Enabling the Gap Between RPA and Process Mining: User Interface Interactions Recorder

DAEHYOUN CHOI, HIND R’BIGUI, AND CHIWOON CHO

1School of Industrial Engineering, University of Ulsan, Ulsan 44610, Republic of Korea
2Digital Enterprise Department, NSOFT Company Ltd., Ulsan 44776, Republic of Korea
Corresponding author: Chiwoon Cho (chiwoon6@ulsan.ac.kr)

This work was supported by the 2022 Research Fund of University of Ulsan.

ABSTRACT Robotic process automation (RPA) is a rapid-emerging technology for process automation that is using software robots to replicate high-volume, manual, repeatable, routine, rule-based, and unmotivating human tasks. The goal is not to replace human workers but to allow them focus more on difficult tasks by delegating their tedious routine tasks to a digital force. RPA tools demonstrated powerful cost-saving and other performances. Nevertheless, one of the main challenges of implementing RPA is the identification of the suitable tasks for automation. Process mining is an emerging technology for process discovery and enhancement based on event log data. Since RPA operates on the user interface level, process mining techniques can play a huge role in deciding the tasks that can be automated. However, process mining requires an event log as input to be used. This paper presents a tool responsible for recording the interactions with user interfaces and generating a UI log that can be used by process mining techniques for deciding the tasks that can be automated with RPA.

INDEX TERMS RPA, robotic process automation, task mining, software robots, UI recorder, digital transformation.

I. INTRODUCTION

Many changes in the global market today along with the rapid development of technologies, has led to the appearance of a new trend called digital transformation today [1]. Digital transformation is defined as the process of changing existing business models as well as creating new ones by implementing today’s digital technologies in the process to meet changing business and market requirements [2]. Many industries and organizations have started taking initiatives to explore new digital technologies to crucially transform their business operations, processes, and management strategies, etc. [3]. Robotic process automation (RPA) is one of these new digital transformation technologies that are rapidly and increasingly drawing the attention of businesses. Robotic process automation tools allow mimicking human digital tasks by providing a virtual workforce in the form of a software bot that automatizes manual, high-volume, repetitive, and routine tasks [4].

Plenty of the traditional Workflow Management (WFM) systems initiatives for automation have been around for many decades but failed because the automation turned out to be too much expensive and due to the complexity of real processes [5], [6]. However, RPA turned out to be cheaper than the traditional automation solutions. RPA has lessened the threshold for process automation. Thus, the recent consideration for Robotic Process Automation has opened a new wave of automation initiatives. RPA is defined by Van der Aalst as “an umbrella term for tools that operate on the user interface of other computer systems in the same way as humans” [7]. In RPA, repetitive tasks performed by people are entrusted to software robots. The most important thing is that RPA bots do not modify or replace any pre-existing information system in the organization. They replace users by interacting with the user interfaces of the same pre-existing information system that human users were using before [5]. By automating repetitive tasks with RPA, people can focus more on difficult tasks and problem-solving. Many benefits related to RPA implementation within industries and organizations have been communicated [8]–[11]. However, the implementation of RPA is still facing many challenges as the research is still new. One of the most important challenges is the determination of business tasks that can be automated with RPA [12]. To automate user tasks with RPA, we need to
know beforehand the tasks that need to be automated. Process mining has been outlined that it can be used to identify the tasks performed by people to be automated [5]. Process mining provides a lot of techniques for process improvement that is using event data stored in today’s information systems. An event log, where each event represents a task executed either by people, a machine, or a system at a particular time; is the starting point of process mining [13]. Process mining techniques exploit these event data to illustrate how people, machines, and organizations are behaving. There are four main categories of process mining. 1) Process discovery techniques automatically discover, from real process event data, the process model which represent the real behavior of the process. 2) Conformance verification techniques consists of determining and diagnosing the deviations between a process model and reality. 3) Performance analysis techniques consists of identifying bottlenecks, reworks, wastes, etc. in the process. 4) Process reengineering techniques allows changing the existing process model. For more information on process mining refer to [13]–[20]. The previous works have proposed [4], [21] methodologies to identify candidate digital tasks for automation with RPA tools that is based on process mining techniques. R’bigui et al. [4] defined a digital task as a task that is performed by a user using a computer by interacting with different graphical user interfaces of various systems and applications. A user interface event log which corresponds to the events accruing while interacting with user interfaces of different applications or systems, is required as input to be able to identify the tasks that can be automated with RPA using process mining techniques. The proposed approaches for detecting routine tasks for automation with RPA using process mining suppose that the UI log already exist or can be recorded. However, existing recording tools do not provide data from which process mining can discover the tasks performed on user interfaces, and existing approaches such as video recording is time-consuming [29]. Thus, the adoption of process mining techniques for RPA is blocked by the absence of tools capable of recording the interactions with the user interface and generating UI logs providing enough information as input for process mining techniques to discover digital tasks that can be automated with RPA. The contribution of this work consists of presenting a tool, namely User Interface Interactions Recorder (UIIR), which fills the gap between robotic process automation and process mining.

