IJICTR
International Journal of Information & Communication Technology Research

Volume 14- Number 2 – 2022 (66 -54)

Implementing Process Mining Techniques to Analyze Performance in EPC Companies

Seyedeh Motahareh Hosseini
Ph.D. Candidate, Systems and Productivity Management Department,
Faculty of Industrial and Systems Engineering,
Tarbiat Modares University
Tehran, Iran
mo.hosseini@modares.ac.ir

Mohammad Aghdasi
Professor, Systems and Productivity Management Department
Faculty of Industrial and Systems Engineering,
Tarbiat Modares University
Tehran, Iran
aghdasim@modares.ac.ir

Babak Teimourpour
Assistant Professor, Systems and Productivity Management Department
Faculty of Industrial and Systems Engineering,
Tarbiat Modares University
Tehran, Iran
teimourpour@modares.ac.ir

Amir Albadvi
Professor, Systems and Productivity Management Department
Faculty of Industrial and Systems Engineering,
Tarbiat Modares University
Tehran, Iran
albadvi@modares.ac.ir

Received: 19 April 2022 – Revised: 21 May 2022 - Accepted: 27 June 2022

Abstract—**The importance of process analysis in engineering, procurement and construction companies (EPC), due to the complexity of the measures, the high level of communication between people, different and non-integrated information systems, as well as the amount of capital involved in these projects is much higher and more challenging. Limited research has been done on exploring business processes in these companies. In this study, in order to better and more accurately analyze the company's performance, three perspectives of process mining (process flow, case and organizational) is analyzed by using the event logs recorded in the supplier selection process. The results of this study led to the identification of challenges in the process, including repetitive loops, duplicate activities, survey of factors affecting the implementation of the process and also examining the relationships between people involved in the project, which can be used to improve the future performance of the company.**

Keywords: Process Mining; Process Mining perspectives; Complex Construction Project Management; EPC Companies.

Article type: Research Article

Publisher: ICT Research Institute

[DOI: 10.52547/itrc.14.2.66]

I. Introduction

The Information systems (ISs) are important from two perspectives, one in digitalization and the other in the relationship between information and physical flows in business processes[1, 2]. With the development and use of (ISs) and increasing the ability to record and extract event logs in organizations, it is possible to analyze performance and review business processes better than before. Business processes are the core of any organization, which needs continuous monitoring and improvement. Extracting information from business processes is the ultimate goal of process mining(PM) [3, 4]. PM is used as a new approach to identify and improve business processes by combining machine learning and business process management (BPM). PM uses data-based techniques, event logs recorded in ISs to discover, monitor, identify deviations and inefficiencies of processes and compare reality with designed models[2]. This data is generated during the execution of business processes in ISs and is used to reconstruct process models[5]. These models are very effective for analyzing, optimizing and support processes against rapid change[6-8]. PM is a new approach that, like a bridge, connects the two areas of data mining and BPM [9-11].

PM extracts knowledge from the reporting of events and presents it as a process model. This technique supplies new tools for a wide range of applications, process discovery, monitoring, and improvement[12]. This technique enables fast detection of business process issues in real time as well as continuous and incremental improvement by analyzing anomalies, bottlenecks and shortcomings in real situations [13].Kadasah has pointed out that objectivity is the main advantage of the mining process because it helps organizations to produce better process models. Organizations can also find exceptions, avoid deviations, and check how well the designed model matches the model that has actually been implemented. On the other hand, due to the high speed of output production, a lot of time savings occur during process analysis [14]. Based on the researches that have been done in recent years on the applications of process mining, the areas that have used this technique the most are: Healthcare, ICT, Manufacturing, Education, Finance, Logistics and Public[15, 16]. However, there is no considerable evidence that the PM is used in EPC companies or construction projects.

Due to the fact that in EPC projects, the costs of delay or loss in project quality are significant, improvement in these projects is very important. According to statistics, many large construction projects have recorded extremely poor performance results, which has mainly led to increased costs[17].In 2015, only 64% of projects were able to achieve their goals, which includes projects of EPC companies. Large projects often face major problems, which affect the performance and quality of projects. Therefore, many companies are looking for ways to improve quality and performance, to avoid mistakes and increase success in project implementation[6, 7]. The changing environment and the presence of different individuals and departments in the implementation of processes in EPC companies has led companies to look for efficient methods to identify the differences between the actual model and the process model by the use of PM[18].

As respects that there are several actors in EPC projects, the dimensions of these projects are very wide and the implementation of these projects is associated with many challenges, the use of PM approach has received less attention. Therefore, in this paper, using event logs recorded in various ISs, including automation systems and other databases, the supplier selection process in a complex construction project in Oil and Gas company with PM techniques has been analyzed in different perspectives.

The rest of the paper is organized as follows. Section II describes theoretical background of this study in the field of PM. Section III explains methodology and Section IV experimental results, while Section V draws some conclusions.

II. Theoretical Background

This part discusses theories related to the research issue, namely PM, and PM perspectives. Each issue is discussed separately.

A. Process Mining

The best definition of PM is provided in the PM Manifesto: "The purpose of PM is to discover, monitor, and improve actual processes by extracting knowledge from event logs"[19]. PM focuses on understanding processes and capturing more important findings by the crossroads between data mining and business process management[20, 21]. Types of PM involves process discovery (focuses on how to build a process model based on recorded event logs), conformance checking (examine the degree of conformity of the model with the actual situation) and process enhancement (modify or develop the model) [2, 19, 22, 23]. In other words, PM is a discipline that causes us to have a better understanding of real-life processes and to better understand the behavior of systems[24].

Due to the great interest in the field of PM in the last decade, PM software tools have been developed and are constantly evolving, such as ProM (Process Mining Framework), (Celonis Gmbh), Disco (Fluxicon), EDS (StereoLOGIC Ltd), Fujitsu (Fujitsu Ltd), etc. Each of these tools has advantages and disadvantages compared to each other.

B. Elements of PM

An event is the most atomic part of a specific process about historical executions that can typically be found in ISs[23]. A case or instance of a process consists of a sequence of events that occur in a single execution of a process. Items in a report are uniquely identified by a specific code. An activity is an operation or part of an item that has been executed[25]. Event reports contain information about each activity performed, its performer, role performer, timestamp and other related data[4].

