**TELKOMNIKA Telecommunication, Computing, Electronics and Control**
Vol. 19, No. 3, June 2021, pp. 781~791
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v19i3.18153


# Mapping log data activity using heuristic miner algorithm in manufacture and logistics company

**Syafrial Fachri Pane, Rolly Maulana Awangga, M. Amran Hakim Siregar, Dinda Majesty**
Applied Bachelor Program of Informatics Engineering, Politeknik Pos Indonesia, Indonesia

**Article Info**

***Article history:***

Received Jul 18, 2020
Revised Nov 6, 2020
Accepted Nov 25, 2020

***Keywords:***

Heuristic miner
Log data
Logistics
Manufacturing
Python
XML

**ABSTRACT**

Strategies for the procurement of goods and services are essential for companies in Indonesia's manufacturing and logistics sectors. The solution to reducing the existing problem is to make a mapping plan, such as verifying documents from each department, so that it takes a long time, resulting in many issues, such as procedural misuse findings. Heuristics miner algorithms get data to form logs that consist of goods and services procurement activities. Processing log data into XML data (data extraction), which produces a dependency model and business and casual matrix (discovery process), then determines the value of fitness and precision (suitability) called the conformity checking phase process. This phase aims to produce a new business (process enhancement phase), which will create a solution to the risk of delay and procedural abuse. The results of each of these processes rank each stage of the procurement of goods and services sequentially and together to provide time-efficient and accurate decisions, resulting in project implementation comparable to the company's business strategy. Implement the heuristics miner algorithm using the Python programming language.



***Corresponding Author:***

Syafrial Fachri Pane
Applied Bachelor Program of Informatics Engineering
Politeknik Pos Indonesia
Sariasih St. No. 54, Sarijadi, Sukasari, Bandung City, West Java, Indonesia
Email: syafrial.fachri@poltekpos.ac.id

## 1. INTRODUCTION

The manufacturing and logistics industry in Indonesia in the second quarter of 2020 developed quite healthy growth. Thus, the coronavirus pandemic's effect that hit all parts of Indonesia did not include being able to disrupt the manufacturing and logistics related sectors in various countries such as China, Japan, and America. Indonesia's most significant investor cooperation partner, who discusses the coronavirus pandemic, thus companies in Indonesia provide business strategies to be able to provide maximum services that are looking for the latest technology-rare measures for renewal, which will happen [1-3]. Manufacturing and logistics companies in Indonesia must have concepts and strategies that align with the target [4]. The company's strategy approves technology's application to improve the performance and quality of service to consumers [5].

Log data activity stores the analysis of the process of procurement of goods and services in the company. The storage capacity of the log data is substantial [6]. Therefore, it requires a long time due to the many verification documents that must be provided to all relevant departments to carry out the procurement of goods and services, resulting in the purchase of products and services experiencing a time delay and many procedural misuses. The process of procuring goods and services in large numbers makes it difficult for

companies to give quick decisions, so it needs to analyze accurate data mining [7-9]. The mechanisms in the procurement of goods and services require analysis to determine and measure how urgent their needs [10]. These mechanisms include implementation procedures, legal rules, specifications, and especially suppliers of goods and services, namely vendor [11]. Selecting vendors helps companies maintain the quality of the needs of goods and services [12, 13]. The process of filling goods and services needed in the activity data log will go through stages such as data extraction, preprocessing, discovery phase, conformity checking phase and placement phase, and analysis and evaluation by producing a new business process diagram model to provide adequate and timely decision direction [14].

The staging process is the application of the miner heuristics algorithm, the characteristics of the miner heuristics algorithm can overcome noise, have a better representation of bias, separate and combine special considerations following the original process and can handle loops [15, 16]. The application of this algorithm is a technology innovation product for companies to provide changes to determine the master schedule up to process order (PO). The programming language used is Python because this programming language is more straightforward, opensource and is a programming language that provides convenience for the process of automation comparing other programming languages [17-19]. This research requires large-scale data export, machine learning, and data analysis using the Python programming language [20, 21]. This study's result applies the miner's heuristic method to provide fast results for companies to determine the procurement process of goods and services sequentially, simultaneously, according to the time needed to accelerate starting the project process.

## 2. RESEARCH METHOD

Base on Figure 1 this research methodology's flowchart, the authors explain the flow of the research process carried out. This research will provide scope/discussion to be well structured and organized and streamline the research because the initial to the final strategy has been determined. The methodology used by the author is the methodology flow of the miner heuristic algorithm. Figure 1 shows the flow of this research methodology.

