워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 방법☆

An Activity Analysis Method for Workflow Visual Verification and Mining

김 광 훈*
Kwang-Hoon Kim

요 약

본 논문에서는 워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 방법을 제안하고, 제안한 액티비티 분석 방법과 관련된 액티비티 분석 도구를 구현한다. 현재 워크플로우 및 비즈니스 프로세스와 관련된 주요한 연구 중 하나는 기존 비즈니스 내 존재하는 프로세스의 문제점을 찾아내고, 개선하는데 있다. 이러한 문제점을 찾아내고 해결하기 위해 제안하는 액티비티 분석 방법은 특정 액티비티가 변경될 때, 임의의 액티비티와 다른 액티비티와의 연관성을 고려하고자 하는 것이다. 워크플로우 표현 방법으로는 ICN(Information Control Nets)을 활용한다. 마지막으로, 본 방법 및 도구의 적용 가능성을 증명하기 위해, 어느 한 특정 기업의 전자결재 프로세스의 경우를 토대로 적용하고자 한다.

Abstract

This paper conceives and implements an activity analysis method as a tool to be used for workflow visual verification and mining. One of the recent issues in the workflow and business process literature is to refine and to improve the deployed workflows and business processes. The activity analysis method proposed in this paper provides a way to fine a set of activities being directly affected by the specific activity that a user tries to change its properties. I would strongly believe that the method can be a useful solution for the dynamic changes and visual verifications problems of workflow models as well as the workflow process mining problems. Finally, to prove the possibility of the proposed method and its applicability, we apply to a workflow model of the electronic approval system run by a real corporation.

1. 서 론

최근 몇 년 동안 국내외의 많은 기업들은 워크플로우 및 BPM(Busniness Process Management)을 도입하였고, 그로 인해, 비즈니스 프로세스의 가시화와 투명화, 비즈니스 프로세스 실행비용의 감소, 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 정의정보가 워크플로우 제정 서비스(Workflow Enactment Service)의 실행할 때 항상 최적화 되어 있거나, 효율적이지 못할 수 있다는 가능성이 제시되면서, 워크플로우 모델 자체를 변경하는 일들이 일어나게 되었다.

현재 워크플로우의 재발견 문제가 이슈화 되면서, 워크플로우 모델 변경에 대한 방법등은 제시되고 있으나, 워크플로우 모델이 변경 시 발생할 수 있는 문제를 피할 수 있는 방법과 관련된 연구는 매우 미흡한 편이다.

본 논문에서는 워크플로우 프로세스 내에 존재

* 종신회원 : 경기대학교 컴퓨터과학과 교수
kwang@kgu.ac.kr
[2008/01/23 투고 - 2008/01/28 심사 - 2008/06/02 심사완료]
☆ 본 논문은 2006학년도 경기대학교 학술연구비 지원에 의하여 수행되었음

하고 있는 임의의 액티비티가 변경될 때, 워크플로우의 제어 흐름에 어떠한 영향을 미칠 수 있는지에 대해 연구하였고, 워크플로우 내 존재하는 액티비티 중 변경될 소지가 있는 임의의 액티비티를 선택하여, 그 액티비티가 제어 흐름에 따라 영향을 줄 수 있는 액티비티가 무엇인지 파악할 수 있는 모델과 알고리즘을 제안하였으며, 액티비티 분석 도구의 구현을 통해 보여준다.