II. RELATED WORK

A. BACKGROUND

RPA tools provides software bots that can mimic human actions performed on a computer while interacting with various user interfaces of different systems ([31], [32]). RPA is defined by IEEE Standards Association [38] as “A preconfigured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software system to deliver a result or service with human exception management.” Traditional definitions define RPA as a tool to be dedicated to being used to replace repetitive, and rule-based tasks ([33], [34]). In consequence, office employees can focus more on difficult tasks and problem solving rather than spending too much time on executing repetitive and tiring tasks. Implementing RPA allows organizations to efficaciously exploit their human resources. Thus, increasing the productivity [35]. However, RPA research is still new, and the implementation of RPA is facing some challenges ([12], [22], [36]). One of the most important challenge is to identify the tasks that need and can be automated with RPA ([12], [4]) before starting to implement RPA. A traditional way to identify the tasks that need to be automated is by knowledge by gathering information based on interviews and checklists [37]. The problem is that this task is time consuming. One way to automate this task is by using process mining. Process mining ([13], [17]) has been proposed to tackle this challenge ([4], [21], [22]). Process mining techniques are capable of extracting knowledge from event logs commonly existing in today’s information systems to discover, monitor, and enhance processes in multiple application fields [39]. Process mining is used to discover the process model of the executed tasks for different purposes. Since RPA can automated repetitive task performed using a computer system, process mining techniques can be used to discover the actions performed on this computer while interacting with different UIs. However, process mining techniques requires an event log as input to discover the executed actions. therefore, the actions performed on computer by interacting with different UIs of different systems need to be recoded and the corresponding log need to be generated.

B. LIMITATIONS OF EXISTING RECORDING TOOLS

Existing UI actions recording tools such as WinParrot, JitBit Macro Recorder, TodayDo can record only low-level actions which refer only to pixels coordinates of mouse clicks which is based on the window size and UI resolution [29]. Some tools like TodayDo and WinParrot can record information related to where the actions have been performed. Nevertheless, actions such as copying, pasting and information related to button clicks types cannot be captured. Some tools also do not record the timestamps, and none of these tools provides a log into a format supported by process mining techniques (e.g., CSV, XES). Concerning RPA tools, they provide recording function to produce the script of the bot, but the problem is that the produced log cannot not be exported nor be readable outside of RPA environment [29]. Accordingly, there is a need to develop tools that can generate UI logs that can be used by process mining techniques to support RPA implementation. Leno et al [29] introduced the first tool that produces a log containing information needed by process mining. The tool supports Excel application and chrome web browser. Our work is like the work presented in [29]. However, we developed more functions such as differentiating between different types of mouse clicks, and other advanced functions. Moreover, our tool supports not only excel and
III. USER INTERFACE INTERACTIONS RECORDING METHODOLOGY

In this section, we discuss the position of UI recorder within the framework of RPA, its architecture, the rules used for recording, and the rule used for simplifying and reducing the generated UI log.

A. USER INTERFACE INTERACTIONS RECORDER’S POSITION WITHIN RPA FRAMEWORK

Robotic process automation tools are capable of automating tasks belonging to a business process. However, employees are performing plenty of and various tasks within an organization. Not all of them need to or can be automated with RPA tools. Thus, the main question is which of the tasks performed by a user worker need to be automated to enable business growth and can be automated with RPA. A term called Robotic Process Mining (RPM) has been introduced in [22] to refer to a category of techniques that enables discovering and analyzing candidate tasks that can be automated with RPA robots from data collected during the execution of user-based tasks. RPM techniques is a subclass of process mining techniques. Process mining allows discovering processes from an event log containing a chronological order of executed tasks recorded with today’s information systems such as Enterprise Resource Planning (ERP), Business Process Management (BPM) systems, etc. While robotic process mining should allow discovering tasks from a graphical user interface log containing a chronological order of executed actions performed on user interfaces of different systems and applications. RPM techniques need to be applied before implementing RPA to identify the tasks suitable for automation with RPA. We proposed in the previous study [4] an approach for RPM consisting of four major steps.