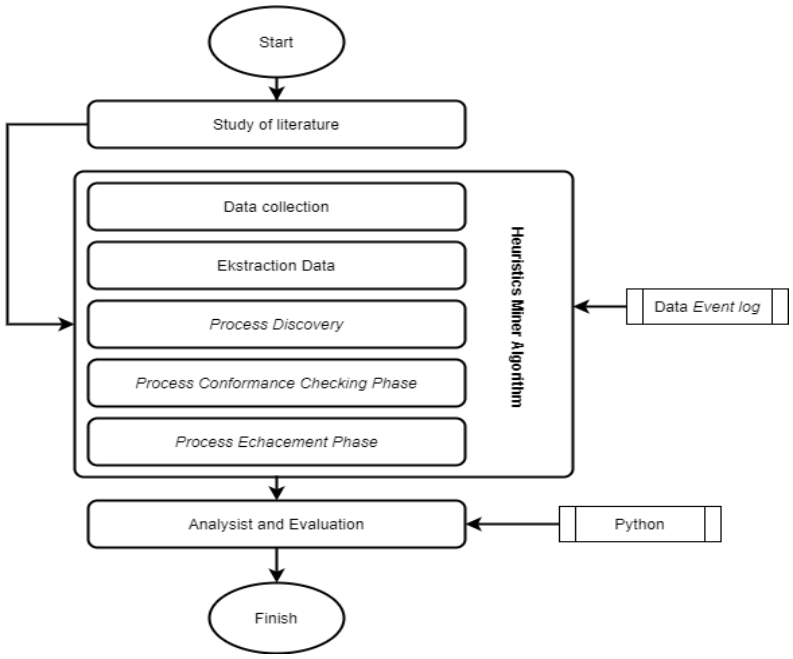


Figure 1. Research method

### 2.1. Study of literature

The current business process for procurement of goods and services has not yet explained in detail the business process model of the heuristic miner algorithm to produce a business process model that can be a suggestion in the process of procuring goods and services [22]. The large and messy data calculation is the primary reference that must be eliminated in the current business process model so that the data entered in the event log data is the correct data calculation to achieve better business goals [23, 24]. The process of

procurement of goods and services has not explained in detail the business processes of the miner's heuristic algorithm to produce business processes that can be suggestions for use [25, 26]. Fraud detection makes it possible to procure goods by mining the event log of goods using the heuristic miner algorithm without calculates the value of the fitness and dependency graph [27]. The threshold value is the main focus as a reference in figuring the value of the dependency graph and displaying the business process model. A University uses this business process model for book procurement [28]. Mapping topological data to analyze business processes requirements using the Python programming language as the best solution in data analysis and machine learning [29, 30].

### 2.2. Implementation of heuristic miner algorithm

#### 2.2.1. Data collection (event log)

Data collection aims to know the history of procurement of goods and services using primary data, namely excel procurement status report data. The data consists of case id, activity id, timestamp, and originator in the form of an event log. The event log data is then analyzed to determine the process of procuring goods and services based on the time sequence. The event log data will be executed using heuristic miner algorithms steps to obtain the value and graph of new business process from the data collection result.

a. Extraction data

Data extraction aims to generate data from the event log into XML data using the heuristic miner method. After getting the data, then the data is tested with the type .xls converted into .xml. Changes to data formats to facilitate coding at a later stage. After obtaining the sequential research data, each .xls data was converted into the .xml data format. version 1.0 with the encoding type "UTF-8". The .xml data structure can be seen in point 3.1.1. event log data extraction.

b. Event log .XML

This .xml log event encoding aims to determine the log event encoding at a later stage. At this stage, after obtaining the event log data extraction format framework. In this research, the .xls event log is converted to .xml form, and can also be converted to another form, namely .mxml.

#### 2.2.2. Process discovery

The objective of the discovery process is to determine the dependency value by calculating the dependency frequency value. The two events use XML event log data to produce a business process model and a simple matrix using the heuristic miner algorithm. The fitness value obtained from two calculations using two different processes flow, i.e., the one received by the discovery process (event log) and the one established before, has a significant difference value. From the fitness value difference of both approaches, it can be seen that the model process has a fitness value of 0.8494.

#### 2.2.3. Process conformance checking phase

The purpose of the phase check process is to determine the fitness value of the casual matrix data using heuristic miners and produce a fitness and precision table where the activities that occur in the event log are related to the initial business model and the dependence of the model on process discovery. Data that has been found according to its frequency can be processed at the conformity check stage with the final result in the form of the value of the enhancement stage to determine the flow of business processes based on performance. In the process conformance checking phase, the value of each activity can be seen and a new business process flow will be formed to determine the activities to be carried out.

#### 2.2.4. Process enhancement phase

The placement stage aims to determine the placement model based on the conformity results from the suitability value table and provide a new business process diagram. At the stage of refining the business model based on the works carried out at the conformance checking process stage, the process flow explanation stage will become a suggestion to apply to business processes based on the information and know the processing time by looking. The initial business process flow makes comparisons based on the explanation of the advice and business processes resulting from the preprocessing process and the four stages in the Yahoo miner heuristic stage.

#### 2.2.5. Analysis and evaluation

The purpose of research and evaluation is to provide conclusions and solutions for procurement of goods and services to reduce the risk of delays. Can provide a dash of business action recommendations, which process flows to use and which are repeated frequently to be accessed and seen from which operations are efficient and practical to do so that business processes do not occur. so that the application of the miner heuristic algorithm can provide solutions to increase effectiveness in carrying out activities that can actually be carried out simultaneously without having to wait for the previous activity to continue the next activity.

