



Institute for Software Research
University of California, Irvine

Discovering Business Processes from User Operation History



Shinobu Saito
University of California, Irvine
shinobus@uci.edu

March 2018

ISR Technical Report # UCI-ISR-18-1

Institute for Software Research
ICS2 221
University of California, Irvine
Irvine, CA 92697-3455
isr.uci.edu

isr.uci.edu/publications

Discovering Business Processes from User Operation History

Shinobu Saito ¹

Institute for Software Research, University of California, Irvine, Irvine CA, USA
shinobus@uci.edu

1 Introduction

Many organizations have begun outsourcing system maintenance to external IT providers, which is referred to as AMO (Application Maintenance Outsourcing). However, business processes, which is one of the most important knowledge needed for maintenance, are often undocumented. A common way to discover business processes is shadowing (i.e., observing & imitating computer users' operation), which is costly. Consequently, it may take dozens of new maintainers over several months to learn the business process.

Process mining is a research field focused on identifying business processes from the raw event data stored in an IT system related to the processes [1]. In process mining, correcting appropriate event data is a significant factor for successful analysis. However, the data stored in a database of an IT system tends to not be ready for process mining. Generally, it is necessary to transform original data into an event log which can then be used to discover business processes.

In this report, we introduce an approach for transforming stored data about computer users' operation history into business processes. To do so, our approach analyzes the database's table that records the history of users' operation to the IT system. We call this an operation history table. In the table, we focus on records representing a process status, which means a specific condition regarding a business process. Based on changes to the value of the process status, we extract activities which took place between two process statuses. We create flows of the activities extracted, and then transform a set of flows into a business process which we are aiming to discover. Additionally, we introduce a result of a case study. In the case, we applied our approach to an industrial IT system which is used by an incorporated shared service center. The center provides support for employees to conduct miscellaneous activities such as purchasing, contracting, and expense reimbursement.

2 Approach

We assume that a database of an IT system related to business processes that we want to discover records the history of user operations over computer screens to the system. Figure 1 describes the relation between a computer user operation and its corresponding record which goes in the operation history table. The user,

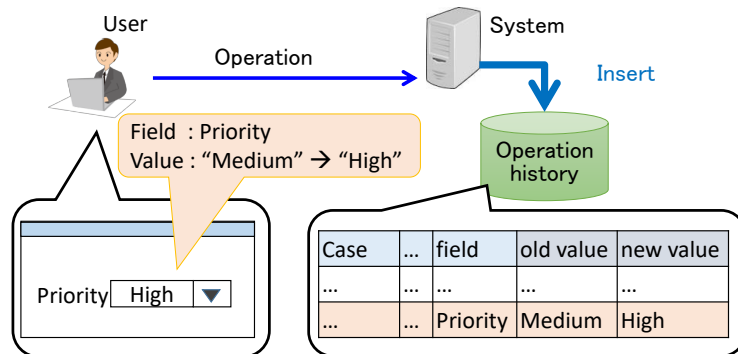


Fig. 1. Computer User Operation and its Corresponding Record in Operation History Table.

¹ ISR Technical Report UCI-ISR-18-1, March 2018

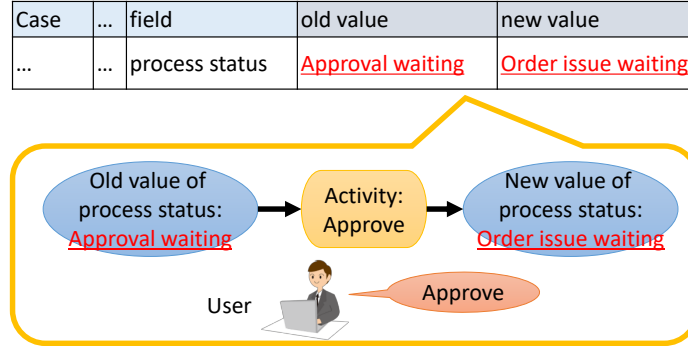


Fig. 2. Activity Extraction from transition of operation status's value.

on the left side of the figure, changes the value of the “Priority” field in the screen from “Medium” to “High”. As shown in the right side of Figure 1, the operation is recorded in the operation history table of the system's database. A new record corresponding to the operation is inserted in the table. In the newly inserted record, the values in the field, old value, and new value columns are “Priority”, “Medium”, and “High”, respectively.

Our approach is composed of three steps: select records, extract activities, and configure activity flows. To begin, we select a set of records which represent statuses of target business processes. For example, as shown in Figure 2, the value of the field column in the record is “process status”, and “Approval waiting” and “Order issue waiting” are set in the old value and new value columns, respectively. In the next step, we extract an activity from the change to the values of process status. From the values of two columns (i.e., old value and new value), we suppose what activity took place during the change of the status. The two columns show change of process status from “Approval waiting” to “Order issue waiting”. We suppose that an approval comes after “Approval waiting” status. Therefore, we can define that an “Approve” activity was conducted by a computer user. Finally, in the third step, we group activities with reference to the value of a Case column as shown in Figure 3. The case represents a particular set of steps for accomplishing a task by computer user(s). We then configure an activity flow using the activities grouped by each case. For example, Figure 3 shows three records grouped by case “1001”. As shown in the right side of the figure, we extract three activities: Approve, Issue order forms, and Send order forms. We then create an activity flow of case 1001 by connecting them in time series order. In similar way, we create one activity flow for each case. Each activity flow is regarded as a part of target business process. Therefore, we configure the business process by making a collection of all activity flows.