Data quality in terms of form and content plays a vital role in the successful implementation of the mining process. Clearing event logs before performing process analysis is a time-consuming, tedious, and of course and very necessary step[25]. Due to the fact that

[DOI: 10.52547/itrc.14.2.66]

data may be stored in different sources of information, defects, out-of-bounds, disruption and even the presence of different levels of detail in the drawings are common problems in data analysis and measures such as extraction and filtering. The data will have a significant impact on better and deeper analysis[16].

C. PM perspectives

PM covers three main perspectives. PM in the perspective of flow control (process) focuses on the sequence of activities with the aim of finding features for all possible paths [2, 19, 26].

The organizational perspective focuses on information about process implementers such as people, systems, maps, or departments and their relationship to each other, which is actually related to the field of social network analysis(SNA)[4]. For better analysis in this perspective, software specific such as Gephi, Pajek can be used. Each social network is represented by a graph in which the nodes represent individuals and each connection between the two nodes indicates the relationship between the two nodes. In social network analysis, various criteria are used, one of the most common of which is the criteria of centrality, and indicates which node has a key position in the whole network. The main criteria of centrality are degree centrality, intermediate centrality and proximity centrality[27].

The case perspective analyzes case properties as the process execution path, executors, and properties of the data elements[19].

In some sources is also mentioned, from the perspective of time, as the fourth perspective[4]. The time perspective is related to the timing and sequence of events. By examining the times recorded in the event logs, bottlenecks can be identified and the processing time of running processes can be predicted[19]. In this study, due to the special nature of time, it has been used as a measure in other perspectives.

III. Methodology

In order to implement the principles of PM, the crisp-DM framework has been used[28]. Based on the steps mentioned in this framework, in the first step, by conducting an interview and reviewing the documents, the company's business is reviewed. In the second step, the data in different databases are collected and analyzed. And then the data preparation and cleaning step is done. In data science related projects, this step is one of the most time-consuming steps. Then, based on the collected data, a process model is created and examined from three perspectives of PM. Finally, the results are shared with process experts and decisions are made to improve the implementation of the process.

This study identifies, extracts, clears, and analyzes data in three process mining perspectives based on information and correspondence related to the procurement process in a complex construction project in an EPC company over several years. In order to study the supply process, the data registered in various databases including automation system, web-based system, documents stored in the network related to the refinery construction project have been used. Each purchase title is considered as a case and the actions taken in each transaction are considered as an activity.

The process under review is a summary of various steps such as submitting a purchase request, reviewing the request by different departments, preparing and approving relevant technical documents, pre-tender steps, reviewing supplier documents, holding tenders, announcing results, and concluding a contract. To carry out this process, different departments such as project site, supply deputy, engineering deputy, plan and program deputy, consultants, Employer and vendors interact with each other.

The collected data are related to the activities of the Employer, three consulting companies, and 270 vendors as out-company actors and 12 managers and 63 employees as in-company actors. This data is related to the stages of announcing the need for equipment by the project manager until it is delivered and approved by the technical inspection department, related to 131 purchases and over 55,000 event logs. The recorded time period includes event logs from 2015 to 2020 and the number of activities surveyed is 129 titles. It should be noted that a number of reports have been deleted due to distorted or incomplete information recorded during the cleanup phase. After achieving the process model, the results are examined in three perspectives of PM. Fig. 1 provides an overview of the research steps.

Purchases are classified into 8 disciplines (mechanical, electrical, (Heating, Ventilating and Air Conditioning) HVAC, process, safety, piping, instrument and fix equipment) and the performance of each of these disciplines is reviewed and compared separately. Event logs used include purchase number, activity title, date, sender, recipient and purchase group. An example of the extracted information is shown in Table 1. Titles of activities include the titles of technical documents and activities performed.

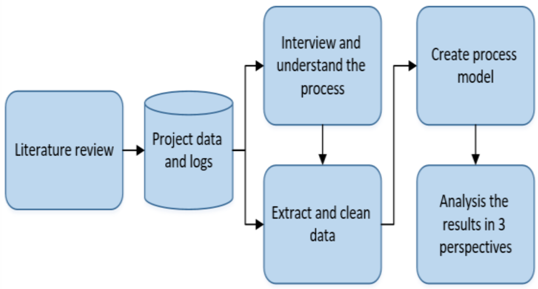


Fig. 1. Methodology of the study.

TABLE I. Sample information extracted

ID	Activity	Time stamp	Send er	Receive r	disciplin e
C1	register is	2017/0 7/08	man5	consulta nt 1	electrical
C 17	diagram	2019/0 9/01	man5	consulta nt 3	instrume nt
C 89	sub vendor list	2018/0 2/25	clerk 3	consulta nt 1	mechanic al

[DOI: 10.52547/itrc.14.2.66]

C 127	VPIS (Vendor Print Index and Schedule)	2018/02/19	consultant 1	man5	process

One of the main software programs used in this research is Disco software version 2.5.0 which is available for free under educational / university license. This software has been used to build the process model. Disco generates a diagram, where nodes represent the resources and the arrows represent a sequence of activities. This sequencing of activities also describes the direction of the interaction between the two resources[4].

The process map of this data is plotted by Disco software in Fig.2. Due to the great complexity of the actual process situation, the details in the process model resulting from the event logs are somewhat invisible. A small part of the process model is magnified. This figure emphasizes the importance of examining the true state of business processes using PM method. Because of great complexity of the actual process map, it is necessary to perform the analysis through simplification or filtering. In Fig.2, activities are recorded in rectangles, and the lines between them represent the flow of the process from one activity to another. Rectangles marked in dark colors indicate activities that have more repetitions than other activities.

An overview of activities in different purchasing disciplines is shown in Table 2. According to the information in Table 2, the ratio of the number of events to the number of purchases is the highest in the groups of mechanical equipment, process and fixed equipment, respectively compared to other groups, and safety equipment is the lowest. In terms of time, the disciplines of mechanical and piping equipment have the highest average time and the discipline of process has the lowest average time.

The other main software used is Gephi software version 0.9.2 for SNA.

IV. PROCESS PERFORMANCE ANALYSIS

In this research, the recorded event logs over several years based on the existing classification in the literature is examined in three perspectives. The dimensions studied in each perspective are determined based on the volume and number of available event logs. In this section, the analysis related to the collection of event logs and in different purchasing disciplines, which are described in section 3, are presented for a more in-depth study. In order to better analyze the company's performance, event logs based on different purchasing disciplines have been reviewed.