### 2.2.6. Script python algorithm heuristic miner

The next stage is to execute the dependency table process known in the calculation of the fitness score and processed into a Python script to obtain a graph of the heuristic miner algorithm results. To produce a heuristic miner algorithm graph diagram, using a Python version of at least 3.2. or above is recommended. At this stage, you are also related to installing the Python environment plugin to support scripts to display graph diagrams that are run through an algorithm script that has been created with the name channel.py.

## 3. RESULTS AND ANALYSIS

Determine the analysis results that produce business processes using the heuristic miner algorithm. In this study, there are four steps of completion based on the event log data, the discovery process, the process conformance check phase, and the process improvement stage. In the event log data, you can see the initial data that will be processed in the heuristic miner algorithm to find out which cases can be done simultaneously.

### 3.1. Event log

Collecting data in the event log must go through a sorting process so that there are no data errors before entering the discovery process. In the event log, there are many activities, as in Table 1. By processing this data, it will result in the discovery of event log data as Table 1. Base on Table 1, there are six examples of activities that exist in the event log. A, B, C, D, E are the identities of the six activities contained in the event log. Use a business process flow to define the data to be processed in Table 1 and match the data from the event log with the heuristic miner's need.

Table 1. Event log

| Description | Case id | Activity id | Originator | Time Stamp |
|---|---|---|---|---|
| Determine the project schedule | Case 1 | A | Department DPP | 9-3-2017:15.01 |
| Vendor data | Case 2 | B | MP | 9-3-2017:15.12 |
| Vendor lose | Case 4 | D | DPPM | 9-3-2017:16.03 |
| Procurement of goods and services | Case 1 | C | MP | 9-3-2017:16.07 |
| Enter the discussion stage | Case 4 | D | MP | 9-3-2017:18.25 |
| Determine project cost | Case 4 | E | MP | 10-3-2017:9.23 |

### 3.1.1. Extraction data event log

After getting the sorted test data, the data of type .xls is changed to .xml. This is done to make coding easier in the next stage, following the event log extraction framework:

```
<?xml version="1.0" encoding="UTF-8"?>
<?mso-application      progid="Excel.Sheet"?><Workbook      xmlns="urn:schemas-microsoft-
com:office:spreadsheet"  xmlns:c="urn:schemas-microsoft-com:office:component:spreadsheet"
xmlns:html="http://www.w3.org/TR/REC-html40"              xmlns:o="urn:schemas-microsoft-
com:office:office"               xmlns:ss="urn:schemas-microsoft-com:office:spreadsheet"
xmlns:x2="http://schemas.microsoft.com/office/excel/2003/xml"       xmlns:x="urn:schemas-
microsoft-com:office:excel" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
……
</Workbook>
```

The data extraction framework in the event log above is an xml data format to determine the dependency value on the event log data.

### 3.1.2. Event log .xml

After getting the sorted test data, the data of type .xls is changed to .xml. This is done to make coding easier in the next stage, following the event log extraction framework:

```
<?xml version="1.0" encoding="UTF-8"?>
<ProcessMap numNodes="8" nodeThreshold="1.0" edgeThreshold="0.2136" discoVersion="2.2.1">
<Layout width="5.0997605" height="1.25"/>
<Nodes size="8">
<Node index="0" activity="">
<Frequency total="2" case="1" start="1" end="1" maxRepetitions="2"/>
<Duration total="0" min="0" max="0" mean="0" median="0"/>
<Layout x="0.0" y="0.4056" width="0.10588733" height="0.1888"/>
</Node>
<Node index="1" activity="activity id">
<Frequency total="1" case="1" start="1" end="1" maxRepetitions="1"/>
<Duration total="0" min="0" max="0" mean="0" median="0"/>
```

```
<Layout x="0.25878865" y="0.4056" width="0.10588733" height="0.1888"/>
</Node>
……..
</ProcessMap>
```

The .xml event log data above is useful for carrying out the discovery stage and determining the dependency value.

### 3.2. Process discovery

After matching and sorting data to determine the data that needs to be collected, in this installation, the heuristic mining algorithm will calculate the initial value in the event log data in Table 1 using the formula:

$$DG = (a;b)|(a \in T \,\Delta\, b \in Ta)\nabla(b \in T \,\Delta\, a \in b)g \quad (1)$$

$$a => w^b = \left(\frac{|a> w^b|-|b>w^a|}{|a> w^b|+|b>w^a|+1}\right) \quad (2)$$

Base on Table 2, After getting the calculation dependency value, the threshold value determines the graph which was obtained by experimentation by taking amounts from 0.0-0.99 compared to the dependency value. Using a small threshold value of 0.88 in the experiment will cause many changes to the data, so that is not suitable. Therefore, this study uses a threshold value of 0.88. Figure 2 shows the process dependency model and the representation of the process model into a casual matrix.

