



# 로봇 프로세스 마이닝: 비전과 과제

볼로디미르 레노 • 아르템 폴리바니 • 말론 뒤마 • 마르첼로 라 로사 • 파브리치오 마리아 마기

접수: 2019년 3월 21일 / 수락: 2020년 2월 17일 / 온라인 게시: 2020년 3월 24일  
저자(들) 2020

**Abstract** 로봇 프로세스 자동화(RPA)는 웹 및 데스크탑 애플리케이션과의 세분화된 상호 작용 시퀀스를 인코딩하는 스크립트를 실행하여 조직에서 반복적인 사무 작업을 자동화할 수 있게 해주는 새로운 기술입니다. 사무 작업의 예로는 파일 열기, 웹 양식에서 필드 선택 또는 스프레드시트에서 셀 선택, 필드 또는 셀에 데이터 복사 붙여넣기 등이 있습니다.

RPA가 광범위한 루틴을 자동화할 수 있다는 점을 감안할 때 처음에 어떤 루틴을 자동화해야 하는지에 대한 질문이 제기됩니다. 이 백서는 이러한 격차를 메우기 위한 로봇 프로세스 마이닝(RPM)이라는 기술 제품군에 대한 비전을 제시합니다. RPM의 핵심 아이디어는 사용자 상호작용(UI) 로그라고도 하는 작업자와 웹 및 데스크탑 애플리케이션 간의 상호작용 로그에서 자동화가 가능한 반복적인 루틴을 발견할 수 있다는 것입니다. 이 문서는 RPM을 뒷받침하는 일련의 기본 개념을 정의하고 RPM 도구가 UI 로그에서 RPA 스크립트를 생성할 수 있도록 하는 처리 단계의 파이프라인을 제시합니다. 이 백서는 또한 계획된 파이프라인을 실현하기 위한 연구 과제에 대해 설명합니다.

Joerg Becker가 두 번 수정한 후 수락했습니다.

V. Leno (&) A. Polyvyanyy M. La Rosa University of Melbourne, Parkville, VIC 3010, Australia 이메일: [vleno@student.unimelb.edu.au](mailto:vleno@student.unimelb.edu.au)

A. Polyvyanyy 이메일: [artem.polyvyanyy@unimelb.edu.au](mailto:artem.polyvyanyy@unimelb.edu.au)  
M. 라 로사 이메일: [marcello.larosa@unimelb.edu.au](mailto:marcello.larosa@unimelb.edu.au)

M. Dumas FM Maggi University of Tartu, Liivi 2, 50409 Tartu, Estonia 이메일: [marlon.dumas@ut.ee](mailto:marlon.dumas@ut.ee)

FM 매기 이메일: [fmmaggi@ut.ee](mailto:fmmaggi@ut.ee)

키워드 로봇 프로세스 자동화 프로세스 마이닝  
로봇 프로세스 마이닝

## 1. 소개

UiPath Enterprise RPA Platform<sup>1</sup> 및 Automation Anywhere Enterprise RPA<sup>2</sup>와 같은 로보틱 프로세스 자동화(RPA) 도구를 사용하면 조직에서 웹 및 데스크탑 애플리케이션과의 세분화된 상호 작용 시퀀스를 인코딩하는 스크립트를 실행하여 반복 작업을 자동화할 수 있습니다(van der Aalst et al 2018). RPA 도구를 사용하여 자동화할 수 있는 일반적인 사무 작업은 이러한 시스템의 사용자 인터페이스를 통해 한 시스템에서 다른 시스템으로 데이터를 전송하는 것입니다. 예를 들어, 그림 1은 웹 기반 학습 정보 시스템으로 하나씩 전송해야 하는 학생 기록이 있는 스프레드시트를 보여줍니다. 이 작업에는 스프레드시트의 각 행에 대해 셀 선택, 선택한 셀의 값을 웹 양식의 해당 필드에 복사, 행이 처리된 후 양식 제출이 포함됩니다. 이와 같은 루틴은 RPA 스크립트로 인코딩되어 RPA 소프트웨어 로봇(또는 줄여서 RPA 봇)이라고도 하는 RPA 도구의 런타임 환경 인스턴스에 의해 실행될 수 있습니다.

많은 사례 연구에서 RPA 기술이

nology는 사무 업무와 관련된 비즈니스 프로세스의 효율성과 데이터 품질을 개선할 수 있습니다(Lacity 및 Willcocks 2016; Aguirre 및 Rodriguez 2017). 그러나 기존 RPA 도구는 광범위한 루틴을 자동화할 수 있지만 처음부터 어떤 루틴이 자동화 대상인지 결정할 수는 없습니다.

<sup>1</sup> <https://www.uipath.com/>.

<sup>2</sup> <https://www.automationanywhere.com/>.

	A	B	C
1	Name	Surname	Country of residence
2	John	Doe	Australia
3	Albert	Rauf	Germany
4	Steven	Richards	Australia
5	Gerard	Dubois	France
6	Audrey	Backer	USA
7	Carl	Gustafsson	Sweden
8	Sarah	Johnson	Australia
9	Andrea	Bolzano	Italy
10	Hannah	Dietmeier	Germany
11	Igor	Honchar	Ukraine
12	Oliver	Duncan	Ireland
13	Terry	Lee	Australia
14	Volodymyr	Leno	Ukraine
15	William	Macdonald	Canada
16	Jorge	Canales	Spain
17	Thomas	Taylor	Australia
18	Jack	Brown	Australia
19	Christina	Esposito	Italy
20	Amelia	Wilson	Australia

### New Record

First Name  
John

Last Name  
Doe

Country of residence  
Australia

☐ International Student

SAVE

(a) Student records spreadsheet (b) New Record creation form

그림 1 웹 양식으로 전송해야 하는 학생 데이터가 포함된 스프레드시트 추출

RPA의 후보 루틴을 식별하기 위한 현재 관행은 현장에서 또는 비디오 녹화를 사용하여 일상 업무를 수행하는 근로자에 대한 인터뷰, 워크스루 및 상세한 관찰을 통한 것입니다(Agaton 및 Swedberg 2018). 이러한 경험적 조사 방법을 통해 분석가는 자동화를 위한 후보 루틴을 식별하고 식별된 루틴 자동화의 잠재적 이점과 비용을 평가할 수 있습니다. 그러나 이러한 방법은 시간이 많이 소요되므로 루틴 수가 매우 많은 조직에서는 확장성 제한에 직면합니다.

이 포지션 백서에서는 작업자와 웹 및 데스크톱 애플리케이션 간의 상호 작용 로그에서 자동화 가능한 루틴을 발견할 수 있는 새로운 종류의 도구, 즉 로봇 프로세스 마이닝(RPM) 도구에 대한 비전을 제시합니다. 구성된 RPM 도구는 필드 또는 셀 선택, 복사 및 붙여넣기, 필드 또는 셀 편집과 같은 이벤트 레코드가 포함된 애플리케이션과의 사용자 상호작용(소위 사용자 상호작용 로그 또는 UI 로그)의 입력 로그를 취합니다. 주어진 UI 로그에서 RPM 도구는 자동화 가능한 루틴과 그 경계를 식별하고, 식별된 각 루틴의 변형을 수집하고, 식별된 변형을 표준화 및 간소화하고, 루틴의 간소화되고 표준화된 변형에 해당하는 실행 가능한 사양을 찾는 것을 목표로 합니다. 출력으로 생성된 루틴은 스크립트로 컴파일하고 RPA 도구에서 실행할 수 있는 플랫폼 독립적인 언어로 정의해야 합니다.

이러한 방식으로 RPM 도구는 분석가가 자동화를 위한 후보 루틴의 체계적인 인벤토리를 작성하는 데 도움이 됩니다. 이 입력은 루틴 수가 너무 많아 수동으로 식별할 수 없는 환경에서 유용합니다. 식별된 후보 루틴은 도메인과 함께 자동으로 파생된 속성(예: 빈도, 루틴의 단계 수, 자동화 가능성)의 조합을 사용하여 잠재적인 이점 및 자동화 비용 측면에서 분석될 것으로 예상합니다.

지식(예: 루틴 자동화의 잠재적인 재정적 이점). RPA의 후보 루틴이 선택되면 RPM은 분석가가 자동화 노력의 시작점으로 사용할 수 있는 루틴(또는 서브루틴)의 실행 가능한 사양을 생성하는 데 도움이 됩니다.

이 문서는 RPM을 뒷받침하는 일련의 개념을 정의하고 RPM 도구가 UI 로그에서 RPA 스크립트를 생성할 수 있도록 하는 처리 단계의 파이프라인을 제시합니다.

이 파이프라인을 기반으로 논문은 연구 과제를 논의하고 이러한 과제를 해결하기 위한 가능한 접근 방식을 지적합니다.

나머지 논문은 다음과 같이 구성된다. 2 장에서는 제안된 RPM 프레임워크를 제시한다. 섹션 3에서는 이 프레임워크를 실현하기 위한 과제와 지침에 대해 설명합니다. 4 절에서는 관련 분야에 대한 RPM을 지정하고, Sect. 5 결론을 도출하고 윤리적 고려 사항을 인정합니다.

## 2RPM 프레임워크

이 섹션에서는 RPM의 컨텍스트와 범위를 명확히 하고 RPM 문제를 상대적으로 독립적인 단계로 분해하는 파이프라인뿐만 아니라 RPM에 대한 개념적 프레임워크를 제안합니다.

### 2.1 맥락과 범위

RPA에 대한 여러 부분적으로 중복되는 정의는 연구 및 산업 문헌에서 찾을 수 있습니다. 예를 들어, Aguirre와 Rodriguez (2017)는 RPA를 '일상적인 작업, 구조화된 데이터 및 결정론적 결과를 포함하는 규칙 기반 비즈니스 프로세스를 자동화'하도록 설계된 소프트웨어 도구 범주로 정의합니다. 한편, van der Aalst et al. (2018)은 RPA를 '인간이 하는 방식으로 다른 컴퓨터 시스템의 사용자 인터페이스에서 작동하는 도구에 대한 포괄적인 용어'로 정의합니다. 반면 Gartner(Tornbohm 2017)는 RPA를 구조화된 데이터에 대해 [if, then, else] 문을 수행하는 도구 클래스로 정의합니다. 일반적으로 사용자 인터페이스 상호 작용의 조합을 사용하거나 API에 연결하여 클라이언트를 구동합니다. 서버, 메인프레임 또는 HTML 코드. RPA 도구는 제어 대시보드에서 스크립트를 실행하기 위해 할당된 런타임과 함께 소프트웨어 로봇을 구동하기 위해 RPA 도구 언어에 프로세스를 매핑하여 작동합니다.

위의 정의에서 세 가지 요소가 나옵니다.

첫째, RPA 도구는 규칙(예: if then-else 규칙)에 의해 구동되고 '결정론적 결과'가 있는 구조화된 데이터와 관련된 일상적인 작업을 자동화하도록 설계되었습니다.

