

워크플로우 마이닝을 위한 프로세스 변형 알고리즘

김학성⁰ 김광훈* 백수기*

동남보건대학 컴퓨터응용과⁰, 경기대학교 대학원 전자계산학과*
amang@dongnam.ac.kr⁰, {kwang, skpaik}@kyonggi.ac.kr*

A Design of Process Transformation Algorithm for Workflow Mining

Hak-Seong Kim⁰ Kwang-Hoon Kim* Su-Ki Paik*

Dept. of Computer & Application, Dongnam Health College⁰

Dept. of Computer Science, Kyonggi University*

요 약

본 논문에서는 워크플로우 마이닝을 위하여 프로세스 변형 알고리즘을 제안한다. 워크플로우 마이닝이란 워크플로우의 모니터링을 이용하여 워크플로우 시스템의 프로세스의 경향을 분석하여 또 다른 정보를 제공하는 기법을 의미한다. 이러한 워크플로우 마이닝을 위한 프로세스의 변형 알고리즘은 빌드 타임에서 정의된 워크플로우 프로세스의 흐름을 마이닝을 위한 형태로 변형하는 과정으로, 데이터의 흐름정보와 프로세스의 흐름정보를 이용하여 정의된 프로세스의 모든 경로를 트리 형태로 바꾸고, 워크플로우 실행 경로인 트리의 각 줄기를 분석하여 워크플로우 마이닝을 실현할 수 있다. 이와 같은 과정은 통하여 추출된 워크플로우 마이닝은 조직 또는 기관의 프로세서 경향을 추측하여 부가적인 정보를 제공 할 수 한다.+

1. 서 론

BPR(Business Process Reengineering)의 개념이 대두되면서부터 비즈니스 프로세서의 개선에 대한 요구가 한층 가열되었다. 이러한 요구로 인해 최근 워크플로우 기술은 사무정보 시스템과 데이터베이스 시스템 분야에서 관심의 대상이 되고 있다. 워크플로우 시스템은 인간, 정보, 컴퓨터 응용 프로그램들 간에 상호작용을 조정해 비즈니스 프로세서를 자동화하여 업무흐름과 프로세서 연관관계를 지원하는 시스템을 의미한다. 또한, 워크플로우 관리 시스템은 사전의 분석과 요약, 정보 중심의 작업과 액티비티의 자동화에 대한 도구의 집합이다[1].

최근 워크플로우 개발 환경의 동향을 살펴보면, 두 가지로 요약될 수 있는데, 첫째는 네트워크에서의 개인 컴퓨팅 환경이 강화되고, 웹과 같은 분산환경에서 수행됨에 따라 워크플로우 시스템은 제어 기술, 데이터 조작기술에서 분산 구조를 갖게 되었고 둘째로 워크플로우 응용분야는 대규모화되고 복잡해짐에 따라 관련된 데이터가 증가하고 시스템에 대한 사용자의 서비스 요구가 증가하고 있다.

최근 B2C/B2B로 대변되는 전자상거래 및 전자시장(E-Market Place)의 활성화가 급속하게 확장됨에 따라 워크플로우를 기반으로 한 기업 내부 업무 흐름의 자동화를 의미하는 B2E의 구축을 더욱 가속화시키고 있으며, 워크플로우 기술 및 시스템은 결국 기업 또는 조직체내의 모든 업무 처리 절차들을 통합 관리하는 인프라구조(Infrastructure)로 인식되고 있다[1,2].

이러한 변화는 단순히 프로세스의 상태정보나 감시정보만을 제공하는 워크플로우 모니터링의 기능에 있어서도 변화가 요구된다. 즉, 전자상거래에서 모니터링의 역할은 업무 프로세스의 분석을 통하여 기업이나 고객에게 또 다른 정보를 제공할 수 있어야 한다.

