

Homework #1 (Scientific Article Summary)

Sep 2, 2022

ID : 2018204002 Name : 박정재Article Title : Robotic Process Mining: Vision and ChallengesAuthors: Volodymyr Leno • Artem Polyvyanny • Marlon Dumas • Marcello La Rosa • Fabrizio Maria Maggi

1. 요약(Highlights)

RPA(Robotic process automation)는 웹 및 데스크톱 애플리케이션과의 세분화된 상호 작용 시퀀스를 인코딩하는 스크립트를 실행하여 조직에서 반복적인 사무 작업을 자동화할 수 있게 해주는 새로운 기술이다. RPA가 광범위한 루틴을 자동화할 수 있다는 점을 감안할 때 처음에 어떤 루틴을 자동화해야 하는지에 대한 질문이 제기되고 이 논문은 이러한 격차를 메우기 위해 사용자 상호작용(UI)에서 자동화가 가능한 반복적인 루틴을 발견할 수 있는 RPM(robotic process mining)이라는 기술 제품군에 대한 비전을 제시한다. 이 논문은 또한 RPM을 뒷받침하는 일련의 기본 개념을 정의하고 RPM 도구가 UI 로그에서 RPA 스크립트를 생성할 수 있도록 하는 처리 단계의 파이프라인을 제시하고 계획된 파이프라인을 실현하기 위한 연구 과제에 대해 설명한다.

2. 서론(Introduction)

2.1 연구 목적

이 논문의 연구목적은 RPM을 뒷받침하는 일련의 개념을 정의하고 RPM 도구가 UI 로그에서 RPA 스크립트를 생성할 수 있도록 하는 처리 단계의 파이프라인을 제시하여 이를 기반으로 연구 과제를 논의하고 이러한 과제를 해결하기 위한 것이다.

2.2 배경

UiPath Enterprise RPA Platform과 Automation Anywhere 엔터프라이즈 RPA와 같은 로봇 프로세스 자동화(RPA) 도구는, 조직이 웹 및 데스크톱 애플리케이션과의 세분화된 상호 작용 시퀀스를 인코딩하는 스크립트를 실행하여 반복적인 작업을 자동화할 수 있다(van der Aalst et al.2018).

많은 사례 연구에서 RPA 기술이 사무 업무와 관련된 비즈니스 프로세스의 효율성과 데이터 품질을 향상시킬 수 있음을 보여주었다(Lacity 및 Willcocks 2016; Aguirre and Rodriguez 2017). 그러나 기존 RPA 도구는 광범위한 루틴을 자동화할 수 있지만 처음부터 어떤 루틴이 자동화 대상인지 결정할 수는 없다. 또한, 경험적 조사 방법을 통해 분석가는 자동화를 위한 후보 루틴을 식별하고 식별된 루틴 자동화의 잠재적 이점과 비용을 평가할 수 있다. 그러나 이러한 방법은 시간이 많이 걸리므로 루틴 수가 매우 많은 조직에서 확장성 제한에 직면한다. 따라서, 이러한 문제들을 새로운 종류의 RPM이라는 도구를 통해 해결하고자 한다.

3. 방법(Method)

3.1 RPM 프레임워크

(1) 맥락과 범위

RPM은 구체적으로 자동화를 위한 후보 루틴의 식별 및 평가와 RPA 봇이 실행할 수 있는 루틴 사양의 검색을 지원하기 위해 사용자 주도 작업을 실행하는 동안 수집된 데이터를 분석하는 기술 및 도구 클래스라고 정의된다. 따라서 정의에서의 내용과 같이 RPM 도구의 주요 데이터 소스는 UI 로그이다. 위의 정의에 따라 RPM은 다음과 같은 세 가지 주요 단계로 구분한다.

[(1) 하나 이상의 작업 실행에 해당하는 UI 로그를 수집하고 전 처리한다. (2) RPA에 대한 후보 루틴을 식별한다. (3) 실행 가능한 RPA 루틴을 발견한다.]

(2) 개념

RPM의 주요 입력은 사전에 기록해야 하는 UI 로그이다. UI 로그를 얻으려면 특정 작업과 관련된 모든 사용자 상호 작용을 기록해야 한다. 그 후, Task trace 모음이 주어지면 RPM의 목표는 여기에서 루틴이라고 하는 여러 Task trace에서 관찰할 수 있는 반복적인 작업 시퀀스를 식별하고 자동화가 가능한 루틴을 식별하는 것이다. 그런 다음 각 루틴에 대해 RPM은 실행 가능한 specification을 추출하는 것을 목표로 한다. 이 routine specification은 초기에 플랫폼 독립적인 방식으로 캡처 된 후 특정 RPA 도구에서 실행 되도록 플랫폼 종속적 RPA 스크립트로 컴파일 될 수 있다.