To conduct step 3 automatically, we implemented a tool, which we call “BP retriever” [2], to generate the visualization of the business process. The tool visualizes not only whole processes but also each activity flow. In addition, the tool counts the number of cases that are transformed to same activity flows. This helps us to recognize how often each activity flow was executed.

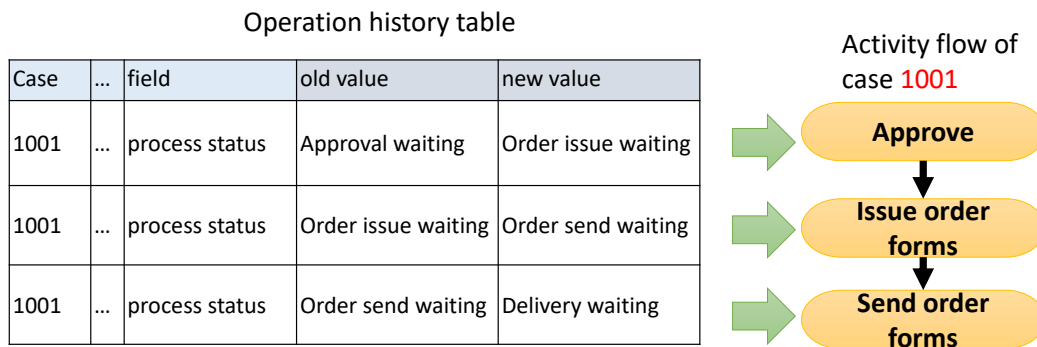


Fig. 3. Transformation of Activity Flow from Operation History Table.

3 Case Study

We discovered a business process for supporting purchasing procedures in an industry. The process is conducted by a shared service center of the industry. All members of the center use an IT system for creating and managing information needed for purchasing, such as quote forms, order forms, and invoices. The system records the history of all users' operations to the system.

In step 1, from the system, we collected 137,678 records for a one year period of 39 users' operations. Among the records, there were 1,960 representing process status. We identified 260 cases from those process status records. In the next step, we defined 19 activities from the change of the values of process status. Table 1 provides a list of activities extracted and the number of executions of those activities. In the final step, the tool identified 152 distinct activity flows, and then portrayed the whole process by collecting all (i.e., 152) activity flows as shown in Figure 4.

In addition, we conducted a frequency analysis to determine the frequency of each of the flows. The five most frequently executed activity flows are depicted in Figure 5. Note that two flows tied for fifth most frequent. We created a line graph of frequency of execution for all activity flows as shown in Figure 6. The total number of executions of the top five activity flows is 60. These activity flows cover about 23% (60/260) of all cases. On the other hand, about 45% (116/260) of the activity flows

Table 1. Activities extracted

Activities (Total number: 19)	No. of Execution
Start	322
Complete scan	137
Receive remaining products	1
Report to dept	1
Pending	34
Confirmed by ac sect	6
Confirmed by supervisor	4
Receive quote forms	71
Receive products	235
File documents	207
File order forms	2
Approved by managers	246
Inspected by dept	194
Inspected by managers	230
Receive inspection sheets	5
Receive invoices	142
Create journalizing data	8
Issue order forms	114
Scan	1
Total	1960

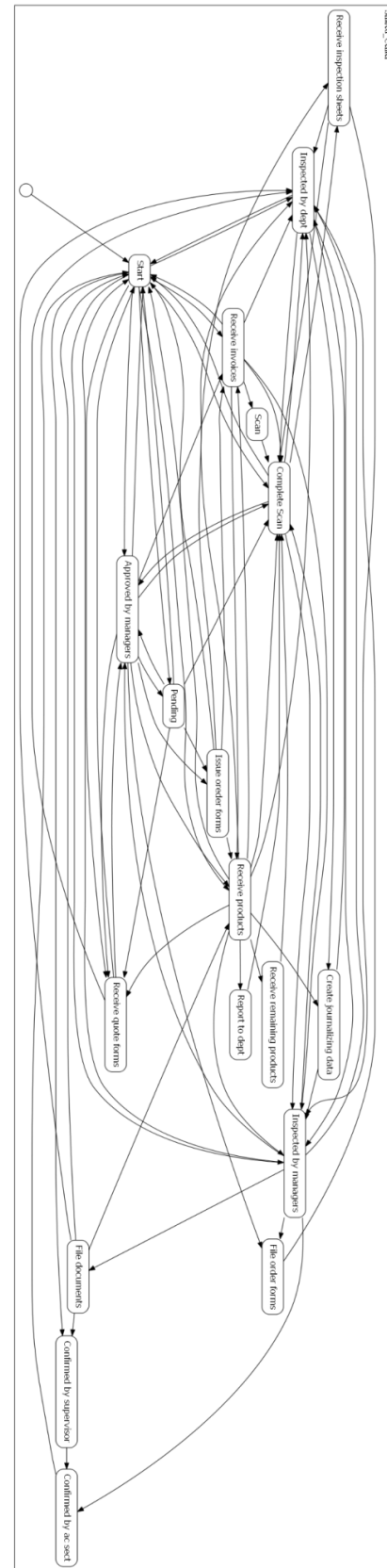


Fig. 4. Whole Business Process Generated.