2. 관련연구

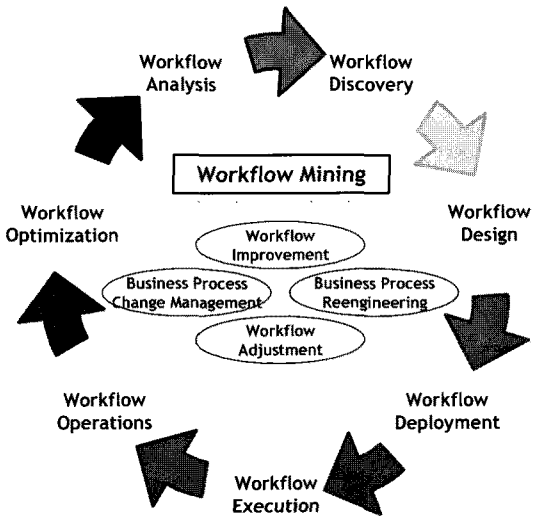
2.1 워크플로우 마이닝 기술

본 연구의 궁극적인 목적은 워크플로우 재 정의에 있으며, 워크플로우 재 정의를 하는 목적은 비효율적인 워크플로우를 정형화 시켜서 분석하고, 개선하는데 있다. 이러한 워크플로우 개선 방법과 관련하여, 워크플로우 마이닝 방법과 관련된 연구가 있었고, 프로세스 모델을 정형적으로 표현해 내기 위해, 여러 가지 방법이 있었으나, 본 연구에서는 ICN 워크플로우 모델링 방법을 사용하여 표현한다.

우리가 본 연구를 통해 목적을 이루고자 하는 워크플로우 마이닝 기술[6, 9, 13, 16]은 많은 연구를 통해 점차 기술적 이론적 정리가 되어가는 실정이다. 그간 연구된 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 1) 워크플로우 마이닝 과정

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 ICN 워크플로우 모델

본 연구에서 워크플로우 마이닝을 위해 워크플로우 모델을 재발견 해 내기 위해, 워크플로우 모델 표현 방법과 관련된 이론으로 ICN(Information Control Nets) 워크플로우 모델링 방법을 이용하고

자 한다. 본 연구에서 기본적으로 사용하는 ICN 워크플로우 모델[1]은 워크플로우를 표현할 수 있는 기호들의 집합으로 기본적으로 액티비티, 저장소, 역할, 수행자, 제어흐름 조건 및 입출력 자료저장소로 구성된다. ICN 워크플로우 모델의 7 튜플은 $\Gamma = (\delta, \gamma, \varepsilon, \pi, \kappa, I, O)$과 같이 표현되며, 각각의 정의는 다음과 같다.

- $\delta = \delta_i \cup \delta_o$
  여기서, $\delta_o : A \rightarrow \mathcal{P}(A)$은 하나의 액티비티를 후행하는 액티비티들의 집합에 연결하는 관계를 나타내며 $\delta_i : A \rightarrow \mathcal{P}(A)$은 하나의 액티비티를 선행하는 액티비들의 집합에 연결하는 관계를 나타내는 관계이다.

- $\gamma = \gamma_i \cup \gamma_o$
  여기서, $\gamma_O : R \rightarrow \mathcal{P}(A)$은 하나의 액티비티를 후속하는 액티비티 집합들을 출력 저장소들의 집합과 연결하는 것 중 하나이며, $\gamma_i : R \rightarrow \mathcal{P}(A)$은 하나의 액티비티를 선행하는 액티비티 집합들을 입력 자료저장소들의 집합과 연결하는 관계를 나타내는 것 중 하나이다.

- $\varepsilon = \varepsilon_a \cup \varepsilon_p$
  여기서, $\varepsilon_a : P \rightarrow \mathcal{P}(A)$는 하나의 액티비티를 조합된 역할들의 집합에 액티비티를 단일 값으로 연결하는 것이고, $\varepsilon_p : A \rightarrow \mathcal{P}(P)$는 조합된 액티비티들의 집합을 역할에 단일 값으로 연결하는 것이다.

- $\pi = \pi_p \cup \pi_c$
  여기서, $\pi_p : C \rightarrow \mathcal{P}(P)$ 조합된 수행자들의 집합을 역할에 단일 값으로 연결하는 것이고, $\pi_c : P \rightarrow \mathcal{P}(C)$는 조합된 역할들을 수행자에 단일 값으로 연결하는 것이다.

- $\kappa = \kappa_i \cup \kappa_o$
  여기서, κ_i는 $a \in A$일 때, $(\delta_i(a), a)$ 사이에서, 제어흐름(T) 조건들의 집합이고, κ_o는 $a \in A$일 때, $(a, \delta_o(a))$ 사이에서, 제어흐름(T) 조건들의 집합이다. 이때 집합 T = {default, or(conditions), and(conditions)}.

- I 는 초기 입력되는 자료 저장소들의 유한 집합이며, ICN의 실행 전에 어떠한 외부의 프로세스에 의해서 로드 되어야 한다.

