

워크플로우 도달가능경로 분석도구

박민재⁰ 원재강 김광훈
 경기대학교 전자계산학과 워크플로우 기술 연구실
 mean222@kyonggi.ac.kr

A Workflow Reachable-path Analyzer

Minjae Park⁰ Jaekang Won Kwanghoon Kim
 Dept. of Computer Science, Kyonggi University

요 약

비즈니스 프로세스의 정의 및 실행이 복잡화되고 대량화 되면서, 비즈니스 프로세스를 좀더 효율적으로 개선하고자 하는 노력이 필요하게 되었다. 본 논문에서는 그러한 노력을 위해 도달가능경로를 이용한 워크플로우 마이닝 기법을 사용하여, 문제를 해결하고자 한다. 따라서, 본 논문에서는 워크플로우 마이닝을 하기 위하여, ICN(Information Control Net) 모델을 기반으로 정의된 프로세스 모델을 도달가능경로를 찾아내기 위하여 WDN(Workflow Dependent Net)으로 구성을 한 후, MWN(Minimal Workflow Net)으로 변환하여, 도달가능경로를 구성할 수 있는 정보를 얻을 수 있도록 하는데 목적을 둔다.

1. 서론

최근 정보기술분야에 있어서 가장 두드러진 변화는 기존의 데이터중심의 정보기술에서 프로세스중심의 정보기술로 빠르게 전이된다는 사실이다. 이러한 정보기술 컴퓨팅환경의 변화는 웨어하우스 및 마이닝 기술에도 영향을 미치고 있다. 즉, 데이터를 중심으로 하는 마이닝 기술에서 프로세스를 중심으로 하는 마이닝 기술로의 전이와 함께 이에 대한 연구 및 개발의 필요성이 요구되고 있다. 따라서, 본 논문은 도달가능경로(Reachable-path)를 이용한 워크플로우 마이닝을 하기 위하여, 워크플로우 마이닝에 대한 개념과 ICN의 개념 그리고, 각 정보를 바탕으로 하여 워크플로우 마이닝을 하기 위한 개념과 방법에 관하여 기술하고 연구한다.

2. 연구배경 및 관련연구

2.1 워크플로우 마이닝의 필요성

비즈니스 프로세스를 모델링 하는 것은 비효율적인 프로세스, 시간 소비와 비용을 고려하는 것이다. 비즈니스 프로세스에 포함된 대부분의 사람들은 그들이 포함된 비즈니스 프로세스의 전반적인 구조에 관한 지식이 거의 없다. 그들은 전형적으로 최소한 자신의 로컬정보만을 알고 있다. 그래서, 존재하는 비즈니스 프로세스의 액티비티들 간 순서를 결정하는 것은 도전과 시간을 필요로 하는 작업이다.[3] 비즈니스 프로세스에서 액티비티의 순서를 발견하는 것은 반복되는 프로세스가 된다. 조합된 정보는 몇 번이고 포함된 사람에 의해 토의되고, 검증될 필요성이 있다. 심지어 올바른 액티비티의 순서가 결정이 되었

다 할지라도, 한 액티비티에서 다음 액티비티로의 제어 흐름과 관련된 비즈니스 규칙은 오류가 있을 수 있다. 종종 비즈니스 전문가도 어떤 도움 없이 전이 조건을 명세 한다는 것은 어려움이 있다. 이러한 이유로 본 논문에서는 이러한 비즈니스 프로세스를 효율적으로 개선하기 위해 워크플로우 마이닝의 개념을 도입한다.

2.2 ICN(Information Control Net)

본 논문에서 제시 하고자 하는 마이닝 기법은 ICN 모델을 기반으로 한다.