Table 2 shows that the row's highest value indicates that the activity caused mostly by the movement. In contrast, the highest value in the column suggests the action that caused the move the most. In the business process model stage, there is a flow of business process activities, an activity paired from the dependency process shown in Figure 2. Base on Table 3, the causal matrix is formed after the dependency graph, in which any ramification can be seen. There are two types/forms of non-observational activities in this causal matrix, namely AND and XOR. AND means that branching of activities can be carried out in parallel, while XOR means that it is allowed only to choose one path in a branching activity. Table 3 shows the comparison of activities a>b and b<a

Table 2. Calculation dependency

| $a = w^b$ | A | B | C | D | E |
|---|---|---|---|---|---|
| A | 0 | 0.8 | 0 | 0 | 0.8 |
| B | -0.5 | 0 | 0.5 | 0.67 | 0 |
| C | 0 | -0.5 | 0 | 0 | -0.5 |
| D | 0 | -0.67 | 0 | 0 | -0.67 |
| E | 0 | -0.5 | 0 | 0 | -0.5 |

Table 3. Casual matrix

| Activity | Input | Output |
|---|---|---|
| A | A | (B v E) ^ (C v E) |
| B | A^B | D ^ E |
| C | A | B ^ C ^ D ^ E |
| D | (B v E) (C v E) | F G H I |
| E | A | J K L |

### 3.3. Process conformance checking phase

The process conformance checking Phase is the process of measuring and examining the trace logs to complete a specific trace log with a logical model process. For all values i, $m_i$, $c_i$ and $r_i$, $p_i$, therefore fitness = 0, f, 1. The following formula helps to find the fitness value. The formula that produces the fitness value in Table 4.

$$f = \frac{1}{2} = \left(1 - \frac{\sum_i^k \; =1 n_i m_i}{\sum_i^k \; =1 n_i c_i}\right) + \frac{1}{2}\left(\frac{\sum_i^k \; =1 n_i r_i}{\sum_i^k \; =1 n_i p_i}\right) \quad (3)$$

The formula above calculates the highest conformity value in the conformity checking process, which is the calculation of log data to produce a sequence of business processes in the procurement of goods and services, as presented in Figure 2.

Base on Table 4, Utilizing the fitness value to measure how high the system's probability of receiving log data with a valid procedure process model for all values of i, $m_i < c_i$, and $r_i < p_i$, therefore fitness = 0<f<1. Obtain a fitness value using the following formula. The following fitness Table 4 in the business process flow: Based on the grooves in Figure 2, which is on the log trace, the fitness value is 0.8494. Meanwhile, based on the initial procedure flow and log trace, the fitness value is 0.8374. Judging from the fitness value derived from the two calculations using two different process flows, namely the process flow obtained from the discovery process (event log) and the previously defined process flow has a considerable difference in value. Based on the differences in the two model's fitness values, the process model with a fitness value of 0.8494. This stage consists of several checks on the value of the process model's suitability flow to the event log. There are several calculations, including calculating the value of fitness (recall), and precision (appropriateness) in Table 5.

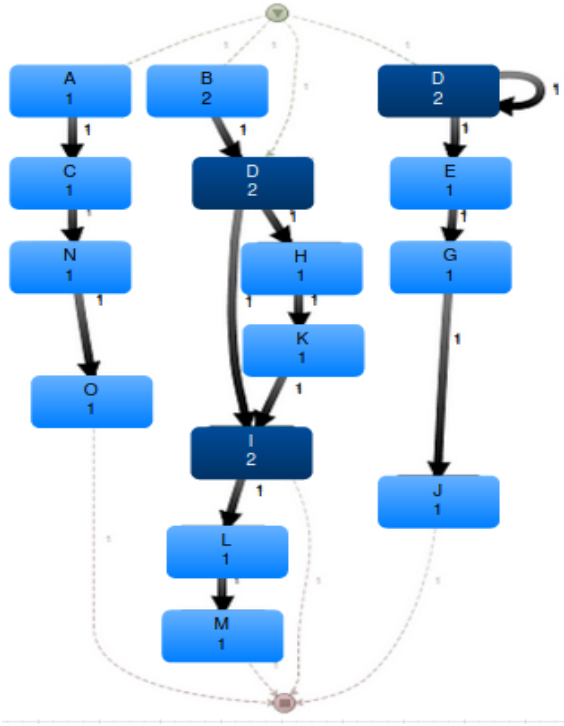


Figure 2. The process of a business model dependency

Table 4. Fitness value

| Dependency Threshold | Relative-to-Best Threshold | Positive Observation Threshold | Number of Connections | Fitness |
|---|---|---|---|---|
| 0.9 | 0.05 | 10 | 333 | 0.6593 |
| 0.9 | 0.05 | 100 | 333 | 0.6725 |
| 0.9 | 0.05 | 200 | 333 | 0.7101 |
| 0.9 | 0.05 | 500 | 328 | 0.8374 |
| 0.9 | 0.05 | 1,000 | 328 | 0.8494 |
| 0.9 | 0.05 | 6,500 | 6,046 | 0.7595 |

Table 5. Appropriateness

| Appropriateness | Log Trace |
|---|---|
| 13 | ABDEGIJKL |
| 12 | ABDFGIKJLN |
| 9 | ABDFGIKJLO |
| 8 | ABDFGH |
| 7 | ABCBCBDFGIJKL |
| 6 | ABCBDEGIJL |
| 5 | ABDFHNO |

### 3.4. Process enhancement phase

The process Enhancement Phase is the stage to improve pre-existing business models based on the results of an analysis of the process conformance checking phase. At this stage, checking for conformity will be discussed in Table 4, which explains the flow of the goods and services procurement process that is shorter, faster, and according to needs. Use the new business process flow to suggest a company's business processes to make it easier to procure goods and services. Table 1 shows a list of activities in procuring goods and services in the event log. Table 5 shows the steps in carrying out these activities. Table 6 shows activity data from the new business process flow, and Figure 3 shows the new business process.

