as acceptable.

2.4.1. Demographic Characteristics

Demographic characteristics of respondents are presented in Table 2.1.

Table 2.1. Demographic Characteristics of Respondents.

		Frequency	**Percentage**
Gender	Female	142	41,5
	Male	200	58,5
	Total	342	100
Marital Status	Single	184	53,8
	Married	158	46,2
	Total	342	100
Age	18-29	159	46,5
	30-39	143	41,8
	40-49	26	7,6
	50+	14	4,1
	Total	342	100
Education	Primary school	9	2,6
	High school	83	24,3
	Associate degree	86	25,1
	University	120	35,1
	Postgraduate	44	12,9
	Total	342	100
Profession	Public employee	121	35,4
	Private sector employee	115	33,6
	Student	85	24,9
	Retired	15	4,4
	Unemployed	6	1,8
	Total	342	100
Income level	0-1500	98	28,7
	1501-3000	166	48,5
	3001-4500	65	19,0
	>4501	13	3,8
	Total	342	100

The data in Table 2.1. show that;

a. Of the total 342 survey participants, %41.5 (142 people) is female and %58.5 (200 people) is male,

b. Of the total 342 survey participants, %53,8 (184 people) is single and %46.2 (158 people) is married,

c. Of the total 342 survey participants, %46,5 (159 people) is between the age of 18-29, %41,8 (143 people) is between 29-39, %7,6 (26 people) is between 39-49, %4,1 (14 people) is above 50,

d. Of the total 342 survey participants, %2,6 (9 people) is primary school graduate, %24,3 (83 people) is high school graduate, %25.1 (86 people) has associate degree, %35,1 (120 people) is university graduate and %12,9 (44 people) is postgraduate,

e. Of the total 342 survey participants, %35,4 (121 people) is public employee, %33.6 (115 people) private sector employee, %24,9 (85 people) is student, %4,4 (15 people) is retired and %1,8 (6 people) is unemployed,

f. Of the total 342 survey participants, %28,7 (98 people)' income is under 1500 tl, %48,5 (166 people)' income is between 1501-3000 tl, %19,0 (65 people)' income is between 3001-4500 tl and %3,8 (13 people)' income is above 4501 tl.

2.4.2. General Findings

The statistical data of the responses of survey participants to statements are shown in Table 2.2.

Table 2.2.: The mean values of the participants' responses to questions.

STATEMENTS	Min	Max	Mean	St. Dev.
Mobile shopping provides information/services that are tailored to my needs.	1	4	1,76	0,782
I can order products that fit my needs through mobile shopping.	1	4	1,80	0,798
Mobile shopping provides me with personalized information.	1	4	1,81	0,798

STATEMENTS	**Min**	**Max**	**Mean**	**St. Dev.**
Mobile shopping provides me with personalized services.	1	4	1,75	0,776
I have no great difficulty using mobile shopping.	1	5	3,88	0,897
I do not need help from other people in using mobile shopping.	1	5	3,83	0,906
I am better able to use mobile shopping than my friends.	1	5	4,01	0,881
I can respond effectively to unexpected events that may occur during mobile shopping.	1	5	3,90	0,907
I have experienced mobile shopping.	1	5	3,75	1,386
I experience easy communication during mobile shopping.	1	5	3,79	1,369
I feel an affinity toward mobile shopping.	1	5	3,76	1,383
Mobile shopping is convenient.	1	5	3,80	1,365
It’s easy to learn the process of using mobile shopping.	1	3	1,58	0,670
Mobile shopping is very easy.	1	3	1,58	0,675
For me, mobile shopping is a simple and easy-to-learn service.	1	3	1,53	0,634
I can find what I need quickly through mobile shopping.	1	3	1,58	0,691
I can do other things while mobile shopping.	1	3	1,53	0,725
Wherever I am, I can obtain the service I want thorough mobile shopping.	1	3	1,57	0,742
It is possible to use mobile shopping anytime, anywhere.	1	3	1,55	0,724
I can access products or services that I need while moving through mobile shopping.	1	3	1,53	0,729
Mobile shopping can be used anytime, anywhere.	1	5	3,22	1,286
Mobile shopping can be used regardless of the location.	1	5	3,23	1,283
Mobile shopping can provide the real-time information that I am interested in.	1	5	3,23	1,286
Mobile shopping is avaliable without time constraints.	1	5	3,23	1,290
I can browse a wide range of products in a short time.	1	5	1,67	0,877
Mobile shopping enables economical shopping.	1	5	1,78	0,909
Mobile shopping can provides me with important	1	5	1,70	0,895

STATEMENTS	Min	Max	Mean	St. Dev.
and valuable information.				
Mobile shopping is convenient and practical because it requires less time and effort.	1	5	1,74	0,894
I use mobile shopping more having fun than for purchasing products.	1	5	1,96	1,048
I enjoy mobile shopping.	1	5	2,02	1,054
I feel pleasure during mobile shopping.	1	5	1,98	1,048
I can feel the exciting shopping atmosphere thorugh mobile shopping.	1	5	2,01	1,064
I frequently purchase products through mobile shopping.	1	5	2,45	0,874
I always use mobile shopping.	1	5	2,48	0,859
I use mobile shopping very often.	1	5	2,46	0,875
I use mobile shopping periodically.	1	5	2,46	0,868

(1: Strongly Agree; 5: Strongly Disagree)

2.4.3. Reliability And Factor Analysis

The internal consistency of the study was calculated using the statistical Cronbach's Alpha coefficient. Cronbach's Alpha values of 0.854, 0.911 and 0,974 in the presence of the research question show that it has a high internal consistency.

To measure the number of sub-dimensions, a factor analysis is applied to survey statements. Factor analysis is generally used to analyze the correlation level of variables with each other. As a result of factor analysis, by means of summarizing data consisting large number of variables, less factor groups are generated with minimum level of data loss [25]. Kaiser-Meyer-Olkin (KMO) sample adequacy criteria is an index that compares observed correlation coefficients to size of partial correlation coefficients for the variables in factor analysis. KMO rate is required to be greater than 0,5. The greater this rate is, the better it is to perform a data set factor analysis [26]. As a result of KMO test applied to survey data, KMO value is found as 0,821. This demonstrates that suitability of variables to the factor analysis is at very good level. Furthermore, provided that p value of Bartlett test is less than 0,05 significance level, one can say that there is enough level of relationship between variables to perform a factor analysis [27]. As the results of both Kaiser-Meyer-Olkin (KMO) sample adequacy test and Bartlett sphericity test are meaningful, data set is found acceptable for factor analysis (KMO=0,821, χ^2Bartlett Test (630)= 17411,178, p=0.000).

Table 2.3. Factors with Regard to Survey Statements

Factors	Items	Loadings	Explanation of Factor (%)	Cronbach's Alpha
Personalization	Mobile shopping provides information/services that are tailored to my needs.	0,910	9,747	0,943
	I can order products that fit my needs through mobile shopping.	0,881		
	Mobile shopping provides me with personalized information.	0,893		
	Mobile shopping provides me with personalized services.	0,912		
Self-Efficacy	I have no great difficulty using mobile shopping.	0,896	9,325	0,933
	I do not need help from other people in using mobile shopping.	0,896		
	I am better able to use mobile shopping than my friends.	0,768		
	I can respond effectively to unexpected events that may occur during mobile shopping.	0,882		
Intimacy	I have experienced mobile shopping.	0,898	10,116	0,984
	I experience easy communication during mobile shopping.	0,863		
	I feel an affinity toward mobile shopping.	0,889		
	Mobile shopping is convenient.	0,885		

Factors	Items	Loadings	Explanation of Factor (%)	Cronbach's Alpha
Simplicity	It's easy to learn the process of using mobile shopping.	0,793	8,653	0,896
	Mobile shopping is very easy.	0,839		
	For me, mobile shopping is a simple and easy-to-learn service.	0,896		
	I can find what I need quickly through mobile shopping.	0,853		
Mobility	I can do other things while mobile shopping.	0,950	10,272	0,971
	Wherever I am, I can obtain the service I want thorough mobile shopping.	0,928		
	It is possible to use mobile shopping anytime, anywhere.	0,939		
	I can access products or services that I need while moving through mobile shopping.	0,936		
Connectivity	Mobile shopping can be used anytime, anywhere.	0,947	11,097	0,993
	Mobile shopping can be used regardless of the location.	0,950		
	Mobile shopping can provide the real-time information that I am interested in.	0,945		
	Mobile shopping is avaliable without time constraints.	0,948		
Utilitarian Value	I can browse a wide range of products in a short time.	0,885	8,870	0,916
	Mobile shopping enables economical shopping.	0,806		
	Mobile shopping can provides me with important and valuable information.	0,840		
	Mobile shopping is convenient and practical because it requires less time and effort.	0,841		

Factors	**Items**	**Loadings**	**Explanation of Factor (%)**	**Cronbach's Alpha**
Hedonic Value	I use mobile shopping more having fun than for purchasing products.	0,911	10,127	0,965
	I enjoy mobile shopping.	0,906		
	I feel pleasure during mobile shopping.	0,882		
	I can feel the exciting shopping atmosphere thorugh mobile shopping.	0,920		
Usage	I frequently purchase products through mobile shopping.	0,962	10,393	0,974
	I always use mobile shopping.	0,953		
	I use mobile shopping very often.	0,948		
	I use mobile shopping periodically.	0,948		

In social sciences, factor analysis is used to test construct validity. However it is required to calculate numerically the reliability of factors obtained via factor analysis and this calculation can be made by using Alpha model. Factors and the statements under them are reliable provided that Cronbach's Alpha value regarding each factor is 0,70 and above [27]. As a result of factor analysis applied to survey data, minimum Cronbach Alpha value is determined as 0.896 and we can say that the factors are reliable.

Table 2.4. Results of Correlation Analysis

	St. Dev.	**PER**	**SEL**	**INT**	**SIM**	**MOB**	**CON**	**UTI**	**HED**	**USA**
PER	0,729	0,808*								
SEL	0,819	-0,353**	0,743*							
INT	1,344	-0,281**	0,517**	0,780*						
SIM	0,582	0,259**	-0,217	-0,217**	0,715*					
MOB	0,699	0,167**	-0,033	-0,122	0,356**	0,880*				
CON	1,272	-0,005	0,273**	0,519**	0,069	0,147**	0,897*			
UTI	0,798	0,192**	0,183**	0,120	0,258**	0,226**	0,100	0,711*		
HED	1,00	0,068	0,206**	0,101	0,058	0,092	-0,236**	0,441**	0,818*	
USA	0,836	-0,038	-0,012	0,160**	0,002	-0,023	0,073	0,193**	0,189**	0,907*

* Square root mean of AVE values are diagonally.