- O 는 마지막으로 출력되는 자료저장소들의 유한 집합이며 ICN의 실행 후에 어떠한 외부의 프로세스에 의해서 이용되는 정보를 포함하고 있다고 가정한다.

2.3 워크플로우 마이닝 기술 동향

워크플로우 마이닝과 관련된 연구 및 개발은 여러 가지 접근법으로 이루어지고 있다. 워크플로우 마이닝 알고리즘과 관련하여, 워크플로우 마이닝 기술은 연구는 [3, 5, 8, 11, 12, 13, 14]와 같은 곳에서 진행되고 있으며, 워크플로우 마이닝 시스템이나 워크플로우 마이닝 도구를 개발하는 측에 중심을 두고 개발한 연구는 [2, 4, 10]에서 이루어지고 있다. [2, 4, 10]의 연구는 워크플로우 감사 데이터 정보를 바탕으로 알고리즘을 적용하여 워크플로우 재발견과 관련된 연구를 진행하고 있고, 본 연구에서는 워크플로우 워크플로우 모델을 통한 재발견 알고리즘을 개발하고, 개발된 알고리즘을 구현을 통해 실제 프로세스에 적용하고자 하고 있다. 본 연구 목표과 비슷한 연구로 워크플로우 마이닝 기술을 처음으로 산업에 적용한 J. Herbsta와 D. Karagiannis는 [4]에서 InWoLvE워크플로우 마이닝 시스템의 실험과 경험으로 주요한 결과를 보여주었다. 그러나, 거의 대부분은 아직 워크플로우 마이닝 기술의 기본 기능들만을 개발하는데 초점을 맞추고 있다. 네델란드 아인트호벤 공과대학의 W.M.P. van der Aalst 교수의 연구 그룹은 논문[13]과 [14]를 통해, 워크플로우 디자인을 지원하기 위한 워크플로우 마이닝 접근법과 알고리즘과 관련된 워크플로우 마이닝의 기본 정의 및 사용을 언급하였다. 그들은 워크플로우 감사 데이터를 통해, 워크플로우의 원래 모습을 찾

고자 하는 연구를 주로 많이 하였다. 또한, 워크플로우 모델링을 하는데 있어, 잘 알려진 모델링 방법인 Petri-Nets을 기반으로, 워크플로우의 기본 모델을 제시하고자 하였다. 물론 잘 알려진 모델링 방법인 Petri-Nets은 워크플로우 프로세스를 모델링하기에 좋은 방법이지만, 본 연구에서는 워크플로우에 좀 더 특화되고, 비교적 그래픽적 표현 방법이 간단하며, 정형적 표현 방법으로, 새로운 방법론을 적용하기에 용이한 ICN 모델링 기법을 사용하였다. [14]에서는 블록 기반의 워크플로우 모델링을 하였다. Pinter와 Goani[11]는 프로세스의 두개의 액티비티가 발생할 때의 정보와 관련해서, 시간 정보를 활용하였다. Cook el al.[5]은 은 통계 정보를 통한 프로세스 마이닝 접근법을 연구하였다. 우리는 앞으로 본 논문에서 제시한 방법과 앞서 언급한 연구자들의 워크플로우 마이닝 접근법의 장단점을 잘 생각하여, 좀 더 효율적인 워크플로우 마이닝 방법을 비교, 개선해 나아가여 할 것이다.