The results of the goods and service procurement process model from the process conformity checking phase produce a new business process flow as a suggestion to improve the existing business process flow in the previous goods and services procurement process and to find out activities in the process of procuring a company's goods, business processes recently has the benefit of minimizing the occurrence of problems in the process of obtaining goods and services, such as delays, out of stock, and other issues in the procurement of goods and services. The actual and related problems that exist today and which can take advantage of the

proposed solution, namely the previous company process flow, see in Table 1, the company's business processes are carried out one by one before the miner's heuristic algorithm is implemented. Solving the business process flow problem must use maps and strategic arrangements that emphasize flow effectiveness, time efficiency, and cost. The problem currently occurring is that there are too many business processes flows, requiring validation from one activity to the next. This underlies the basis of research using a miner heuristic algorithm with the ability to log events or business processes. This research implements the heuristic miner algorithm as a basis for forming a new business process flow in the company based on activities that can be seen in Table 6 using the python programming method as an application to process data and produce business process images which can be seen in Figure 3. Python is used because of excellence in analyzing scientific data.

Table 6. Enhancement phase

| No | Activity | Description | Processing Time |
|---|---|---|---|
| A | Determine the project schedule | Dating | Optional according |
| B | Vendor data | All data vendor | One working day |
| C | Vendor lose | Searching | One working day |
| D | Need for the procurement | Specification | Optional according |
| E | Discussion stage | Negotiation | Specification agreement |
| ……. | ……. | ……. | ……. |
| N | Signature contract | Approval of the project holder | One working day |
| O | New vendor validation | Vendor turnover | One working day |

### 3.5. Analysis and evaluation

The purpose of the test is to find out whether the logic functions in Figure 4 using the Python programming language is running well or not. Using the Python programming language in this study is the right decision. Large amounts of data and complex algorithms will make this research challenging to obtain maximum results when choosing a programming language that is unable to process big data. Python programming language is a programmer's choice when going to process big data, data analysis, data science, and learn machine learning. Python programming language is the perfect language for processing big data.

The test went well and got the results as needed. The testing process produces business process outputs, as shown in Figure 4. When the goods and service procurement process starts, the company can carry out activities A and activity B. To save time and costs in the process of procuring goods and services. The company can carry out activity A simultaneously with activity C or activity E. To find out activities A, B, C, and so on, see Table 6. Grouping activities based on the flow of new business processes in Figure 3 dan analysis of activities in Table 6. The circle allows a company to carry out these activities to save time and costs simultaneously.

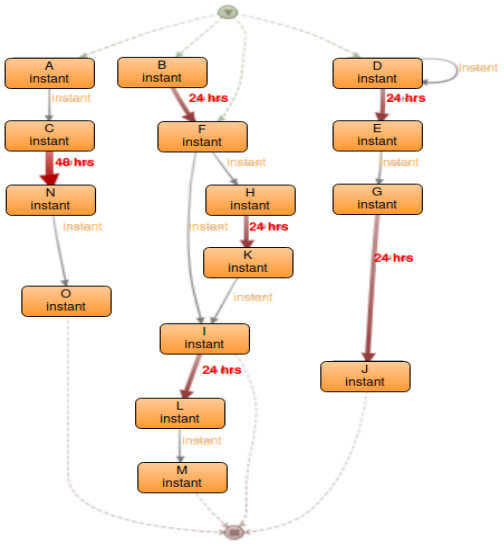


Figure 3. The process of a business model based on performance

Previous research on solving manufacturing and logistics problems using a heuristics miner algorithm did not focus on business processes for the procurement of goods and services that hinder a company from carrying out the procurement process. Many previous studies researched log data on companies, fraud in

running the business process. Still, very few have examined optimizing the procurement business process to become more precise and fast in terms of time and cost. This research superior in solving the problem of improper procurement of goods and service and business processes that do not work by the provisions.