1) UI INTERACTION RECORDING

This step consists of recording the interactions (i.e., actions), which is based on mouse clicks and the keyboard, of a human user with different applications such as web, desktop, system, application, etc. while performing their administrative tasks. This step is the major scope of this study. Fig. 1 shows the position of UI interactions recording within the pipeline of RPA framework.

2) UI LOG TRANSFORMATION AND FILTERING

This step consists of transforming the generated UI log into a log supported by process mining tools. UI log transformation rules are defined in [4]. The rules differ based on the type of the action performed. For more details on the transformation rules refer to [4]. This step also addresses the filtering of the generated UI log. While performing their tasks using a computer, employees can do other actions also which are not related to work on a computer such as opening personal email or sending SNS messages. This type of actions needs to be filtered to keep only actions relevant to work.

3) TASK DISCOVERY

Each task performed using a computer is composed of a set of performed actions. Thus, during this stage, the conducted tasks need to be identified or discovered based on the sequence or the chronological order of the performed actions. The process discovery category techniques of process mining [23]–[27] allow us to do this job based on UI log data.

4) CANDIDATE TASKS SELECTION

After discovering all tasks performed while interacting with different applications and systems, the tasks that need to
and can be automated need to be determined. The identifications of candidate tasks can be done using different methods for instance criteria-based selection such as periodicity, frequency, etc.

After candidate tasks are appropriately identified in the last step of RPM, RPA can be implemented by creating software robots in charge of executing the selected tasks. As can be seen, UI interactions recording can be considered the most important step is it is the starting point of all the pipeline of RPA Framework. Without a UI log, process mining techniques cannot be used to enable RPA.

B. USER INTERFACE INTERACTIONS RECORDER ARCHITECTURE

User Interface interaction recorder (UIIR) records the actions performed on (i) chrome web browser, (ii) windows applications such as interaction with windows folders, notepad, (iii) and Microsoft applications such as Excel, PowerPoint, and Word. We developed a plugin for recording actions conducted on the web browser as well as a windows program which records keyboard usage and mouse clicks performed on windows and Microsoft applications. Both plugin and windows program are surveying the events of the performed actions and then sending the information to the logging component for generating and updating the UI interactions log in real-time.

Before starting the recording, the target user needs to sign in into the recording tool UIIR with his id and password to differentiate between users using for example the same computer. The recording of the actions performed on web browser and those performed on windows and Microsoft applications starts automatically after sign in. The logs are stored directly to a server database as shown in Fig. 2. After stopping the recording, one integrated UI log, which integrates both the log generated from the web and the log generated from windows program, is generated. The log can be any time downloaded by user id and by date. Then the generated raw UI log is reduced and simplified with a filtering program that we developed as well. Fig. 3 shows the architecture of the recorder UIIR.

UIIR tool allows recording all the data involved in the context of each captured action. For instance, for an action performed in web browser, UIIR records the information about the URL link, the button clicked, active UI, the data entered, etc., and for an action performed in an Excel file, it captures information about the path of the spreadsheet, the active sheet, the cell and its value, the button clicked, etc. The tool generates a log in the format of Excel file which can be converted easily into and CSV file that is one of the formats required by process mining tools for further analysis.

C. USER INTERFACE INTERACTIONS RECORDING RULES

To generate UI logs that can be processed with process mining techniques to identify the digital tasks that can be/need to be automated with RPA, the recorder tool should ensure that the recorded information is suitable for this analysis. First, two main questions need to be answered when developing the user interface interactions recorder and logger: (1) What should be recorded? and (2) what should not be recorded? The tool should record only meaningful and value adding actions and data. This work presents a set of rules applied to record actions performed on different UIs.

1) R1. RECORDING SIGNIFICANT ACTIONS

To perform specific digital tasks, users perform many actions on different applications and UIs. However, while interacting with these systems and UIs, not all performed actions are significant.

There are two types of actions: actions performed with the mouse and actions performed with the keyboard. Each of these actions is performed on a specific user interface. The actions that can be performed with a mouse are moving the mouse, right click, left click, and scrolling. For instance, actions of scrolling and moving the mouse are not meaningful as they do not impact the outcome of a task. Right and left clicks can be irrelevant based on what has been clicked and based on which UI the clicking action is performed. For instance, clicking on the background of a desktop or of a website, etc. is not meaningful. Hence, this type of actions should not be captured. However, for instance, button clicks are relevant actions that need to be recorded. We define below...
an example of a set of actions that are essential and need to be captured in the log.

a: MOUSE CLICKS ACTIONS
We defined six types of mouse clicks that can be a part of a performing task and should be captured: button clicks, checkbox clicks, text field clicks, URL link clicks, selection related clicks, and general clicks (e.g., menu) actions. The recorder should differentiate between all these type of mouse clicks.

b: COPY AND PASTE ACTIONS
These two actions can be performed using only mouse clicks (i.e., left click + copy/paste button click) as well as they can be performed using the keyboard (i.e., copy = ctrl + c, paste = ctrl + v). Performing the copy action is preceded by a selection action with the mouse which allows selecting the content to be copied, and the paste action is preceded with a mouse click such as text field click which allows specifying the place where the copied content should be pasted. The actions performed by typing ctrl + c/ctrl + v in the keyboard are converted into copy and paste actions respectively, and the actions performed by left click action + copy/paste button click action are converted into one action copy and paste respectively.