이러한 기능은 워크플로우 마이닝을 통하여 제공될 수 있고 워크플로우 마이닝은 프로세스 변형, 프로세스 모니터링, 프로세스 분석, 데이터 추출의 단계와 같이 4단계로 이루어진다. 본 논문에서는 워크플로우 마이닝을 위한 첫 번째 단계로서 Build Time에서 정의된 프로세스를 변형하기 위한 알고리즘을 제안한다.

본 논문의 구성은 2장에서는 관련 연구로 워크플로우 모니터링에 관한 사항을 언급하고, 워크플로우 프로세스 모니터링에서 추가로 수집해야 하는 정보를 제시하고, 3장에서는 워크플로우 마이닝의 필요성에 대하여 기술하고 4장에서는 마이닝을 하기 위한 프로세스의 변형 과정 및 알고리즘을 제시하고, 5장에서 결론을 맺는다.

2. 관련연구

2.1 프로세스 모니터링

프로세스 모니터는 특정 프로세스 모델의 현재 또는 축적된 프로세스의 상태를 나타낸다. 각 프로세스 인스턴스에 대해 액티비티들의 상태나 관련 정보들을 모니터링 한다. 그 형태는 액티비티의 처리시간의 양, 현재 진행중인 프로세스의 수, 액티비티당 작업의 최소, 최대량 같은 것이다.

2.2 워크플로우 모니터링

일반적으로 워크플로우 모니터링이라 함은 워크플로우 시스템에서 일어나는 이벤트를 추적하고, 기록하는 능력을 말한다. 워크플로우 모니터링 서비스는 시스템 수준의 모니터링부터 사용자 수준 혹은 응용 프로그램 수준의 모니터링 서비스까지 매우 다양하다.

2.3 워크플로우 모니터링 기능

워크플로우 모니터링 기능은 정보 처리부, 모니터링 정보 수집부, 모니터링 정보 표현부와 같이 3가지 기능

을 수행한다. 모니터링 정보 수집부는 워크플로우 실행 부에게 정보를 요청하고, 상태정보나 이벤트 정보를 반환 받는 인터페이스를 가진다. 정보 처리부는 수집된 정보 중 처리가 요구되는 정보를 처리하거나 분석하는 역할을 담당한다. 마지막으로, 정보 표현부는 모니터링 서비스를 받는 곳에서 쓰인다. 수집되거나 처리된 정보를 사용자에게 보다 인식되기 쉬운 방법으로 표현하는 역할을 담당한다[3].

3. 마이닝의 필요성

3.1 비즈니스 프로세스 모델링

비즈니스 프로세스를 모델링하는 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 액티비티들간 순서를 결정하는 것은 도전과 시간을 필요로 하는 작업이다[4].

비즈니스 프로세스에서 구성원과의 면접 같은 기술들은 로컬 정보를 수집하는데 사용되어야 한다. 이러한 로컬 정보는 비즈니스 프로세스의 전반적인 구조를 포함하는 포괄적인 정보로 옮기고, 작업들의 전체적인 순서를 정한다.

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 되었다 할지라도, 한 액티비티에서 다음 액티비티로의 제어 흐름과 관련된 비즈니스 규칙은 오류가 있을 수 있다. 우리는 종종 비즈니스 전문가도 어떤 도움 없이 전이 조건을 명세 한다는 것은 어려움이 있다.

이러한 이유로 본 논문에서는 이러한 비즈니스 프로세스를 정확히 정의하기 위해 워크플로우 마이닝의 개념을 도입하였다.

3.2 워크플로우 시스템의 동작 과정

워크플로우의 경향을 분석하기 위해 워크플로우 시스템의 동작에 대해 자세히 살펴본다. 워크플로우 시스템은 그림 1과 같이 동작한다.