A.Process Perspectives

In this perspective, by using recorded events, causal relationships of activities at the time of execution and process modeling tools, the process model is illustrated. In the general review of event logs related to the company under study, by applying the parameters "number of repetitions" and "case repetition" in the process map, the activities and routes that had the highest number of repetitions were identified and presented in Tables 3 and 4.

[DOI: 10.52547/itrc.14.2.66]

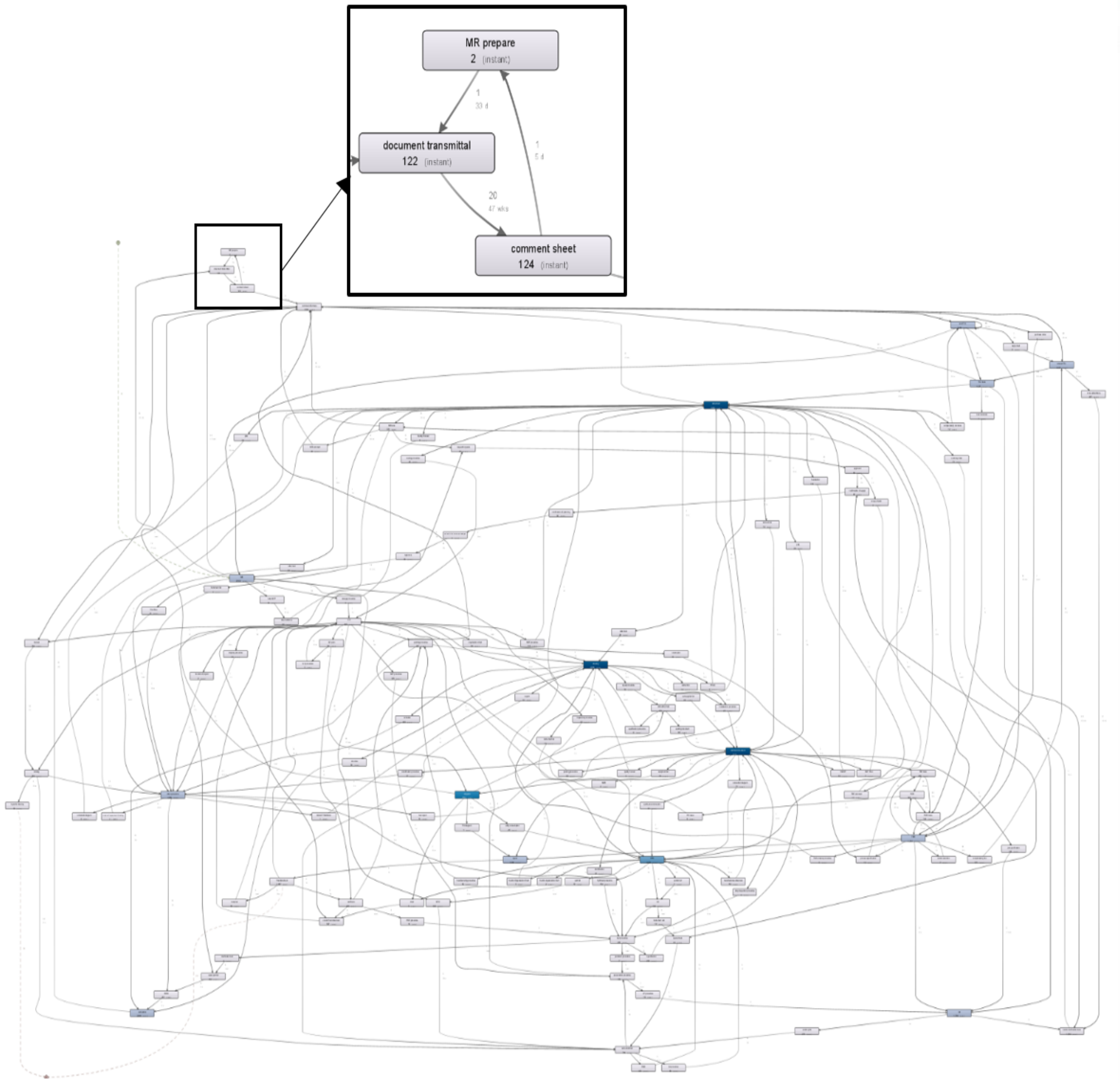


Fig. 2. The actual status of the purchasing process

TABLE II. GENERAL INFORMATION ABOUT PURCHASING DISCIPLINES

Discipline	Cases	Activity	Events	*Events / Cases*	Start	End	Median time (months)	Mean time (months)
electrical	29	77	10127	350	2015/11/07	2020/05/03	40.1	36.5
fix equipment	7	66	3427	490	2015/11/07	2020/05/05	40.5	35.7
HVAC	15	53	5436	363	2015/11/07	2020/05/04	35.8	32.1
instrument	26	89	10930	420	2015/11/07	2020/05/04	42.9	40.2
mechanical	14	100	11638	831	2015/11/07	2020/05/05	48	46.6
piping	24	71	8164	340	2015/11/03	2020/05/05	44.1	43.5
process	7	73	3892	556	2015/11/07	2020/05/02	15.1	24.6
safety	7	74	1271	182	2015/11/07	2020/05/03	38.7	35.6

TABLE III. THE MOST REPETITIVE ACTIVITIES

Activity	Number of repetitions	Case frequency	Maximum repetition
drawing	4507	67	956
Technical doc	5412	78	698
diagram	3660	43	556
layout	1341	24	393
data sheet	4938	83	311
calculation	1880	54	254
list	1660	46	251

TABLE IV. PROCESS PATHS WITH THE MOST REPETITION

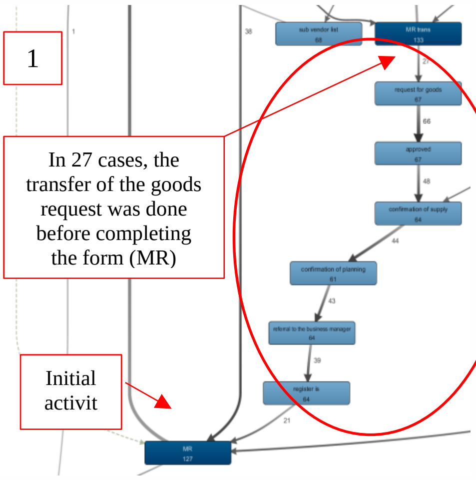
Process paths	Number of repetitions	Case frequency	Maximum repetition
(Material Requisition)MR-MR	1002	118	33
Answer (Technical Clarification Letter)TCL-TCL trans	458	101	16