2) R2. RECORDING RELEVANT DATA
Open, click button, copy etc. are the name of the recorded actions which is performed by a user. Considering only the name of the actions conducted is not enough to identify the performed task. Information related to what has been opened, what URL has been opened, which folder has been opened, what is the path of the opened folder, which excel sheet has been opened, which cell has been modified, which button has been clicked, what content has been copied and pasted, what content is entered with the keyboard, etc. is necessary to be able to extract the performed task and need to be captured. Moreover, timestamps are the time at which the actions are performed. Hence, timestamp information is also essential in order to identify the order of actions. In conclusion, besides the performed actions, the recorder should also capture the data that supports them.

3) R3. DIFFERENT APPLICATIONS ‘UIS RECORDING
A task consists of a set of actions. One task can be performed using different user interfaces such as web-based applications and systems, Microsoft applications such as Excel, word, ppt, etc., windows applications such as folders, etc. The interactions between different UIs to perform a task need to be recorded. For instance, filtering and copying data from an Excel sheet and pasting it in a web-based system such as ERP. As can be seen, this task is performed through filter, copy, and paste actions and through two user interfaces Excel, and web-based system.

4) R4. PRIVACY AWARE RECORDING
The goal of recording the interactions of users with different user interfaces of various systems and applications is to generate a log that will be analyzed to discover the tasks that has been performed and identify the ones that can be and need to be automated with RPA. The generated log will be analyzed using different techniques such as process mining by managers who will decide the tasks to be automated. Since, all interactions will be recorded, private and personal data also can be recorded. Therefore, there is a need to protect users ‘privacy. In this recorder, we took into consideration some of the privacy issues. For instance, all entered passwords are not recorded as they are, but the entered passwords are recorded with the word “password”.

D. UI LOG FILTERING – SIMPLIFICATION
Since, UIIR recorder records the performed actions in detail, the generated log needs to be simplified. Therefore, we developed another tool for filtering and simplifying the generated UI log. We define below some examples of simplification and filtering that can be performed by the filtering tool.

1) KEYBOARD ENTERING SIMPLIFICATION
Recording some content and values entering with the keyboard can also be meaningful. however, each letter or...
number entered with the keyboard is recorded as one action. For instance, the actions of opening the following URL: www.google.com by typing it with the keyboard will be indeed recorded not in one action but into 14 actions, which means 14 rows will be generated in the log (e.g., \{w\} in one row, \{.\} in one row, etc.). After filtering the raw log with the developed filtering tool, the 14 actions are simplified into one action recorded in one row where the content of the action is recorded as www.google.com.

2) MOUSE CLICKS SIMPLIFICATION AND FILTERING

Every mouse click is composed of a set of \{pressed, released\} which means two actions/two rows are recorded. When we press the mouse, it is recorded in one row and when we release the mouse, it is recorded in one row. All released clicks are deleted to keep only meaningful rows.

Moreover, the set of \{pressed, released\} can be the result of a single click as can also be the result of a selection. The only difference is the position of the press and the release. Thus, the log can be simplified with the tool such that if the position of the pressing is equivalent to the position of the release, then the action is converted to a click action and if the position of the press is different than the position of the release, then the action is converted to a selection action.

3) SIMPLIFICATION RELATED TO COPY AND PASTE ACTIONS

When copying and pasting a content with CTRL+C, and CTRL+V, the actions are recorded in two rows respectively. The tool converts the CTRL+C into “copy” recorded in one row and converts the CTRL+V into “paste” recorded in one row.
4) REDUNDANT ACTIONS FILTERING
Log reducer tool filters also redundant rows (i.e., deletes rows having identical information in all columns), double copying, copying actions without the corresponding pasting actions.

IV. RESULT AND DISCUSSION
This section presents a case study to demonstrate the User Interface Interactions Recorder (UIIR) and to demonstrate that the generated log is useful for process mining investigations.