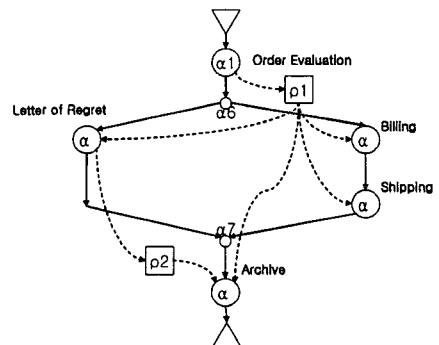


그림 1. ICN기반 주문처리 관계를 표현한 프로세스

ICN은 사무실(Office)의 개념을 일련의 관련된 프로시저(Procedures)의 집합으로 정의하며 이러한 프로시저는

선후관계가 존재하는 액티비티들의 집합으로 표현된다. ICN은 그림형태로 프로시저, 액티비티, 저장소(Repositories), 선후관계를 나타내는 제어흐름(Control Flow)과 데이터흐름(Data Flow)을 표현한다. ICN 제어흐름 그래프는 큰 원으로 표현되는 일련의 액티비티와 작고 빈 원으로 표현되는 OR노드, 작고 채워진 원으로 표현되는 AND노드, 그리고 이러한 노드들을 연결하는 선(edge)로 구성된다. 화살표(Arc)는 실선(Solid)과 점선(Dashed)으로 표현되는데 이들은 노드들간의 선후관계 및 자료저장소와의 입출력을 표현한다.

그림 1에서 표현하고 있는 전체 프로시저의 이름은 Order Process이며, 각 과정을 담당하는 액티비티는 Order Evaluation, Letter of Regret, Billing, Shipping, Archive($\alpha 1, \alpha 2, \alpha 3, \alpha 4, \alpha 5$)이다. 그리고 $\alpha 6, \alpha 7$ 은 OR분기를 나타내는 노드이며, $p 1, p 2$ 는 저장소를 나타낸다. 또한 직선형태로 된 실선은 제어의 흐름을 나타내고 곡선형태의 실선은 정보의 흐름을 나타낸다. [4]

3. 도달가능경로를 이용한 워크플로우 마이닝

본 논문에서 활용하고자 하는 워크플로우 마이닝 기법은 프로세스 정의 정보와 프로세스 실행 시 생성된 인스턴스 정보를 바탕으로 구축된 웨어하우스를 기반으로 한 도달가능경로를 이용한 워크플로우 마이닝 기법이다.

3.1 도달가능경로(Reachable-path)

실제 프로세스를 정의한 정보 중 실제 프로세스 실행 경로는 분기 등이 일어났을 경우 각 분기로 가는 경우의 수는 많이 다를 수 있다. 즉, 프로세스 실행 경로는 많은 경우에 한 두 측면에서만 집중적으로 일어날 수 있다. 그림 2의 경우와 같이 프로세스의 분기가 3가지 경우 인스턴스 발생이 이루어진 정보를 바탕으로 살펴보게 되면, 한 쪽 실행경로로만 프로세스의 흐름이 치우쳐져 있는 경우를 볼 수 있다. 이렇게 프로세스의 실행경로가 한쪽으로 치우쳐져 인스턴스가 발생 할 경우 프로세스 모델 정보 그대로 프로세스 인스턴스를 발생시키는 것은 매우 비효율적이다. 이런 경우 프로세스 모델 정보와 워크플로우 로그정보를 바탕으로 프로세스를 개선시킬 수 있다.

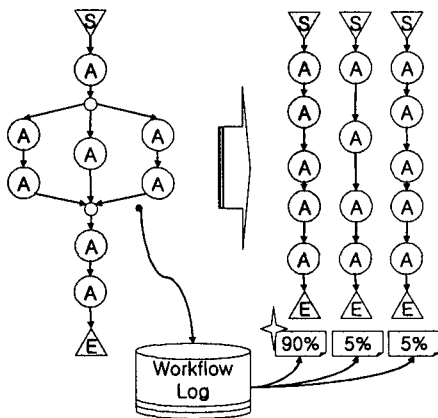


그림 2. 도달가능경로

3.2 WDN(Workflow Dependent Net)

도달가능경로를 이용한 워크플로우 마이닝을 하기 위하여, ICN모델을 바탕으로 WDN(Workflow Dependent Net)을 구성하는 일은 필수적이다.