[DOI: 10.52547/itrc.14.2.66]

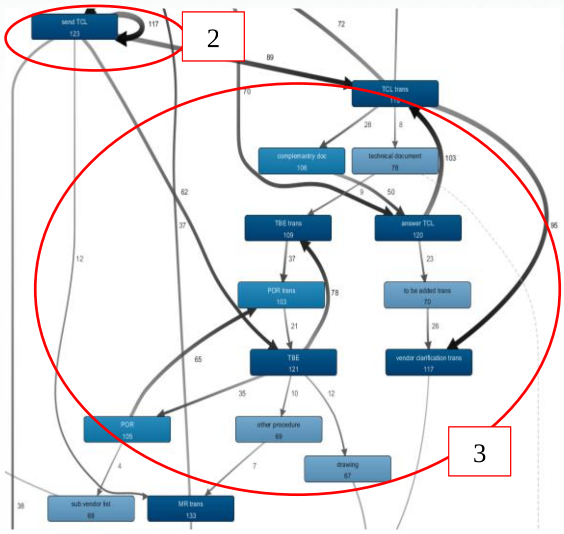
Technical offer trans-send TCL	452	93	13
TCL trans- vendor clarification trans	213	93	7
Send TCL-TCL trans	192	88	7

From the items presented in Tables 3 and 4, important results such as the identification of bottlenecks and routes with high repetition can be achieved. Like the goods request route - the goods request, which due to the high number of repetitions and the return loop, there is a need to analyze the cause and correct the performance.

By applying the filter of repetitive activities and removing other activities from the process map, the following results can be achieved: 1- In most cases, the purchasing process begins with the activity of requesting goods. According to the purchase process document, the initial activity must be to complete the purchase request form. 2- Transferring Technical clarification letter (TCL) activity - Transferring TCL is an activity that has been seen in 117 purchases and has been repeated 670 times. The maximum number of repetitions in one purchase was 26! 3- In the purchase process document, the order of goods requesting activities, technical evaluation and purchase order request are defined as 1- product request, 2- technical evaluation and 3- purchase order request, respectively, and also it is to be expected in sending their documents, respectively. While in the resulting process map, this order is not observed in all cases and the design and implementation of the process needs to be reviewed. These are shown in Fig. 3(a and b).



(a)



(b)

Fig. 3. (a): Deviations in the process and (b): Deviations in the process

In Table 5, the most frequent activities are compared based on the criteria of maximum repetition and frequency of repetition in different purchasing disciplines. As can be seen in Table 5, many of the most repetitive activities are common to most purchasing disciplines, and this may indicate that activities that are difficult to perform, and more duplicates are usually common and include topics such as technical documents, information documents, technical drawings, and so on.

Table 6 shows the parts of the process that have the longest execution times in the eight purchasing disciplines and are compared with each other. As can be seen in Table 6, the demand section and the activities that follow are common to most purchasing disciplines and are one of the most time-consuming routes. The duplicate loops observed in different purchasing disciplines in terms of process flow are shown in Table 7 for better comparison. As can be seen in Table 7, many of the loops created are common to most of the disciplines. This indicates that the challenging parts of the process are the same in most purchasing disciplines.

B. Case Perspectives

The goal in this perspective is to focus on process inputs such as sample types, units and individuals, different sources and different processing times, and the effectiveness of each of these variables. One of the key variables in analyzing event logs in this process, performance analysis, is based on different purchasing

TABLE V. REPETITIVE ACTIVITIES BASED ON THE CRITERION OF MAXIMUM REPETITION AND FREQUENCY OF REPETITION IN DIFFERENT PURCHASING DISCIPLINES

discipline	criterion	activity	number	activity	number	activity	number	activity	number
mechanical	maximum repetition	Technical document	535	Drawing	403	Data sheet	302	Diagram	189
	frequency of repetition	Technical document	1807	Drawing	1731	Data sheet	1330	List	574
piping	maximum repetition	Technical document	247	Data sheet	152	VPIS	115	Final data book	108

[DOI: 10.52547/itrc.14.2.66]

	frequency of repetition	Data Sheet	752	VPIS	716	Other procedure	615	MR	600
instrument	maximum repetition	Data sheet	374	Technical document	335	List	270	Diagram	192
	frequency of repetition	Data sheet	1586	Technical document	683	VPIS	677	Send TCL	629
electrical	maximum repetition	Diagram	557	Layout	393	Technical document	298	MTO (Make To Order)	144
	frequency of repetition	Diagram	2014	Layout	779	Technical document	776	Send TCL	581
fix equipment	maximum repetition	Drawing	956	Calculation	222	Data sheet	156	Diagram	72
	frequency of repetition	Drawing	1181	Data sheet	381	Calculation	261	VPIS	142
HVAC	maximum repetition	Technical document	724	Data sheet	298	Diagram	241	Spare part list	127
	frequency of repetition	Technical document	1461	Data sheet	623	Diagram	421	Schedule	255
process	maximum repetition	Drawing	463	Technical document	340	Calculation	254	Data sheet	149
	frequency of repetition	Drawing	659	Technical document	503	Calculation	389	Diagram	271
safety	maximum repetition	Data sheet	87	Technical document	68	Send TCL	35	Final data book	28
	frequency of repetition	Send TCL	180	Data sheet	139	Technical offer trans	110	Technical document	94

TABLE VI. THE PATHS OF THE PROCESS THAT HAVE THE MOST TIME IN DIFFERENT PURCHASING DISCIPLINES

Discipline	Path 1	Path 2	Path 3	Path 4	Path 5
mechanical	Drawing- technical document	TBE(Technical Bid Evaluation)-ITP(Inspection and Test Plan)			
piping	*MR-MR	*Register IS-MR trans	*MR-ITP	*POR(Purchase Order Request) trans-MR	*MR-technical offer trans
instrument	TBE- sub vendor list	POR trans-TBE	*MR-MR	TCL trans-data sheet	
electrical	*Send TCL-MR	*Drawing-MR	POR trans-ITP	*MR trans-request for goods	*MR –MR trans
fix equipment	*Data sheet-MR	ITP-data sheet	*MR-MR trans	*MR trans-MR	TBE-TBE trans
HVAC	*Technical document-MR	*MR trans-MR	*MR-MR	*MR-MR trans	
process	Register IS-technical offer trans				
safety	*MR trans- technical offer trans	*Referral to business manager-MR	Sub vendor list-technical document	*MR-MR trans	
legend	Priority 1: More than 40 weeks	Priority 2: 30-40 weeks	Priority 3: 15-30 weeks	Priority 4: 10-15 weeks	Priority 5: Less than 10 weeks
Paths marked with * are paths that are repeated in more than one purchasing discipline.					