The case study consists of (1) logging to a web-based online shopping system, (2) downloading all orders, (3) opening the excel sheet of the orders, (4) filtering the orders based on delivery completed status, (5) copying filtered orders, and (6) pasting them in a new excel sheet called completed orders. UIIR tool records the performed actions. The generated log can be downloaded by selecting the date or the period as shown in Fig. 4. After downloading the produced log, it is filtered with the simplification tool as shown in Fig. 5.

A fragment of the generated log after filtering is shown in Fig. 6. To test the produced UI log, it needs to be used as input by process discovery techniques. For this purpose, we used the Disco tool [30], a process mining tool that is based on the fuzzy algorithm [24] that allows discovering process models from an event log. In our case, the aim is to discover the performed task that consists of a set of actions performed using the keyboard and the mouse on different user interfaces of different applications.

Fig. 7 illustrates the model of the performed task that has been discovered automatically with a process discovery technique from the recorded UI log depicted in Fig. 6. The model shows the sequence of the actions performed while interacting with a web-based sales system and excel application interfaces.

Based on the result, we can see that the produced model is understandable. It shows that the performed actions are discovered in the correct chronological order. Also, the information provided in every discovered action, provides a full understanding of what action has been performed, on which
system or application the action is performed, which button has been clicked, which content has been entered with the keyboard, which folder and link address are involved, etc.

This work shows that the UI log generated by the user interface interactions recorder (UIIR) and simplified with the filter tool can successfully be used by process mining techniques to discover the actions performed while interacting with different UIs of different systems and applications.

V. CONCLUSION

RPA is a new and trending topic for automating repetitive and routine tasks. One of the most important challenges of RPA implementation is the determination of business tasks that can be automated with RPA. We need to know beforehand the tasks that need to be automated. Process mining can play a huge role in identifying the tasks performed by people to be automated. However, the adoption of process mining techniques for RPA is blocked by the absence of tools capable of recording the interactions with the user interface and generating UI logs providing enough information that can be used as input for process mining techniques to discover digital tasks that can be automated with RPA. In this work, we presented a tool that can fill the gap between robotic process automation and process mining. We defined the rules that allow capturing relevant information and data needed by process mining techniques to correctly discover the performed tasks with a user interface.

The tool has been tested and evaluated with a case study and the results show that the UI log generated by the user interface interactions recorder can successfully be used by process mining techniques to discover the actions performed while interacting with different UIs of different systems and applications. However, this topic is still new. In the future work, we will perform more further investigations and evaluations to enhance the capability of the tool in capturing meaningful interactions.

AUTHOR CONTRIBUTION

Hind R‘bigui and Daehyoun Choi contributed to the main idea and the methodology of the research. Hind R‘bigui designed the experiment, performed the simulations, and wrote the original manuscript. Hind R‘bigui contributed significantly to improving the technical and grammatical contents of the manuscript. Hind R‘bigui and Chiwoon Cho reviewed the manuscript and provided valuable suggestions to further refine the manuscript.

REFERENCES

[1] A.-M. Stjepić, L. Ivančić, and D. S. Vugec, “Mastering digital transformation through business process management: Investigating alignments, goals, orchestration, and roles,” J. Entrepreneurship, Manage., Innov., vol. 16, no. 1, pp. 41–73, 2019, doi: 10.7341/20201612.

[2] M. P. Bach, M. Spremčić, and D. S. Vugec, “Integrating digital transformation strategies into firms: Values, routes and best practice examples,” in Management and Technological Challenges in the Digital Age, P. N. Melo and C. Machado, Eds. Boca Raton FL, USA: CRC Press, 2018, pp. 119–140.

[3] C. Matt, T. Hess, and A. Benlian, “Digital transformation strategies,” Bus. Inf. Syst. Eng., vol. 57, no. 5, pp. 339–343, 2015.
DAEHYOUN Choi received the Ph.D. and master’s degrees in industrial engineering from the University of Ulsan, South Korea. He has a lot of industry experiences and is currently the CEO of NSOFT Company Ltd.

HIND R’BGUI received the Ph.D. degree in industrial engineering from the University of Ulsan, South Korea, and the State Engineering degree in industrial engineering from the National School of Applied Sciences of Fes, Morocco. She is currently working as a Consultant and Solution Architect of SIEMENS Smart Factory Solutions, NSOFT Company Ltd., South Korea. Her research interests include process mining, process modeling, BPM, robotic process automation, smart factory, and digital transformation.

CHIWOON CHO is currently a Full Professor in industrial engineering with the University of Ulsan, South Korea. He has a lot of industry experiences, including Hyundai Heavy Industries, LG CNS, and Samsung SDS.