** Correlations statistically significant at 0.01 levels. (2-tailed)

2.4.4. Testing the Developed Model and Hypotheses with the Structural Equation Model

A confirmatory factor analysis has been made via LISREL structural equation model. The goodness of fit statistics are as follows: chi-square (χ^2) value=1891.15, p=0,00; Degrees of Freedom= 565; χ^2/sd= 3,34; Root Mean Square Error of Approximation-(RMSEA)=0.08; Goodness of Fit Index (GFI)=0.76; Adjusted Goodness of Fit Index (AGFI)=0.72; Comparative Fit Index (CFI)=0.93; Normed Fit Index (NFI)=0.90; Root Mean Square Residual (RMR)=0.05 and Standardized Root Mean Square Residual (SRMR)=0.05. Values derived from the structural equation model and the acceptability criteria of the goodness of fit statistics [24] are shown in below.

Table 2.5. Values Derived from the Structural Equation Model and the Acceptability Criteria of the Goodness of Fit Statistics

Goodness of fit Index	**Values Derived from the Model**	**Acceptability Criteria**
Chi-Square (χ^2)/ sd	3,34	≤ 3 perfect fit
GFI	0,76	Acceptable fit
RMSEA	0,08	≤ 0,08 Good fit
RMR	0,05	≤ 0,05 perfect fit
SRMR	0,05	≤ 0,05 perfect fit
CFI	0,93	≥ 0,90 Good fit
NFI	0,90	≥ 0,90 Good fit

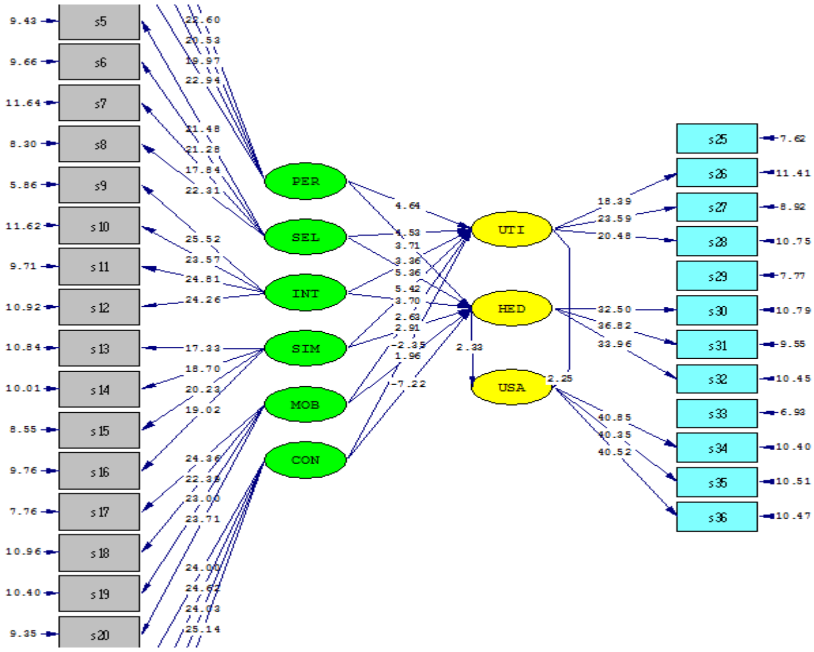


Fig. 2.1. The Results of Measurement Model

Table 2.6 Results of the Structural Equation Analysis

Dependent Variable	Independent Variable	Path Coefficients	t Values
(UTI)	(PER)	0,27	4,64
	(SEL)	0,29	4,53
	(INT)	0,23	3,36
	(SIM)	0,33	5,42
	(MOB)	0,15	2,63
	(CON)	-0,14	-2,35
(HED)	(PER)	0,21	3,71
	(SEL)	0,34	5,36
	(INT)	0,25	3,70
	(SIM)	0,17	2,91
	(MOB)	0,11	1,96
	(CON)	-0,44	-7,22
(USA)	(UTI)	0,13	2,25
	(HED)	0,13	2,33

When both the path diagram derived from the second-order confirmatory factor analysis and t values are examined, t values and standardized solution values are seen to be meaningful with 0,05 reliability level. The T values and path coefficients from the measurement model are shown in Figure 2.1.

When the goodness of fit statistics in Table 2.5 and the results of the structural equation analysis in Table 2.6 are taken into consideration, the model, which is used for investigating the relationship between personalization, self-efficacy, intimacy, simplicity, mobility, connectivity factors and m-shopping use (Usage) through the mediating effect of m-shopping value (Hedonic, Utilitarian), which is tested by LISREL structural equation model, is found satisfying in terms of significance and reliability, through which the fact that this model is acceptable has been tested.

2.4.5. Results of T-tests and ANOVA/Welch tests

Research model dimensions are tested through independent t-test and one way ANOVA/Welch tests. First, the dimensions were tested by independent samples t–test according to the participants' gender. The test results show that there is significant difference in the variables Intimacy and Usage according to the participants'gender.

Table 2.7 T-Test Results According to Gender

Variables/ dimensions	Gender	N	Mean	Std. Dev.	p value (Sig.)
Personalization	Female	142	1,737	0,745	0,367
	Male	200	1,810	0,717	
Self-Efficacy	Female	142	3,93	0,798	0,613
	Male	200	3,88	0,835	
Intimacy	Female	142	3,51	1,536	0,040
	Male	200	3,96	1,158	
Simplicity	Female	142	1,56	0,613	0,924
	Male	200	1,57	0,561	

Mobility	Female	142	1,54	0,696	0,953
	Male	200	1,54	0,703	
Connectivity	Female	142	3,00	1,299	0,060
	Male	200	3,38	1,231	
Utilitarian Value	Female	142	1,65	0,760	0,196
	Male	200	1,77	0,823	
Hedonic Value	Female	142	1,95	0,969	0,577
	Male	200	2,01	1,026	
Usage	Female	142	2,34	0,896	0,030
	Male	200	2,54	0,782	

The difference in variables is tested through One Way Variance Test (ANOVA/Welch) according to participants' age. In the first step of One Way Variance Test, the equation of variances has to be tested. If the variances are homogene the ANOVA test should be used, and if the variances are not homogene the Welch test should be used [27]. The homogeneity and One Way Variance Analysis Tests show that; there is significant difference in the variables named as Self-Efficacy, Intimacy, Mobility, Utilitarian and Hedonic Value according to the participants' age.

Table 2.8 One Way Variance (Anova / Welch) Test results According to Age

Variables/ Dımensıons	Homogeneity test P value (sig.)	p value (Sig.)		Result
		Anova	Welch	
Personalization	0,411	0,059	-	H_o accepted
Self-Efficacy	0,000	-	0,000	H_o rejected
Intımacy	0,000	-	0,000	H_o rejected
Simplicity	0,681	0,686	-	H_o accepted
Mobility	0,000	-	0,025	H_o rejected
Connectivity	0,000	-	0,480	H_o accepted

Utilitarian Value	0,516	0,007	-	H_o rejected
Hedonic Value	0,004	-	0,000	H_o rejected
Usage	0,009	-	0,560	H_o accepted

3. Conclusion

With the advance in mobile devices, customers had the advantage of shopping anywhere and anytime. Mobile devices enables customers to shop anytime, anywhere, thereby providing them with new value. This study provides an empirical analysis of the relations between m-shopping characteristics and use through the mediating effect of m-shopping value, and the results have important theoretical and practical implications.

In this study we analysed the relationships between personality (personalization and self-efficacy), usability (intimacy and simplicity) and technological (mobility and connectivity) factors and m-shopping use through the mediating effect of m-shopping value. The results show that m-shopping has relations with factors such as personalization, self-efficacy, intimacy, simplicity, mobility, and connectivity. The shopping value which customers have while m-shopping was both utilitarian and hedonic.

First, in contrary to other researches about m-shopping, this study not only focuses on the characteristics of m-shopping but also provides the literature by verifying the factors influencing m-shopping value and use. Because of this the study could be a basis for future research. In this study we also examined the moderating effect of user tendencies. The results show that user tendencies had moderating effects on the relationships between personalization, self-efficacy, intimacy, simplicity, mobility, connectivity and hedonic value/utilitarian value. The results also have important practical implications for m-shopping service providers. As seen in structural equation analysis personalization, self-efficacy, intimacy, simplicity and mobility had significant positive effects on both utilitarian value and hedonic value. Because of this m-shopping service providers should;

- concentrate on consumers' ability to personalize their m-shopping experience,
- use new methods for promoting consumers' self-efficacy to improve their m-shopping value,
- focus on the intimacy of customers while shopping to increase the m-shopping value,
- maximize the simplicity of m-shopping because the simpler the m-shopping experience, the greater the m-shopping value is,
- enhance the mobility of m-shopping to promote the m-shopping value,

In this study the results also show that connectivity had significant negative effects on both utilitarian value and hedonic value. This suggests that if customers can't get high-quality connectivity, their m-shopping value will be effected negatively. The results show that utilitarian value and hedonic value had significant positive relationships with m-shopping use. This means that one of the main tasks of m-shopping providers should be to increase both of utilitarian value and hedonic value.

In the literature there are a few researches studying the relations of personality, usability and technological factors with m-shopping value and use and this study was made only in Istanbul province. These are the limitations for our research. There may be other factors in relation with m-shopping value and use. With the research of literature new factors may be found and included in further researches.

REFERENCES

[1] Gartner press releases (2010a) http://www.gartner.com/it/page.jsp?id=1455314 (11.03.2016).

[2] Gartner press releases (2010b) http://www.gartner.com/it/page.jsp?id=1466313 (10.03.2016).

[3] Aldhaban, Fahad, (2012), Exploring the Adoption of Smartphone Technology: Literature Review, Proceedings of PICMET '12: Technology Management for Emerging Technologies. pp.2758-2770.

[4] https://www.emarketer.com/public_media/docs/eMarketer_ Mobile_Commerce_ Roundup.pdf (10.03.2016).

[5] http://www.idc.com/getdoc.jsp?containerId=prUS25988815 (06.03.2016).

[6] Rose, Susan, Neil Hair, and Moira Clark. (2011). Online Customer Experience: A Review of The Business-to-Consumer Online Purchase Context, International Journal of Management Reviews, Vol.13, No. 1, pp. 24-39.