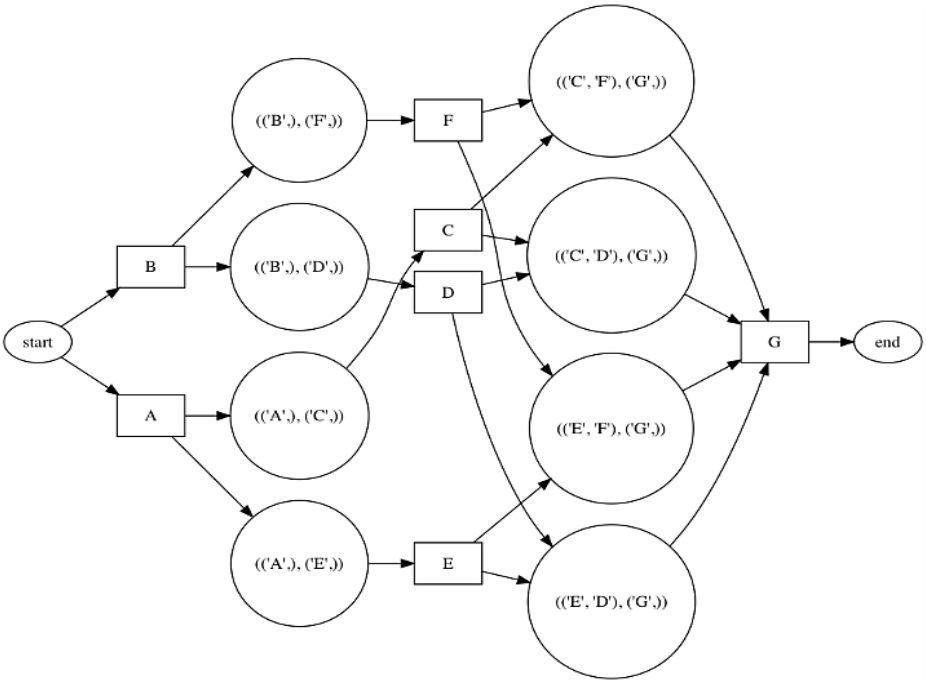


Figure 4. Analysis and evaluation

### 3.6. Script Python algorithm heuristic miner

To generate dependency data in Figure 2 and Figure 3, algorithm execution can be carried out in the Python programming language, with the Python plug-in graphviz script to produce the dependency structure. The following is a heuristic miner algorithm Python script:

```
import graphviz as gv
def apply(log, input_file, output_file):
    satu = set()
    dua = []
    cs = []
    malmo = []
    par = []
    xl = []
    yl = []
    ti = []
    to = []
    satu, dua, cs, malmo, par = build_ordering_relations(log)
    xl, yl, ti, to = make sets(log, satu, dua, cs, malmo)
    print "all tasks:", satu
    print "direct followers:", dua
    print "causalities:", cs
    print "no_causalities:", malmo
    print "parallels:", par
    print "x list:", xl
    print "y list:", yl
    print "initial tasks:", ti
    print "terminal tasks:", to
    build_petrinet(satu, yl, ti, to, output_file)
def build_ordering_relations(log):
    satu = set([item for sub in log for item in sub])
    dua = get_direct_followers(log)
    cs = get_causalities(satu, dua)
    malmo = get_no_causalities(satu, dua)
    par = get_parallels(satu, dua)
```

```
    return satu, dua, cs, malmo, par
def make_sets(log, satu, dua, cs, malmo):
    xl = make_xl_set(satu, dua, cs, malmo)
    yl = make_yl_set(xl)
    ti = make_ti_set(log)
    to = make_to_set(log)
    return xl, yl, ti, to
def get_direct_followers(log):
    dua = []
    for trace in log:
        for index, event in enumerate(trace):
            print index, event
            if index != len(trace)-1:
                if (event, trace[index+1]) not in dua:
                    dua.append((event, trace[index+1]))
    return dua
……
    pn.render(output_file)
```

## 4. CONCLUSION

Do an Analysis of the business process of procuring goods and services for a company to obtain large amounts of data. Processing big data uses a heuristic miner algorithm, which consists of several critical stages. Such as processing event log data (extraction data event log and event log .xml), discovery process, process conformity check phase, process enhancement phase, and analysis and evaluation. Data processing and data mining exploration using the Python programming language. The data study results are in the form of a new business process model that provides recommendations for a company. The business process for procurement of goods and services runs on time and according to company procedures.

## ACKNOWLEDGEMENTS

Syafrial Fachri Pane, Rolly Maulana Awangga, M. Amran Hakim Siregar, Dinda Majesty, Politeknik Pos Indonesia, Indonesia. This research, sponsored by the Ministry of Education of the Republic of Indonesia, is gratefully acknowledged.