둘째, RPA 도구는 API를 통해 액세스할 수 있는 다른 작업(어쨌든 자동화된 작업) 외에도 사용자 상호작용과 관련된 작업을 실행할 수 있습니다. 셋째, RPA에서

도구를 사용하여 스크립트를 지정하고 이러한 스크립트를 실행하는 소프트웨어 봇을 작동(즉, 제어 대시보드를 통해 실행 및 모니터링)할 수 있습니다.

이러한 요소를 종합하여 RPA를 사용자가 구조화된 데이터, 규칙, 사용자 인터페이스 상호 작용 및 API를 통해 액세스할 수 있는 작업과 관련된 결정적 루틴을 지정할 수 있도록 하는 도구 클래스로 정의합니다. 이러한 루틴은 제어 대시보드를 통해 작동되는 소프트웨어 봇에 의해 실행되는 스크립트로 인코딩됩니다.

제어 대시보드를 사용하는 방법에 따라 두 가지 RPA 사용 사례를 구분할 수 있습니다. 유인 및 무인(Tornbohm 2017). 유인 사용 사례에서 봇은 사용자에게 의해 트리거됩니다. 실행하는 동안 유인 봇은 사용자에게 데이터를 제공/수신할 수 있습니다. 또한 이러한 사용 사례에서 사용자는 봇의 스크립트를 단계별로 실행하거나 봇을 중지하거나 스크립트 실행 중에 개입할 수 있습니다. 유인 봇은 동적 입력(예: 루틴 중에 수집된 입력)이 필요한 루틴, 사람의 판단이 필요한 일부 결정 또는 확인이 필요한 경우 또는 루틴에 예기치 않은 예외가 발생할 가능성이 있고 다음을 수행하는 것이 중요한 루틴에 적합합니다. 이러한 예외를 감지합니다. 스프레드시트 형식의 송장의 데이터를 재무 시스템에 입력하는 것은 유인 RPA에 적합한 루틴의 한 예입니다. 이 설정에서 일부 유형의 오류는 재무

결과.

반면 무인 RPA 봇은 사람의 개입 없이 스크립트를 실행하고 실행 중에 입력을 받지 않습니다. 무인 RPA 봇은 모든 실행 경로(예외 포함)를 잘 이해하고 코드화할 수 있는 결정적 루틴을 실행하는 데 적합합니다. 일련의 복사-붙여넣기 작업을 통해 사용자 인터페이스를 통해 한 시스템에서 다른 시스템으로 레코드를 복사하는 것은 무인 봇이 실행할 수 있는 루틴의 한 예입니다.

위의 관점에서 RPA를 특정 유형의 프로세스 자동화 기술로 분류할 수 있습니다. 즉, BPMS(비즈니스 프로세스 관리 시스템), 문서 워크플로 시스템 및 기타 유형의 워크플로 자동화 도구(Dumas et al. 2018). 한편으로는 RPA와 다른 한편으로는 BPMS 및 워크플로 시스템 간의 주요 차이점은 RPA는 자동화된

상호 작용이 애플리케이션의 UI 또는 API와 함께 수행되는 단계(두 경우 모두 단계가 자동화됨). 대조적으로, BPMS 및 워크플로 시스템은 자동화된 작업과 수동 작업의 조합을 포함하는 프로세스를 자동화하도록 설계되었습니다. 이와 관련된

BPMS와 워크플로 시스템은 여러 유형의 참가자(예: 역할, 그룹)가 수행하는 여러 작업으로 구성된 중간 간 프로세스를 자동화하도록 설계되었습니다. 한편, RPA 도구는 개별 작업에 해당하는 더 작은 루틴을 자동화하기 위해 개발되었습니다.

프로세스 또는 정보 시스템에서 청구서 또는 학생 기록 생성과 같은 작업 내 단계까지 포함됩니다. 따라서 RPA 도구와 BPMS는 상호 보완적입니다. BPMS는 RPA 도구를 트리거하여 주어진 단계를 수행할 수 있습니다.

프로세스.

RPA 도구는 광범위한 루틴을 자동화할 수 있으므로 다음과 같은 질문이 제기됩니다. RPA를 사용하여 조직에서 어떤 루틴을 자동화할 수 있습니까?

우리는 이 질문에 답하는 RPM 도구3라는 도구 클래스를 구상하고 있습니다, 특히, 우리는 자동화를 위한 후보 루틴의 식별 및 평가 및 RPA에 의해 실행될 수 있는 루틴 사양의 발견을 지원하기 위해 사용자 주도 작업을 실행하는 동안 수집된 데이터를 분석하는 기술 및 도구 클래스로 RPM을 정의합니다. 봇. 이러한 맥락에서, 사용자 주도 작업은 사용자(예: 비즈니스 프로세스의 작업자)와 하나 이상의 소프트웨어 응용 프로그램 간의 상호 작용을 포함하는 작업입니다. 따라서 RPM 도구의 주요 데이터 소스는 UI 로그입니다.

위의 정의에 따라 RPM에서 세 가지 주요 단계를 구분합니다. (1) 하나 이상의 작업 실행에 해당하는 UI 로그를 수집하고 전처리합니다. (2) RPA에 대한 후보 루틴 식별; (3) 실행 가능한 RPA 루틴 발견

## 2.2 개념

RPM에 대한 주요 입력은 미리 기록되어야 하는 UI 로그입니다. UI 로그는 하나 이상의 응용 프로그램(웹 및 데스크톱 응용 프로그램 포함)에서 생성된 이벤트와 관련하여 단일 워크스테이션에서 단일 사용자가 수행하는 이벤트의 타임스탬프가 지정된 시퀀스입니다. 여기에서 실행 예제로 사용하는 UI 로그의 예는 표 1에 나와 있습니다.

이 예의 각 행은 하나의 이벤트에 해당합니다(예: URL "<https://www.unimelb.edu.au>"에 액세스, 버튼 "새 기록" 클릭 등). 각 이벤트는 이벤트 유형(예: 클릭 버튼, 텍스트 필드 편집), 타임스탬프 및 기타 정보(예: 버튼 레이블,

<sup>3</sup> 일부 상용 및 오픈 소스 도구 개발자는 작업 마이닝이라는 용어를 RPM을 참조하는 데 사용합니다(예: PM4Py 도구 세트 <http://pm4py>). [pads.rwth-aachen.de/task-mining/](http://pads.rwth-aachen.de/task-mining/).

<sup>4</sup> RPA 루틴이 RPA 봇을 통해 자동화되면 네 번째 단계는 이 봇을 모니터링하여 봇을 조정, 다시 구현 또는 폐기해야 할 수 있음을 알리는 이상 현상이나 성능 저하 이벤트를 감지하는 것입니다. 실용적인 관점에서 볼 때 이 단계는 수동으로 개발된 봇과 RPM 기술을 사용하여 개발된 봇 모두에 관련이 있으므로 이전 세 단계와 직교합니다. 또한 이전 작업에서는 기존 프로세스 마이닝 도구가 모니터링 목적으로 RPA 봇이 생성한 로그를 분석하는 데 적합하다는 것을 보여주었습니다(Geyer-Klingenberg et al. 2018).

표 1 UI 로그의 예

타임스탬프	이벤트 유형	원천	인수 1	인수 2	인수 3
1 2019-03-03T19:02:18	파일 열기	파일 체계	파일 이름: student_data.xls		
2 2019-03-03T19:02:23	URL 웹으로 이동		URL: "https://www.unimelb.edu.au"		
3 2019-03-03T19:02:26	클릭 버튼 웹		레이블: "새 레코드"		
4 2019-03-03T19:02:28	셀로 이동	워크시트 SheetName: Sheet1		주소: A2	값: "남자"
5 2019-03-03T19:02:31	텍스트 클릭 필드	편물	레이블: "이름"	값: ""	
6 2019-03-03T19:02:37	텍스트 필드 편집 웹		레이블: "이름"	값: "존"	
7 2019-03-03T19:02:40	URL 웹으로 이동		URL: "https://www.산만.com"		
8 2019-03-03T19:07:33	이메일 이메일 열기	고객	보낸 사람: "student@abc.com" 메시지: "코스 코디네이터님,..."		
9 2019-03-03T19:07:40	이메일 버튼 클릭	고객	레이블: "답장"		
10 2019-03-03T19:07:48	텍스트 필드 편집 이메일	고객	레이블: "메시지"	값: "학생 여러분, 귀하의 요청은 처리되었습니다"	
11 2019-03-03T19:07:50	이메일 버튼 클릭	고객	레이블: "보내기"		
12 2019-03-03T19:07:55	URL 웹으로 이동		URL: "https://www.unimelb.edu.au"		
13 2019-03-03T19:08:02	텍스트 클릭 필드	편물	레이블: "성"	값: ""	
14 2019-03-03T19:08:05	텍스트 필드 편집 웹		레이블: "성"	값: "Do3"	
15 2019-03-03T19:08:08	텍스트 클릭 필드	편물	레이블: "성"	값: "Do3"	
16 2019-03-03T19:08:12	텍스트 필드 편집 웹		레이블: "성"	값: "Doe"	
17 2019-03-03T19:08:17	텍스트 클릭 필드	편물	레이블: "국가 거주"	값: ""	
18 2019-03-03T19:08:21	텍스트 필드 편집 웹		레이블: "국가 거주"	값: "오스트레일리아"	
19 2019-03-03T19:08:28	클릭 버튼 웹		레이블: "자장"		
2019-03-20 03T19:08:35	클릭 버튼 웹		레이블: "새 레코드"		
21 2019-03-03T19:08:38	셀로 이동	워크시트 SheetName: Sheet1		주소: A3	값: '앨버트'
22 2019-03-03T19:08:39	복사	워크시트 내용: "앨버트"			
23 2019-03-03T19:08:40	복사	워크시트 내용: "앨버트"			
24 2019-03-03T19:08:42	텍스트 클릭 필드	편물	레이블: "이름"	값: ""	
25 2019-03-03T19:08:43	반죽	편물	값: "앨버트"		
26 2019-03-03T19:08:44	텍스트 필드 편집 웹		레이블: "이름"	값: "앨버트"	
27 2019-03-03T19:08:47	셀로 이동	워크시트 SheetName: Sheet1		주소: B3	값: '라우프'

표 1 계속

타임스탬프	이벤트 유형	원천	인수 1	인수 2	인수 3
2019-03-28 03T19:08:49	복사	워크시트 내용: "라우프"			
2019-03-29 03T19:08:52	텍스트 클릭 필드	편물	레이블: "성"	값: ""	
30 2019-03-03T19:08:53	반죽	편물	값: "라우프"		
31 2019-03-03T19:08:54	텍스트 필드 편집 웹		레이블: "성"	값: "라우프"	
32 2019-03-03T19:08:58	셀로 이동	워크시트 SheetName: Sheet1		주소: C3	값: "독일"
33 2019-03-03T19:09:01	복사	작업 시트 내용: "독일"			
34 2019-03-03T19:09:03	텍스트 클릭 필드	편물	레이블: "국가 거주"	값: ""	
35 2019-03-03T19:09:04	반죽	편물	값: "독일"		
36 2019-03-03T19:09:05	텍스트 필드 편집 웹		레이블: "국가 거주"	값: "독일"	
37 2019-03-03T19:09:09	체크박스	편물	레이블: "인터네셔널 학생"		
38 2019-03-03T19:09:14	클릭 버튼 웹		레이블: "저장"		
...	...	...	...	...	...