[7] Safeena, Rahmath, Hundewale, Nisar, and Kamani, Abdullah (2011). Customer's Adoption of Mobile-Commerce:A Study on Emerging Economy, International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 1, No. 3, pp. 228-233.

[8] Elbadrawy, R. and Aziz, R. A. (2011). Resistance to Mobile Banking Adoption in Egypt: A Cultural Perspective, International Journal of Managing Information Technology, vol. 3 No. 4, pp. 9-21.

[9] Chen, Lisa Y. (2013). The Quality Of Mobile Shopping System And Its Impact On Purchase Intention And Performance, International Journal of Managing Information Technology, Vol.5, No.2, pp.23-32.

[10] Ahmed, T. T. (2012). Electronic Shopping Behavior In Mobile Commerce Context: An Empirical Study. In Conference proceedings of the 12th Information Technology Conference held in Saudi Arabia.

[11] Hung, M.C., Yang, S.T., & Hsieh, T.C. (2012), An Examination Of The Determinants Of Mobile Shopping Continuance. International Journal of Electronic Business Management, 10(1), pp.29-37.

[12] Yang, K., & Kim, H. (2012), Mobile Shopping Motivation: An Application Of Multiple Discriminant Analysis. International Journal of Retail & Distribution Management, 40(10), pp.778-789.

[13] Lee, C. and J. P. Shim (2006), An Empirical Study on User Satisfaction with Mobile Business Applications Use and Hedonism, Journal of Information Technology Theory and Application (JITTA), 8:3, pp.57-74.

[14] Voropanova, Ekaterina (2015), Conceptualizing Smart Shopping With A Smartphone: Implications Of The Use Of Mobile Devices For Shopping Productivity And Value, The International Review of Retail, Distribution and Consumer Research, Vol. 25, No. 5, pp.529–550.

[15] Changsu Kim, Wen Li & Dan J. Kim (2015), An Empirical Analysis of Factors Influencing Mshopping Use, International Journal of Human-Computer Interaction , 31, pp.974–994.

[16] Chae M. and Kim J. (2004), Size And Structure Matter To Mobile Users: An Empirical Study Of The Effects Of Screen Size, Information Structure, And Task Complexity On User Activities With Standard Web Phones, Behav. Inf. Technol. 23, pp.165–181.

[17] A. Ghose, A. Goldfarb, S.P. Han (2012), How Is The Mobile Internet Different? Search Costs And Local Activities, Inf. Syst. Res. 24, pp. 613–631.

[18] Kusuma Gede Wira, Syafiie Idrus, Atim Djazuli (2013), The Influence of Hedonic Shopping Motivations on Buying Decision with Gender as Dummy Variable: (A Study on Consumers at the Hardy's Mall Singaraja, Buleleng Regency, Indonesia), European Journal of Business and Management, Vol.5, No.31, pp.240-247.

[19] Koçak Ahmet (2013), The Effect Of Hedonic Shopping Values And Price In Different Age, Gender And Income Groups For Mobile Phone", Master Thesis, p.7.

[20] Babin, B.J., Darden, W.R. ve Griffin, M. (1994), Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value, Journal of Consumer Research, 20, pp.644-656.

[21] Bakırtaş Hülya, Bakırtaş İbrahim, Çetin Mümin Atalay (2015), Effects of Utilitarian and Hedonic Shopping Value and Consumer Satisfaction on Consumer Behavioral Intentions, pp. 91-98.

[22] Batra, R. ve Ahtola, O.T. (1990), Measuring the Hedonic and Utilitarian Sources of Consumer Attitudes, Marketing Letters, 2, pp.159-170.

[23] Holbrook, M.B. ve Hirschman, E.C. (1982), The Experiential Aspects of Consumption: Consumer Fantasies, Feeling and Fun, Journal of Consumer Research, 14, pp.508-522.

[24] Çokluk Ö., Şekercioğlu, G., Büyüköztürk, Ş., (2012), Sosyal Bilimler İçin Çok Değişkenli İstatistik SPSS ve LISREL Uygulamaları. 2.Baskı, Ankara, Pegem Akademi, p.266.

[25] Gegez, E., Pazarlama Araştırmaları (2005), 1.Baskı, İstanbul, Beta Basım Yayım, p.288.

[26] Kalaycı, Ş. (2010), SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, 5.Baskı, Ankara, Asil Yayın, p.322.

[27] Durmus, B., Yurtkoru, E.S., Cinko, M., (2011), Sosyal Bilimlerde SPSS'le Veri Analizi. 4.Baskı, Beta Yayınları, İstanbul, p.133.

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.80-89
Makine Mühendisliği/Mechanical Engineering

EFFECT OF MISSING PATH AND FORCE ON VIBRATION RESPONSE PREDICTION

Hakan UÇAR

Design Project Office, Istanbul Naval Shipyard, Pendik, Istanbul, Türkiye
hkucar@gmail.com

Date of Receive: 11.12.2015 **Date of Acceptance: 02.03.2016**

ABSTRACT

Accurate prediction of the vibration responses is an essential requirement during the design and optimization stages of mechanical structures. For this reason, a vibration response prediction (VRP) methodology, based on a set of frequency response functions (FRF) and operational responses, is established. In this methodology, defining the sources and transmission paths of the structure accurately is critical. Since it is based on the FRFs of the system, unconsidered sources and/or missing path result in inaccurate predictions. The goal of this paper is to present the effect of a missing path on the prediction results. In accordance with this purpose, a three degree-of-freedom system will be investigated analytically and VRP methodology will be applied along with the calculated receptances and displacements of the system.

EKSIK PATİKA VE KUVVETİN TİTREŞİM CEVABI ÖNGÖRÜSÜNE ETKİSİ

ÖZ

Mekanik yapılarınd dizayn ve optimizasyon aşamalarında yapı üzerindeki titreşimi doğru olarak öngörülmesi önemli bir ihityaçtır Bu nedenle, operasyonel cevap ve frekans cevap fonknsiyonuna (FCF) bağlı olan bir titreşim cevabı öngörüsü (TCÖ) metodu kullanılmaktadır. Bu metotta,titreşim kaynaklarını ve iletim patikalarını doğru olarak tanımlamak çok kiritiktir. Metot sistemin FCF'lerine bağlı olduğundan, hesaba katılmayan kaynak ve/veya eksik patikalar doğru olmayan öngörülere neden olacaktır. Bu makalenin amacı eksik patikanın öngörü son uçlarına olan etkisini sunmaktır. Bu amaç doğrultusunda üç sebrstlik dereceli bir system analitik olarak incelenecek ve TCÖ metodu hesaplanan deplasman ve FCF'ler ileuygulanacaktır.

Keywords: vibration response, matrix inversion, cross-coupling, 3 DOF systems.
Anahtar Kelimeler: Titreşim cevabı, matris tersine çevirme, bağıl etki, üç serbestlik dereceli system.

1. INTRODUCTION

Accurate prediction of the vibration responses is an essential requirement during the design and optimization stages of mechanical structures. The first step of the vibration response prediction (VRP) is to identify the exciting forces acting on the structure. The force data is usually not available as the direct measurement of these forces is impractical or almost impossible. Therefore, the identification of these forces by using the vibration data has been attracted a lot of attention. Verheij [1] presented dynamic stiffness method which seems to be the most basic way, especially for structures having elastomer components. However, accurate complex dynamic stiffness data is seldom available and even if there is, it is only valid for a given load condition. Hillary and Ewins [2] investigated the problem of sinusoidal force identification of a cantilever beam using a least-squares method. Dascotte and Desanghere [3] presented a methodology based on the Bayesian force identification procedure. Janssens et al. [4] investigated the use of an equivalent force method to determine the sound transmission in ships. Transmissibility method has been used to predict the vibration response at a point of interest without identifying the forces [5-8]. This approach is much simpler and faster compared to other methods but unconsidered potential cross-coupling between the paths lead to incorrect predictions. Matrix inversion method has been developed in the early 1980s [9, 10]. This technique uses an inverted matrix composed of the frequency response functions (FRFs) and a vector of measured vibration responses. In this method, defining the sources and paths accurately is critical. Thus, the decoupling program is usually created in order to avoid missing paths. A missing path means that a force is acting on the structure at a given point but no path input (e.g. an acceleration signal) is measured at that point [8]. Missing paths result in inaccurate prediction of the total force and hence, the vibration response. The goal of this paper is to present the effect of a missing path on the prediction results. For this reason, a three degree-of-freedom system will be investigated analytically and the matrix inversion method will be applied based on the calculated receptances and displacements of the system.

2. THEORY OF VIBRATION RESPONSE PREDICTION

In a dynamic system, the vibration response of a point strongly depends on not only the corresponding point force but also the other operating forces. This effect is called cross-couplings and should be considered by including all FRFs between path inputs, as in Equation (1).

$$\bar{X}_1 = F_1 H_{11} + F_2 H_{21} + \cdots + F_n H_{n1} \qquad (1)$$

where *n* is the number of path inputs or internal forces, *F* is the exciting force and *H* is FRF between the force and the path inputs.

A square FRF matrix, *n x n*, is created since the number of forces and responses are equal to each other. By taking the inverse of this FRF matrix and multiplying the vector of vibration responses, the exciting forces are identified, as shown in Equation (2).

$$\{F_i(\omega)\} = [H_{ij}(\omega)]^{-1}\{\bar{X}_i(\omega)\} \qquad (2)$$

where *i=j* which denotes the number of paths and forces.

Vibration response at the point of interest, *k* is predicted once the exciting forces acting on each path are calculated. Assuming that the system is linear and time-invariant, partial contributions of each path to the response are added up and the vibration response at the target is calculated, as seen in Equation (3).

$$\bar{X}_k(\omega) = \sum_{i=1}^{n} \bar{X}_{ki}(\omega) = \sum_{i=1}^{n} F_i(\omega) H_{ki}(\omega) \qquad (3)$$

3. THREE DEGREE-OF-FREEDOM (DOF) SYSTEM

In order to illustrate the vibration response prediction based on the matrix inversion method and the transmissibility approach, consider the following 3 DOF system.

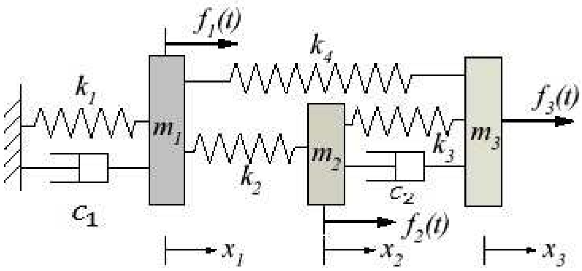


Figure 1. Three DOF system [11]

The degrees of freedom of a lumped mass system are equal to the number of vibrating mass points and the number of natural frequencies. In this system, there are three coordinates defining the movement for each moving mass; x_1, x_2, and x_3. The moving masses m_1, m_2, and m_3 are connected to each other through a series of spring and viscous dampers with the coefficients k_1, k_2, k_3 and c_1, c_2, respectively. The system is excited by the external forces f_1, f_2 and f_3.