3.2.1 WDN 자료구조

WDN을 구성하기 위한 최소한의 자료구조는 액티비티와 제어흐름 정보로 볼 수 있다. 각 정보는 ICN을 구성하고 있었던 정보와 상이하지만 WDN으로 구성하기 위해선 다음과 같은 자료구조의 형태 및 구조를 이루고 있는 정보가 반드시 필요하다. Activity 클래스는 각 액티비티의 아이디를 나타내는 activityID, 액티비티의 이름을 나타내는 activityName, 그리고 WDN을 구성할 때 트리의 분기를 구하고자 사용할 수 있는 액티비티의 타입 값인 activityType로 이루어진다. Transition 클래스는 트랜지션 클래스의 ID값인 transitionID와 트랜지션으로 제어되는 앞, 단 액티비티인 inputActivity와 outputActivity로 구성된다.

- 액티비티의 자료구조


```
class Activity{
    String activityID;
    String activityName;
    String activityType;
}
```
- 제어흐름의 자료구조


```
class Transition{
    String transitionID;
    Activity inputActivity;
    Activity outputActivity;
}
```

액티비티와 제어흐름은 위 자료구조를 바탕으로 구성된 모델링 정보의 제어흐름과 제어흐름의 입력 액티비티와 출력 액티비티의 상태를 고려하여 WDN의 구조를 구성한다.

3.2.2 ICN 기반의 WDN

WDN(Workflow Dependent Net)은 ICN의 컨디션 전이 정보를 중심으로 구성하게 된다. 조건 분기(or-split 과 or-join) 또는 병렬 분기(and-split 과 and-join)의 시점을 중심으로 각 분기점에서 노드의 깊이를 늘여가는 방법으로 구성한다. 각 액티비티와 제어흐름은 트리의 형태로 구성되게 되며, 분기된 노드와 노드의 깊이를 통해, 워크플로우 마이닝에 필요한 정보를 얻어내고자 한다.

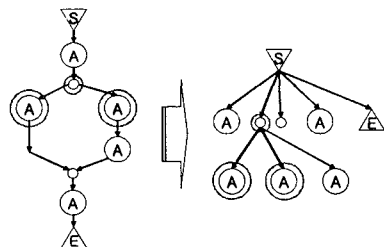


그림 3. Order-Process의 Workflow Dependent Net

위 그림 3은 ICN으로 구성되었던 Order-Process를 Workflow Dependent Net으로 변환한 것을 보여준다.

3.3 MWN(Minimal Workflow Net)

MWN은 ICN을 기반으로 한 프로세스 정의 정보를 가지고, WDN을 구성한 트리 중 실제로 분기 시점과 각 분기 시점에서 꼭 필요하다고 판단되었던 액티비티들만을 가지고 구성할 수 있다. 이전의 Order-Process로 예를 들면 ICN정보를 통해 WDN을 뽑아낸 정보를 통해, 깊이가 가장 깊었던, 즉 or-split 상태에서 뿔어 나온 두 액티비티를 중심으로 하여 MWN을 구성할 수 있다.

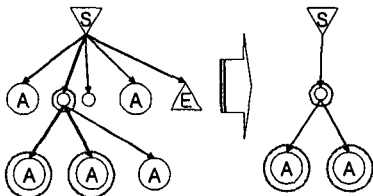


그림 4. Order-Process의 Minimal Workflow Net

그림 4는 3.2절의 그림 3에서 보여준 WDN을 이루고 있는 정보 중 or-split 정보를 중심으로 얻어낸 MWN을 보여준다.