3. 워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 방법

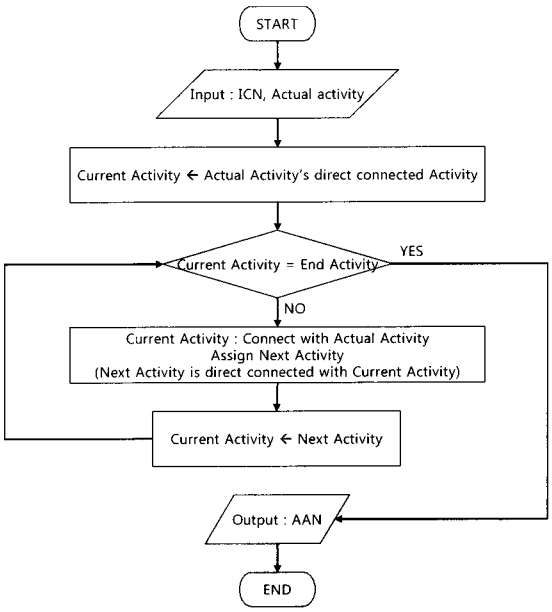
본 연구에서 우리는 워크플로우 프로세스 내에 존재하는 임의의 액티비티와 같은 워크플로우 프로세스 내에 존재하는 액티비티들 간에 존재하는 제어흐름 정보를 파악하고, 분석하여, 워크플로우를 재발견해 내고자 한다. 액티비티간에 존재하는 제어흐름 정보를 재발견해 내기 위한 방법으로 액티비티 분석 알고리즘인 AA(Activity Analysis)-알고리즘을 제안하고, AA-Nets(Activity Analysis Nets) 모델로 정형화 시켜 표현한다. 또한 제안된 알고리즘을 구현하여 구현 예로써 보여준다.

3.1 워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 방법

본 연구는 기본적으로 워크플로우 마이닝을 위한 액티비티 분석 방법을 제안하고, 구현한다. 본 연구에서 제안하는 액티비티 분석 방법의 기반은 AA-알고리즘과 AA-Nets 이면 각각은 다음과 같이 나타낼 수 있다.

3.1.1 AA-알고리즘

본 연구를 통해 제안하고 있는 간단한 액티비티 분석 알고리즘인 AA(Activity Analysis)-알고리즘은 그림 2에서 표현한 바와 같이 매우 간단한 흐름으로 이루어져 있다. 입력 값으로 분석하고자 하는 ICN 모델과 ICN 모델내 존재하는 임의의 액티비티를 갖는다. 입력 값으로 갖는 임의의 액티비티는 제어흐름 정보를 파악하기 위한 즉, 워크플로우 리모델링 또는 리디자인 등의 과정을 거치고자 할 때, 임의의 한 액티비티가 제어흐름상 프로세스 내 존재하는 다른 액티비티와의 관계를 파악해야하는데, 그 과정에서 분석하게 될 액티비티이다. 이렇게 두가지의 입력 값을 가지고 기본적인 시작을 하게 되며, ICN 모델 내에서 선정된 한 액티비티는 직접적으로 제어흐름과 관련



(그림 2) AA(Activity Analysis)-알고리즘

된 다음 액티비티를 연결하고, 연결된 액티비티를 통해 제어흐름을 파악하고, 연결해 나아가 마지막 액티비티까지 검색하게 되면, 선정된 임의의 액티비티와 제어흐름상 관련된 모든 액티비티를 찾아낼 수 있는 AA-Nets(Activity Analysis Nets) 모델을 구성해 낼 수 있다.

3.1.2 AA-Nets

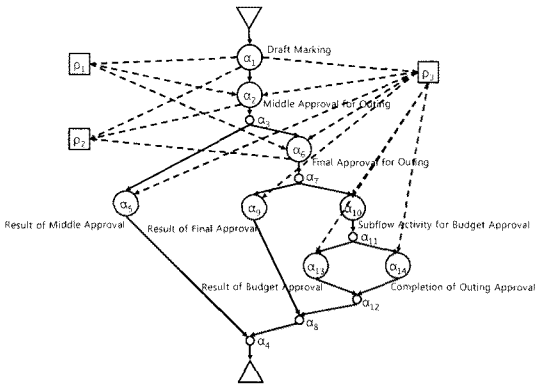
본 연구를 통해 제안하고 있는 AA-알고리즘을 통해 얻어진 워크플로우 재발견 모델을 우리는 본 연구를 통해 마찬가지로 제안하고 있는 AA-Nets 으로 표현해 내고자 한다. AA-Nets 은 T = {σ, τ, I, O}와 같이 4개의 튜플로 이루어져 있고, 각각 다음과 같이 표현할 수 있다.