TABLE VII. REPETITION LOOPS IN PROCESS EXECUTION IN DIFFERENT PURCHASING DISCIPLINES

Discipline	Loop1	Loop2	Loop3	Loop4	Loop5	Loop6	Loop7
mechanical	Drawing-Technical document	Data sheet-Drawing	Answer TCL -To be added trans-Vendor clarificati on trans	Answer TCL - To be added trans-Technical offer trans	Answer TCL - Send TCL - TCL trans-Technical offer trans-Answer TCL	MR-MR trans	POR-POR trans
piping	MR-Request for goods-Approved	TCL trans- Vendor clarification trans-Complementary doc					

[DOI: 10.52547/itrc.14.2.66]

instrument	MR-Technical offer trans-Send TCL	Technical offer trans-Send TCL-Answer TCL-TCL trans	TBE-TBE trans-POR-POR trans				
electrical	Send TCL - Technical document-To be added trans	Technical offer trans-Send TCL-Data sheet	Send TCL-TBE-complementary doc	POR trans-POR	Technical offer trans-Send TCL-Answer TCL-TCL trans-Vendor clarification trans	Technical offer trans-Send TCL-TBE-TBE trans-POR trans-ITP	
fix equipment	Answer TCL-complementary doc-To be added trans-Vendor clarification trans	Answer TCL-complementary doc	MR-MR trans	Data sheet-Sub vendor list-ITP	TBE-TBE trans-POR-POR trans		
HVAC	MR trans-MR	Technical offer trans-Send TCL - TCL trans	TCL trans-Vendor clarification trans-complementary doc-Answer TCL	TCL trans-Vendor clarification trans -Answer TCL	Technical offer trans-Send TCL - TCL trans-To be added trans	TBE- POR-POR trans	Drawing-ITP-Data sheet
process	MR trans-MR	Answer TCL- TBE-TBE trans-Complementary doc					
legend	100-125 repetitions	75-99 repetitions	50-74 repetitions	25-49 repetitions	1-24 repetitions		

disciplines. Part of this analysis is reviewed in Section 4.1 in terms of process flow, and the other part in Section 3-4 in organizational perspective.

Another influential factor in the analysis is the number of events in the purchase case, which is very variable and different. Purchases are divided into four groups based on the number of event logs and the performance of each of these groups is analyzed, which results as follows:

1- Purchases with less than 50 events, including 6% of purchases and about 1% of events: This filter includes 8 purchases and 25 activities that are done in the period between 140 days to 2 years and 51 days. These low-event purchases are purchases that have completed the steps from the start of the procurement process to the completion of the technical evaluation. The point to consider these items is that the purchase time of these items is longer than expected. The irregularities mentioned in the previous analyses and the repetition of an activity, especially the request for goods, can also be seen. Another point is that most of these purchases were not made at the end of the 2019 period, and it is possible that these purchases were stopped for some reasons or information related to them is not available.

2- Purchases with the number of events between 51-51, including 38% of purchases and 13% of events: This filter includes 51 purchases and 70 activities and the time of these activities varies between 298 days to more than 4 years. The middle standard for doing these things is buying for more than 30 months. Of the 51 purchases mentioned, only 12 have completed the steps of requesting the goods, technical evaluation and requesting the purchase order and have entered other parts of the process. In these cases, the purchase of more details of events related to the preparation of other technical documents required to conclude the contract and conduct technical inspections can also be seen.

3- Purchases with the number of events between 500-500, including 36% of purchases and 33% of events: This filter includes 48 purchases and 83 activities and the time of these activities is changeable between 2 years and 45 days to 4 years and 9 days. The average criterion for doing these things is buying for more than 42 months. Out of 48 purchases, 18 are still involved in the process of requesting goods, technical evaluation and requesting purchase orders, but 30 have passed these steps. Of course, considering that the total number of activities is 124, it can be seen that these steps have been done by repeating some activities.

4- Purchases with the number of events between 1999-501, including 18% of purchases and 52% of events: This filter includes 25 purchases and 100 activities and the time of these activities is between 2 years and 152 days to 4 years and 8 days. The middle parameter of doing these purchases is more than 44 months. Out of 25 purchases, 18 are in the final stages of the purchase process.

The high dispersion between the numbers of events in different cases of purchasing in the period from 12 to 1972 can be due to various reasons, such as incomplete information related to some cases or excessive and

[DOI: 10.52547/itrc.14.2.66]

repetitive repetition of some activities. From the general results of analyzing the number of events and the progress of the process in these four categories, it can be concluded that in fact, the repetition of activities is an integral part of the data related to this project.

The performance of project managers is another parameter affecting the performance of the process. The tenure of project managers is recorded in Table 8.

TABLE VIII. REVIEW THE PERFORMANCE OF PROJECT MANAGERS

Project manager	start	end	Ongoing cases	Ongoing activities
man8	2015/11/03	2017/08/23	96%	19%
man7	2017/07/17	2018/05/15	86%	30%
Man2	2018/05/17	2020/05/05	88%	48%

From the review of Table 8, it can be seen that during the tenure of the project manager (eight managers), almost all purchase titles were active, but the amount of activities performed and not completed indicates that the follow-up of activities as Not effective.

Similarly, it can be concluded that during the tenure of Project Manager 7 and Project Manager 2 (Manager 7), the performance has been somewhat better. Of course, the performance of the project manager can't be considered the only effective factor in the progress of the project and should be considered other factors such as changes in the executive staff of the Employer, changes in society in different time periods, sanctions, political and economic problems.

