

Looking for the Why in Event Logs for Robotic Process Automation

Antonio Martínez-Rojas

University of Seville. Computer Languages and Systems Department. E.T.S. Ingeniería Informática. Avenida Reina Mercedes, s/n, 41012, Seville, Spain

Abstract

The concept of Robotic Process Automation (RPA) has gained relevant attention in both industry and academia. RPA raises a way of automating mundane and repetitive human tasks requiring less intrusiveness with the IT infrastructure. Besides traditional user interviews and process document analysis, a common practice starts by observing the behavior of humans with the information systems while they perform the process to be automated. This sequence of human interactions with the user interface (i.e., mouse clicks and keystrokes) is stored in logs for later analysis. Analyzing these interactions brings significant benefits when conducting RPA projects. Nonetheless, some decision-based behaviors of humans require additional information to be explained. For example, a human may reject an invoice because some field is missing on a form. However, there is no interaction with that field since such information is not stored in the log. Therefore, this Ph.D. elaborates on a method to obtain additional information based on screenshots collected during the process execution. Some features are extracted from the screenshots to enrich the log, which is later used for classifying human decisions in a machine-and-human-readable form. The proposed method can be applied to generate advanced support in the RPA projects, e.g., producing an enhanced process analysis, supporting the robot development, or generating predictions and simulations. The approach has been validated using synthetic data where promising results were obtained.

Keywords

Robotic Process Automation, Process Discovery, Task mining, Decision Model Discovery

1. Research problem and motivation

In the last decade, the industry has embraced Robotic Process Automation (RPA) as a new process automation level that focuses on tackling structured and repetitive tasks quickly and efficiently. Thus, a digital workforce is enabled to mimic human employees' behavior. This approach sharply contrasts with other paradigms for process automation that consists of the orchestration of application programming interfaces (APIs) of the software [1]. In turn, RPA implies a lower level of intrusiveness since this type of software sits on top of the information technology infrastructure of a company instead of being part of such infrastructure [2, 3]. It is acknowledged that a successful RPA adoption goes

 amrojas@us.es (A. Martínez-Rojas)

 0000-0002-2782-9893 (A. Martínez-Rojas)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

beyond just simple cost savings but also contributes to improvements in terms of agility and quality [4, 5, 6].

Most RPA projects start by observing human workers performing the process, which is later automated. More precisely, terms like Robotic Process Mining [7], Task Mining [8], or Desktop Activity Mining [9] have been coined by the RPA community to exploit the UI logs, i.e., a series of timestamped event logs (e.g., mouse clicks and keystrokes) obtained by monitoring user interfaces. These methods are very convenient in helping analysts to identify candidate processes to robotize, their different variants, and their decision points efficiently [10]. However, a traditional user interface log is limited to explaining all human behavior, e.g., a decision may be motivated by a form field even though it is never directly interacted with. Therefore, human behaviors (i.e., decision points) are unexplainable by current proposals.

This problem is accentuated in the context of Business Process Outsourcing (BPO), where the processes being executed are hosted on external systems. Connections to these systems are typically made via secure connections through virtualized environments (e.g., Citrix or TeamViewer). These types of connections only allow collecting raw images of the monitored screen, i.e., screenshots, rather than the structure of the information being processed (e.g., the DOM tree of a website). This context needs some support to manage screenshots in the lifecycle that existing proposals do not cover.

Therefore, this Ph.D. intends to address these challenges based on the following premises: (1) it is possible to discover processes from a UI log, (2) it is possible to extract useful features from screenshots, and (3) it is possible to extract the reasons why decisions of the processes are made. In this context, we rely on the following **RQ** to give rise to this research: *How does image analysis improve RPA support?*.

2. Research plan and methodology

As the research question is very generic, here we specify it in 4 sub-questions:

RQ1: Are the images displayed on the screen relevant to the analysis of processes with Robotic Process Mining (RPM)?

RQ2: What alternatives exist to incorporate screen information into RPM?

RQ3: How can screen information be exploited in the early stages of the RPA life cycle?

RQ4: What effects would it have on the analysis and further stages of the RPA life cycle?

In answering these questions, this project proposal is planned following the Design Science methodology [11] and is organized into five main phases (P) subdivided into tasks (T). The methodology proposes that 3 of the 5 phases should be covered extensively (i.e., in-depth).

