패턴기반 워크플로우 마이닝 기법

박민재[O] 유혁재 안형진 원재강 김광훈
경기대학교 전자계산학과 워크플로우 기술 연구실
{mjpark[O], hjyoo, hjahn, jkwon, kwang}@kyonggi.ac.kr

Pattern Based Workflow Mining Method

Minjae Park[O] Hyuckjae Yoo Hyungjin Ahn Kwanghoon Kim
Dept. of Computer Science, Kyonggi University

요 약

비즈니스 프로세스를 연구하는데 있어 최근의 이슈는 비즈니스 프로세스가 복잡화되고 대량화 됨에 따라, 비즈니스 프로세스를 좀더 효율적으로 개선하고자 하는 데에 맞춰져 있으며, 이러한 동향에 힘입어 워크플로우 마이닝이라는 연구분야가 생겨나게 되었다. 이에, 본 논문에서는 워크플로우 패턴을 기반으로 한 워크플로우 마이닝 기법에 관하여 연구하고, 기술한다. 첫째로, 기본적으로 워크플로우 마이닝 기술과 그의 근원이 되는 워크플로우 로그에 관하여 간략히 설명한다. 그리고, 패턴기반 워크플로우 마이닝을 하기 위한 기본적인 워크플로우 패턴에 관하여 기술하고, 패턴 기반 워크플로우 마이닝을 하기 위한 기법으로 몇 가지 알고리즘을 제시한다.

1. 서론

최근 몇 년 동안 국내외의 많은 기업들은 워크플로우 및 BPM을 도입하였고, 그로 인해, 비즈니스 프로세스의 가시화와 투명화, 비즈니스 프로세스 실행비용의 감소화, 비즈니스 프로세스의 가속화, 자동화, 생산성 향상, 비즈니스 프로세스의 복잡도 개선, 비즈니스 프로세스 실행 오류의 감소, 비즈니스 프로세스의 투명한 관리 및 외부 비즈니스와의 조정으로 기업 품질 자체를 향상시켜 고객 만족도를 증가 시켰으며, 변화에 대한 민첩한 대응력 및 유연성을 갖추었다. 하지만, 정의 시에 정의된 프로세스 정의정보와 워크플로우 제정 서비스(Workflow Enactment Service)의 실행 시 발생하는 프로세스 인스턴스와의 불일치가 일어난다는 점을 발견하게 되었다. 이는 워크플로우 정의 프로세스 모델이 실제로 일어나는 비즈니스 프로세스를 반영한 최적화 된 모델이 아니라는 것을 의미하며, 따라서, 프로세스 모델은 개선이 필요하다는 것을 의미한다.

2. 관련연구

프로세스의 정의정보와 실행정보 사이의 불일치를 해소하고, 프로세스를 개선하기 위한 목적으로 워크플로우 웨어하우징, 워크플로우 마이닝 기술이 등장하였다.

2.1 워크플로우 마이닝 기술

워크플로우 마이닝이란 용어는 워크플로우를 연구하는데 있어서 더 이상 생소한 용어는 아니다. 이미 국외에서는 워크플로우 및 프로세스 마이닝이라는 주제로 깊이 있는 연구가 진행되고 있다. 하지만, 국내의 워크플로우 및 프로세스 마이닝에 관한 연구가 아직 미약함으로 다음과 같이 간단히 정리하고 넘어가고자 한다. 워크플로우 마이닝에 관한 정의와 그 의미를 살펴보면 다음과 같이 나타낼 수 있다.

- 워크플로우 마이닝이란 실행 시에 발생한 워크플로우 로그 데이터로부터 정의시점에 알려지지 않은 프로세스의 정보를 추출해 내는 프로세스 발견 방법이다.
- 워크플로우 마이닝은 프로세스를 발견하거나 재발견하는 방법, 그 기술 자체만을 의미하는 것이 아니며, 비즈니스 프로세스의 문제를 이해하고, 이러한 문제를 해결하기 위하여 여러 가지 정보기술을 적용하는 포괄적인 과정을 의미한다.

현재 워크플로우 마이닝(Workflow Mining)은 프로세스 발견 또는 재발견(Process Discovery or Rediscovery)이라고도 불리며, 그 이외에 프로세스 마이닝(Process Mining)이라고도 불린다.