셀의 값 등), 페이로드라고 하는 충분한 수행된 활동을 재구성합니다. 예를 들어 버튼 클릭을 참조하는 이벤트, 저장하는 것이 중요합니다. 이 버튼의 고유 식별자(예: 요소 식별자 또는 페이지에서 고유한 경우 이름). 마찬가지로 필드 편집을 참조하는 이벤트의 경우 식별자 해당 필드에 할당된 새 값뿐만 아니라 필수 속성. 같은 유형의 이벤트는 일반적으로 페이로드에서 동일한 속성을 특징으로 합니다. 일반적으로 다른 소스 응용 프로그램에서 기록된 이벤트 페이로드에 다른 속성을 포함합니다. 예를 들어, 스프레드시트(예: Excel 스프레드시트)에 의해 생성된 이벤트에는 스프레드시트 이름 및 관련된 셀 또는 셀 범위의 위치인 반면 웹 기반 이벤트는 해당 웹에 의해 특성화됩니다. 관련된 HTML 요소의 페이지, 이름 및/또는 식별자. UI 로그의 이벤트는 시간순으로 정렬됩니다. 그들의 타임 스탬프에. 일부 이벤트는 다음으로 집계될 수 있습니다. 더 높은 수준의 행동. 예를 들어 두 개의 이벤트 셀로 이동 및 셀 내용 복사는 셀을 복사합니다. UI 로그를 얻기 위해 관련된 모든 사용자 상호 작용 특정 작업을 기록해야 합니다. 이 녹음 절차는 여러 세션을 포함하여 장기 실행될 수 있습니다. 사용자가 여러 인스턴스를 수행하는 경우 작업 시간

작업을 차례로 수행합니다. 그러한 세션 동안, 작업자는 여러 가지 작업을 수행해야 합니다. 같거나 다른 유형. 실행으로 사용된 UI 로깅 예제에 해당하는 작업의 실행을 설명합니다. 스프레드시트에서 웹으로 학생 데이터 전송 연구 정보 시스템의 한 형태. 웹 양식 학생의 이름, 성 등의 정보가 필요합니다. 이름과 거주 국가. 거주 국가의 경우 호주가 아닌 경우 사용자는 한 단계를 더 수행해야 합니다. 학생으로 등록해야 함을 나타냅니다. 국제 학생. 작업의 각 실행은 작업 추적으로 표시됩니다. 실행 중인 예제에는 다음에 속하는 두 개의 추적이 있습니다. 새 레코드 생성 작업. 로그에서 우리는 그것을 볼 수 있습니다 사용자는 두 번에 새 레코드 생성을 수행했습니다. 다른 방법들. 첫 번째 경우에는 양식을 작성했습니다. 수동으로, 두 번째 경우에는 데이터를 복사했습니다. 워크시트에서 해당 항목에 붙여넣습니다. 필드. 작업 추적 모음이 주어지면 RPM의 목표는 수행할 수 있는 반복적인 일련의 작업을 식별합니다. 여기에서는 루틴이라고 하는 여러 작업 추적에서 관찰되며, 자동화가 가능한 루틴을 식별합니다. 각각 이러한 루틴에서 RPM은 실행 가능한 사양(여기서는 루틴 사양이라고 함)을 추출하는 것을 목표로 합니다.

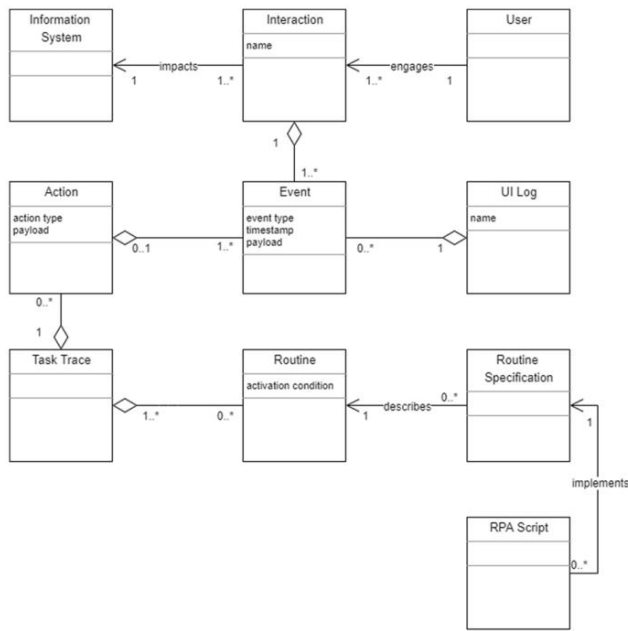


그림 2 RPM 개념의 클래스 다이어그램

이 루틴 사양은 초기에 플랫폼 독립적인 방식으로 캡처된 다음 특정 RPA 도구에서 실행될 플랫폼 종속적 RPA 스크립트로 컴파일될 수 있습니다.

요약하면 그림 2 는 클래스 다이어그램 캡처를 나타냅니다 . 위의 개념과 그 관계.

### 2.3RPM 파이프라인

앞서 언급했듯이 RPM의 세 가지 주요 단계는 다음과 같습니다. (1) UI 로그 수집 및 사전 처리; (2) 후보 루틴 식별; (3) 실행 가능한 일상적인 발견.

RPM의 목표를 달성하는 데 필요한 단계에 대한 보다 자세한 보기를 제공하기 위해 첫 번째 단계를 기록 단계 자체와 세 가지 사전 처리 단계, 즉 관련 없는 이벤트 제거(노이즈 필터링), 로그 분할로 분해합니다. 루틴 추적으로 변환하고 결과 루틴 추적을 단순화합니다. 그런 다음 두 번째 단계를 단일 단계로 매핑하고 세 번째 단계를 두 단계로 분해합니다. 플랫폼 독립 루틴 사양의 발견과 후자를 플랫폼별 사양(스크립트)으로 컴파일합니다. 세 단계를 단계로 분해하는 과정은 그림 3에 나와 있는 RPM 파이프라인에 요약되어 있습니다. 다음에서 이 파이프라인의 각 단계에 대해 설명합니다.

UI 로그 기록에는 양식에서 필드 선택, 필드 편집, 데스크톱 응용 프로그램 열기 또는 웹 페이지 열기와 같은 낮은 수준의 UI 이벤트 캡처가 포함됩니다. UI 로그 기록은 플러그인 또는 확장 메커니즘을 통해 작업자가 사용하는 소프트웨어 애플리케이션(웹 브라우저 포함)을 계속하여 달성할 수 있습니다. 로그

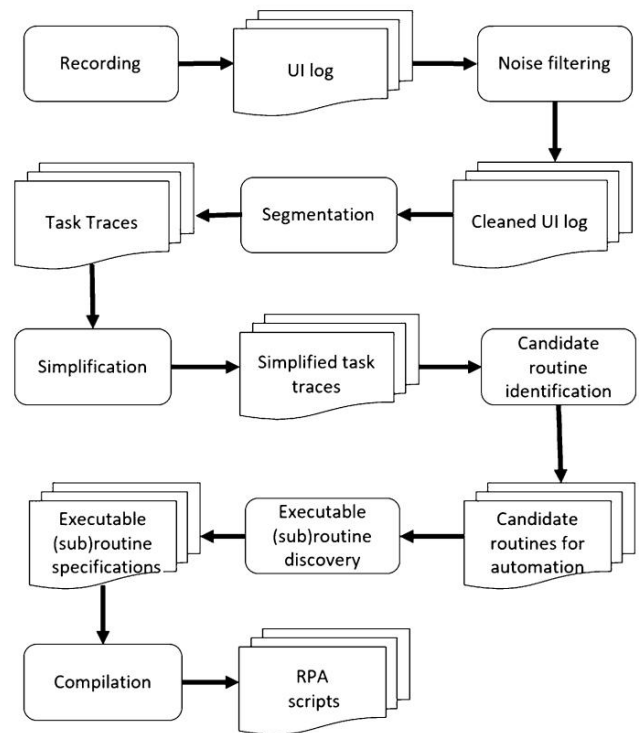


그림 3 RPM 파이프라인

이러한 플러그인 또는 확장에 의해 수집된 정보는 일정 기간 동안 사용자가 하나 이상의 작업을 실행한 것에 해당하는 원시 UI 로그를 생성하기 위해 병합될 수 있습니다. 이 원시 로그는 일반적으로 RPM에 적합하도록 사전 처리를 거쳐야 합니다.

도 2 에 도시된 바와 같이 , UI 로그는 어떠한 동작의 실행에 속하지 않는 이벤트를 포함할 수 있으며, 여기서는 노이즈라고 한다. 예를 들어 사용자가 작업 실행 중에 방해받거나 주의를 산만하여 문제의 작업과 관련이 없는 활동을 수행할 때 소음이 발생할 수 있습니다(예: 이메일에 답장하기 위해 학생 기록 전송 일시 중지). 따라서 파이프라인의 첫 번째 단계(기록 단계 후)는 작업에 속하지 않는 이벤트(노이즈 필터링)를 식별하고 필터링하는 데 전념하므로 자동화해서는 안 됩니다. 실행 중인 예제에서 이벤트 7(<https://www.distracton.com> 방문)과 이벤트 8-11(이메일에 회신)은 노이즈의 예입니다.

노이즈 필터링된 UI 로그가 주어지면 다음 문제는 작업 추적의 경계를 식별하는 것입니다. 우리는 이 문제를 세분화라고 부릅니다. 특히, 분할의 목적은 작업 실행을 나타내는 연속적인 동작의 시퀀스를 식별하는 것입니다. 분할 입력은 단일 이벤트 시퀀스를 포함하는 UI 로그이고 출력은 각각 작업 실행을 나타내는 추적 집합입니다. 노이즈 필터링과 분할이 얹혀 있음을 관찰합니다. 작업 추적의 경계를 식별하여 어떤 이벤트가

작업의 일부가 아니므로 소음을 나타냅니다. 분할은 여러 가지 방법으로 수행할 수 있습니다. 예를 들어 도메인 지식을 사용하여 수행하거나 UI 로그를 엔터프라이즈 시스템에서 기록된 트랜잭션 데이터와 결합하여 작업의 시작 및 종료 이벤트를 식별할 수 있습니다(Linn et al. 2018).