C. *Organizational Perspective*

In this section, event logs related to all actors are examined from an organizational perspective and based on the most common criteria of centrality, including Degree centrality, Betweenness centrality, and Closeness centrality, which are briefly described below. The degree of centrality indicates the connections and links between people in the network. Degree centrality is used to determine the importance of nodes and to represent the structural position of actors in a network. Central actors have a higher credibility or centrality than other actors[29].

If a node is the only connection point in the path of other nodes, this node is very important and has a high betweenness centrality[30]. An actor with a high betweenness centrality has control over the network and acts as a mediator in the network, because other actors depend on this actor to communicate with each other[31].

Closeness centrality measures the proximity of the path length from one node to the other nodes[29].Actors with a higher closeness centrality are actors who are in direct contact with many actors on the network. In other words, centrality of closeness stands for convenience and ease of connection between one node and other nodes[30]. The social network diagram resulting from the connections between the actors active in the process based on the Degree centrality parameter is shown in Figure 4. To reduce the complexity of the graph, nodes whose degree of centrality is calculated to be zero have been removed. According to Figure 4, the roles of the actors are compared based on three criteria of centrality and in eight shopping disciplines in Tables 9 and 10.

Based on the designed process, the results of Tables 9 and 10 have been largely expected. Because people such as engineering deputy and Procurement deputy (with the task of approving and directing the process), consultants (with the task of controlling and providing advice in the process) and document control (DCC)experts (with the task of registering and distributing information), are among the most influential people in the network.

A person like the order and contract manager who is responsible for the final review, assigning the code and sending the purchase request to other departments, is also present in the network as a person who has several communication channels. Some experts, according to the characteristics of each purchasing group, have multiple communication channels to better advance the process.

Given that the documents and information are distributed among the experts after the approval of the engineering and procurement deputies, it is clear that these people have a very important role in the distribution of information. DCC experts and document control consultants have also been ranked high in this criterion because they are responsible for distributing another part of the documents. Of course, the low presence of project managers and contract managers is one of the ambiguous points in communication network analysis. Also, the lack of communication between engineering experts with consultants, weak communication between the engineering and procurement deputies with their experts in these two fields, etc. are also among the issues that can be improved performance of the communication network between the actors in the network.

[DOI: 10.52547/itrc.14.2.66]

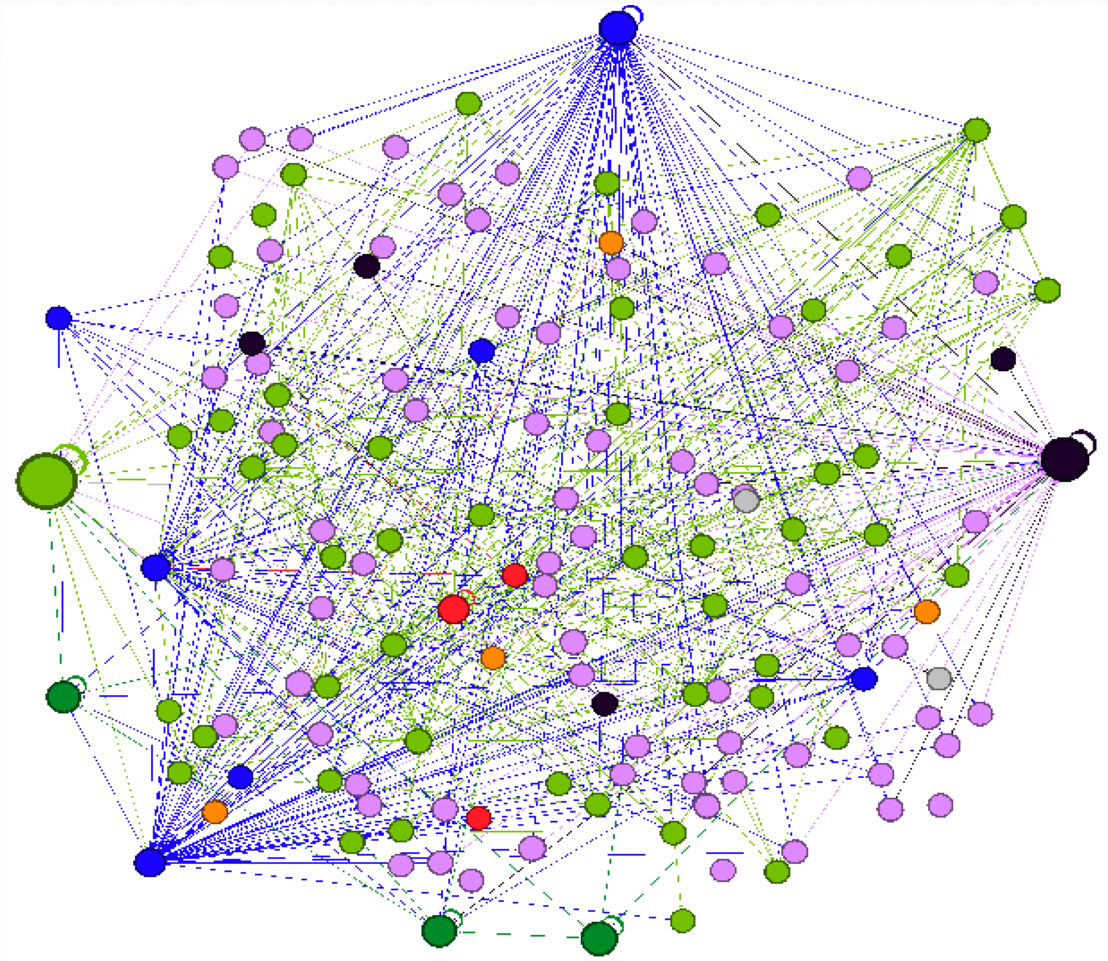


Fig. 4. Social network graph based on the Degree Centrality

Color legend: Vendor: Purple, Engineering: Green - Document Control: Blue - Project Management: Black - Consultant: dark green- procurement: Orange - Quality Control: Red , Other Sections: Gray