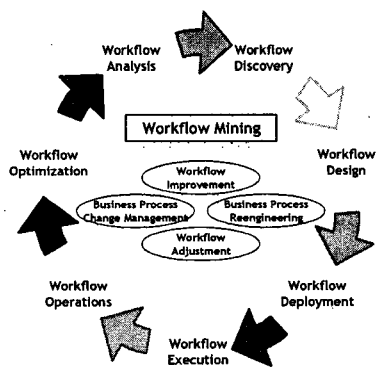

그림 2. 워크플로우 마이닝

일련의 워크플로우 마이닝 과정은 워크플로우 발견, 워크플로우 디자인, 워크플로우 전개, 워크플로우 실행, 워크플로우 운용, 워크플로우 최적화, 워크플로우 분석 단계의 반복으로 이루어질 수 있으며, 이러한 워크플로우 마이닝의 과정을 거치면서, 워크플로우의 개선, 비즈니스 프로세스 리엔지니어링, 비즈니스 프로세스 변화 관리, 워크플로우 조정 등을 이룰 수 있다.

2.2 워크플로우 로그

워크플로우 마이닝 기술 연구의 근간을 이루고 있는 것은 크게 워크플로우 로그와 워크플로우 마이닝 알고리즘이라고 볼 수 있다.

워크플로우 로그는 워크플로우 관리 시스템에서 운용되는 이벤트에 의해 발생하는 몇 가지 상태의 변화에 따라 남겨질 수 있는 데이터이다. 일반적인 워크플로우는 각각의 생성, 시작, 중지, 재개, 취소, 완료 이벤트들에 의해 생성, 준비, 활성화, 중지, 종료, 완료의 기본적인 상태변화를 일으킬 수 있다.

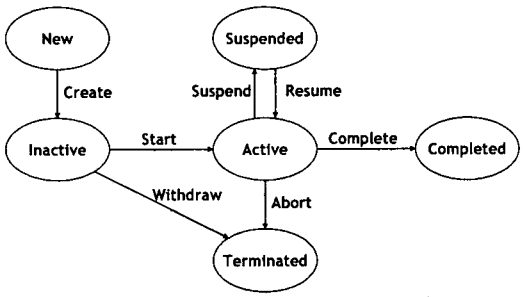

그림 3. 워크플로우의 기본 상태 변화

워크플로우 엔진에서 운용되는 워크플로우 제정서비스의 몇 가지 이벤트에 의해 발생한 워크플로우 로그는 의미상의 일정한 구성요소를 가지고 있으며, 정해진 프로세스 모델에 관하여 서로간의 통합이 필요하다. 일정한 워크플로우 로그 스키마에 의 의해 통합된 워크플로우 로그는 각각의 목적을 가지고 있는 워크플로우 마이닝 도구에 의해 분석된다.

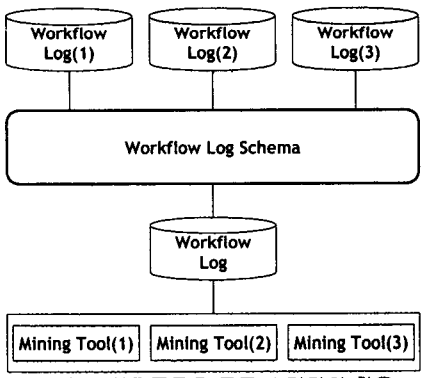

그림 4. 워크플로우 로그 스키마의 활용

워크플로우 로그를 통합하고 활용하기 위한 워크플로우 로그 스키마는 워크플로우 엔진에서 발생할 수 있는 경우의 일반화를 통해 일정한 데이터 형식의 구조로써 구성될 수 있다. 일반적으로 크게 워크케이스(프로세스 인스턴스), 런타임 액티비티(액티비티 인스턴스), 워크아이템으로 이루어질 수 있으며, 각각의 구성요소 또한 구성요소에 관련한 식별 값과 인스턴스를 발생시키는 부모개체의 식별 값, 그리고 일어나게 한 이벤트의 구성요소 및 이벤트 발생 시간정보 등을 필수적으로 필요로 하며, 구성 된다.