(3) RPM 파이프라인

1. Recording: 필드 선택, 필드 편집, 데스크톱 응용 프로그램 열기 또는 웹 페이지 열기와 같은 낮은 수준의 UI 이벤트 캡처와 같은 UI 로그를 기록한다.
2. Noise Filtering: 어떤 작업에도 속하지 않는 이벤트(노이즈)를 식별하고 필터링하는 작업이다. (자동화해서는 안 된다.)
3. Segmentation: task trace의 경계 즉, 작업 실행을 나타내는 연속적인 작업의 시퀀스를 식별하는 것이다. 입력은 단일 이벤트 시퀀스를 포함하는 UI 로그이고 출력은 각각 작업 실행을 나타내는 trace 집합이다.
4. Simplification: 이벤트를 상위 수준 작업으로 집계하는 작업이 포함된다. 이러한 방식으로 task trace은 훨씬 더 간결하게 대상 언어로 번역될 수 있다.
5. Candidate routine identification: 여러 task trace(루틴)에서 발생하는 반복적인 작업 시퀀스를 추출하고 자동화가 가능한 작업을 식별하는 것을 목표로 한다. 이 단계의 결과는 자동화 가능성에 따라 순위가 매겨진 자동화 또는 반자동화 루틴 세트이다.
6. Executable (sub)routine discovery: 각 후보 루틴에 대해 이 단계는 루틴의 인스턴스가 트리거 되어야 하는 시기를 나타내는 활성화 조건과 해당 루틴내에서 수행되어야 하는 작업을 지정하는 routine specification을 식별한다.
7. Compilation: routine specification의 작업을 대상 RPA 도구의 스크립팅 언어로 된 명령으로 매핑하여 실행 가능한 스크립트를 생성한다.

3.2 과제 및 지침

1. Recording: 이 단계의 주요 과제는 기록해야 하는 작업을 식별하는 것이다. 작업자가

웹 양식을 선택할 때 논리적 입력 요소 수준에서 루틴을 학습하기 위해 웹 페이지 수준에서 이벤트를 기록해야 한다.

2. Noise Filtering: 이 단계의 주요 과제 중 하나는 작업에 기여하는 이벤트에서 노이즈를 분리하는 것이다. 가능한 솔루션은 프로세스 실행 중 어디에서나 발생할 수 있는 혼란스러운 이벤트를 노이즈로 처리하는 것이다.

3. Segmentation: 자동화를 위한 후보 루틴을 식별하기 전에 UI 로그를 trace로 분할하여 각 trace가 작업의 한 실행에 해당하도록 해야 한다.

4. Simplification: 이벤트가 작업에 속하더라도 여전히 중복될 수 있다. 그러므로 작업의 의미론적 경계와 페이로드를 형성하는 속성을 식별하는 것입니다.

5. Candidate routine identification: (이 단계는 2개의 하위 단계로 분해될 수 있다.) 1) 일상적인 추출: 루틴 실행을 나타내는 반복적인 순차적 패턴의 식별 및 추출을 목표로 한다.

2) 자동화 가능한 루틴 식별: 자동화에 적합한 루틴을 식별하는 것이다.

6. Executable (sub)routine discovery: 동일한 루틴을 수행하는 여러(대체) 방법이 있을 수 있다. 또 다른 과제는 루틴의 각 작업 내에서 발생하는 데이터 변환을 발견하는 것이다.

7. Compilation: routine specification이 주어지면 컴파일 단계는 specification을 구현하는 실행 가능한 RPA 스크립트를 생성하는 것을 목표로 한다.

4. 결과(Results)

해당 논문의 연구는 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 제시했다. 이 비전을 구체화하기 위한 첫 번째 단계로 이를 파이프라인으로 분해하고 파이프라인의 각 구성 요소를 구현하기 위해 극복해야 할 과제를 스케치했다. 또한 이러한 문제를 해결하기 위한 몇 가지 지침을 제공했다. 이 논문에서 제안된 RPM 파이프라인은 RPA 봇이 end-to-end 방식으로 실행할 수 있는 루틴의 발견에 중점을 둔다.

5. 토의(Discussion)

(1) 가치와 기여점

이 연구의 가치는 노출된 RPM의 비전이 자동화 가능한 루틴을 발견하는데 중점을 뒀다는 것이다. 또한 로봇 프로세스 발견 외에도 RPM 분야는 RPA 봇의 성능 마이닝과 같은 보완적인 문제와 질문을 포함할 것으로 예상된다.

(2) 앞으로 다뤄야 할 연구

제안된 RPM 파이프라인 너머의 주요 과제는 부분적으로 결정적인 루틴을 발견하는 방법이다. 완전한 결정적 루틴은 모든 경우에 end-to-end 방식으로 실행될 수 있지만 봇이 루틴 실행 중에 수집하는 입력 데이터 및 기타 데이터를 고려할 때 봇이 루틴을 결정적으로 계속할 수 없는 지점에 도달하면 부분적으로 결정적인 루틴을 중지할 수 있다. 따라서, 루틴을 결정적으로 계속(또는 시작)할 수 없는 조건을 발견하는 것은 RPM의 주요 과제이다.