4. 구현

다음 그림 5은 워크플로우 모니터링, 마이닝의 시나리오를 보여준다. 워크플로우 엔진은 프로세스 정의 정보를 실행시키면서 워크플로우 로그 정보를 남긴다. 남겨진 워크플로우 로그 정보는 각 정보와 웨어하우스에 정의된 매핑관계에 따라 매칭의 과정을 거쳐 워크플로우 웨어하우스를 구축하게 되고, 구축된 정보를 바탕으로 하여, 프로세스 마이닝을 할 수 있으며, 마이닝 정보와 프로세스 인스턴스 정보 자체를 통한 통계 정보 등을 얻을 수 있다.

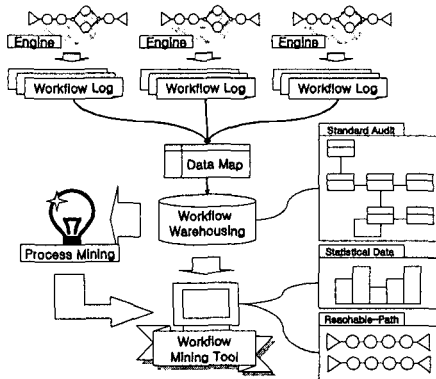


그림 5. 워크플로우 마이닝

Reachable-path를 고려한 워크플로우 마이닝을 하기 위해서는 비즈니스 프로세스를 모델링한 모델링 정보 및 모델링 된 정보를 해석할 수 있는 기술과 마이닝 알고리즘

을 적용하여 실제 구현할 수 있는 기술, 그리고 프로세스 개선에 도움이 되는 프로세스 인스턴스의 성능 평가등을 고려한 프로세스 데이터 웨어하우스 구축 기술 등이 필수적이다. 워크플로우 마이닝 도구는 ICN으로 모델링된 모델링 정보를 통하여, WDN을 구성하고, 구성된 WDN으로 프로세스 인스턴스의 상태정보를 알아내어 도달가능경로를 분석하는데 필요한 최소 정보인 MWN을 구성한다. 다음은 마이닝 도구에서 Order-Process의 ICN모델을 통해 WDN과 MWN을 구성하여 구현한 구현 예를 보여준다.

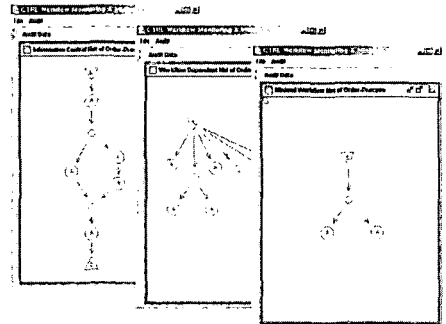


그림 6. Order Process의 ICN, WDN, MWN의 구현

5. 결론 및 향후 연구방안

본 논문에서는 워크플로우 마이닝의 필요성과 마이닝을 하기 위하여, ICN으로 모델링 된 프로세스 모델 정보를 가지고, WDN을 구성하는 자료구조와 구성 방법, 그리고 MWN을 구성하는 방법 등에 대하여 기술 하였고, 구성하였으며, 실제 구현한 구현 예를 보여주었다. 향후 연구로는 본 연구에서 기술한 기술들을 실제 도달가능경로를 얻어 낼 수 있는 기반으로 하여, 실제 적용 가능한 워크플로우 마이닝 시스템을 설계 및 구현하고자 한다.

Acknowledgement

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

참고문헌

- [1] Clarence A. Ellis and Gary J. Nutt, " The Modeling and Analysis of Coordination Systems" , University of Colorado/Dept. of Computer Science Technical Report, CU-Cs-639-93, Jan 1993
- [2] Kwang-Hoon Kim, Clarence A. Ellis, "Workflow Reduction for Reachable-path Rediscovery", IEEE Proceeding of the ICDM Workshop 2003, 2003.11
- [3] Rakesh Agrawal, Dimitrios Gunopulos Mining Process Models From Workflow Logs, IBM Almaden Research Center, San Joes, CA, USA
- [4] 류재광, 김광훈, "실시간 협업지원 그룹 ICN 에디터의 설계 및 구현", 한국인터넷정보학회 논문지, 2권 5호, pp.1 - 7, 2001.1