- P1. **Explicate Problem:** A validation is proposed to answer the RQ1, verifying that the problem is significant for the community and, therefore, an interesting

contribution considering the needs of the scientific community and the industry. A Delphi [12] study is proposed for this phase.

- **P2. Define Requirements** (in-depth): Definition of solution requirements, related to RQ3. (T.2.1) Solution requirements that encompass process discovery in image environments (i.e., processing, cleaning, and feature extraction based on a set of screenshots obtained from human monitoring). (T.2.2) The requirements of a solution that receives as input the output of the previous subtask, to be able to explain the decision points in a machine-and-human-readable form.
- **P3. Design and Develop Artefact** (in-depth): Design and development of what is defined in P2. Which is related to RQ2 and RQ3. (T.3.1) The architecture, algorithms, and technologies to be used to address the phases of the proposed method (cf. Section 3) will be defined. This involves studying tools and algorithms such as ProM, Disco (for process discovery), Canny, Sobel, Scikit-Image, Keras, Keras-OCR¹, Scikit-learn, PyTorch (for image processing), RPA-Logger², Splyrix or Splytech (for user behavior monitoring). (T.3.2) Implementation of the solution designed in T.3.1.
- **P4. Demonstrate Artefact** (in-depth): Demonstration of the artifact developed in P3 taking as reference the protocol defined in [13], widely applied in software engineering. Two experimentations are intended to be developed. (T4.1) Using synthetic data that allows to refine the artifact. (T4.2) Using real data that allows us to bring the proposal as close as possible to a final prototype. That is trying to support a solution to RQ3 and RQ4.
- **P5. Evaluate Artefact**: Validating the application of the proposal deployed in a real industrial context and analyzing the feedback from the users. The use case will be designed to be compatible with use in a BPO environment, which is the clearest example of the use of virtualized systems. That finally completes the answer to RQ3 and RQ4.

3. Approach

In this section, a method to enable advanced RPA support is described (cf. Fig. 1). This method proposes an image-based decision model discovery system for virtualized environments that offers RPA support. At a glance, the most representative phases of the approach are:

1. **Behavior monitoring** to obtain a UI log. This UI log should include a screenshot for each event, e.g., using a tool such as [14]. This phase is already being extensively covered in previous investigations [14, 15]. However, the current research would require an adaptation to capture more sources of information, e.g., images.
2. **Discover processes** from the UI log to build the process model that best represents

¹<https://github.com/faustomorales/keras-ocr>

²<https://gitlab.com/ajramirez/rpa-logger>

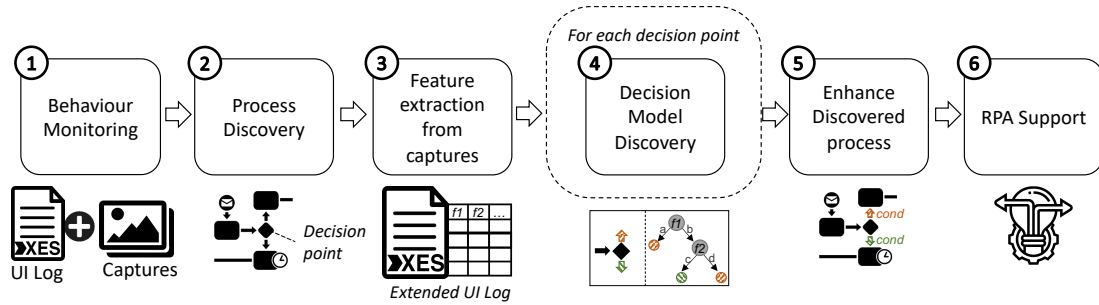


Figure 1: Proposed method for RPA support through explainable decisions from UI logs.

the captured human behavior, e.g., using [10]. The phase makes explicit decision points but lacks further information regarding how a decision is made.

Similar to the previous phase, the current state-of-the-art already provides suitable mechanisms to conduct the discovery phase [7, 10]. However, they are required to be adapted according to the extensions being performed in phase 1.