작업 추적에는 최종 결과에 영향을 주지 않는 이벤트가 포함될 수 있습니다. 그러한 사건은 낭비입니다. 을 위한

예를 들어, 작업 추적에는 중복 이벤트가 포함될 수 있습니다(예: 동일한 필드에서 Ctrl-C를 연속으로 두 번 누르면 번만 수행하는 것과 같은 효과가 있음). 또 다른 유형의 낭비는 결합과 관련이 있습니다. 예를 들어 텍스트 필드에 입력한 다음 필드의 내용을 삭제하고 다른 것을 입력하는 것입니다. 실행 중인 예제에서 이벤트 13, 14 및 22는 과잉 처리 낭비를 나타냅니다. 따라서 파이프라인에는 폐기물 식별 및 제거를 목표로 하는 단순화 단계가 포함됩니다. 단순화 단계에는 이벤트를 상위 수준 작업으로 집계하는 작업이 포함됩니다. 이러한 방식으로 작업 추적은 훨씬 더 간결하고 간결하여 대상 언어로 번역하기가 더 쉽습니다.

단순화된 작업 추적 세트가 주어지면 다음 단계는 자동화를 위한 후보 루틴을 식별하는 것입니다. 이 단계는 여러 작업 추적(일명 루틴)에서 발생하는 반복적인 작업 시퀀스를 추출하고 자동화가 가능한 작업을 식별하는 것을 목표로 합니다. 이 단계의 결과는 자동화 가능성에 따라 순위가 매겨진 자동화 또는 반자동화 루틴 세트입니다(예: 실행 빈도 및 길이에 따라).

자동화를 위한 후보 루틴이 식별된 후 다음 단계는 실행 가능한 (서브) 루틴 검색입니다. 각 후보 루틴에 대해 이 단계는 루틴의 인스턴스가 트리거되어야 하는 시기를 나타내는 활성화 조건(표 1의 이벤트 3 및 20)과 해당 루틴 내에서 수행되어야 하는 작업을 지정하는 루틴 사양을 식별합니다.

실행 가능한 (서브)루틴 검색 단계는 플랫폼 독립적인 루틴 표현으로 이어지며, 그런 다음 최종 컴파일 단계를 통해 특정 RPA 도구를 대상으로 하는 스크립트로 컴파일될 수 있습니다. 이 단계에서는 루틴 사양의 작업을 대상 RPA 도구의 스크립팅 언어로 된 명령으로 매핑하여 실행 가능한 스크립트를 생성합니다.

생성된 봇은 유인 또는 무인 설정에서 실행할 수 있습니다. 유인 설정에서 루틴 사양에서 추출된 활성화 조건이 주어지면 조건이 충족될 때 루틴을 수행하기 위한 "준비"에 대해 사용자에게 알릴 수 있습니다. 실행 중에 일시 중지할 수 있으므로 필요한 경우 사용자가 약간의 수정을 수행한 다음 작업을 재개할 수 있습니다. 무인 설정에서 봇은 사람의 개입 없이 독립적으로 작동합니다.

실행 중인 이 RPM 파이프라인을 시연해 보겠습니다.  
예(표 1):

노이즈 필터링. 이벤트 e7, e8, e9, e10 및 e11은 노이즈입니다. 로그에서 필터링해야 합니다.

분할. 실행 중인 예제에서 캡처한 작업의 주요 목표는 학생의 새 레코드를 만드는 것입니다. 따라서 작업 추적의 종료 이벤트는 "저장" 버튼을 클릭한 결과로 이러한 레코드가 실제로 생성되는 것입니다. 따라서 두 가지 작업 추적이 있습니다.

- 추적 1: e1, e2, e3, e4, e5, e6, e12, e13, e14, e15, e16, e17, e18, e19;
- 추적 2: e20, e21, e22, e23, e24, e25, e26, e27, e28, e29, e30, e31, e32, e33, e34, e35, e36, e37, e38;

단순화. 추적 1의 이벤트 e13 및 e14와 추적 2의 이벤트 e22는 낭비이므로 제거해야 합니다. 세 가지 가능한 이벤트 병합이 있습니다.

- 이벤트 {e5, e6}, {e15, e16} 및 {e17, e18}은 페이로드  $p = \{\text{Label, Value}\}$ 인 텍스트 필드에 쓰기 작업으로 병합될 수 있습니다.

- 이벤트 {e24, e25, e26}, {e29, e30, e31} 및 {e34, e35, e36}은 페이로드  $p = \{\text{Label, Value}\}$ 인 텍스트 필드에 붙여넣기 작업으로 병합될 수 있습니다.

- 이벤트 {e21, e23}, {e27, e28} 및 {e32, e33}은 페이로드  $p = \{\text{SheetName, Address, Content}\}$ 인 복사 셀로 병합될 수 있습니다.

후보자 이상 식별. 웹 양식 필드 수정과 관련된 작업은 두 가지 모두에서 발생합니다.

흔적. 따라서 해당하는 일련의 작업이 루틴을 구성합니다. 추적 1에는 몇 가지 작업이 포함되어 있습니다.

자동화할 수 없는 반면(사용자가 수동으로 양식 작성) Trace 2는 자동화 가능한 작업으로만 구성됩니다.

실행 가능한 (서브)루틴 발견. 추출된 루틴의 활성화 조건은 Click 버튼 "New Record"(실행 예제의 e3 및 e20)입니다. 그림 4는 New Record Creation 루틴 사양을 나타냅니다.

편집. 그런 다음 루틴 사양이 RPA 스크립트로 컴파일됩니다. 여기에서 사양 모델의 각 단계는 대상 RPA 도구의 언어로 특정 명령으로 "번역"됩니다. 그림 5는 발견된 루틴 스펙에서 생성된 스크립트의 예를 제공합니다.

### 3 과제 및 지침

그림 3에 제시된 RPM 파이프라인의 각 단계는 연구 과제를 야기합니다. 다음으로 이러한 과제 중 일부에 대한 개요를 제공하고 이를 해결하기 위한 접근 방식을 제안합니다.

녹음. 이 단계의 주요 과제는 기록해야 하는 작업을 식별하는 것입니다. 동일한 동작(예: 마우스 클릭)이 중요하거나 관련이 없을 수 있습니다.



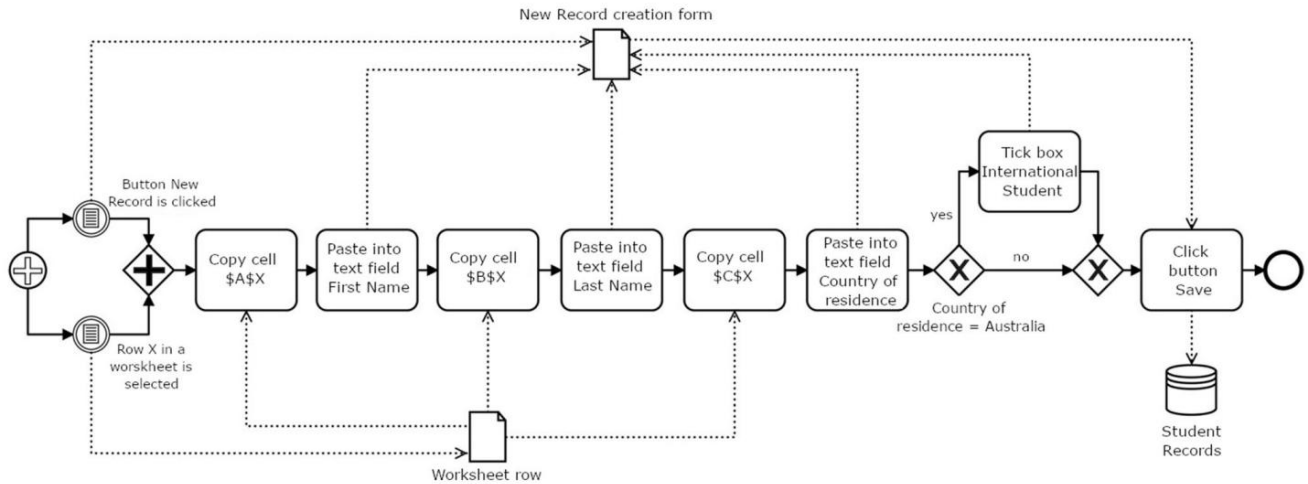


그림 4 New Record Creation 루틴 사양

```

1  Assign value of selected row to variable "$X$"
2  Excel: Get value of cell "A$X$" and assign to
   variable "$Clipboard$"
3  Assign value of Clipboard to variable "$FirstName.value$"
   in "http://www.unimelb.edu.au"
4  Excel: Get value of cell "B$X$" and assign to
   variable "$Clipboard$"
5  Assign value of Clipboard to variable "$LastName.value$"
   in "http://www.unimelb.edu.au"
6  Excel: Get value of cell "C$X$" and assign to
   variable "$Clipboard$"
7  Assign value of Clipboard to variable "$CountryOfResidence.value$"
   in "http://www.unimelb.edu.au"
8  If $CountryOfResidence.value$ Not Equal To (<>) "Australia" Then
9    Mouse Click: Left Button on "International Student"
   in "http://www.unimelb.edu.au"
10 End If
11 Mouse Click: Left Button on "Save"
   in "http://www.unimelb.edu.au"

```

그림 5 새 레코드 생성 스크립트

주어진 맥락. 예를 들어, 버튼에 대한 마우스 클릭은 중요한 이벤트이지만 웹 페이지의 배경에 대한 마우스 클릭은 관련이 없는 이벤트입니다. 또한 작업자가 웹 양식을 선택할 때 논리적 입력 요소(예: 필드) 수준에서 루틴을 학습하기 위해 웹 페이지 수준(문서 개체 모델 - DOM)에서 이벤트를 기록해야 합니다. 화면 해상도, 창 크기 등에 따라 달라지는 픽셀 좌표 수준

JitBit Macro Recorder,<sup>5</sup> TinyTask,<sup>6</sup> 및 WinParrot<sup>7</sup> 과 같은 기존 UI 이벤트 기록 도구는 픽셀 좌표를 참조하여 사용자가 수행한 모든 작업을 너무 낮은 수준의 세분성으로 저장합니다(예: 좌표 748,365에서 마우스 클릭). 결과적으로 이러한 도구에 의해 생성된 UI 상호 작용 로그는 유용한 루틴을 추출하는 데 적합하지 않습니다. RPA 도구(예: UiPath Enterprise RPA Platform 및 Automation Anywhere Enterprise RPA)는 기록 기능을 제공합니다. 그러나 이 기능은 RPA 스크립트를 기록하기 위한 것입니다. 이러한 도구는 캡처하지 않습니다

이러한 값은 RPA 스크립트 생성과 관련이 없기 때문에 다른 필드의 값에 대한 세부 정보입니다. 따라서 RPM에 필요한 UI 로그를 기록하려면 새로운 기록 도구 제품군이 필요합니다.