3. 패턴기반 워크플로우 마이닝 기법

워크플로우 로그 정보로부터 프로세스 정보를 알아 내기 위한 방법으로 본 논문에서는 패턴을 기반으로한 워크플로우 마이닝 기법을 제안한다.

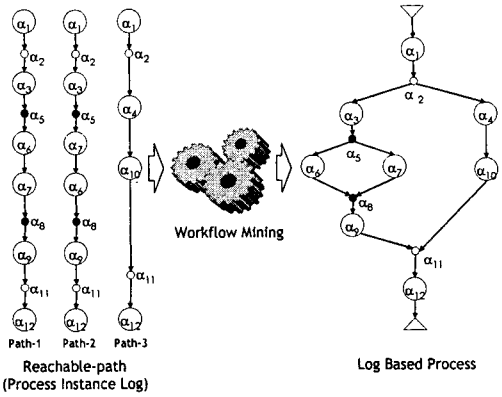

그림 5. 패턴 기반 워크플로우 마이닝

패턴기반 워크플로우 마이닝이란 워크플로우 프로세스 인스턴스에서 발생하는 워크플로우 프로세스 인스턴스의 로그정보로부터 워크플로우 패턴을 찾아내어 프로세스를 완성해 나아가는 방법이라고 할 수 있다.

3.1 워크플로우 기본 패턴

워크플로우 프로세스는 기본적으로 몇 가지 워크플로우 패턴을 바탕으로 이루어진다. 워크플로우 프로세스를 이루는 기본 패턴으로는 ' Sequence' , ' Parallel Split' , ' Synchronization' , ' Exclusive Choice' , ' Simple Merge' , 그리고, ' Loop' 패턴이 있다. 다음 표 1은 각 워크플로우 기본 패턴과 그것을 그래픽적으로 표현한 ICN 그래프를 보여준다.

표 1. 워크플로우 기본 패턴

기본 패턴	ICN Graph
Sequence	
Parallel Split	
Synchronization	
Exclusive Choice	
Simple Merge	
Loop	

본 논문에서는 워크플로우 프로세스 인스턴스 로그로부터 워크플로우 기본 패턴을 찾아내어 워크플로우 마이닝을 하기 위한 기법에 관하여 기술한다.

3.2 패턴기반 워크플로우 마이닝 알고리즘

런타임에서 발생하는 워크플로우 로그 정보로부터 워크플로우 프로세스를 다시 만들어내기 위하여 본 논문에서 제시하고자 하는 방법은 패턴을 기반으로한 워크플로우 마이닝 방법이다.

패턴을 기반으로하는 워크플로우 마이닝의 가장 기본적인 시나라오는 다음과 같다.

첫째로, 프로세스 인스턴스는 기본적으로 ' Sequence' 한 정보로써 출발하게 된다. 다음으로 ' Sequence' 를 가진 프로세스 인스턴스 정보가 계속적으로 발생하게 될 시, 각각의 ' Sequence' 를 구성하고 있는 액티비티들의 집합이 ' OR-패턴' 인지 ' AND-패턴' 인지를 구별하여 프로세스 액티비티들을 추가, 갱신의 행위를 반복하여, 더 이상의 입력 프로세스 인스턴스 로그가 없을 때 최종적인 로그 기반 프로세스를 생성하는 방법을 나타낸다.

다음 표 2는 워크플로우 인스턴스 로그 정보를 입력으로 하여 각각의 패턴을 분류해서 프로세스의 액티비티를 추가하고 프로세스를 갱신해 나가는 알고리즘을 나타낸다.

표 2. 패턴기반 워크플로우 마이닝 알고리즘

```
Input (Process Instance Log)
While (Activities exist Process Log)
    //Pattern classification and processing
    If (Sequence-Pattern)
        Sequence activities add
        Processes refine
    Else If (OR-Pattern)
        Exclusive Choice activity add
        New activities set add
        Simple Merge activity add
        Processes refine
    Else If (AND-Pattern)
        Parallel Split activity add
        Sequence activities set change
        Synchronization activity add
        Processes refine
    End If
End While
```

위 표 2에서 표현한 알고리즘의 구성요로서 반드시 필요한 것은 각각의 패턴을 인식해 내내는 것이다. 다음 그림 6과 같은 정보를 가질 수 있다면, OR에 관련된 ' Exclusive Choice' 와 ' Simple Merge' 패턴이 적용 가능하다는 것을 나타낸다.