TABLE IX. COMPARISON OF THE THREE CRITERIA OF CENTRALITY IN THE WHOLE EVENT LOGS

Degree Centrality		**Closeness Centrality**		**Betweenness Centrality**	
actor	grade	actor	grade	actor	grade
Procurement deputy	19585	Manager 6 (Order and Contract Manager)،	1	Procurement deputy	6340.422
Engineering deputy	19430	Consultant 2.DCC	0.729	Consultant 2.DCC	4136.766
Consultant 2.DCC	9337	Employee 2 (DCC Expert)	0.669	Engineering deputy	1145.40
Consultant3	6214	Procurement deputy	0.577	Project Manager 8	296.323
Consultant2	5242	Engineering deputy	0.536	Consultant2	266.279
Consultant1	4864	Consultant2	0.508	Employee 2 (DCC Expert)	214.498
Vendor37	3451	Employee 7 (DCC Expert)	0.505	Consultant1	143.928
Employee 2 (DCC Expert)	2287	Vendor239	0.483	Employer	86
Vendor19	1627	Vendor259	0.468	Consultant3	68.38
Vendor161	1378	Consultant3	0.390	Project Manager 7	46.67
Indicates people who have more power and influence in the network, and other people refer to them as a reference.		Representative people who have high flexibility and use many communication channels.		People who have a role and dominance in the network and are the connection point between different parts.	

[DOI: 10.52547/itrc.14.2.66]

TABLE X. COMPARISON OF THREE CENTRALITY CRITERIA BASED ON PURCHASING DISCIPLINES

Discipline	Degree Centrality	Closeness Centrality	Betweenness Centrality
mechanical	Procurement deputy, Consultant 2.DCC, Engineering deputy, Vendor 19, Employee 2 (DCC Expert)	Manager 6 (Order and Contract Manager), Employee 12 (Electricity Manager), Employee 35 (material expert), Employee 38 (Rotating Equipment Expert)	Employee 26 (Fixed Equipment Expert), Procurement deputy, Employee 7 (DCC Expert), Employee 17 (engineering expert), Employee 54 (mechanic expert), Employee 47 (mechanic expert)
piping	Engineering deputy Procurement deputy, Consultant3, Consultant2, Employee 2 (DCC Expert)	Employee 8 (Process Expert), Manager 6 (Order and Contract Manager), Employee 55 (Process Expert), Employee 43 (Mechanical Equipment Expert), Employee 57 (DCC Expert)	Procurement deputy, Engineering deputy، Employee2 (DCC Expert), Consultant 2.DCC
instrument	Engineering deputy, Procurement deputy, Consultant2, Consultant 2.DCC, Consultant3	Manager 6 (Order and Contract Manager), Employee 53 (Piping Design Expert), Employee 21 (Civil Manager)	Employee 7 (DCC Expert), Procurement deputy, Engineering deputy, Employee 18 (instrumentation expert)
electrical	Engineering deputy, Procurement deputy, Consultant 2.DCC, Consultant3 ،Consultant2, Consultant1	Manager 6 (Order and Contract Manager), Employee 57 (DCC Expert)	Procurement deputy, Engineering deputy, Consultant 2.DCC
fix equipment	Procurement deputy, Vendor12, Consultant 2.DCC, Engineering deputy	Manager 6 (Order and Contract Manager), Employee 18 (instrumentation expert), Employee 8 (process expert), Employee 16 (piping manager), Employee 12 (electrical manager), Employee 57 (DCC expert)	Procurement deputy ،Employee 7 (DCC Expert), Engineering deputy, Employee 13 (fixed equipment expert)
HVAC	Vendor37, Procurement deputy, Consultant 2.DCC, Engineering deputy	Manager 6 (Order and Contract Manager), Employee 12 (electrical manager), Employee 18 (instrumentation expert), Employee 50 (civil engineer expert), Employee 57 (DCC expert), Employee 21 (civil engineer manager)	Procurement deputy, Consultant 2.DCC, Vendor37, Employee 47 (Mechanical Equipment Expert),
process	Procurement deputy, Vendor71, Engineering deputy, Consultant 2.DCC	Employee 18 (instrumentation expert), Employee 21 (Civil manager), Manager 6 (Order and Contract Manager), Employee 12 (electrical manager), employee 24 (control and instrumentation expert), employee 50 (civil engineer expert)	Employee 8 (Process Expert), Employee 17 (Stress Analysis Expert), Employee 31 (Engineering Coordinator)
safety	Engineering deputy, Procurement deputy, Consultant 2	Manager 6 (Order and Contract Manager), Employee 57 (document control expert), Employee 54 (expert in charge of mechanics)	Procurement deputy, Engineering deputy ،Employee 7 (DCC Expert),Vendor 43
review	According to this criterion, most of the engineering deputy, supply deputy, consultants and two document control staff are present.	Indicates the presence of people with the role of mediator and connecting different sectors	Presence of engineering deputy, procurement deputy and DCC experts in most purchasing disciplines

V. CONCLUSION AND DISCUSSION

Due to the long duration of this project, there have certainly been many challenges and problems in various dimensions in the implementation of the project. The most important concerns are challenges at the highest levels of project management to the expert level, imposing sanctions and creating problems in contracts with vendors, incorrect implementation of procedures defined by individuals, poor and inadequate communication between actors and frequent mistakes. Obviously, due to the wide scope of the project, the unwritten nature of all effective conditions and the passage of time, not all of these factors can be examined. In this study, using the reporting of events recorded in information systems, document control system, we have tried to analyze some of these issues from three perspectives of the process.

From the results of the analysis, it can be seen that the challenges of the common index between the purchasing disciplines include the existence of loops and high repetition in performing activities, especially in the process of preparing technical and engineering documents (e.g. product request, technical evaluation, order request purchasing) as well as poor communication and lack of direct communication between consultants and the experts and lack of direct and appropriate communication between procurement and engineering deputies with experts in these two fields.

These problems have ultimately led to increased time and consequently increased costs, reduced company credit, reduced employee morale and motivation, and so on. Therefore, it is recommended to reconsider the process and strengthen the appropriate and practical communication between the actors in future projects.

REFERENCES

[1] Li, F., *Leading digital transformation: three emerging approaches for managing the transition.* International Journal of Operations & Production Management, 2020.