Using the free body diagrams for each mass and analyzing the forces acting on them by applying Newton's Second Law of motion, the equation of motion for the system in matrix notation is as follows;

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}\begin{Bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{Bmatrix} + \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & -c_2 \\ 0 & -c_2 & c_2 \end{bmatrix}\begin{Bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{Bmatrix} + \begin{bmatrix} k_1+k_2+k_4 & -k_2 & -k_4 \\ -k_2 & k_2+k_3 & -k_3 \\ -k_4 & -k_3 & k_3+k_4 \end{bmatrix}\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} f_1 \\ f_2 \\ f_3 \end{Bmatrix} \quad (4)$$

As shown in Equation (4), the matrix containing the masses which is multiplied by a vector of accelerations is called mass matrix, *M*, the matrix containing the damping coefficients which is multiplied with the vector of velocity is called the structural damping matrix, *C* and similarly, the matrix containing the spring coefficients which is multiplied with the vector of displacements is called the stiffness matrix, *K*.

$$[M]\{\ddot{x}\} + [C]\{\dot{x}\} + [K]\{x\} = \{f\} \quad (5)$$

Removing the damping effects and considering the free vibration, Equation (5) becomes;

$$[M]\{\ddot{x}\} + [K]\{x\} = \{0\} \quad (6)$$

The solution of the undamped system can be expressed as a set of harmonic functions, which is simply amplitude multiplied by a complex exponential.

$$\{x\} = \{\emptyset\}\, e^{j\omega_n t} \quad (7)$$

The term $e^{j\omega_n t}$ is a mathematical expression to solve differential equations by using the phasor form for the solution. After substituting the solution to the matrix;

$$([K] - \omega_n^2 [M])\{\emptyset\}e^{j\omega_n t} = \{0\} \quad (8)$$

In Equation (8), the exponential part, $e^{j\omega_n t}$ cannot be zero for finite *t*. Thus, the equation is reduced to the problem often encountered in mathematics which is called the eigenvalue problem, as in Equation (9)

$$([K] - \omega_n^2 [M])\{\emptyset\} = \{0\} \quad (9)$$

In this eigenvalue problem, ω_n^2 contains the eigenvalues of the system, namely natural frequencies and $\{\emptyset\}$ is the eigenvector representing the mode shapes. After solving the eigenvalue problem, the modal matrix is constructed with the eigenvectors and denoted by Φ. Modal transformation is carried out by defining the relationship between the physical coordinate, *x* and modal coordinate, *q*.

$$\{x\} = [\Phi]\{q\} \tag{10}$$

The modal matrix is used to transform into modal coordinates and uncouple the equations of motion. Using the orthogonality, diagonal modal mass, damping and stiffness matrices are created as;

$$= [\Phi]^T[M][\Phi] \quad [C_q] = [\Phi]^T[C][\Phi] \quad [K_q] = [\Phi]^T[K][\Phi] \tag{11}$$

Force vector is also transformed into modal coordinates;

$$\{f_q\} = [\Phi]^T\{f\} \tag{12}$$

Substituting Equation (10) and (11) into Equation (5), the modal equation of motion for the damped system is determined as follows;

$$[M_q]\{\ddot{q}\} + [C_q]\{\dot{q}\} + [K_q]\{q\} = \{f_q\} \tag{13}$$

After solving Equation (13), the results are back-transformed to physical coordinates and a solution exists of the form $\{x\}$ depending on $\{f\}$. The relationship between $\{x\}$ and $\{f\}$ can be expressed as;

$$\{x\} = [\alpha]\{f\} \tag{14}$$

where $[\alpha]$ is the receptance FRF matrix and defines the frequency response of the model. The individual component of the receptance matrix, $\alpha_{jk}(\omega)$ is defined as follows;

$$\alpha_{jk}(\omega) = \frac{x_j(\omega)}{f_k(\omega)} \tag{15}$$

where j represents the number of DOF and k denotes the number of forces. Individual FRF parameter, $\alpha_{jk}(\omega)$ can be calculated by using the following formula based on the modal matrix elements.

$$\alpha_{jk}(\omega) = \sum_{r=1}^{N} \frac{(\phi_{jr})(\phi_{kr})}{(k_r - \omega^2 m_r) + i(\omega c_r)} \tag{16}$$

where k_r, m_r and c_r are the indices of the modal mass, damping, and stiffness matrices, respectively.

3.1. Vibration Response Prediction Study

Turning now to the VRP study based on the 3-DOF system in Figure 1, the study is composed of three parts. In the first part, x_3 is aimed to be predicted when there are just two forces applied on m_1 and m_2. In this case, there is a *(2x2)* FRF receptance matrix and a *(2x1)* vector of vibration responses as displacement vector. x_3 is aimed to be predicted by applying matrix direct inversion method along with Equation (2) and (3). The following parameters are used in this section.

Table 1 Three DOF system properties

m_1	1 kg	c_1	1 Ns/m	k_1	$1.4\text{x}10^5$ N/m	f_1	20 N
m_2	0.5 kg	c_2	1 Ns/m	k_2	$2.8\text{x}10^5$ N/m	f_2	10 N
m_3	0.2 kg			k_3	$1\text{x}10^3$ N/m	f_3	0 N
				k_4	$7\text{x}10^4$ N/m		

According to the comparison between calculated and predicted response, x_3 shown in Figure 2, it can be stated that VRP based on the matrix inversion method poses accurate prediction if all sources and paths are considered.

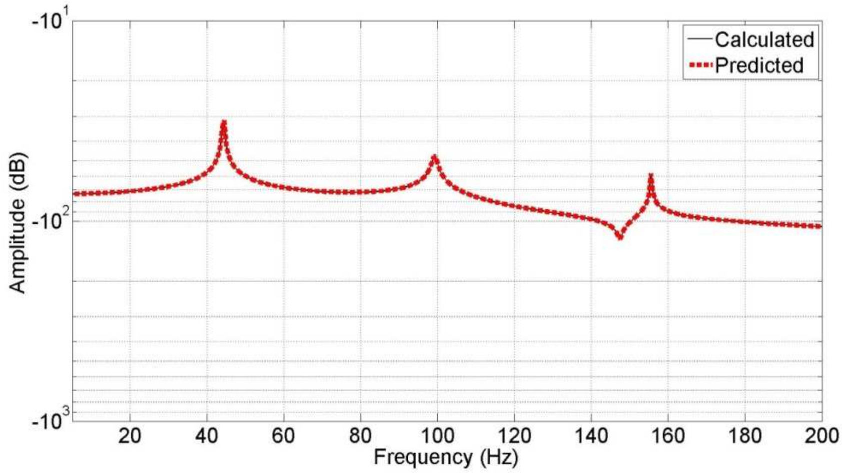


Figure 2. Vibration response, x_3

The question is now what if any path and one of the sources are not taken into account. In the second part of the study, an external force, f_3 is applied on m_3 with a magnitude of 10 N. The response, x_3 is predicted in such a case that third DOF and the external force, f_3 are not considered. Although the force also contributes to the target response, it is not taken explicitly into the calculations. In this respect, the predicted response, x_3 is shown in Figure 3.

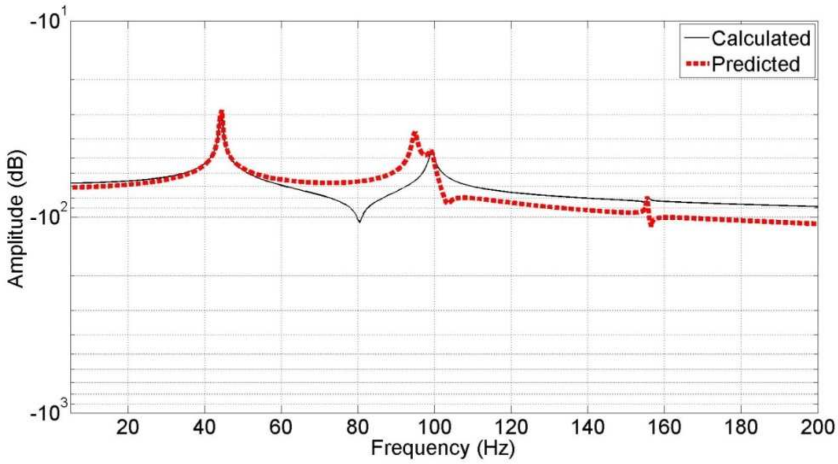
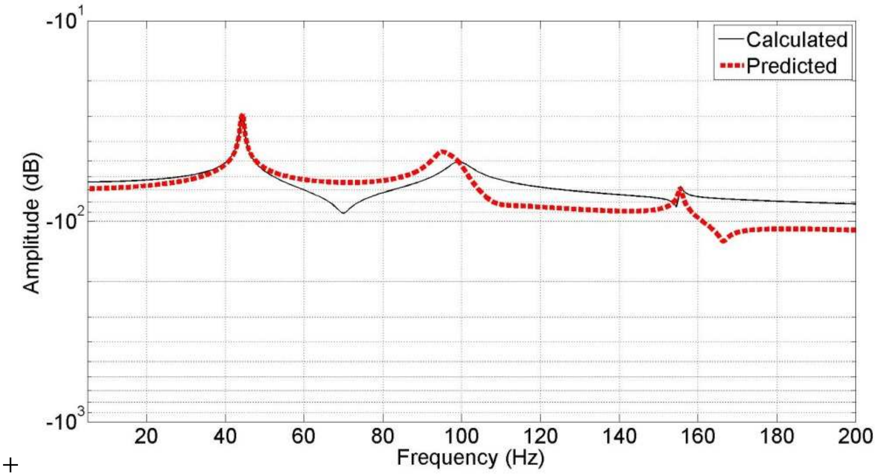


Figure 3. Vibration response, x_3 (without considering f_3 = 10 N)

As shown in Figure 3, unconsidered source and path result in remarkable error up to 40 dB. In the third case study, f_3 is increased up to 20 N and the response at x_3 is predicted without considering f_3, as shown in Figure 4.