3. **Feature extraction** to transform the screenshots into objective and actionable knowledge, e.g., the presence of specific buttons or text. These features are automatically included as attributes (i.e., columns) in the events of the UI log. For this extraction, several proposals exist, such as screen scraping algorithms [16] or AI techniques, e.g., Keras-OCR. Primarily, this proposal will focus on applying neural networks following the approach of [17, 18]. We assume that this feature extraction will greatly increase the horizontal size of the UI log since a large number of additional columns will be added. Therefore, it is expected that a noise reduction technique will be necessary to discern relevant information on the screen from that which is superficial. For example, analyzing the UI designs (i.e., how the UI is constructed), the user attention (i.e., which parts of the UI are relevant for the user), or the user behavior (i.e., how the user interacts and navigates through the UIs).
4. **Discovering decision models** from the log enriched with the extracted features into a machine-and-human-readable form. The discovering process is addressed for each decision point of the process model. Herein, the extended UI log is transformed into a dataset, which is prepared to train an explainable classifier such as decision trees. Motivated by the work of [19], what is applied to traditional logs, the UI log will be converted to a dataset. To do this, each case in the UI log will generate a line in the dataset that will be labeled with the decision that is made at the decision point.
5. **Enhance discovered process** by incorporating new information into the process model. This requires the development of a new process modeling language for RPA, or the extension of an existing one, with two main objectives: (1) to offer a better understandability of the process model for the human, and (2) to use the formality

of the language to add technical information to be able to automate or systematize the RPA support tasks.

6. Provide **RPA Support** using the new modeling language. Similarly, as SmartRPA [20] does, but covering those image-based contexts where SmartRPA does not offer support. This support can be reflected in the following applications. First, in an automatic development of robots, based on the extracted process model. Second, generating predictions about the decisions that robots should make before they are made, since richer information about the process is available. Third, offering simulation scenarios extending the possibility of RPA testing automation outlined in [21]. And lastly, offering graphical support to visually represent what are the features on which the decisions of the process are based.

Although the proposed method and the application of the different techniques together represent a novelty at the research level, there are existing works related to each specific phase of this proposal. In the case of **behavioral monitoring**, there are several industrial keylogger solutions to monitor human behavior [22, 23, 24, 25]. However, they only store keystrokes and mouse clicks, in contrast to [14] keylogger that additionally extracts screenshots. In the field of image **feature extraction**, some existing proposals allow to identify and classify GUI (Graphical User Interface) components within a screenshot [26, 17]. GUI components are atomic graphical elements with predefined functionality, which are displayed within the GUI of a software application [17]. In this Ph.D. specific knowledge of these areas is applied to obtain enriched logs from processes to be automated.

Focusing on **process discovery** proposals related to this work, Agostinelli et al. [20] and Leno et al. [7] cover the complete RPA lifecycle from event capture to automatic generation of scripts for process automation and monitoring. Their way to capture data is based on an *Action Logger*, which captures only parts of the activity on the system through plugins. Thus, although they are focused on keyboard and mouse events, they also capture the DOM tree on events captured through the web browser. Unlike these approaches, the present work focuses on virtualized environments, where screenshots are the main source of information and there is no access to deeper elements such as the DOM tree. Furthermore, it focuses on the early stages of the RPA lifecycle since it is hypothesized that the more effort put into those stages, the better results will be obtained in subsequent ones.

Considering **decision model discovery** proposals, Rozinat and van der Aalst [19] use decision trees to analyze the choices made in terms of data dependencies affecting the routing of a case. However, this approach does not offer the possibility to show graphically to a non-expert user why a decision has been made. Moreover, this solution has not been validated in RPA contexts. Furthermore, Leno et al. [27] present an algorithm that generates "association rules" between the events that occurred and the results or decisions obtained. Nonetheless, the method of capturing information is based on a plugin, similarly to the aforementioned *Action Logger* which does not capture the information that the user generates outside the context of the plugin. In contrast, the present work relies on capturing the complete activity in the user interface. Thus, all interaction performed by the user is recorded to support the process discovery phase.

4. Contributions to BPM Research

This research contributes to BPM research by providing an entirely image-based approach, which provides a new source of information for the study of business processes. This consists of a more effective and comprehensive discovery of human behavior based on the extraction of features from the screenshots to enrich the UI log to be analyzed. Previously, there were decision points that were not discovered or whose reasons were wrong discovered, resulting in erroneous implementations. The latter is mitigated by applying this approach, which increases the capabilities of the analysis phase and, thus, the subsequent phases of the RPA lifecycle. Besides that, it also contributes in areas such as process mining or decision model discovery, where its application is immediate. Subsequently, this approach increases the current degree of automation, so that automatable processes that were previously not automatically discoverable, now they are.