## REFERENCES

[1] E. Budiyanti, "The Impact of the Corona Virus on the Indonesian Trade and Tourism Sector (in Indonesia: Dampak Virus Corona Terhadap Sektor Perdagangan Dan Pariwisata Indonesia)," *Kaji. Bid. Ekon. Dan Kebijak. Publik*, vol. XII, no. 4, pp. 19-24, 2020.
[2] A. Mu, "Public Procurement Strategy for Accelerating the Economic Recovery," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689-1699, 2019.
[3] K. Joachim and N. Luciano, "Lessons from the COVID-19 Situation: Rethinking Global Supply Chain Networks and Strengthening Supply Management in Public Procurement in Germany," *IUBH Discuss. Pap.*, pp. 1-12, 2020.
[4] Singh R. K., Modgil S., & Acharya P., "Assessment of supply chain flexibility using system dynamics modeling," *Global Journal of Flexible Systems Management*, vol. 20, no. 1, pp. 39-63, 2019.
[5] S. F. Pane, R. M. Awangga, and B. R. Azhari, "Qualitative evaluation of RFID implementation on warehouse management system," *TELKOMNIKA Telecommunication Computing Electronics and Control,* vol. 16, no. 3, pp. 1303-1308, 2018.
[6] E. R. Mahendrawathi, S. O. Zayin, and F. J. Pamungkas, "ERP Post Implementation Review with Process Mining: A Case of Procurement Process," in *Procedia Computer Science*, vol. 124, pp. 216-223, 2017.
[7] Andrews, Robert, *et al*., "Quality-informed semi-automated event log generation for process mining," *Decision Support Systems,* vol. 132, 2020.
[8] Y. A. Effendi and R. Sarno, "Implementation of the semantic web in business process modeling using Petri nets," *2018 Int. Conf. Inf. Commun. Technol. ICOIACT 2018*, 2018, pp. 741-746.
[9] H. R'Bigui and C. Cho, "The state-of-the-art of business process mining challenges," *Int. J. Bus. Process Integr. Manag.*, vol. 8, no. 4, pp. 285-303, 2017.
[10] Eriksson, Per Erik, "Procurement strategies for enhancing exploration and exploitation in construction projects," *Journal of Financial Management of Property and Construction,* vol. 22, no. 2, 2017.
[11] Noorzai, Esmatullah, "Performance Analysis of Alternative Contracting Methods for Highway Construction Projects: Case Study for Iran." *Journal of Infrastructure Systems,* vol. 26, no. 2, 2020.
[12] L. Prihatini, A. Malik, and R. T. Komara, "Evaluasi kegagalan penyedia jasa konstruksi dalam proses pengadaan jasa konstruksi dengan sistem elektronik (in Indonesia: Evaluation of the failure of construction service providers in the process of procuring construction services with an electronic system)," *J. Online Mhs. Fak. Tek. Univ. Riau*, vol. 4, no. 1, pp. 1-10, 2017.

[13] M. Y. H. Setyawan, R. M. Awangga, and N. A. Lestari, "K-Nearest neighbor algorithm on implicit feedback to determine SOP," *TELKOMNIKA Telecommunication Computing Electronics and Control,* vol. 17, no. 3, pp. 1425-1431, 2019.
[14] A. Rozinat and W. M. P. Van Der Aalst, "Conformance testing: Measuring the fit and appropriateness of event logs and process models," *Conformance testing: Measuring the fit and appropriateness of event logs and process models,* 2005.
[15] Mahananto, "Conformance Checking Process Mining SAP Modul SD (Sales and Distribution) dengan Metode Heuristic Miner Alexander," *J. Sisfo*, vol. 09, no. 02, pp. 1-12, 2020.
[16] Dos Santos Garcia, Cleiton, *et al.* "Process mining techniques and applications–a systematic mapping study," *Expert Systems with Applications,* vol. 133, pp. 260-295, 2019.
[17] R. M. Awangga, S. F. Pane, K. Tunnisa, and I. S. Suwardi, "K means clustering and meanshift analysis for grouping the data of coal term in puslitbang tekMIRA," *TELKOMNIKA Telecommunication Computing Electronics and Control,* vol. 16, no. 3, pp. 1351-1357, 2018.
[18] Virtanen, Pauli, *et al*., "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nature methods,* vol. 17, no. 3, pp. 261-272, 2020.
[19] Jeon, Junhyeok, and Hyun Uk Kim. "Setup of a scientific computing environment for computational biology: Simulation of a genome-scale metabolic model of Escherichia coli as an example," *Journal of Microbiology,* vol. 58, no. 3, pp. 227-234, 2020.
[20] G. Tauzin, *et al.*, "giotto-tda: A Topological Data Analysis Toolkit for Machine Learning and Data Exploration," *J. Mechine Learn. Resear*, vol. 1, pp. 1-5, 2020.
[21] J. Brittain, M. Cendon, J. Nizzi, and J. Pleis, "Data Scientist's Analysis Toolbox: Comparison of Python, R, and SAS Performance," *SMU Data Sci. Rev.*, vol. 1, no. 2, 2018.
[22] S. Azka, P. Angelina, P. Kurniati, and I. A. S. Si, "Process Mining Pada Proses Pengadaan Barang dan Jasa Dengan Menggunakan Algoritma Heuristic Miner (Studi Kasus: Unit Logistik Telkom Engineering School) (in Bahasa: Process Mining in the Process of Procurement of Goods and Services Using the Heuristic Miner Algorithm (Case Study: Logistics Unit Telkom Engineering School))," *e-Proceeding Eng.*, vol. 2, no. 1, 2015, pp. 1332-1338.
[23] R. S. Mangunsong, A. P. Kurniati, and M. K. Sabariah, "Analisis dan Implementasi Process Mining dengan Algoritma Heuristic Miner studi kasus: event logs Rabobank Group ICT Netherlands (Analysis and Implementation of Process Mining with the Heuristic Miner Algorithm. Case study: event logs Rabobank Group ICT Netherlands)," *e-Proceeding Eng.*, vol. 2, no. 1, 2015, pp. 1681-1687.
[24] R. Sowmya and K. R. Suneetha, "Data Mining with Big Data," in *Proceedings of 2017 11th International Conference on Intelligent Systems and Control, ISCO 2017*, 2017.
[25] Y. A. Effendi and R. Sarno, "Time-based α+ miner for modelling business processes using temporal pattern," *TELKOMNIKA Telecommunication Computing Electronics and Control,* vol. 18, no. 1, pp. 114-123, 2020.
[26] F. Mannhardt, M. De Leoni, and H. A. Reijers, "Heuristic mining revamped: An interactive, data-Aware, and conformance-Aware miner," in *CEUR Workshop Proceedings*, 2017.
[27] D. Rahmawati, M. Ainul Yaqin, and R. Sarno, "Fraud detection on event logs of goods and services procurement business process using Heuristics Miner algorithm," *Proc. 2016 Int. Conf. Inf. Commun. Technol. Syst. ICTS 2016*, 2017, pp. 249-254.
[28] I. A. Harin Veradistya Maharani, Angelina Prima Kurniati, "Process Mining in Book Procurement Process with Heuristic Miner Algorithm (Case Study: Telkom University Library) (in Indonesia: Process Mining pada Proses Pengadaan Buku dengan Algoritma Heuristic Miner (Studi Kasus: Perpustakaan Universitas Telkom))," *e-Proceeding Eng.*, vol. 2, no. 1, 2015, pp. 1551-1561.
[29] J. Murugan and D. Robertson, "An Introduction to Topological Data Analysis for Physicists: From LGM to FRBs," arXiv:1904.11044v1, 2019.
[30] R. Semeraro, A. Magi, and J. Hancock, "PyPore: A python toolbox for nanopore sequencing data handling," *Bioinformatics*, 2019, doi: 10.1093/bioinformatics/btz269.