+ */+

Figure 4. Vibration response, x_3 (without considering f_3 = 20 N)

According to the analytical case studies, it is shown that the resulting error is proportional to the amplitude of the omitted force and the error increases as 2 dB when the force is doubled. Since each path is connected to each other, cross-coupling effects become more of an issue. Missing paths are one of the most important inaccuracies in the VRP study. Therefore, this issue should be given utmost importance before performing VRP study.

4. SUMMARY AND CONCLUSIONS

In this study, the effect of missing path on the prediction of vibration response was discussed with the methodology of matrix inversion. Three analytical case studies were executed. Throughout the set of these analytical case studies, it is determined that unconsidered source and missing path result in remarkable error up to 40 dB. Furthermore, the resulting error in prediction depends on the amplitude of the missing force and the overall error increases up to 2 dB when the missing force is doubled. As a consequence, cross-coupling effects become more of an issue since each path is connected to each other and missing paths are one of the most important inaccuracies in the VRP study. Therefore, this issue should be given utmost importance before performing VRP study.

5. REFERENCES

[1] Verheij, J. W., (1982), Multipath Sound Transfer from Resiliently Mounted Shipboard Machinery, TNO, Delft, Netherlands

[2] Hillary, B., (1983), Indirect Measurement Of Vibration Excitation Forces, *PhD Thesis,* Imperial College London, Dynamics Section, London.

[3] Desanghere, G., Identification *of external forces based on frequency response measurements,* Proc. 8th Int. Seminar on Modal Analysis, Leuven, Belgium, 1983.

[4] Janssens, M. H. A., Verheij, J. W., and Thompson, D. J., (1999), *The use of an equivalent forces method for the experimental quantification of structural sound transmission in ships*, Journal of Sound and Vibration, 226(2), pp. 305-328

[5] Ribeiro, A. M. R., Silva, J. M. M., and Maia, N. M. M., (2000), *On the generalisation of the transmissibility concept,* Mechanical Systems and Signal Processing, 14(1), pp. 29-35.

[6] Maia, N. M. M., Silva, J. M. M., and Ribeiro, A. M. R., (2001), *The transmissibility concept in multi-degree-of-freedom systems*, Mechanical Systems and Signal Processing, 15(1), pp. 129-137.

[7] Ribeiro, A. M. R., Fontul, M., Maia, N. M. M., and Silva, J. M. M., *Further developments on the transmissibility concept for multiple degree of freedom systems*, Proceedings of the XIth International Conference on Vibration Engineering, Timisoara, Romania, 2005.

[8] Gajdatsy, P., Janssens, K., Desmet, W., and Van der Auweraer, H., (2010), *Application of the transmissibility concept in transfer path analysis,* Mechanical Systems and Signal Processing, 24(7), pp. 1963-1976.

[9] Roesems, D., *Analysis of structure borne noise transmission paths,* Proc. 6th Int. Seminar on Modal Analysis, Leuven, Belgium , 1981.

[10] Verheij, J. W., *Experimental procedures for quantifying sound paths to the interior of road vehicles,* Proc. 2nd Int. Conference on Vehicle Comfort, pp. 483-491, Bologna, Italy, 1992.

[11] Efunda, *Multiple degree - of - freedom*, 2015.

Journal of Naval Science and Engineering
2016, Vol. 12, No.1, pp.90-104
Temel Bilimler/Basic Sciences

SOLUTION OF CONTACT PROBLEM USING "MIXED" MLPG FINITE VOLUME METHOD WITH MLS APPROXIMATIONS

Cengiz ERDÖNMEZ

Turkish Naval Academy, Tuzla, Istanbul, Turkiye
cerdonmez@dho.edu.tr

Date of Receive: 24.12.2015 **Date of Acceptance: 27.03.2016**

ABSTRACT

Meshless methods are became an alternative to most popular numerical methods used to solve engineering problems such as Finite Difference and Finite Element Methods. Because of element free nature, problems are solved using meshless methods depending on the general geometry and conditions of the problem. Mixed Meshless Local Petrov-Galerkin (MLPG) approach is based on writing the local weak forms of PDEs. Moving least squares (MLS) is used as the interpolation schemes. In this study contact analysis problem is modelled using Meshless Finite Volume Method (MFVM) with MLS interpolation and solved for beam contact problem. Meshless discretization and linear complementary equation of the 2-D frictionless contact problems are described first. Then the problem is converted to a linear complementary problem (LCP) and solved using Lemke's algorithm. An elastic cantilever beam contact to a rigid foundation is considered as an example problem.

"KARIŞIK" MLPG SONLU HACİMLER YÖNTEMİ İLE MLS YAKLAŞTIRMASI KULLANILARAK TEMAS PROBLEMİNİN ÇÖZÜMÜ

ÖZ

Ağsız yöntemler son yıllarda Sonlu Farklar ve Sonlu Elemanlar Yöntemleri gibi mühendislik problemlerini çözmek için kullanılan en popüler sayısal yöntemlere alternatif haline gelmiş durumdadır. Ağsız yöntemlerin eleman bağımsız yapısı gereği problemlerin çözümleri yalnızca çözümün yapılacağı geometri ve problemin koşullarına bağlıdır. Karışık Ağsız Yerel Petrov-Galerkin (MLPG) yaklaşımı Kısmi Diferansiyel Denklemlerin (PDEs) yerel zayıf formlarının yazılması temeline dayanmaktadır. Hareketli En Küçük Kareler (MLS) yöntemi interpolasyon şeması olarak kullanılmaktadır. Bu çalışmada Ağsız Sonlu Hacimler Yöntemi (MFVM) ile MLS interpolasyon şeması birlikte kullanılarak temas analizi problemi modellenmiş ve kiriş temas problemi için çözülmüştür. Ağsız ayrıklaştırma ve 2-D sürtünmesiz temas problemlerinin doğrusal tamamlayıcı denklemleri ilk olarak açıklanmıştır. Daha sonra problem doğrusal tamamlayıcı probleme (LCP) dönüştürülmüş ve Lemke algoritması kullanılarak çözülmüştür. Örnek problem olarak rijit bir temele temas halindeki elastik bir konsol kiriş problemi ele alınmıştır.

***Keywords**: Mixed Meshless Local Petrov-Galerkin, Moving least squares, Meshless Finite Volume Method, Linear Complimentary Problem, Cantilever beam contact.*

***Anahtar Kelimeler:** Karışık Ağsız Yerel Petrov-Galerkin, Hareketli En Küçük Kareler, Ağsız Sonlu Hacimler Yöntemi, Doğrusal Tamamlayıcı Problem, Konsol kiriş teması.*

1. INTRODUCTION

Due to the complex nature of the engineering problems, scientists are interested in numerical methods to model and find solutions for engineering problems from different disciplines. In computational mechanics both Finite Difference and Finite Element Method (FEM) found a very important place to solve problems. Mostly FEM is preferred and many commercial programs are based on the theory of FEM to solve wide variety of engineering problems. But due to the element based nature, solution of the problems brings some difficulties and researches are made researches on new alternatives from the beginning of the 1990's. As a result number of mesh free methods are developed and still there are a lot of studies are continuing on this area.

The meshless methods overcoming the drawback of mesh-based methods, such as mesh-generation and poor solutions. Many mesh free methods are developed based on global weak forms, such as Smooth Particle Hydrodynamics (SPH) and the element-free methods which require certain meshes or background cells. Contrary to these methods, the meshless local Petrov Galerkin (MLPG) approach pioneered by Atluri (2004) [1] and his colleagues is based on writing the local weak forms of partial differential equations (PDEs), on overlapping local subdomains. The integration of the weak-form is also performed within the local sub-domains; thus negating any need for any kind of meshes or background cells, making the MLPG approach a truly meshless method. MLPG method is used in many areas such as elasto-statics [2-5], elasto-dynamics [6], fracture mechanics [7], fluid mechanics and etc.

The MLPG approach gives oppurtunity to select trial and test functions, define the size and shape of local sub-domains, and has the ability to use various unsymmetric and symmetric weak forms of the PDEs. As a test function Heaviside function can be used in symmetric-weak forms to eliminat domain integrals.

Meshless Finite Volume Method (MFVM), using the Meshless Local Petrov-Galerkin (MLPG) “Mixed” approach, is developed for solving elastostatic problems and labeled as MLPG-5 in [8]. In MLPG mixed approach MLS are used as the interpolation schemes. The MLPG local weak form is written for the equilibrium equations over the local sub-domains, and the Heaviside test function is used. Xiao et.al. solved the 2D contact problem using MLPG with RBF [9]. They implement the meshless linear complementary formulation which can be solved by using Lemke's algorithm.

In this study solution of the contact problem of elastic cantilever beam using Mixed MLPG Finite Volume Method (FVM) and Linear Complementary Problem (LCP) is modeled and solved. Meshless discretization and linear complementary equation of the 2-D frictionless contact problems is described. The problem is converted to LCP and solved using Lemke's algorithm. An elastic cantilever beam contacting a rigid foundation is considered as an example problem. The beam is modeled as a thin body in the plane stress state. The result found by using MFVM of beam contact problem is compared with the results available in the literature.

2. MESHLESS APPROXIMATIONS

Radial Basis Functions are proposed to interpolate large sets of multivariate data before by many researchers. But lately it has been shown that a fast and accurate approximation method for large sets of multivariate data can be accomplished. An alternative to radial basis function interpolation and approximation is the so-called moving least squares method (MLS). In the traditional moving least-squares (MLS) method the amount of work is shifted. There is no large system to solve. Instead, for every evaluation one needs to solve a small linear system to find the coefficients of the moving local approximant, and then evaluate a summation [10].

MLS is generally considered to be one of the best methods to interpolate random data with a reasonable accuracy, because of its completeness, robustness and continuity. A function $u(\mathbf{x})$ can be approximated over a number of scattered local points $\{x_i\}$, (i = 1, 2,...,n) as,

$$u(\mathrm{x}) = \mathrm{p}^T(\mathrm{x})\mathrm{a}(\mathrm{x}), \quad \forall \mathrm{x} \in \Omega_s \tag{1}$$

where $\mathrm{p}^T(\mathrm{x}) = [p_1(\mathrm{x}), p_2(\mathrm{x}), ..., p_n(\mathrm{x})]$ is a monomial basis of order n and a(x) is a vector containing coefficients, which are functions of the global Cartesian coordinates $[x_1, x_2, x_3]$, depending on the monomial basis. They are determined by minimizing a weighted discrete L_2 norm, defined, as:

$$\begin{aligned} J(\mathrm{x}) &= \sum_{i=1}^{n} w_i(\mathrm{x})\left[\mathrm{p}^T(\mathrm{x}_i)\mathrm{a}(\mathrm{x}) - \hat{u}_i\right]^2 \\ &\equiv \left[\mathrm{P}\cdot\mathrm{a(x)\text{-}\hat{u}}\right]^T \mathrm{W}\left[\mathrm{P}\cdot\mathrm{a(x)\text{-}\hat{u}}\right] \end{aligned} \tag{2}$$

where $w_i(\mathrm{x})$ are the weight functions and $\hat{u}_i$ are the fictitious nodal values.