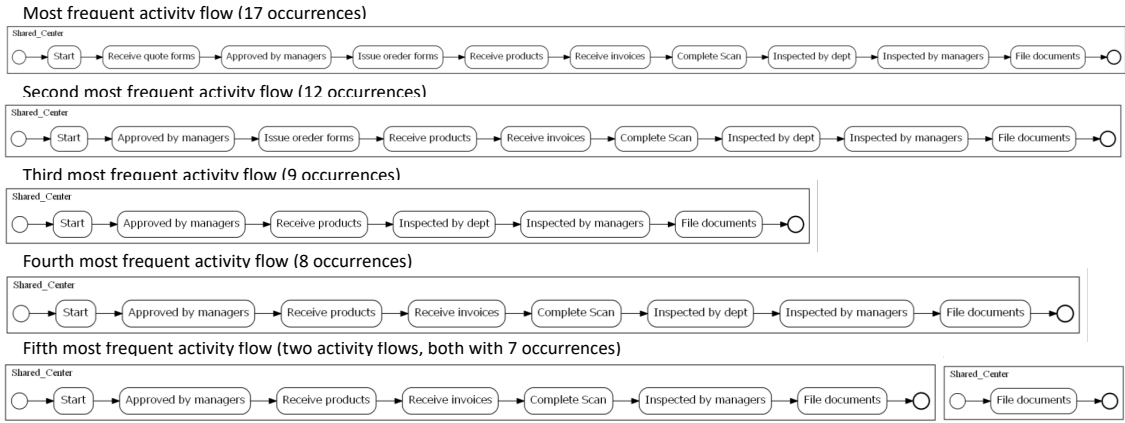


Fig. 5. Top Five Activity Flows Based on Frequency of Execution.

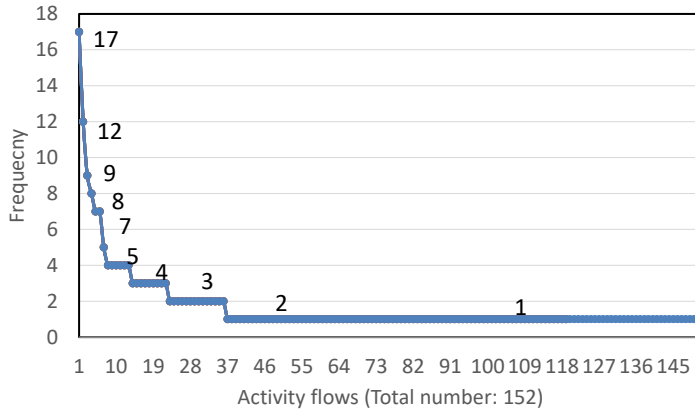


Fig. 6. Line Graph of Frequency of Execution.

were executed only once. Figure 7 depicts the business process as a collection of the five most frequent activity flows. In comparison to the 19 activities in the whole process (Figure 4), the process in Figure 7 is comprised of only ten activities, and is very helpful in comprehending the characteristics of the business process.

4 Summary

This report introduced an approach for discovering business processes from an operation history of computer users. We also conducted a case study of an industrial process for purchasing procedures, applying our approach. In the future, we plan to verify and evaluate the generated processes from the case study.

References

1. W. M. P. van der Aalst, and S. Dustdar, "Process Mining Put into Context, " IEEE Internet Computing, vol. 16, no. 1, pp. 82-86, Jan. 2012.
2. BP retriever: <https://github.com/BPretreiver/>

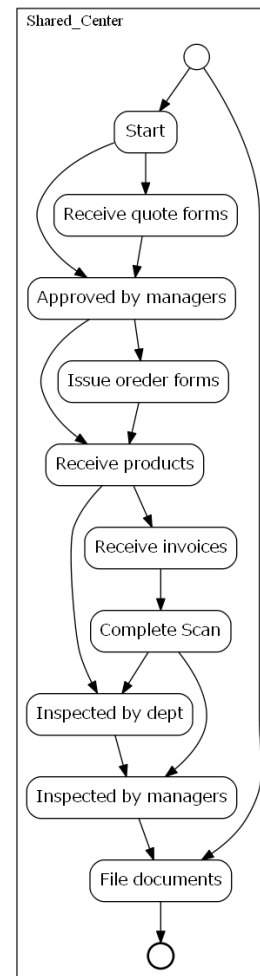


Fig. 7. Collection of Five Most Frequent Activity Flows.