## BIOGRAPHIES OF AUTHORS

**Syafrial Fachri Pane** is a lecturer and researcher in the department of informatics engineering from Higher Education in Indonesia Polytechnic Pos Indonesia Bandung. He was gaining a degree: Associate Degree in Informatics Engineering from Pos Polytechnic of Indonesia, Bachelor of Informatics Engineering from Pasundan University (Indonesia) in 2012, and Magister of Informatics Engineering from Bina Nusantara University (Indonesia). Research interests, teaching, professional experience, and so on. His research conducted in the field of Data Scientist, Data Analysis, Big Data In addition to being a researcher, I am an assessor of BNSP (National Agency for Professional Certification) in the field of Database licensed by the Indonesian government, and assumes the position of head Information Communication and Technology (ICT) in Polytechnic Pos Indonesia. I have an international Big Data certification license, which is Big Data Professional Enterprise from APMG.

**Rolly Maulana Awanggais**, an IT professional in the field of Machine Learning, Internet of Things, Social Network Analysis, and also Geospatial-Intelligence has been implemented in the application of the Agenda of the President of the Republic of Indonesia and the National Policy Support System of the State Secretariat of the Republic of Indonesia. After graduating high school, he was Continuing in college STT Telkom in 2004. There are many from organization and committee, from internal to external campus. During college (2008-2010) working in marketing and sales division Telkom DivreIII Jabar Banten in Competitive Intelligence section as Community Content Developer. After he graduates from a bachelor in 2010, he immediately got a scholarship from faculty to continue master studies in IT Telkom. His master's field in informatics engineering with engineering software engineering. During the master's program, he had been a lecturer in the Department of Information Management at Telkom Polytechnic for one semester simultaneously also became lecturer professional in the Information Technology Faculty of Telkom Institute of Technology for one year and even acting as Account Manager of IT Project under the Vice-Rector for Cooperation. After master graduated in 2013, to fulfill his passion as a professional in the IT field, he founded the company and became an Expert in some IT Consultant in Bandung. At the same time, the community of Saung IT also established a non-profit organization for the development of IT education in the city. Become a journalist in Pikiran Rakyat and Volunteer ICT West Java. Currently active as a Lecturer in D4 Polytechnic Studies Program Pos Indonesia.

**Muhammad Amran Hakim Siregar** is graduated majoring in informatics engineering from the college of Politeknik Pos Indonesia Bandung (2019). Now, I work in a kidney specialty hospital, Ny. RA. Habibie as head of the Information Technology Department. His research conducted in the field of Information Engineering and Logistics. In addition to being a researcher, also a provider of consulting services and web-based application creation. More information on the homepage: https://www.linkedin.com/in/muhammadamran/

**Dinda Majesty** is a student majoring in informatics engineering from the college of Politeknik Pos Indonesia Bandung. His research conducted in the field of Data Scientist and Machine Learning. She is a member of IRC (Informatics Research Center) of Applied Bachelor Program of Informatics Engineering in Politeknik Pos Indonesia.