최근 연구에서 Leno et al. (2019)은 UI 로그를 RPM에 적합한 형식으로 기록하는 도구를 도입했습니다. 이 도구는 UI 작업(필드 선택, 필드 편집, 클립보드에 복사 또는 붙여넣기)뿐만 아니라 이러한 작업과 관련된 값(예: 편집 이벤트 후 필드 값)도 기록합니다. 이 도구는 MS Excel 및 Google 크롬을 지원합니다. 또한 이 도구는 중복 이벤트(예: 붙여넣기 없이 이중 복사, 내용 수정 또는 복사 없이 Excel에서 셀 간 탐색)를 제거하여 기록된 UI 로그를 단순화합니다. 그러나 이러한 도구의 적용 가능성은 UI 이벤트를 수신하고 이러한 이벤트에 의해 소비 및 생성된 데이터에 액세스하기 위한 API를 제공하는 데스크톱 애플리케이션으로 제한됩니다. 보다 일반적인 솔루션을 개발하려면 Ramirez et al.에 요약된 대로 애플리케이션 스크린샷에서 UI 이벤트 및 관련 데이터를 감지하기 위해 이 후자의 접근 방식을 OCR 기술과 결합해야 할 수 있습니다.

(2019); Linn et al. (2018).

노이즈 필터링. 이 단계의 주요 과제 중 하나는 작업에 기여하는 이벤트에서 노이즈를 분리하는 것입니다. 가능한 솔루션은 프로세스 실행 중 어디에서나 발생할 수 있는 혼란스러운 이벤트로 노이즈를 처리하는 것입니다. 이러한 혼란스러운 이벤트를 필터링하는 기술은 Tax et al. (2019). 그러나 노이즈가 작업의 특정 상태 또는 상태 집합으로 끌리는 경우(예: 작업 시작 또는 끝 부분), 위에서 언급한 것과 같은 기술이 이를 발견하지 못하여 결과적으로 필터링하지 못할 수 있습니다. 또한 일부 이벤트는 동일한 작업을 수행할 수 있는 다른 방식으로 인해 잘못 제거될 수 있으며 혼란스러운 이벤트 시퀀스로 잘못 나타날 수 있는 것을 유도할 수 있습니다. 이것은 다음과 같이 피할 수 있습니다.

<sup>5</sup> <https://www.jitbit.com/macro-recorder/>.

<sup>6</sup> <https://www.tinytask.net/>. <http://>

<sup>7</sup> [www.winparrot.com/](http://www.winparrot.com/).



프로세스의 데이터 관점, 즉 작업 및 이벤트에 의해 조작되는 데이터 개체의 값을 고려합니다.

데이터 개체를 보면 동일한 속성 값을 공유하거나(예: 워크시트에서 값을 복사한 다음 웹 양식에 붙여넣기) 동일한 소스/원점(예: 모든 동일한 웹 사이트에서 수행된 작업). 데이터 속성 및/또는 값을 공유하지 않거나 일반적인지 않은 소스에서 발생하는 이벤트는 노이즈를 구성할 가능성이 높습니다.

분할. 원시 형식의 UI 로그는 세션 중에 기록된 하나의 단일 이벤트 시퀀스로 구성됩니다. 이 세션 동안 사용자는 하나 또는 여러 작업의 여러 실행을 수행했을 수 있습니다. 즉, UI 로그에는 사용자에 의한 특정 실행 순서를 반영하는 어떤 순서로 작업과 이벤트가 혼합된 여러 작업에 대한 정보가 포함될 수 있습니다. 또한, 예를 들어 다른 워크스테이션에서 작업하는 여러 사용자가 작업을 수행하는 경우 동일한 작업이 여러 로그에 '확산'될 수 있습니다. 따라서 자동화를 위한 후보 루틴을 식별하기 전에 UI 로그를 추적으로 분할하여 각 추적이 작업의 하나의 실행에 해당하도록 해야 합니다.

일부 시나리오에서는 Linn et al.에서 제안한 것처럼 기업 정보 시스템에 의해 기록된 트랜잭션 데이터와 UI 로그를 결합하여 분할을 수행할 수 있습니다. (2018). 예를 들어 실행 중인 예제에서 '저장' 버튼을 누른 후 현재 작업 추적의 끝점을 표시하는 레코드 만들기 이벤트가 생성될 수 있습니다. 그러나 이 접근 방식의 문제는 이러한 트랜잭션 데이터가 작업에 대한 제한된 정보만 제공할 수 있다는 것입니다.

RPM의 세분화 문제는 웹 세션 마이닝의 문제와 유사합니다. 웹 로그 마이닝 분야에서 널리 연구되었습니다(Liu 2007). 여기서 입력은 클릭 스트림 집합이고 목표는 사용자가 참여하는 세션을 추출하는 것입니다. 목표를 달성하기 위한 웹 애플리케이션. 세션 식별에 대한 대부분의 기존 접근 방식은 RPM에도 사용할 수 있습니다. 그러나 웹 조직 특성을 기반으로 하기 때문에 웹 상호 작용 컨텍스트에서만 사용할 수 있습니다. 예를 들어, 그들이 사용하는 핵심 개념 중 하나는 동일한 세션의 이전 페이지에서 페이지에 도달해야 한다는 것입니다. 따라서 RPM 분할의 문제 중 하나는 작업이 일반적으로 서로 다른 시스템과 응용 프로그램에서 수행되며 웹 브라우저는 이러한 응용 프로그램 중 하나일 뿐입니다. 다른 접근 방식은 시간 기반 휴리스틱을 사용하여 세션의 총 지속 시간 또는 최대

두 이벤트 사이에 허용된 시간 차이. 그러나 이 접근 방식은 사용자가 작업을 수행할 때 다른 활동에 포함될 수 있으므로 신뢰할 수 없습니다. 또한 작업은 일반적으로 일괄 처리로 수행되며, 이는 올바른 작업에 시간 기반 휴리스틱을 사용하는 데 어려움을 증가시킵니다.

작업의 경계 식별. 예를 들어 스프레드시트에서 데이터를 복사하여 웹 양식을 채우는 작업을 가정해 보겠습니다. 스프레드시트의 각 행에 대해 사용자는 새 양식을 만들고 해당 행의 셀에서 필요한 데이터를 복사하여 해당 텍스트 필드에 붙여넣은 다음 제출 버튼을 누르고 작업을 다시 시작합니다. 이 예에서 두 개의 다른 작업 사이의 시간 차이는 시간 차이보다 작을 수 있습니다.

동일한 작업의 이벤트 간에 발생하여 잘못된 분할로 이어집니다.

UI 로그 분할의 문제는 프로세스 마이닝에 사용되는 이벤트 로그에서 상관이 없는 이벤트의 상관 관계와도 관련이 있습니다(Bayomie et al. 2019, 2016; Ferreira and Gillblad 2009). 그러나 이 문제는 제한적인 설정으로 광고되었습니다. 특히, Ferreira와 Gillblad (2009)는 프로세스(우리의 경우 루틴)에 주기/반복이 없을 때 문제를 해결한 반면(Bayomie et al. 2016, 2019) 프로세스 모델이 입력으로 제공된다고 가정합니다. 루틴 사양이 알려져 있음을 의미합니다. 또한 Ferreira and Gillblad (2009) 및 Bayomie et al. (2016)은 다소 부정확한 결과를 생성하는 것으로 나타났지만 RPM은 루틴을 부정확하게 복제하면 특히 무인 봇이 사용되는 상황에서 많은 비용이 드는 오류를 초래할 수 있다는 점을 감안할 때 높은 수준의 신뢰도로 루틴을 식별하려고 합니다.

단순화. 이벤트가 작업에 속하더라도 여전히 중복될 수 있습니다. 예를 들어, 사용자가 텍스트 필드를 실제로 채운 다음 다시 채워야 하는 경우입니다. 이 경우 텍스트 필드를 처음 채운 시점에 속하는 이벤트는 중복됩니다. 컨텍스트에 따라 동일한 이벤트가 루틴의 필수 부분이거나 중복될 수 있습니다. 따라서 (Conforti et al. 2017)과 같은 기존 주파수 기반 필터링 접근 방식은 이 문제를 해결하는 데 적용할 수 없습니다. 가능한 솔루션 중 하나는 순차적 패턴 마이닝 기술을 사용하여 주류 행동의 일부인 이벤트와 이상값 이벤트를 구별하는 것입니다(Sani et al. 2017). 그러나 일부 이벤트는 작업 실행 중에 거의 볼 수 없는 경우 이상값으로 잘못 처리될 수 있습니다. 요약된 문제는 의미론적 필터링이 필요합니다. 이벤트 그룹은 더 높은 의미의 동작으로 결합될 수 있습니다. 여기서 문제는 작업의 의미론적 경계와 페이로드를 형성하는 속성을 식별하는 것입니다.

후보자 루틴 식별. 이 단계는 2개의 하위 단계로 분해될 수 있습니다. 1) 일상적인 추출; 2) 자동화 가능한 루틴 식별. 제시된 각 하위 단계는 고유한 문제에 직면해 있습니다.

첫 번째 하위 단계는 루틴 실행을 나타내는 반복적인 순차적 패턴의 식별 및 추출을 목표로 합니다. 여기서 문제 중 하나는 루틴을 실행하는 동안 사용자가 다른 작업을 수행할 수 있다는 것입니다.

루틴의 일부가 아닌 작업. 루틴을 식별할 때 이러한 작업은 무시되어야 합니다. 이와 관련하여 순차적 패턴 마이닝 기술, 특히 갭 패턴(Liao and Chen 2013)으로 작업하는 기술을 사용할 수 있습니다. 또 다른 문제는 때때로 루틴을 구성하는 작업이 임의의 순서로 수행될 수 있다는 것입니다(예: 웹 양식을 채울 때). 따라서 자주 발생하는 패턴을 식별하기가 어렵습니다. 한 가지 가능한 솔루션은 Bialy et al.에 표시된 대로 추상화 메커니즘을 사용하는 것입니다. (2019). 대안적인 접근 방식은 이벤트가 발생하는 순서를 고려하지 않는 알파벳 반복(Bose and van der Aalst 2009)과 같은 패턴의 보다 유연한 개념이나 Leno et al. (2020).

두 번째 하위 단계의 주요 목표는 자동화에 적합한 루틴을 식별하는 것입니다. 발견된 루틴은 다음과 같은 경우 자동화 후보로 간주됩니다.

반자동 또는 완전 자동화가 가능합니다. 이러한 맥락에서 과제는 루틴이 자동화 가능한지 여부를 식별하는 방법입니다. Geyer-Klingenberg et al. (2018), 저자는 작업의 자동화 가능성을 평가하는 방법을 설명합니다. 작업 실행 빈도는 자동화의 주요 기준으로 제시됩니다. 그러나 작업이 빈번한 경우 자동화 가능하다는 보장은 없습니다.

Lacity와 Willcocks (2016)는 Telefonica의 사례 연구 맥락에서 작업이 자동화 대상인지 여부를 결정하기 위한 높은 수준의 지침을 제안합니다. 그러나 이 작업은 자동화 가능한 루틴의 식별을 자동화하는 데 중요한 자동화 가능한 작업에 대한 공식적이고 정확한 정의를 제공하지 않습니다. 실제로 주요 과제는 효율적인 알고리즘을 설계하여 대량의 UI 로그에서 이러한 루틴을 식별할 수 있도록 충분히 정확한 방식으로 루틴을 RPA에 적합하게 만드는 요소를 공식적으로 특성화하는 방법입니다. 한 가지 가능한 해결책은 결정론의 개념을 사용하는 것입니다. 루틴 수

루틴에 속한 모든 이벤트가 결정적으로 활성화되고 이전 작업에서 생성된 데이터를 사용하는 경우 자동화됩니다(예: 텍스트 필드에 수동으로 입력하는 것은 비결정적 작업의 예). 여기서 문제는 수행 중인 비결정적 작업을 반영하는 비결정적 이벤트를 UI 로그에서 식별하는 것입니다.