정의된 워크플로우는 워크플로우 관리 시스템으로 인스턴스 하나를 발생한다. 발생된 인스턴스는 각 액티비티를 수행하기 위해 수행자에게 작업을 할당하고 할당된 작업이 종료가 되면 다시 시스템으로 그 정보를 넘겨준다. 이것을 반복해서 한 인스턴스의 마지막 액티비티가 종료되면 해당 프로세스 인스턴스는 종료가 되고 이 과정에 수행자는 시스템은 저장소에 현재 진행되고 있는 인스턴스에 대해 감시 추적 정보를 저장하게 된다. 이러한 정보는 수행자가 수행 도중에 현 프로세스에 대한 정보를 얻고자 할 때 모니터링을 이용하여 얻을 수 있고, 프로세스 디자이너나 시스템 관리자는 전체 프로세스에 대한 추적 정보를 요구할 수도 있다[3].

현재 대부분의 모니터링은 시스템 또는 프로세스 관리자와 수행자를 위한 모니터링 시스템이다. 프로세스 디자이너에게 유용하게 사용될 데이터를 제공하기 위해서는 워크플로우 마이닝을 통해 프로세스의 경향을 분석하여 이상적인 프로세스를 정의할 수 있도록 설계를 하여야 한다.

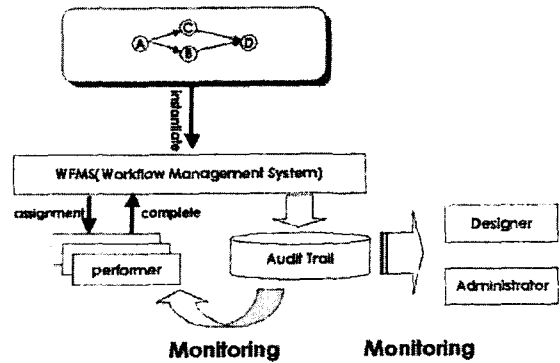


그림 1 워크플로우 시스템 동작과정

4. 워크플로우 마이닝을 위한 프로세스 변형 알고리즘

4.1 워크플로우 마이닝 요구 데이터

워크플로우 마이닝을 위한 데이터는 모니터링 정보 즉 추적 감시 정보만으로 충분치 않다. 추적 감시 정보만으로 프로세스의 경향을 찾기 어렵기 때문에 본 논문에서 워크플로우 관리 시스템에서 끌어올 데이터를 두 가지의 형태로 나누었다. 즉, 워크플로우의 데이터 흐름 정보를 의미하거나, 워크플로우 내부에서 일어나는 관련데이터 (Relevant Data)를 의미하는 내부적인 데이터와, 워크플로우의 프로세스 흐름 정보를 의미하는 외부데이터이다. 즉 외부데이터는 외부에서 일어나는 전이정보를 의미하는데 이는 프로세스가 시작하기 전에 모든 경로에 대한 정보를 저장하여, 어떠한 경로를 통해 워크플로우가 진행이 되었는지에 대해 정보를 수집하고 저장하고 있다.

4.2 워크플로우 마이닝 단계

워크플로우에서 발생하는 데이터를 바탕으로 프로세스에 관련된 새로운 데이터를 발생하는 과정으로 이를 위해 앞에서 정의된 데이터를 바탕으로 알고리즘을 통해 새로운 데이터를 생성하고, 워크플로우 로그 정보와 비교하여 프로세스에 관련된 새로운 데이터를 생성한다. 그림 2는 워크플로우 마이닝의 전체적인 과정을 보여주고 있다.

4.3 프로세스 변형 알고리즘 설계

프로세스의 변형은 빌드 타임에서 정의된 워크플로우의 흐름을 마이닝을 위한 형태로 변형하는 것이다. 마이닝의 목적은 프로세스의 경향을 찾아내는 것이다. 이는 모든 프로세스의 경로를 찾아내는 것이 중요하다. 따라

서 정의된 프로세스를 모든 경로로 나열하고자 한다. 따라서 프로세스를 트리 형태로 바꾸고, 트리의 한 가지들 한 경로로 찾는 변형(Transformation)을 하고자 한다.