One may obtain the shape function as,

$$u(\mathrm{x}) = \mathrm{p}^T(\mathrm{x})\mathrm{A}^{-1}(\mathrm{x})\mathrm{B}(\mathrm{x})\hat{\mathrm{u}} \equiv \Phi^T(\mathrm{x})\hat{\mathrm{u}}, \quad \forall \mathrm{x} \in \partial\Omega_x \tag{3}$$

where matrices A(x) and B(x) are defined as,

$$\mathrm{A(x)} = \mathrm{P}^T\mathrm{WP}, \quad \mathrm{B(x)} = \mathrm{P}^T\mathrm{W}, \quad \forall \mathrm{x} \in \partial\Omega_x . \tag{4}$$

The weight function in Eq. (2) defines the range of influence of node I. Normally it has a compact support. A fourth order spline weight function is used which is defined as,

$$w_I(\mathrm{x}) = \begin{cases} 1 - 6\left(\dfrac{d_I}{r_I}\right)^2 + 8\left(\dfrac{d_I}{r_I}\right)^3 - 3\left(\dfrac{d_I}{r_I}\right)^4, & 0 \le d_I \le r_{I,} \\ 0 & , \quad d_I \ge r_I, \end{cases} \tag{5}$$

where $d_I = |\mathrm{x}\text{-}\mathrm{x}_I|$ is the Eucludian distance from node x_I to point x, r_I is the size of the support for the weight function w_I and thus determines support of the node x_I.

3. MLPG FINITE VOLUME METHOD

The equations of balance of linear and angular momentum can be written as:

$$\sigma_{ij,j} + f_i = 0; \quad \sigma_{ij} = \sigma_{ji}; \quad (\ )_{,i} \equiv \frac{\partial}{\partial \xi_i} \tag{6}$$

where σ_{ij} is the stress tensor corresponding to the displacement field u_i, f_i is the body force. The boundary conditions are given by,

$$\begin{aligned} u_i &= \overline{u}_i, \quad on \quad \Gamma_u \\ t_i &\equiv \sigma_{ij} n_j = \overline{t}_i, \quad on \quad \Gamma_t \end{aligned} \tag{7}$$

where $\overline{u}_i$ and $\overline{t}_i$ are the prescribed displacements and tractions respectively on the corresponding parts of the boundary and n_i is the unit outward normal to the boundary on the correponding points of Γ. A generalized local weak form of the differential equation (6) over a local sub-domain Ω_s, can be written as:

$$\int_{\Omega_s} \left(\sigma_{ij,j} + f_i\right) v_i d\Omega = 0 \tag{8}$$

where u_i and v_i are the trial and test functions, respectively. By applying the divergence theorem, imposing the traction b.c. and using the Heaviside function as the test function, the local symmetric weak form of Eq. (8) can be found as in [11],

$$-\int_{L_s} t_i d\Gamma - \int_{\Gamma_{su}} t_i d\Gamma = \int_{\Gamma_{st}} \overline{t}_i d\Gamma + \int_{\Omega_s} f_i d\Omega . \tag{9}$$

With the constitutive relations of an isotropic linear elastic homogeneous solid, the tractions in Eq. (9) can be written in term of the strains:

$$t_i = \sigma_{ij} n_i = E_{ijkl} \varepsilon_{kl} n_j \tag{10}$$

where, $E_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right)$ with λ and μ being the Lame's constants. The strains are independently interpolated as,

$$\varepsilon_{kl}(\mathrm{x})=\sum_{K=1}^{N}\Phi^{(K)}(\mathrm{x})\varepsilon_{kl}^{(K)}. \tag{11}$$

With Eqs. (10) and (11), one may discretize the local symmetric weak-form of Eq. (9), as,

$$-\sum_{K=1}^{N}\left[\int_{L_s}\Phi^{(K)}(\mathrm{x})E_{ijkl}n_j d\Gamma\right]\varepsilon_{kl}^{(K)} - \sum_{K=1}^{N}\left[\int_{L_{su}}\Phi^{(K)}(\mathrm{x})E_{ijkl}n_j d\Gamma\right]\varepsilon_{kl}^{(K)} = \int_{\Gamma_{st}}\bar{t}_i d\Gamma + \int_{\Omega_s} f_i d\Omega. \tag{12}$$

The number of the variables reduced by transforming the strain variables back to the displacement variables via the collocation methods. The interpolation of displacements can also be accomplished by using the same shape function, for the nodal displacement variables, and written as,

$$u_i(\mathrm{x})=\sum_{J=1}^{N}\Phi^{(J)}(\mathrm{x})u_i^{(J)}. \tag{13}$$

The strain-displacement relations are given by,

$$\varepsilon_{kl} = \frac{1}{2}\left(u_{k,l} + u_{l,k}\right). \tag{14}$$

With the displacement approximation in Eq. (13), the two sets of nodal variables can be transformed through a linear algebraic matrix:

$$\varepsilon_{kl}^{(I)} = H_{klm}^{(I)(J)}u_m^{(J)} \tag{15}$$

which is reduced to the same number of nodal displacement variables by using the transformation. Data preparation, quadrature techniques and post processing issues are detailed in reference [11].

4. CONTACT PROBLEM OF A CANTILEVER ON RIGID FOUNDATION USING LCP

An elastic cantilever beam, in the plane stress state, contacting a rigid foundation is considered. The boundary and constraint conditions are illustrated in Fig. (1). Assume the beam, with a constant rigidity EI, is subjected to a uniform loading *f*. If the initial gap between the beam and the rigid foundation is given as δ^0, then the location of the contact interface computed from thin beam theory without shear deformation effect [9] is;

$$l_c = \sqrt[4]{\frac{72EI\delta^0}{f}} . \quad (16)$$

In order to verify the prediction by the present method, the thick beam theory [12] should be used as a reference.

In this paper the MLPG FVM is used with the MLS approximation and the problem is modeled using the LCP and solved. The beam is placed over a foundation with an initial gap g_i is configured in Figure (1). The boundary conditions can be easily seen from the Figure (1).

The beam parameters are: E=30,000, v=0.3, D=6, L=48, f=1, δ^0=0.01. The cross-section of the beam is assumed as rectangular with k=0.85, $G = \left(E/2(1+v)\right)$ and A=6. The location of contact by thin beam theory is l_c=24.97 and by thick beam theory [9] is l_c=21.37.

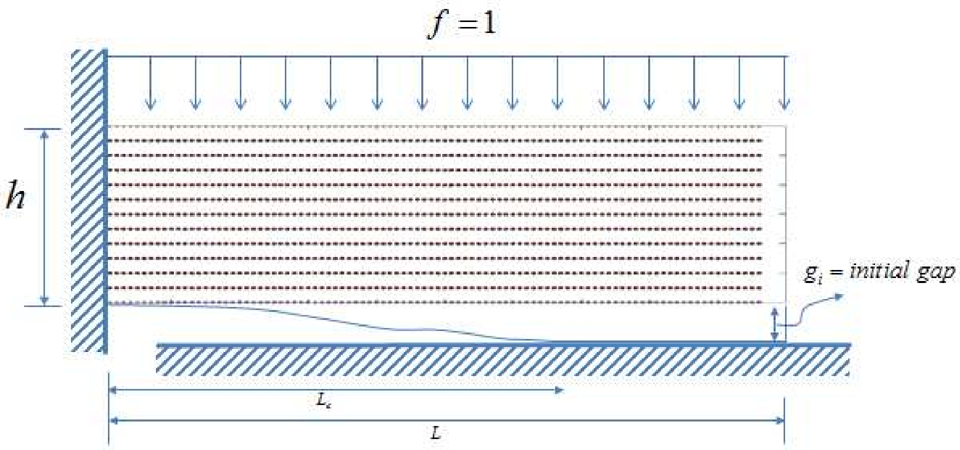


Figure 1: Contact problem of a cantilever on rigid foundation.

The contact problem is modeled using the MLPG-FVM. For each body the following equation has to be satisfied:

$$\sigma_{ij,j}(u^e)+b_i^e=0 \text{ in } \Omega^e \text{ (e=1,2)}. \quad (17)$$

The weak form of the equation (17) is found using the MLPG mixed method which gives the following equation for two body;

$$-\int_{\Gamma_{si}^e} t_i^e v^e d\Gamma - \int_{\Gamma_{st}^e} \overline{t_i}^e v^e d\Gamma \geq 0, \text{ (e=1,2)}. \quad (18)$$

By applying meshless method to the equation (18), the following discrete equation for each node is found;

$$\sum_{j=1}^{n} K_{ij}u_j - f_i \geq 0 \quad (19)$$

where,

$$K_{ij}=-\int_{\Gamma_{si}^e} t_i^e v^e d\Gamma, \quad f_i=\int_{\Gamma_{st}^e} \overline{t_i}^e v^e d\Gamma. \quad (20)$$

Finally the following matrix system is obtained for two contacted elastic body by collecting the equations obtained from each local sub-domain $\Omega_s^{(i)}$ and without any element assembly one can write;

$$\sum_{e=1}^{2}\left(K^e u^e - f^e\right) \geq 0 . \tag{21}$$

In order to obtain Linear Complementary Equation (LCE), we can write the stiffness matrix in partitioned matrix form as:

$$K^e = \begin{bmatrix} K_{ii}^e & K_{ic}^e \\ K_{ci}^e & K_{cc}^e \end{bmatrix} \tag{22}$$

Subscript '*c*' refers the contact surface, and subscript '*i*' refers to other nodes. If we use eqn. (22) in eqn. (21) and eliminate we can write eqn. (21) in a small form where we have just the unknown values of *u* along the contact surface. In this aspect we can write equation (21) as follows:

$$K_{ii}^e u_i^e + K_{ic}^e u_c^e = f_i^e , \tag{23}$$

$$\sum_{e=1}^{2}\left(K_{ci}^e u_i^e + K_{cc}^e u_c^e\right) \geq 0 . \tag{24}$$

We can find u_i from equation (23) and then put into equation (24) so we will find the following relation:

$$\sum_{e=1}^{2}\left(K_p^e u_c^e - f_p^e\right) \geq 0 \tag{25}$$

where,

$$K_p^e = K_{cc}^e - K_{ci}^e [K_{ii}^e]^{-1} K_{ic}^e , \tag{26}$$

$$f_p^e = -K_{ci}^e [K_{ii}^e]^{-1} f_i^e . \tag{27}$$

The initial gap is defined as $g_0 = u_{c1} + u_{c2}$ in Γ^c. So the following transformation based on this assumption, the gap is defined as;

$$\begin{Bmatrix} u_c^1 \\ u_c^2 \end{Bmatrix} = \begin{bmatrix} I & 0 \\ -I & I \end{bmatrix} \begin{Bmatrix} u_c^1 \\ g \end{Bmatrix}. \tag{28}$$