발생할 수 있는 비결정적 동작과 관련된 문제 중 하나는 자동화 가능한 하위 루틴을 포함하여 부분적으로 자동화 가능한 루틴을 식별하는 것입니다.

루틴 중간 어딘가에서 비결정적 동작이 발생하면 이 동작은 루틴을 두 개의 자동화 가능한 하위 루틴으로 분할합니다. 우리는 또한 모든 루틴이 자동화될 가치가 있는 것은 아님을 관찰했습니다. 일부 자동화

루틴은 다른 루틴의 자동화보다 훨씬 더 많은 이점을 가져올 수 있습니다. 따라서 일상적인 자동화의 비용-편익 분석은 RPA에서 중요한 작업입니다 (Lacity and Willcocks 2016).

실행 가능한 일상 검색. 루틴 세트가 주어지면 실행 가능한 루틴 발견은 데이터로 강화된 제어 흐름 모델의 형태로 전체 루틴 세트를 나타내는 루틴 사양을 구성하는 것으로 구성됩니다.

흐름. 여기서 문제는 동일한 루틴을 수행하는 여러(대체) 방법이 있을 수 있다는 것입니다. 예를 들어, 다른 작업자가 동일한 루틴을 다르게 수행할 수 있습니다. 따라서 루틴 사양을 찾을 때 루틴이 트리거되어야 하는 모든 전제 조건과 루틴의 효과(사후 조건)를 캡처하는 데 집중해야 합니다. 이것은 사전 조건과 관찰된 루틴의 효과가 주어진 루틴 사양에 의해 커버되는 정도를 캡처하는 루틴 사양에 대한 전용 품질 측정을 요구합니다. 또한 두 개의 서로 다른 루틴이 동일한 효과를 내는 경우 최적의 루틴을 식별하는 것이 중요합니다.

루틴의 최상의 대안 변형을 찾는 것은 실행 가능한 루틴 발견에서 어려운 일입니다.

일부 반복 루틴은 특정 조건에서만 트리거될 수 있습니다. 예를 들어 구매 주문이 '소매-EU' 유형인 경우 특정 EU 규정을 준수하기 위해 특정 일련의 작업이 수행되며 이 일련의 작업은 자동화할 수 있는 반복적인 루틴에 해당합니다. 반면에 주문이 'retail-US' 유형인 경우 다른 루틴이 수행됩니다.

또는 "retail-EU" 유형의 주문 처리는 실행 가능한 프로세스 모델을 통해 캡처할 수 있는 몇 가지 지정된 단계 시퀀스를 따르는 반면 "retail-CN"의 경우 주문 처리 임시 절차이며 규칙성을 찾을 수 없습니다. 따라서 "retail-EU" 주문 처리는 실행 가능한 모델을 통해 자동화할 수 있지만 "retail-CN" 주문 처리는 자동화할 수 없습니다. 최근 작업(Bosco et al. 2019)은 RIPPER와 같은 규칙 마이닝 기술을 사용하여 주어진 루틴을 자동화할 수 있는 조건을 찾는 아이디어를 제시했습니다. 그러나 실제 RPM 시나리오에서 이러한 기술의 적용 가능성은 아직 테스트되지 않았으며 확장성과 견고성 문제를 제기할 가능성이 있습니다.

이 단계의 또 다른 과제는 루틴의 각 작업 내에서 발생하는 데이터 변환을 발견하는 것입니다.

실제로 RPA 봇이 루틴의 작업을 재현하도록 하려면 루틴의 입력 매개변수 또는 루틴의 이전 작업 매개변수에서 각 작업의 매개변수를 계산하는 방법을 봇의 스크립트에 인코딩해야 합니다. 최근 연구(Bosco et al. 2019)는 실행 가능한 루틴의 발견에서 이 단계가 '예시를 통해' 데이터 변환의 자동화된 발견을 위한 기존 방법을 사용하여 구현될 수 있다고 제안합니다(Abedjan et al. 2016; Jin et al. 2017). 그러나 이러한 방법은 확장성 문제가 있습니다. 또한 그 범위(즉, 발견할 수 있는 변환 유형)는 다소 제한적입니다. 따라서 이 분야의 새로운 발전

RPM의 맥락에서 데이터 변환을 적용하려면 데이터 변환의 자동 검색이 필요합니다.

편집. 루틴 사양이 주어지면 컴파일 단계는 사양을 구현하는 실행 가능한 RPA 스크립트를 생성하는 것을 목표로 합니다. 이 단계에서는 루틴 실행 중에 관련된 애플리케이션 요소(예: 웹 양식의 버튼 또는 텍스트 필드)를 올바르게 식별해야 합니다. 예를 들어 웹 페이지에서 버튼을 클릭하는 동작을 실행 가능한 명령으로 변환할 때 이 버튼을 나타내는 HTML 요소를 식별하고 해당 DOM 위치를 추출해야 합니다. 이러한 정보는 기록 단계에서 로거에 의해 기록될 수 있습니다. 그러나 때때로 이 정보가 누락될 수 있습니다. 예를 들어, 일부 웹 요소(예: 링크)에는 페이지에서 해당 요소를 찾는 데 사용할 수 있는 식별자가 없습니다. 웹 사이트가 동적으로 생성되고 많은 양의 중첩 컨테이너로 구성된 경우 요소의 정확한 위치를 추출하는 것은 매우 어렵습니다. 따라서 API 없이 사용자 정의 응용 프로그램으로 작업할 경우 이벤트 유형을 올바르게 식별하지 못할 수 있습니다. 이러한 이유로 요소에 대한 능동적인 인식이 필요합니다. 이와 관련하여 OCR과 같은 기술이 사용될 수 있지만 여기서 문제는 기록된 작업의 의미를 보존하고 실행 중에 관련된 모든 데이터를 캡처하는 것입니다.

#### 4 관련 업무

RPA 도구를 통한 자동화를 위한 후보 루틴의 발견은 지금까지 거의 탐구되지 않은 문제입니다. 최근 작업(Leopold et al. 2018)은 자동화가 가능한 작업을 참조할 수 있는 비즈니스 프로세스(예: 작업 지침)의 텍스트 설명에서 구절을 식별하는 접근 방식을 스케치했습니다. 그러나 이 접근 방식은 자연어 분석의 복잡성으로 인해 부정확한 결과를 초래할 수 있습니다. 또한 적절한 품질과 완전성에 대한 텍스트 문서가 필요하며 작업이 문서화된 대로 정확하게 수행된다고 가정합니다. 실제로 작업자는 예외 및 변형을 처리하기 위해 완전히 문서화되지 않은 단계를 수행할 수 있습니다. 따라서 작업 지침에 따라 자동화 가능한 것처럼 보일 수 있는 작업이 실제로는 자동화할 수 없는 것으로 판명될 수 있습니다. 관련 작업의 또 다른 본문에는 기본값 또는 예측 값을 사용하여 웹 양식을 자동 완성하기 위한 접근 방식이 포함됩니다(Hermens and Schlimmer 1994). 이러한 접근 방식은 수동 양식 작성 중에 사용자에게 도움이 되지만 RPA 도구와 같은 방식으로 루틴을 자동화하지는 않습니다.

위의 작업 외에도 이 백서에서 제시한 RPM 비전은 다양한 행동 모델을 발견하고자 하는 데이터 마이닝의 다른 하위 분야와 관련이 있습니다.

로그의 종류. 아래에서는 RPM과 이러한 세 가지 필드, 즉 프로세스 마이닝, 웹 사용 마이닝 및 사용자 인터페이스 로그 마이닝 간의 관계에 대해 논의합니다.

프로세스 마이닝. RPM은 프로세스 마이닝 분야의 확장으로 위치할 수 있습니다(van der Aalst 2016).

특히, RPA 루틴을 발견하는 것은 프로세스 마이닝 분야에서 널리 연구되어 온 APD(Automated Process Discovery) 문제와 밀접한 관련이 있습니다(Augusto et al. 2019). APD 기술의 목적은 엔터프라이즈 시스템의 작업 실행을 기록하는 이벤트 로그에서 비즈니스 프로세스 모델을 발견하는 것입니다.

APD 알고리즘의 상당 부분은 제어 흐름 관점에서 프로세스 모델을 발견하는 데 중점을 둡니다.

APD 알고리즘의 이 하위 집합은 프로세스 작업에 의해 입력으로 사용되어 출력으로 생성되는 데이터나 분기 조건을 평가하기 위해 프로세스 실행 엔진에서 사용하는 데이터를 고려하지 않습니다. APD 기술의 또 다른 하위 집합은 데이터 기반 분기 조건(de Leoni et al. 2013)과 특정 조건에서만 유지되는 제어 흐름 관계(Mannhardt et al. 2017)가 있는 프로세스 모델을 발견하는 문제를 목표로 합니다. 이러한 후자의 기술은 RPA 루틴을 발견하기 위한 기술을 개발하기 위한 출발점을 제공합니다. 실제로, RPA 루틴을 발견하려면 루틴을 트리거하는 활성화 조건과 같은 루틴 내 조건을 발견해야 합니다. 다른 APD 기술은 시뮬레이션 모델 발견에 중점을 둡니다(Martin et al. 2016). 후자 유형의 모델은 비즈니스 프로세스 시뮬레이터에 대한 입력으로 제공될 수 있으며, 이는 확률론적 의미에서 실행합니다.

프로세스 마이닝 분야의 풍부한 작업에도 불구하고 비즈니스 프로세스 실행 엔진에 배포하거나 컴파일할 준비가 된 실행 가능한 프로세스 모델을 발견하는 기술을 인식하지 못합니다. 특히, 우리는 자동으로 발견된 프로세스 모델에서 데이터 변환(즉, 입력과 출력 간의 매핑)을 발견하려고 시도하는 자동화된 프로세스 발견에 대한 작업을 알지 못합니다. 그러나 이러한 데이터 변환은 프로세스 실행 엔진 또는 RPA 도구에서 실행할 수 있는 프로세스 모델을 발견하는 데 필수적입니다.

프로세스 마이닝에 사용되는 UI 로그와 이벤트 로그에는 유사점이 있습니다. 특히 두 가지 유형의 로그는 각 레코드가 사용자의 작업(또는 작업) 실행을 참조하도록 타임스탬프가 지정된 레코드로 구성됩니다. 또한 각 레코드에는 하나 이상의 속성-값 쌍으로 구성된 페이로드가 포함될 수 있습니다. 일부 상용 프로세스 마이닝 공급업체는 RPM 관련 기능을 제공하기 위해 UI 로그와 비즈니스 프로세스 이벤트 로그 간의 유사성을 이용했습니다. 예를 들어 Minit8 프로세스 마이닝 도구는 일부 RPM 작업을 지원하기 위해 다단계 프로세스 검색 기능을 제공합니다. 특히, 주어진 이벤트 로그

<sup>8</sup> <https://www.minit.io/>.