In addition, some other areas benefit from this approach such as (1) testing of robots, (2) checking conformance of the process models to be replicated, (3) tracking and monitoring the execution of robots in production environments for the same purpose, or (4) ensuring that service level agreements are met.

5. Project status and challenges

Existing results related to this approach acknowledge its suitability for supporting the RPA lifecycle. Specifically, in [10] a method is proposed to support the analysis of human behavior in scenarios that highly depends on screen captures. Herein, an algorithm is proposed to (1) efficiently identify similar activities in a UI log based on the fingerprints of the screen captures and, (2) discover the underlying process model based on process mining and noise filtering techniques. Later on, [14] formalizes a cross-platform keylogger with a distributed architecture that can be used to generate and manage the UI logs of several workers working in the same processes. This logger addresses the needs of the first phase of the suggested method (cf. Fig. 1) while the image analysis proposal covers the second one. Different algorithms for image recognition are being evaluated in the third phase (i.e., feature extraction). Although these algorithms belong to the Machine Learning area, our initial results indicate that they are appropriate for carrying out this task. However, their suitability depends on the information in the screen capture, e.g., the layout like single or double columns, the source like a web form or PDF document, etc. In addition, we conduct the fourth phase based on previous results. More precisely, we build upon previous work in the area of *Configurable Business Process Models* [28] that generate decision trees for each decision point and, afterward, a questionnaire to help to make the decision. Currently, our research is being based on a first version of the framework that supports this method³ focused on feature extraction and decision model discovery phases [29]. Promising results are being obtained there, and they seem to be appropriate for RPA as well.

The next identified challenges are: (1) generate synthesized data respecting a given

³<https://github.com/a8081/melrpa>

process model in order to validate the proposal, (2) study the user's attention (e.g. gaze analysis) for noise reduction in the UI log, to select the relevant information from all the features extracted from screenshots, and (3) perform tests with explainable algorithms other than trees to compare the results of decision models discovery.

Lastly, these challenges aim to fully automate the RPA lifecycle using new sources of information like images [30]. This final goal is ambitious and will require a gradual increase in the organization's digital maturity, so until the point of total automation is reached, it will be necessary to consider the paradigm of human-in-the-loop [31] so that automatic techniques and human intervention coexist.

Acknowledgments

This research is part of the project PID2019-105455GB-C31 funded by MCIN/AEI/10.13039/501100011033. The author of this work is currently supported by the FPU scholarship program, granted by the Spanish Ministry of Education and Vocational Training (FPU20/05984) and by his Ph.D. supervisors, Andrés Jiménez Ramírez and José González Enríquez.

References

- [1] W. M. P. van der Aalst, M. Bichler, A. Heinzl, *Robotic Process Automation*, *Business & Information Systems Engineering* 60 (2018) 269–272. doi:10.1007/s12599-018-0542-4.
- [2] C. Frank, *Introduction To Robotic Process Automation*, *Institute for Robotic Process and Automation* (2015) 35.
- [3] L. Willcocks, M. Lacity, *A New Approach to Automating Services*, *MIT Sloan Management Review* 58 (2016) 40–49.
- [4] A. Asatiani, E. Penttinen, *Turning robotic process automation into commercial success - Case OpusCapita*, *Journal of Information Technology Teaching Cases* 6 (2016) 67–74. doi:10.1057/jittc.2016.5.
- [5] C. Capgemini, *Robotic Process Automation - Robots conquer business processes in back offices* (2017).
- [6] M. Lacity, L. Willcocks, *What Knowledge Workers Stand to Gain from Automation*, *Harvard Business Review* (2015).
- [7] V. Leno, A. Polyvyanyy, M. Dumas, M. La Rosa, F. M. Maggi, *Robotic Process Mining Vision and Challenges*, *Business & Information Systems Engineering* (2020). doi:10.1007/s12599-020-00641-4.
- [8] L. Reinkemeyer., *Process Mining in Action. Principles, Use Cases and Outlook*, Springer, 2020.
- [9] C. Linn, P. Zimmermann, D. Werth, *Desktop activity mining-a new level of detail in mining business processes*, in: *Workshops der INFORMATIK 2018-Architekturen, Prozesse, Sicherheit und Nachhaltigkeit*, Köllen Druck+ Verlag GmbH, 2018.