[2] Zerbino, P., A. Stefanini, and D. Aloini, *Process science in action: A literature review on process mining in*

[DOI: 10.52547/itrc.14.2.66]

business management. Technological Forecasting and Social Change, 2021. 172: p. 121021.
[3] Van der Aalst, W.M., *Challenges in business process mining*. Applied Stochastic Models in Business and Industry (to appear), 2010.
[4] Alvarez, C., et al., *Discovering role interaction models in the Emergency Room using Process Mining*. Journal of biomedical informatics, 2018. 78: p. 60-77.
[5] Pika, A., et al., *Using big data to improve safety performance: an application of process mining to enhance data visualisation*. Big Data Research, 2021. 25: p. 100210.
[6] Dahlgaard, J.J. and S.M. Dahlgaard ‐ Park, *Lean production, six sigma quality, TQM and company culture*. The TQM magazine, 2006.
[7] THAMRIN, D.A.F., *Six Sigma Implementation and Integration within Project Management Framework in Engineering, Procurement, and Construction Projects-A Case Study in a Southeast Asian Engineering, Procurement, and Construction Company*. 2017.
[8] Grisold, T., et al., *Using process mining to support theorizing about change in organizations*. 2020.
[9] Song, M. and W.M. Van der Aalst, *Towards comprehensive support for organizational mining*. Decision support systems, 2008. 46(1): p. 300-317.
[10] Van Der Aalst, W., *Service mining: Using process mining to discover, check, and improve service behavior*. IEEE transactions on services Computing, 2012. 6(4): p. 525-535.
[11] Van Der Aalst, W., *Spreadsheets for business process management: Using process mining to deal with "events" rather than "numbers"?* Business Process Management Journal, 2018.
[12] Rashnavadi, Y., et al., *Business Process Discovery from Emails: Text Classification and Process Mining-A Case Study of Procurement Process*. Innovatus, 2022. 5(1): p. 1-10.
[13] Siek, M. and R. Mukti. *Business process mining from e-commerce event web logs: Conformance checking and bottleneck identification*. in *IOP Conference Series: Earth and Environmental Science*. 2021. IOP Publishing.
[14] Kadasah, E.A., *PROCESS MINING IMPLEMENTATION BENEFITS FOR ORGANIZATIONS USING ERP SYSTEMS*.
[15] dos Santos Garcia, C., et al., *Process mining techniques and applications–A systematic mapping study*. Expert Systems with Applications, 2019. 133: p. 260-295.
[16] Dakic, D., et al., *BUSINESS PROCESS MINING APPLICATION: A LITERATURE REVIEW*. Annals of DAAAM & Proceedings, 2018. 29.
[17] Flyvbjerg, B., N. Bruzelius, and W. Rothengatter, *Megaprojects and risk: An anatomy of ambition*. 2003: Cambridge university press.
[18] Ruan, X., et al., *Knowledge integration process in construction projects: a social network analysis approach to compare competitive and collaborative working*. Construction management and economics, 2012. 30(1): p. 5-19.
[19] Aalst, W.v.d., et al. *Process mining manifesto*. in *International conference on business process management*. 2011. Springer.
[20] Eggers, J. and A. Hein. *Turning Big Data into Value: A Literature Review on Business Value Realization from Process Mining*. in *ECIS*. 2020.
[21] PUDJIANTO, W., *PROCESS MINING IN GOVERNANCE, RISK MANAGEMENT, COMPLIANCE (GRC), AND AUDITING: A SYSTEMATIC LITERATURE REVIEW*. Journal of Theoretical and Applied Information Technology, 2021. 99(18).
[22] Van der Aalst, W.M., *Process mining: data science in action*. 2016: Springer.
[23] Claes, J. and G. Poels, *Merging event logs for process mining: A rule based merging method and rule suggestion algorithm*. Expert Systems with Applications, 2014. 41(16): p. 7291-7306.
[24] Helm, E., et al., *Towards the use of standardized terms in clinical case studies for process mining in healthcare*. International journal of environmental research and public health, 2020. 17(4): p. 1348.
[25] Suriadi, S., et al., *Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs*. Information systems, 2017. 64: p. 132-150.
[26] Perimal-Lewis, L., et al., *Application of process mining to assess the data quality of routinely collected time-based performance data sourced from electronic health records by validating process conformance*. Health informatics journal, 2016. 22(4): p. 1017-1029.
[27] Liao, P.-C., et al., *The relationship between communication and construction safety climate in China*. KSCE Journal of Civil Engineering, 2014. 18(4): p. 887-897.
[28] Wirth, R. and J. Hipp. *CRISP-DM: Towards a standard process model for data mining*. in *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. 2000. Manchester.
[29] Hossain, L., *Communications and coordination in construction projects*. Construction management and economics, 2009. 27(1): p. 25-39.
[30] Zhang, J. and Y. Luo. *Degree centrality, betweenness centrality, and closeness centrality in social network*. in *Proceedings of the 2017 2nd International Conference on Modelling, Simulation and Applied Mathematics (MSAM2017)*. 2017.
[31] Chowdhury, A.N., P.H. Chen, and R.L. Tiong, *Analysing the structure of public–private partnership projects using network theory*. Construction management and economics, 2011. 29(3): p. 247-260.

Seyedeh Motahareh Hosseini received her M.Sc. degree in Industrial Engineering from Tehran University, School of Industrial Engineering. She is currently a Ph.D. candidate at the Faculty of Industrial and Systems Engineering, Tarbiat Modares University. Her research interests include Business Process Management and Process Mining.

Mohammad Aghdassi is a Professor in the Faculty of Industrial and Systems Engineering at Tarbiat Modares University. His field of research is Business Process Management, Reengineering and Organizational Transformation Management (BPR, BPMS, BPM), Organizational Design, Knowledge Management Social Development, Knowledge Management Systems.

Babak Teimourpour obtained his Ph.D. in Industrial Engineering from Department of Industrial Engineering, Tarbiat Modares University (TMU), Tehran, Iran. He teaches Ph.D. and M.Sc. level courses. His research interests include Data Mining and Social Network Analysis. His team won the Iran Data Mining Cup in 2010.

[DOI: 10.52547/itrc.14.2.66]

Amir Albadvi is a Professor of Information Systems. He received his Ph.D. from LSE in London and eventually earned his spot as technology thought leader and researcher in IT transformation and e-Strategy. His recent research focuses on Digital Ecosystems. He has more than 250 publications and over 2000 citations in different scientific journals.

[DOI: 10.52547/itrc.14.2.66]