If we write equation (25) in expanded form we found the following one;

$$K_p^1 u_c^1 - f_p^1 + K_p^2 u_c^2 - f_p^2 \geq 0\,. \tag{29}$$

It can be written in matrix form by making some arrangements;

$$\begin{bmatrix} K_p^1 + K_p^2 & 0 \\ -K_p^2 & K_p^2 \end{bmatrix} \begin{Bmatrix} u_c^1 \\ g \end{Bmatrix} \geq \begin{Bmatrix} f_p^1 - f_p^2 \\ f_p^2 \end{Bmatrix} \tag{30}$$

$$u_c^1 = (K_p^1 + K_p^2)^{-1} \left[K_p^2 g + f_p^1 - f_p^2 \right] \tag{31}$$

$$\left[K_p^2 - K_p^2 (K_p^1 + K_p^2)^{-1} K_p^2 \right] g + K_p^2 (K_p^1 + K_p^2)^{-1} \left[f_p^2 - f_p^1 \right] - f_p^2 \geq 0\,. \tag{32}$$

If we denote *K* and *f* as follows;

$$\begin{aligned} K &= \left[K_p^2 - K_p^2 (K_p^1 + K_p^2)^{-1} K_p^2 \right] \\ f &= K_p^2 (K_p^1 + K_p^2)^{-1} \left[f_p^2 - f_p^1 \right] - f_p^2 \end{aligned}, \tag{33}$$

equation (32) becomes to the following form:

$$Kg + f \geq 0\,, \tag{34}$$

which can be solved by using Lemke’s algorithm as a LCE problem.

5. RESULTS OF THE BEAM CONTACT PROBLEM

The beam contact problem modeled using the MLPG Finite Volume Method as in equation (34) and solved using Lemke's algorithm as a LCE problem. Distribution of the displacement along the beam and the contact force is presented in Figure (2) and Figure (3) respectively.

As it can be seen from Figure (1), the line of contact of the surface is represented by L_c. Figure (2) shows the displacement of the beam and it is observed that the displacement distribution along the beam is compatible with the line of contact along the beam surface. Figure (3) shows the distribution of the contact force along the beam.

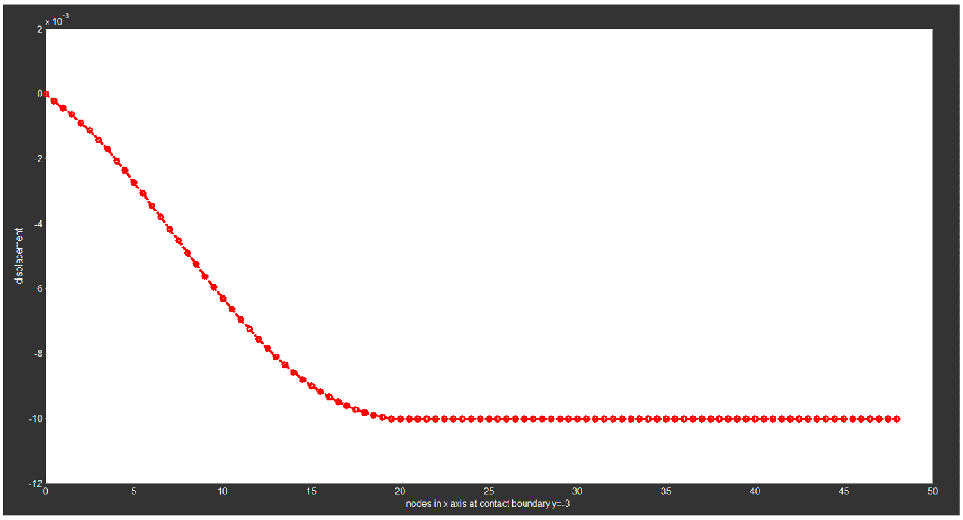


Figure 2: Distribution of the displacement along the beam.

In reference [9] it has been explained that, the location of contact (transition point) by thin beam theory is *lc*=24.97 and by thick beam theory is *lc*=21.37. Both MQ and TPS are studied in reference [9] and the calculated results with MQ shows closer to the thick beam result. The simulated contact force along the interface by using TPS with $\eta = 4$ gives a better result. In this study, it has been shown that the contact force along the interface is in good agreement with the thick beam theory. This result is found by using mixed MLPG FVM and using the LCP solution of the beam problem.

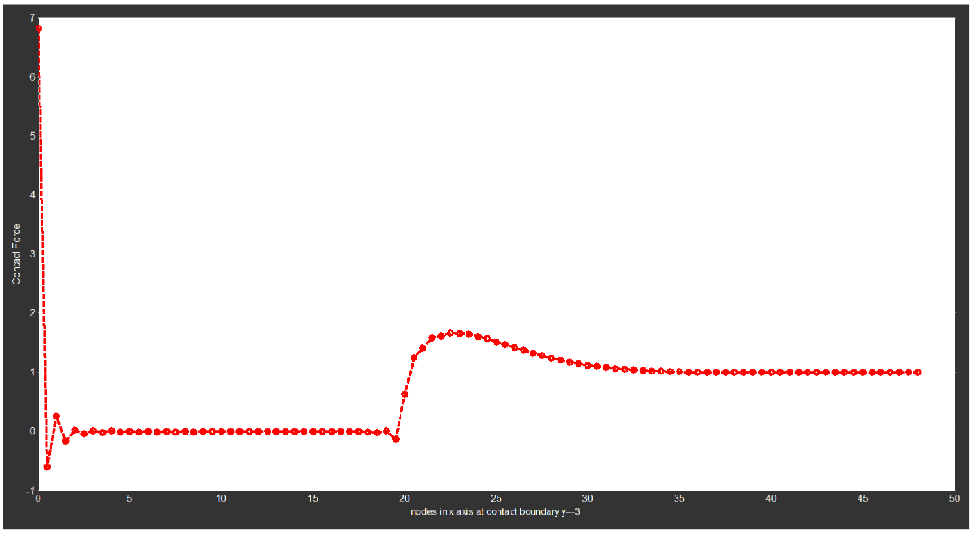


Figure 3: Contact force along the interface of a cantilever beam using MLS.

6. CONCLUSION

In this paper first of all the meshless approximations using MLS is mentioned. Then the general theory of the "mixed" MLPG Finite Volume Method is used to model the balance of linear and angular momentum. Contact problem of a cantilever beam on rigid foundation is modeled using the MLPG-FVM. For each body a contact equation is written and the inequality is solved by using Lemke's algorithm as a LCE problem.

Distribution of the displacement along the beam and the contact force in agreement with the theory given in literature. The earlier studies in the literature reports that, the location of contact (transition point) by thin beam theory is lc=24.97 and by thick beam theory is lc=21.37. Both MQ and TPS interpolating schemes are studied before by other researchers shows closer results to the thick beam theory. In this study, it has been shown that the contact force along the interface is in good agreement with the thick beam theory. This means that the proposed "mixed" MLPG FVM using LCP with MLS approximation gives good results and could be used to solve contact problems.

REFERENCES

[1] Atluri S.N.: *The Meshless Local Petrov Galerkin (MLPG) Method for Domain & Boundary Discretizations*, Tech Science Press, (2004), 680 pages.

[2] Li Q., Shen S., Han Z.D., Atluri S.N.: Application of Meshless Local Petrov-Galerkin (MLPG) to Problems with Singularities, and Material Discontinuities, in 3-D Elasticity, CMES: *Computer Modeling in Engineering & Sciences*, vol. 4 no. 5, (2003), pp. 567-581.

[3] Han Z.D., Atluri S.N.: On Simple Formulations of Weakly-Singular Traction & Displacement BIE, and Their Solutions through Petrov-Galerkin Approaches, CMES: *Computer Modeling in Engineering & Sciences*, vol. 4 no. 1, (2003) pp. 5-20.

[4] Han Z.D., Atluri S.N.: Truly Meshless Local Petrov-Galerkin (MLPG) Solutions of Traction & Displacement BIEs, CMES: *Computer Modeling in Engineering & Sciences*, vol. 4 no. 6, (2003), pp. 665-678.

[5] Han Z.D., Atluri S.N.: Meshless Local Petrov-Galerkin (MLPG) approaches for solving 3D Problems in elasto-statics, CMES: *Computer Modeling in Engineering & Sciences*, vol. 6 no. 2, (2004), pp. 169-188.

[6] Han Z.D., Atluri S.N.: A Meshless Local Petrov-Galerkin (MLPG) Approach for 3-Dimensional Elasto-dynamics, CMC: *Computers, Materials & Continua*, vol. 1 no. 2, (2004), pp. 129-140.

[7] Atluri S.N., Kim H.G., Cho J.Y.: A Critical Assessment of the Truly Meshless Local Petrov Galerkin (MLPG) and Local Boundary Integral Equation (LBIE) Methods, *Computational Mechanics*, 24:(5), (1999), pp. 348-372.

[8] Atluri S.N., Shen, S.: The meshless local Petrov-Galerkin (MLPG) method: A simple & less costly alternative to the finite element and boundary element methods. CMES: *Computer Modeling in Engineering & Sciences*, vol. 3, no. 1, (2002), pp. 11-52.

[9] Xiao, J.R., Gama, B.A., Gillespie, Jr J.W., Kansa, E.J.: Meshless solutions of 2D contact problems by subdomain variational inequality and MLPG method with radial basis functions, *Engineering Analysis with Boundary Elements,* 29, (2005), 95–106.

[10] Fasshauer G. E., Approximate moving least-squares approximation: A fast and accurate multivariate approximation method, *in Curve and Surface Fitting*: SaintMalo 2002, A. Cohen, J.-L. Merrien, and L. L. Schumaker (eds.), Nashboro Press, (2003), 139–148.

[11] Atluri S.N., Han Z.D., Rajendran A.M., A new implementation of the meshless finite volume method, through the MLPG "mixed" approach, CMES: *Computer Modeling in Engineering & Sciences,* vol.6, no.6, (2004), pp.491-513.