- [10] A. Jimenez-Ramirez, H. A. Reijers, I. Barba, C. Del Valle, A method to improve the early stages of the robotic process automation lifecycle, in: International Conference on Advanced Information Systems Engineering, Springer, 2019, pp. 446–461.
- [11] P. Johansson, E. Perjons, Design science, An introduction to Design Science. Springer 10 (2014) 978–1.
- [12] N. C. Dalkey, The Delphi method: An experimental study of group opinion, Technical Report, RAND CORP SANTA MONICA CA, 1969.
- [13] P. Brereton, B. Kitchenham, D. Budgen, Z. Li, Using a protocol template for case study planning, in: 12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12, 2008, pp. 1–8.
- [14] J. M. López-Carnicer, C. del Valle, J. G. Enríquez, Towards an opensource logger for the analysis of rpa projects, in: International Conference on Business Process Management, Springer, 2020, pp. 176–184.
- [15] V. Leno, A. Polyvyanyy, M. La Rosa, M. Dumas, F. M. Maggi, Action logger enabling process mining for robotic process automation, in: Proceedings of the Dissertation Award, Doctoral Consortium, and Demonstration Track at 17th International Conference on Business Process Management,(BPM 19), Vienna, Austria, 2019, pp. 124–128.
- [16] J. Bisbal, D. Lawless, B. Wu, J. Grimson, Legacy information systems: Issues and directions, IEEE software 16 (1999) 103–111.
- [17] K. Moran, C. Bernal-Cárdenas, M. Curcio, R. Bonett, D. Poshyvanyk, Machine learning-based prototyping of graphical user interfaces for mobile apps, IEEE Transactions on Software Engineering 46 (2018) 196–221.
- [18] Z. Feng, J. Fang, B. Cai, Y. Zhang, Guis2code: A computer vision tool to generate code automatically from graphical user interface sketches, in: International Conference on Artificial Neural Networks, Springer, 2021, pp. 53–65.
- [19] A. Rozinat, W. M. van der Aalst, Decision mining in prom, in: International conference on business process management, Springer, 2006, pp. 420–425.
- [20] S. Agostinelli, M. Lupia, A. Marrella, M. Mecella, Automated generation of executable rpa scripts from user interface logs, in: International Conference on Business Process Management, Springer, 2020, pp. 116–131.
- [21] A. Jiménez-Ramírez, J. Chacón-Montero, T. Wojdyski, J. Gonzalez Enriquez, Automated testing in robotic process automation projects, Journal of Software: Evolution and Process (2020) e2259.
- [22] Spyrix Inc, Spyrix. parental & employees monitoring software, Available at www.spyrix.com, Last accessed May 2022.
- [23] Bestxsoftware, Best free keylogger, Available at bestxsoftware.com/es, Last accessed May 2022.
- [24] Spytech Software and Design, Inc, Spytech, providing computer monitoring solutions since 1998, Available at www.spytech-web.com/spyagent.shtml, Last accessed May 2022.
- [25] Randhawa, A., Blackcat keylogger, Available at <https://github.com/jayrandhawa/Keylogger>, Last accessed May 2022.
- [26] Z. Xu, X. Baojie, W. Guoxin, Canny edge detection based on open cv, in: 2017 13th

- IEEE international conference on electronic measurement & instruments (ICEMI), IEEE, 2017, pp. 53–56.
- [27] V. Leno, A. Augusto, M. Dumas, M. La Rosa, F. M. Maggi, A. Polyvyanyy, Identifying candidate routines for robotic process automation from unsegmented ui logs, in: 2020 2nd International Conference on Process Mining (ICPM), IEEE, 2020, pp. 153–160.
- [28] A. Jiménez-Ramírez, I. Barba, B. Weber, C. Del Valle, Automatic generation of questionnaires for supporting users during the execution of declarative business process models, in: Business Information Systems, Springer International Publishing, Cham, 2014, pp. 146–158.
- [29] A. Martínez-Rojas, A. Jimenez Ramirez, J. Gonzalez Enríquez, H. Reijers, Analysing variable human actions for robotic process automation, in: International Conference on Business Process Management,(BPM 22), 2022. (In press).
- [30] A. Jiménez-Ramírez, Humans, processes and robots: a journey to hyperautomation, in: International Conference on Business Process Management, Springer, 2021, pp. 3–6.
- [31] R. C. Ruiz, A. J. Ramírez, M. J. E. Cuaresma, J. G. Enríquez, Hybridizing humans and robots: An rpa horizon envisaged from the trenches, Computers in Industry 138 (2022) 103615.