실행 작업과 UI 로그를 기록하면 Minit은 2단계 프로세스 맵을 생성할 수 있습니다. 첫 번째 수준은 엔터프라이즈 시스템에서 추출한 로그에 기록된 작업을 보여줍니다. 각 작업은 UI 작업과 제어 흐름 관계를 보여주는 2단계 프로세스 맵으로 확장될 수 있습니다. 이러한 방식으로 이 도구는 내부 구조가 비교적 단순하여 잠재적으로 자동화될 수 있는 작업의 (시각적) 식별을 지원합니다.

그러나 작업이 완전히 결정적인(하위) 루틴을 포함하는지 여부를 결정할 수 없으며 결정적인 루틴의 실행 가능한 사양을 생성할 수도 없습니다. 또한 도구는

UI 로그의 이벤트와 비즈니스 프로세스 이벤트 로그의 이벤트, 즉, RPM 파이프라인의 분할 단계를 다루지 않습니다.

또 다른 상용 도구인 Kryon Process Discovery<sup>9</sup>는 사용자가 하나 이상의 애플리케이션에서 작업을 수행하는 동안 찍은 스크린샷과 함께 UI 로그를 분석하여 RPA의 후보 루틴을 식별합니다. 그러나 Kryon이 식별하는 후보 루틴은 사용자가 입력한 실제 데이터 값에 따라 자동화할 수도 있고 그렇지 않을 수도 있습니다. 특정 단계에서 입력된 데이터 값을 이전에 관찰된 속성의 값에서 결정할 수 없는 경우 사용자가 외부 데이터 소스(UI에서 관찰되지 않음) 또는 자신의 도메인 지식에서 입력을 제공하고 있음을 의미합니다. 따라서 루틴의 해당 단계는 자동화할 수 없습니다. 즉, 이 도구에서 자동화의 후보 날짜로 식별된 모든 루틴을 자동화할 수 있는 것은 아닙니다.

한편으로는 UI 로그와 프로세스 마이닝에 사용되는 이벤트 로그 간에는 유사점이 있지만 4가지 주목할 만한 차이점이 있습니다. 첫째, 이벤트 로그는 더 높은 수준의 추상화에서 이벤트를 캡처합니다.

특히 이벤트 로그의 기록은 일반적으로 구매 주문 확인 또는 학생 기록 전송과 같은 비즈니스 프로세스 내의 전체 작업 실행을 나타냅니다. 이러한 작업은 UI 로그에 기록될 수 있는 하위 수준 작업의 구성으로 볼 수 있습니다. 예를 들어, 학생 기록 전송 작업에는 한 자원에서 다른 자원으로 학생과 관련된 기록(이름, 성, 주소, 과정 세부 정보)을 복사하는 여러 작업이 포함될 수 있습니다.

둘째, UI 로그에는 케이스 식별자(또는 프로세스 인스턴스 식별자)라는 개념이 포함되어 있지 않지만 이벤트 로그에는 일반적으로 포함됩니다. 즉, UI 로그의 이벤트는 명시적으로 상관 관계가 없으며 이러한 이유로 섹션에서 설명한 대로 분할해야 할 수 있습니다. 2.3. 셋째, 이벤트 로그의 레코드에는 해당 작업을 실행하는 동안 사용되거나 생성된 모든 입력 또는 출력 데이터가 포함되지 않는 경우가 많습니다. 예를 들어, 학생 기록 전송 작업 실행에 해당하는 이벤트 로그의 기록은 해당 학생의 모든 속성을 포함하지 않을 수 있습니다.

(예: 그들의 주소). 반면 UI 로그에 있는 모든 입력 및 출력 속성은 RPM을 위해 필요합니다. UI 로그에 일부 입력 또는 출력 속성이 누락된 경우 결과 루틴 사양이 불완전하여 결과 RPA 봇이 루틴을 올바르게 수행하지 못합니다. 네 번째 차이점은 이벤트 로그는 일반적으로 분석 목적으로 명시적으로 기록되는 것이 아니라 정보 시스템에서 실행되는 트랜잭션의 부산물로 획득된다는 것입니다.

후자의 특징은 이벤트 로그가 위에서 논의한 바와 같이 누락된 속성과 누락된 이벤트를 포함하여 불완전성을 겪을 가능성이 더 높다는 것을 수반합니다. 예를 들어, 병원의 환자 치료 과정에서 환자가 직접 응급실에 도착했을 때는 실제 환자가 응급실에 도착한 것을 기록하지 않고 구급차를 통해 환자가 도착했을 때 기록하는 경우가 있습니다. 즉, 이벤트 로그에 이벤트가 있는지 여부는 정보 시스템이 이벤트를 기록하도록 설계되었는지 여부와 작업자가 실제로 기록하는지 여부에 따라 다릅니다. 반면에 UI 로그는 분석 목적으로 특별히 기록되므로 UI 기록 도구의 기능에 따라 모든 관련 이벤트를 수집할 수 있습니다.

웹 사용 마이닝. 웹 사용 마이닝은 웹 응용 프로그램과의 사용자 상호 작용을 캡처하는 클릭 스트림과 같은 웹 데이터의 순차적 패턴을 발견하고 분석하려고 합니다(Srivastava et al. 2000). 이러한 데이터를 분석하면 웹 기반 응용 프로그램의 기능을 최적화하고 사용자에게 개인화된 콘텐츠를 제공하며 웹 페이지에 대한 가장 효과적인 논리적 구조를 찾는 데 도움이 될 수 있습니다(Liu 2007).

웹 사용 마이닝은 RPM과 유사한 세분성 수준에서 데이터로 작동합니다. 또한 웹 로그 마이닝에서 조작된 데이터는 상관 관계가 없는 경우가 많습니다. 즉, 특정 세션에 대한 작업을 명시적으로 할당하지 않고 여러 세션에서 수행되는 일련의 작업을 나타냅니다.

이러한 유사성을 감안할 때 웹 사용 마이닝 기술은 RPM 파이프라인을 실현하기 위한 출발점을 제공할 수 있습니다.

예를 들어, 웹 로그에서 세션을 추출하기 위한 웹 마이닝 기술은 위에서 논의한 분할 문제를 해결하도록 조정될 수 있습니다. 반면에 웹 사용 마이닝 기술은 자동화를 위한 후보 루틴을 찾는 문제를 다루지 않습니다.

또한 RPM은 웹 애플리케이션에 국한되지 않는다는 점에서 웹 사용 마이닝과 다릅니다.

사용자 인터페이스 로그 마이닝. RPM은 사용자 인터페이스 로그 마이닝 주제와도 관련이 있습니다. 데스크톱 도우미의 맥락에서 TaskTracer 및 TaskPredictor와 같은 연구 제안은 사용자가 수행하는 현재 작업을 식별하고 한 작업과 다른 작업 사이의 전환을 감지하기 위해 데스크톱 응용 프로그램에서 생성된 사용자 인터페이스 로그를 분석하는 문제를 해결했습니다(Shen et al. 2007, Dragunov et al. 2005). 이 분야의 다른 관련 작업에는

<sup>9</sup> <https://www.kryonsystems.com/process-discovery/>.

데스크톱 앱 사용자 인터페이스 로그에서 작업 식별 및 분류 문제(Oliver et al. 2006; Rath et al. 2010)와 시끄러운 사용자 인터페이스 로그에서 빈번한 작업 시퀀스를 추출하는 문제(Dev and Liu 2017) (자동화를 위한 후보 루틴을 구성할 수 있음). 이전에 인용된 연구와 관련하여 RPM의 참신함은 작업의 입력 및 출력(예: 클립보드에 복사하거나 붙여넣은 데이터, 셀에 입력된 데이터)을 포함하는 로그를 분석하여 실행 가능한 루틴 사양을 발견하려는 것입니다. 관련 데이터 없이 일련의 작업을 순전히 고려하는 것과는 대조적입니다.

## 5. 결론

우리는 RPA 도구를 사용하여 자동화할 수 있는 루틴을 식별하기 위해 IT 시스템과의 세분화된 사용자 상호 작용의 UI 로그를 분석할 수 있는 새로운 클래스의 프로세스 마이닝 도구, 즉 RPM 도구에 대한 비전을 제시했습니다. 이 비전을 구체화하기 위한 첫 번째 단계로 우리는 이를 파이프라인으로 분해하고 파이프라인의 각 구성 요소를 구현하기 위해 극복해야 할 과제를 스케치했습니다. 또한 이러한 문제를 해결하기 위한 몇 가지 지침을 제공했습니다.

제안된 RPM 파이프라인은 RPA 봇이 중단 간 방식으로 실행할 수 있는 루틴의 발견에 중점을 둡니다. 이 가정은 제한적입니다. 실제로 루틴은 경우의 특정 하위 집합에 대해 자동화될 수 있지만 모든 경우에 대해 자동화되지는 않습니다(즉, 자동화는 부분적으로만 달성할 수 있음). 제안된 RPM 파이프라인 너머의 주요 과제는 부분적으로 결정적인 루틴을 발견하는 방법입니다.

완전한 결정적 루틴은 모든 경우에 중단 간 실행될 수 있지만 봇이 루틴 중에 수집하는 입력 데이터 및 기타 데이터를 고려할 때 루틴을 결정적으로 계속할 수 없는 지점에 봇이 도달하면 부분적으로 결정적 루틴을 중지할 수 있습니다. 실행. 예를 들어, 스프레드시트나 기업 시스템에서 구매 주문 기록을 복사하는 동안 봇은 이 주문이 중국에서 온 것을 감지하고 이러한 주문을 처리하는 방법을 모르거나 PO 번호(빈 셀), 따라서 진행할 수 없습니다. 루틴을 결정적으로 계속(또는 시작)할 수 없는 조건을 발견하는 것은 RPM의 주요 과제입니다.

이 백서에서 노출된 RPM의 비전은 자동화 가능한 루틴을 발견하는 데 중점을 두고 있습니다. 이는 우리가 예측하는 광범위한 RPM 작업 세트, 즉 로봇 프로세스 발견의 한 가지 작업일 뿐입니다. 로봇 프로세스 발견 외에도 RPM 분야는 RPA 봇의 성능 마이닝과 같은 보완적인 문제와 질문(예: '성공 또는

주어진 루틴을 수행할 때 봇의 결함 비율은?", "어떤 패턴이 봇 실패와 상관 관계가 있거나 인과 관계가 있습니까?"뿐만 아니라 "봇 또는

봇 행동의 효과가 비정상적이므로 수동 검사 및 수정이 필요합니까?".

감사의 말 이 연구는 호주 연구 위원회(Grant DP180102839)와 유럽 연구 위원회(PiX 프로젝트)의 지원을 받았습니다.