[12] Kikuchi N, Oden JT. Contact problems in elasticity: a study of variational inequalities and finite element methods. Philadelphia: SIAM; 1988.

JOURNAL OF NAVAL SCIENCE AND ENGINEERING
PUBLISHING RULES

Submission of Papers: Submission of a paper implies that it is has not been published previously, it is not under consideration for publication elsewhere and that if accepted, it will not be published elsewhere in the same from, in English or in any other language, without the written consent of the publisher. Authors are requested to submit an electronic copy of their original works to given **e-mail address** below.

E-mail Address:
Lt.Commander Emrah YAMAKOĞLU
eyamakoglu@dho.edu.tr

Types of Contributions: Original papers, review articles; technical notes; short communications; book reviews; letters to the editor; extended reports of conferences and meetings.

Manuscript Preparation

General : Manuscripts must be prepared in MS Word, single-spaced with justify. Font: Times New Roman 12 points. Margins: left 4.5cm., right 3.5 cm, top 5.5 cm and bottom 7 cm. header 3.25 cm. footer 6 cm. gutter 0. Paper type: A4. Numbers should be on the middle of bottom of page with 1,2,3 etc format.

The information about authors should be prepared separately from main text and should include;

Title
Name of authors
Authors' associations
postal address
e-mail of the authors

Body of Text: Follow this order when typing manuscripts: Title, Authors, Abstract, Keywords, Main text, Appendix (if any), References, Authors' short bios.

Title: Title should reflect objectives of the paper clearly and should be easily understandable, should not exceed 12 words.

Abstracts: Each paper should have an abstract with 100-200 words, reporting concisely on the purpose and results of the paper.

Paper Length: The size of the manuscript should not exceed 10 printed pages.

Keywords: Author must provide some keywords (between 3 and 10) that will be used to classify the paper.

Units: The SI system should be used for all scientific and laboratory data.

References: Referring to references should be in the text by using reference numbers in parenthesis (for example: [2]). Using footnotes for references is not allowed.

List of References: Publications cited in the text should only be presented in a list of references following the text of the manuscript. References should be listed according to their orders in which they are cited in the manuscript, i.e., [1] cited first, [2] cited next, and so on.
Examples:

[1] Hollstein T., Halgamuge S.K., and Glesner M., (1996). *Computer-Aided Design of Fuzzy Systems Based on Generic VHDL Specifications*, IEEE Trans. on Fuzzy Systems, vol. 4 (4), pp.403-417. **(journal paper)**

[2] Johnson N., and Jajodia S., *Steganalysis: The Investigation of Hidden Information*, Proc. IEEE Information Technology Conference, Syracuse, NY, USA, Sep. 1998, pp. 113-116. **(conference proc.)**

[3] Okamoto G., (1999). Smart Antenna Systems and Wireless LANs, Kluwer Academic Publishers, Boston. **(books or thesis)**

Abbreviations and Acronyms: Standard abbreviations and acronyms should be used for each related discipline. Acronyms should be identified at the first occurrence in the text. Abbreviations and acronyms may also be attached to main text as an appendix.

Equations and Formulas: Equations and formulas should be numbered consecutively. These numbers must be shown within parenthesis which are aligned to right.

Figures and Tables: Figures and tables should be numbered consecutively. In the text refer to figures and tables should be made by typing "Figure 1" or "Table 1" etc. A suitable title should be assigned to each figure and table and this title must be separately listed with their reference numbers.

Copyright: Authors are responsible for obtaining permission from the copyright owner (usually the publisher) to use any quotations, illustrations, or tables from another source. The Journal of Naval Sciences and Engineering does not accept any responsibility if authors did not follow Copyright Law requirements. Authors submitting articles to the Journal agree that if accepted, copyright of the article, including the right of reproduction in all forms and media, will belong exclusively to the Journal.

Evaluation Process: All manuscripts are reviewed and forwarded to peer review committee if manuscripts match with journal's minimum publishing rules. At least two members of peer review committee evaluate submitted works and send their reports to editorial board. If editorial board receives any recommendation on evaluated papers from peer review committee, it may require necessary revision from author.

DENİZ BİLİMLERİ VE MÜHENDİSLİĞİ DERGİSİ
YAYIN KURALLARI

Yazıların Gönderilmesi: Dergiye gönderilen makaleler başka bir yerde yayımlanmamış ya da yayımlanmak üzere gönderilmemiş olmalıdır. Yayımlanması istenilen yazılar aşağıdaki adrese 3 kopya kağıda basılı olarak, aynı zamanda e-mail adresine de dijital olarak gönderilmelidir. İngilizce metninin yayınlanmasına karar verilen makalelerden yazar tarafından Türkçe kopyası da gönderilen metinler, ingilizce fasikülden ayrı olarak basılacak olan Türkçe fasiküle de konulacaktır. Dergide yer alması amacıyla gönderilecek makalelerin, Nisan sayısı için 05 Ocak, Kasım sayısı için de en geç 05 Eylül tarihine kadar Deniz Bilimleri ve Mühendisliği Enstitüsüne ulaştırılması gerekmektedir.

Adres:
Öğ.Bnb.Emrah YAMAKOĞLU
Deniz Harp Okulu Komutanlığı, DEBİM
34942 Tuzla/ İSTANBUL

E-mail adresi : eyamakoglu@dho.edu.tr

Yazı Türleri : Dergide; orijinal yazılar, derlemeler, kitap incelemeleri, teknik notlar ve editöre mektuplar yayımlanır.

Yazım Kuralları

Genel bilgiler : Yazılar; Microsoft Word'de tek satır aralığı, Times New Roman, 12 punto; marjlar: alt 7 üst 5.5 cm; sağ 3,5 sol 4.5 cm; kağıt ölçüsü: A4 formatında hazırlanmalı ve sayfaları alt ortada 1,2,3 şeklinde numaralandırılmalıdır. Dergiye gönderilecek yazıların yazarlarına ait kimlik ve kurum bilgileri, yazıdan kolaylıkla ayrılabilir şekilde ayrı bir sayfaya hazırlanmalı ve şu bilgileri içermelidir;

Yazının başlığı
Yazar(lar)ı
Yazar(lar)ın bağlı bulundukları kuruluşlar
Yazar(lar)ın kısa biyografisi
Birinci Yazarın fax numarası, posta adresi ve e-mail adresi

Önerilen ayrıntılı yazı formatı www.dho.edu.tr. adresinde mevcuttur.

Yazı Yapısı : Yazı sırası ile aşağıdaki bölümlerden oluşur. Başlık,yazarlar, özet, anahtar kelimeler, ana yazı, ek (varsa), referanslar.

Başlık : Başlık; açık, net, anlaşılır olmalı ve 12 kelimeyi geçmemelidir.

Özet (Abstract) : Yazı, 100-200 kelimelik, yazının amacını ve sonuçlarını anlatan özeti içermelidir.

Sayfa Sayısı : Yazılar 20 sayfayı geçmemelidir.

Anahtar Kelimeler : Yazıyı sınıflandırmaya yarayacak, anahtar görevi yapan 3-10 kelime yer almalıdır.

Birimler : Yazının uluslararası alanlarda da kolay izlenebilir ve anlaşılabilir olması için Uluslararası Bilimsel Notasyona ve SI (System International) birim sistemine uygun olarak hazırlanması gerekir.
Referans : Referanslar dipnotlar ile değil metnin içinde açılacak parantezlerde yer alan referans numarası ile ([2] gibi) yapılmalıdır.
Referans Listesi : Makalenin sonunda sadece yazıda gönderme yapılan eserler yer almalı, düzenleme şu sıra ile olmalıdır: Yazarın soyadı ve adı. Yayın yılı. Yazının adı. Yayınlandığı yerin (kitap/dergi/seminer) adı. Yayınevi, basım (kitap/dergi)/seminer yeri. Sayfa numaraları.
Örnek :
Yıldırım, Cemal (1995). Bilimin Öncüleri. Tübitak Popüler Bilim Kitapları Dizisi:9. Ümit Matbacılık Ankara. Sayfa: 12-14.

Notasyon ve Kısaltmalar : İlgili bilim alanının standart notasyon ve kısaltmaları kullanılmalı, yeni notasyonlar ise metin içinde ilk geçtiği yerde tanımlanmalıdır. Gerekli durumlarda, kullanılan notasyon ve kısaltmalar bir ek olarak konulabilir.
Denklem ve Formüller : Denklem ve formüller ardışık olarak numaralandırılmalı ve bu numaralar sağa dayalı parantez içinde yazılmalıdır. Metin içinde denklem ve formüllere parantez içinde yazılan numaraları ile atıfta bulunulmalıdır. Metin arasında verilmesi uygun olmayan kapsamlı formüller şekil olarak hazırlanmalıdır.
Şekiller ve tablolar : Şekiller ve tablolar, ardışık olarak numaralandırılmalıdır. Bunlara metin içinde "Şekil 1" veya "Tablo 1" şeklinde atıfta bulunulmalıdır. Her bir şekil ve tablo için uygun bir başlık kullanılmalı ve bunlar ayrı birer ek olarak şekil ve tablo numaralarıyla birlikte listelenmelidir.
Telif Hakları : Yazılarda 5846 sayılı Fikir ve Sanat Eserleri Yasasının hükümlerine aykırı iktibas (alıntı) bulunması halinde tüm sorumluluk yazar(lar)a ait olacaktır. Yayımlanmasına karar verilen yazıların telif hakkı Deniz Bilimleri ve Mühendisliği Dergisine aittir.

Yazıların Değerlendirilme Süreci : Dergi Yayın Kurulu yayımlanmak için gelen yazıların yayımlanıp yayımlanmamasına ve dergide hangi statüde yayımlanacağına karar verir. Yayın Kurulu gerekli gördüğü takdirde yazının sahibinden değişiklik isteme hakkına sahiptir.

Gönderilen yazıların basımı öncesinde hakem görüşü alınır. Yayın kuruluna gönderilen yazıların dergide yayımlanabilmesi için yayın kurulunca, belirlenen yayım kurallarına uygun görülmesi ve Hakem Heyeti tarafından kabul edilmesi gerekir. Yazılar hakem kurulunda yer alan en az iki hakem tarafından değerlendirilir. Yazı hakemlerin değerlendirme sürecinden geçtikten sonra, eğer varsa, düzeltmeleri yapması için yazara geri gönderilir. Düzeltmeler yazarlarca kendilerine verilen süre içinde tamamlanmalıdır.