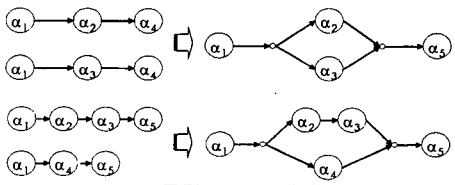

그림 6. OR-패턴

다음 표3은 ' OR-패턴' 의 경우를 구분해 낼 수 있는 방법으로 새로운 액티비티 집합을 발견했을 때, 새로운 액티비티 집합의 전후 액티비티를 인식하여, 이전 프로세스에서 존재하는 액티비티일 경우에 전후 액티비티 사이에 이루어진 OR 관련 액티비티 집합이라는 것을 알아 낼 수 있는 알고리즘이다.

표 3. OR-패턴을 구분하기 위한 알고리즘

```
//OR-Pattern
If (exist ' new Activities set' ?)
    If (exist before and after activity of ' new activities set' ?)
        New activities set is ' or activities set'
        between before and after activity of new activities set
    End If
End If
```

또한 그림 7과 같은 정보를 가질 수 있다면 AND에 관련된 ' Parallel Split' 과 ' Synchronization' 패턴이 적용가능 하다는 것을 나타낸다.

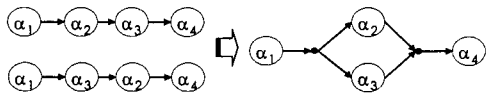

그림 7. AND-패턴

또한 표 4은 ' AND-패턴' 의 경우를 구분해 낼 수 있는 방법으로 한 프로세스 인스턴스의 액티비티들의 집합의 순서와 또 다른 프로세스 인스턴스의 액티비티들의 집합이 순서가 다를 때, 각 순서가 다른 액티비티들의 집합의 전후 액티비티가 같다면, AND 관련 액티비티들의 집합이라는 것을 나타내는 알고리즘이다.

표 4. AND-패턴을 구분하기 위한 알고리즘

```
//And-Pattern
If (exist difference of processes sequence?)
    If (exist before and after activity of different sequence of
     activities set?)
        Different sequence of activities set is ' and activities
         set'
    End If
End If
```

4. 결론 및 향후 연구방안

본 논문에서는 워크플로우 패턴을 기반으로 한 워크플로우 마이닝 기법에 관하여 제시하였다. 패턴기반 워크플로우 마이닝을 하기 위한 방법으로, ' 패턴기반 워크플로우 마이닝 알고리즘' 을 제시하였고, 그에 관련된 ' OR-패턴 구분 알고리즘' 과 ' AND-패턴 구분 알고리즘' 을 제시하였다.

향후 연구로는 ' AND-패턴' ,' OR-패턴' 뿐만 아니라, ' Loop' 에 관련된 패턴인식에 관한 알고리즘을 고려하고자 한다. 그리고, 본 논문에서 제시한 알고리즘을 중심으로 워크플로우 마이닝 시스템의 컴포넌트를 구성하고자 한다. 또한, 실제 워크플로우 프로세스 인스턴스 로그 정보를 기반으로 본 논문에서 제시한 알고리즘을 통해 패턴기반 워크플로우 프로세스 모델을 자동화하여 만들어보고자 하며, 프로세스의 개선에 얼마만큼 어떻게 도움이 될 수 있는지에 관하여 연구하고자 한다.

Acknowledgement

본 연구는 정보통신연구진흥원 정보통신 기초기술연구지원사업(03-기초-0005)의 지원으로 수행되었음.

참고문헌

[1] W. van der Aslst and et el. Workflow Mining: Which processes can be rediscovered? Technical Report, Department of Technology Management, Eindhoven University of Technology, 2002.

[2] Kwanghoon Kim and Clarence A. Ellis: Workflow Reduction for Reachable-path Rediscovery. Proceedings of the ICDM2003 WORKSHOP: Foundations and New Directions in Data Mining, Melbourne, Florida, USA, November, 2003

[3] 김광훈, "워크플로우 기술 I", TTA 저널, 85호, pp.107 - 118, 2003.01

[4] Workflow Patterns Home Page. http://www.tm.tue.tue.nl/it/research/patterns.