오픈 액세스 이 기사는 원본 저자(들)와 출처에 대한 적절한 크레딧을 제공하는 한 모든 매체 또는 형식으로 사용, 공유, 각색, 배포 및 복제를 허용하는 Creative Commons Attribution 4.0 International License에 따라 사용이 허가되었습니다. 크리에이티브 커먼즈 라이선스에 대한 링크를 제공하고 변경 사항이 있는지 표시합니다. 이 기사의 이미지 또는 기타 제3자 자료는 자료에 대한 크레딧 라인에 달리 명시되지 않는 한 기사의 크리에이티브 커먼즈 라이선스에 포함됩니다. 자료가 기사의 크리에이티브 커먼즈 라이선스에 포함되어 있지 않고 의도한 사용이 법적 규정에 의해 허용되지 않거나 허용된 사용을 초과하는 경우 저작권 소유자로부터 직접 허가를 받아야 합니다. 이 라이선스의 사본을 보려면 <http://creativecommons.org/licenses/by/4.0/>.

## 참고문헌

- Abedjan Z, Morcos J, Ilyas IF, Ouzzani M, Papotti P, Stonebraker M(2016) Dataxformer: 강력한 변환 검색 시스템. In: 2016년 5월 16-20일 핀란드 헬싱키에서 열린 ICDE 2016 데이터 엔지니어링에 관한 32차 IEEE 국제 회의. IEEE Computer Society, pp 1134-1145
- Agaton B, Swedberg G(2018) 비즈니스 프로세스 적합성을 평가하기 위한 평가 및 개발 방법 로봇 프로세스 자동화를 위한 설계 연구 접근 방식. 석사 논문, Chalmers University of Technology
- Aguirre S, Rodriguez A(2017) 로봇 프로세스 자동화(RPA)를 사용한 비즈니스 프로세스 자동화: 사례 연구. In: 엔지니어링 응용 프로그램(WEA)에 대한 4차 워크숍의 절차. Springer, New York, pp 65-71
- Augusto A, Conforti R, Dumas M, La Rosa M, Maggi FM, Marrella A, Mecella M, Soo A(2019) 이벤트 로그에서 프로세스 모델 자동 검색: 검토 및 벤치마크. IEEE Trans Knowl Data Eng 31(4):686-705
- Bayomie D, Awad A, Ezat E(2016) 주기적인 비즈니스 프로세스 실행에서 레이블이 지정되지 않은 이벤트를 연관시킵니다. In: 첨단 정보 시스템 공학(CAISE)에 관한 제 28회 국제 회의의 절차. Springer, New York, pp 274-289
- Bayomie D, Di Ciccio C, La Rosa M, Mendling J(2019) 프로세스 마이닝을 위한 이벤트 케이스 상관 관계에 대한 확률론적 접근. In: 개념적 모델링에 관한 국제 회의(ER2019). Springer, New York
- Bialy Y, Yehezkel A, Roseberg E, Smutko A(2019) 낮은 수준의 사용자 작업에서 마이닝을 처리합니다. In: BPM 2019 워크숍의 절차. 스프링거, 뉴욕
- Bosco A, Augusto A, Dumas M, La Rosa M, Fortino G (2019) 사용자 상호 작용 로그에서 자동화 가능한 루틴을 검색합니다. In: 비즈니스 프로세스 관리 포럼(BPM 포럼)의 절차. Springer, New York
- Bose RPJC, van der Aalst WMP(2009) 프로세스 마이닝의 추상화: 패턴의 분류. In: Dayal U, Eder J, Koehler J,



- Reijers HA (eds) 비즈니스 프로세스 관리. BPM 2009, 스프링거, 하이델베르크
- Conforti R, La Rosa M, ter Hofstede AHM(2017) 비즈니스 프로세스 이벤트 로그에서 드물게 발생하는 동작을 필터링합니다. *IEEE Trans Knowl Data Eng* 29(2):300–314 de Leoni M, Dumas M, Garc 'a-Ban~uelos L (2013) 비즈니스 프로세스 실행 로그에서 분기 조건 발견. In: *Proceedings of the 16th International Conference on Fundamental approach to software engineering – (FASE)*, pp 114–129 Dev H, Liu Z (2017) 애플리케이션 로그에서 빈번한 사용자 작업 식별. In: *Proceedings of IUI 2017*. Springer, New York, pp 263–273 Dong G, Pei J(2007) 시퀀스 데이터 마이닝, *Advances in Database Systems*의 33권. Kluwer Dragunov AN, Dietterich TG, Johnsrude K, McLaughlin MR, Li L, Herlocker JL(2005) Tasktracer: 멀티태스킹 지식 근로자를 지원하는 데스크탑 환경. In: St. Amant R, Riedl J, Jameson A (eds) *Proceedings of the 10th International Conference on Intelligent User Interfaces, IUI 2005, San Diego, California, USA, 2005년 1월 10–13일*. ACM, pp 75–82 Dumas M, La Rosa M, Mendling J, Reijers HA (2018) *비즈니스 프로세스 관리의 기초*, 2nd edn. Springer, New York
- Ferreira DR, Gillblad D(2009) 레이블이 지정되지 않은 이벤트 로그에서 프로세스 모델 발견. In: *비즈니스 프로세스 관리(BPM)에 관한 제7회 국제 회의의 절차*. Springer, New York, pp 143–158 Geyer-Klingenberg J, Nakladal J, Baldauf F, Veit F(2018) 프로세스 마이닝 및 로봇 프로세스 자동화: 완벽한 일치. In: *Proceedings of the Dissertation Award, Demonstration, and Industrial Track at BPM 2018*. CEUR-WS.org, pp 124–131 Hermens LA, Schlimmer JC(1994) 반복적인 양식 완성을 위한 기계 학습 견습생. *IEEE Expert* 9(1):28–33 Jin Z, Anderson MR, Cafarella MJ, Jagadish HV(2017) Foofah: 예시를 통한 데이터 변환. In: Salihoglu S, Zhou W, Chirkova R, Yang J, Suciu D (eds) *데이터 관리에 관한 2017 ACM 국제 회의, SIG MOD 회의 2017, 미국 일리노이주 사카고, 2017년 5월 14일–19일*. ACM, pp 683–698
- Lacity Mary, Willcocks Leslie P(2016) 로봇 프로세스 자동화 텔레포니카 O2에서 MIS Q 실행 15(1)
- Leno V, Polyvyanyy A, La Rosa M, Maggi FM (2019) 작업 로그: 로봇 프로세스 자동화를 위한 프로세스 마이닝을 활성화합니다. In: *비즈니스 프로세스 관리 데모 트랙의 진행*. CEUR
- Leno V, Dumas M, Maggi FM, La Rosa M, Polyvyanyy A (2020) 관련 데이터 조건이 있는 선언적 프로세스 모델의 자동 검색. *Inf Syst* 89:101482
- Leopold H, van der Aa H, Reijers HA (2018) 텍스트 프로세스 설명에서 로봇 프로세스 자동화를 위한 후보 작업 식별. In: *BPMDS 및 EMMSAD 절차*. Springer, New York, pp 67–81 Liao VC-C, Chen MS (2013) 생물학적 서열의 모티프에 대한 효율적인 마이닝 갭 순차 패턴. *BMC Syst Biol* 7(S– 4):S7
- Linn C, Zimmermann P, Werth D(2018) 데스크톱 활동 마이닝 - 마이닝 비즈니스 프로세스의 새로운 수준의 세부 정보입니다. In: *Workshops der INFORMATIK 2018 – Architekturen, Prozesse, Sicherheit und Nachhaltigkeit*, pp 245–258
- Liu B(2007) 웹 사용 마이닝. In: *웹 데이터 마이닝: 하이퍼링크 탐색. 콘텐츠 및 사용 데이터*. Springer, 베를린, pp 449–483 Mannhardt F, de Leoni M, Reijers HA, van der Aalst WMP(2017) 데이터 기반 프로세스 검색 – 이벤트 로그에서 조건부로 드물게 발생하는 동작을 나타냅니다. In: *고급 정보 시스템 공학에 관한 29차 국제 회의의 절차*. Springer, New York, pp 545–560
- Martin N, Depaire B, Caris A(2016) 비즈니스 프로세스 시뮬레이션 모델 구축에서 프로세스 마이닝 사용 – 현장 구조화. *버스 정보 시스템 영어* 58(1):73–87
- Oliver N, Smith G, Thakkar C, Surendran AC (2006) SWISH: 창 제목 및 전환 기록의 의미론적 분석. In: Paris C, Sidner CL (eds) *Proceedings of the 11th International Conference on Intelligent User Interface, IUI 2006, Sydney, Australia, 2006년 1월 29일–2월 1일*. ACM, pp 194–201 Ramirez AJ, Reijers HA, Barba I, Del Valle C(2019) 로봇 프로세스 자동화 라이프사이클의 초기 단계를 개선하는 방법. In: *고급 정보 시스템 공학(CAiSE)에 관한 제31회 국제 회의의 절차*. Springer, New York, pp 446–461 Rath AS, Devaurs D, Lindstaedt SN (2010) 컴퓨터 바탕 화면에서 자동 사용자 작업 감지에 영향을 미치는 요인 연구. In: Wolpers M, Kirschner PA, Scheffel M, Lindstaedt SN, Dimitrova V (eds) *Sustaining TEL: 혁신에서 학습 및 실습으로 – 기술 강화 학습에 관한 5차 유럽 회의, EC-TEL 2010, 스페인 바르셀로나, 9월 28일–10월 1, 2010*. Springer, New York, pp 292–307
- Sani MF, van Zelst SJ, van der Aalst WMP(2017) 조건부 행동 확률을 사용하여 이상값을 필터링하여 프로세스 검색 결과를 개선합니다. In: *비즈니스 프로세스 관리 워크숍 진행*. Springer, New York, pp 216–229 Shen J, Li L, Dietterich TG(2007) 데스크톱 사용자의 작업 전환에 대한 실시간 감지. In: *Veloso MM (ed) IJCAI 2007, 인공지능에 관한 20차 국제 공동 회의, 인도 하이데라바드, 2007년 1월 6–12일*, pp 2868–2873 Srivastava J, Cooley R, Deshpande M, Tan PN(2000) 웹 사용 마이닝: 웹 데이터에서 사용 패턴을 발견하고 적용합니다. *SIGKDD Explor Newsl* 1(2):12–23 Tax N, Sidorova N, van der Aalst WMP(2019) 혼란스러운 활동을 필터링하여 이벤트 로그에서 보다 정확한 프로세스 모델을 발견합니다. *J Intell Inf Syst* 52(1):107–139 Tornbohm C(2017) 로봇 프로세스 자동화 소프트웨어용 Gartner 시장 가이드. 보고서 G00319864, Gartner van der Aalst WMP(2016) 프로세스 마이닝 – 실행 중인 데이터 과학, 2nd edn. Springer, New York van der Aalst WMP, Bichler M, Heinzl A(2018) 로봇 프로세스 자동화. *버스 정보 시스템 엔지니어* 60(4):269–272