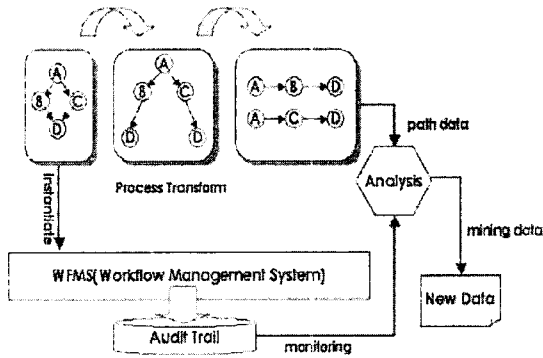


그림 2 워크플로우 마이닝 단계

트리는 노드 구조로 표현되는데, 다음 표 1과 같이 노드는 노드의 이름, 부모 노드, 자식 노드로 구성된다.

표 1 노드 구조체

Node {		
String	name	//node이름
Node	parent	//부모 이름
NodeSequence	children	//자식노드의 집합
} // 구조체		

프로세스에 나타나는 액티비티의 종류는 시작 액티비티, 끝 액티비티, 일반 액티비티, 분기 액티비티, 집합 액티비티이다. 따라서 트리를 구성하기 위하여 이 액티비티의 종류에 따라 법칙을 세워야 한다.

시작 액티비티는 트리의 시작을 의미하므로, 항상 루트를 구성하고, 끝 액티비티는 최종노드(leap node)를 구성하게 된다. 일반 액티비티는 경로를 찾기 위해서는 불필요하고, 또한 트리의 복잡도를 줄이기 위해 생략한다. 분기 액티비티는 경로를 결정하는 중요한 액티비티이다. 분기 액티비티의 차수는 분기되는 경우의 수를 의미하며, 분기수 만큼 반복하여 끝 액티비티까지 검색하고, 트리를 구성한다.

표 2 트리 변형 알고리즘

ProcessToTree (Input Activity A, Node N)	
begin	
if A is start Activity	
begin	
set A to N	

```

N is ProcessToTree(Next, N)
end
else if A is normal Activity or join Activity
begin
  N is ProcessToTree(Next, N)
end
else if A is split Activity
begin
  for 1 to the number of split do
  begin
    set Nextto N1
    parent of N1 is N
    N1 is ProcessToTree(Next of Next, N1)
    add N1 to children of N
  end
end
else if A is end Activity
begin return N end
return N
end
    
```

5. 결 론

본 논문은 워크플로우 모니터링의 응용 분야로 워크플로우 마이닝에 대한 기법을 소개하고 워크플로우 마이닝을 위하여 프로세스 변형 알고리즘을 제안 하였다.

향후 연구과제로서 빌드타임에서 정의된 프로세스를 본 논문에서 제안된 알고리즘을 이용하여 트리 형태로 변형 후, 각 경로를 나열한 다음, 모니터링을 통해 추출된 추적 정보와 비교 분석하여, 프로세스의 경향을 추출하는 것이다. 이 정보는 워크플로우를 정의하는 사람에게 유용한 정보로 쓰이며, 더 나아가 전자상거래의 CRM에도 이용할 수 있는 정보로 변형이 된다.

6. 참고문헌

[1]Frank Leymann, Dieter Roller, Production Workflow Concepts and Techniques, Prentice Hall PTR 2000

[2]Thomas M. Koulopoulos, The Workflow Imperative, published by Van Nostrand, a division of International Thomson Publishing Inc, 1995

[3]H S Hong , B S Lee, K H Kim, S K Paik, A Web-based Transaction Workflow Monitoring System WISE 2000 in Hong Kong Proceeding , IEEE PRESS

[4]Rakesh Agrawal, Dimitrios Gunopulos Mining Process Models From Workflow Logs, IBM Almaden Research Center, San Joes, CA, USA