- $\sigma = \sigma_i \cup \sigma_o$
  여기서, σ_i : A→℘(A)는 분석된 엑티비티를 선택한 엑티비티의 집합에 연결하는 관계를 나타내며 σ_o : A→℘(A)는 선택한 액티비티와 연결되는 엑티비티들의 집합에 연결하는 관계를 나타낸다.
- $\tau = \tau_i \cup \tau_o$
  여기서, τ_i는 각 호 $(\sigma_i(a), a)$ 연관 관계의 집합이며, τ_o는 각 호 $(a, \sigma_o(a))$ 연관 관계의 집합이다.
- I 는 초기 입력되는 자료 저장소들의 유한 집합이다.
- O 는 마지막으로 출력되는 자료저장소들의 유한 집합이다.

3.2 워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 방법 적용 예

3.2.1 특정기업의 전재결재 프로세스 모델

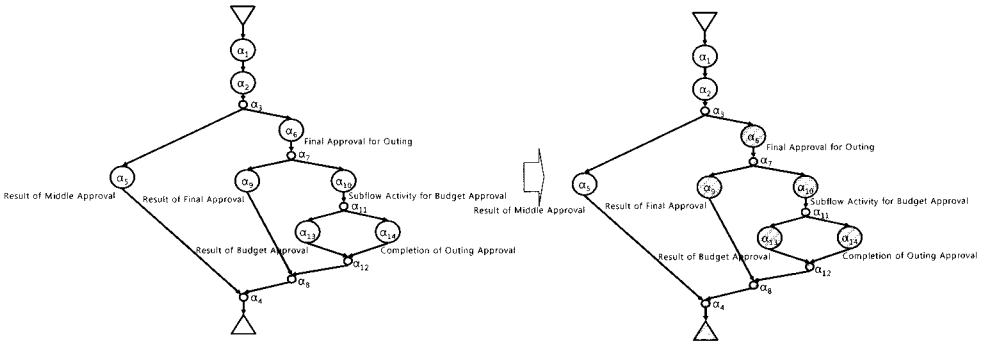
본 논문에서는 워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 방법의 적용 예제로써 특정기업의 전자결재(간이 출장) 프로세스[17] 모델을 중심으로 설명하고자 한다. 전자결재 프로세스 ICN 모델의 시각적 표현은 다음 그림 3과 같이 할 수 있고, 정형적 표현은 표 1과 같이 할 수 있다.



(그림 3) 특정기업의 전자결재(간이 출장) 프로세스의 시각적 ICN 표현

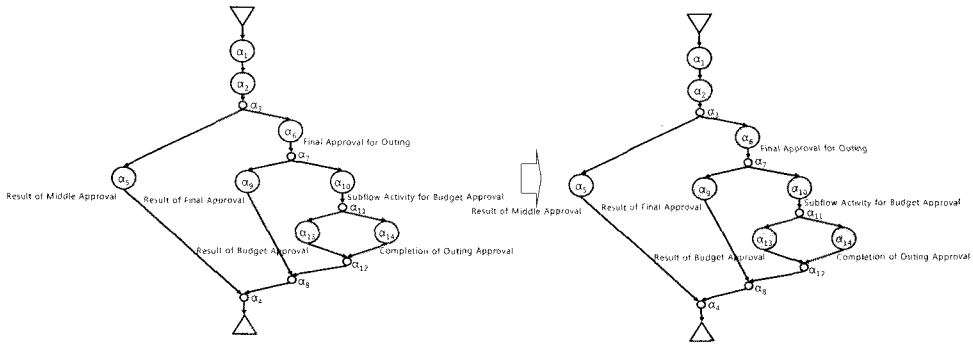
(표 1) 특정기업의 전자결재(간이 출장) 프로세스의 정형적 ICN 표현

Γ = (δ, γ, ε, π, κ, I, O) over A, R, P, C, T	//ICN
A = {α_1, α_2, α_3, α_4, α_5, α_6, α_7, α_8, α_9, α_{10}, α_{11}, α_{12}, α_{13}, α_{14}, α_I, α_F}	//Activities
R = {ρ_1, ρ_2, ρ_3}	//Repositories
P = {ω_1, ω_2, ω_3, ω_4, ω_5}	//Roles
C = {η_1, η_2, η_3, η_4, η_5, η_6, η_7, η_8, η_9, η_{10}}	//Actors
T = {d(default), or_1('reject'), or_2('accept')}	//TransitionConditions
I = {0}	//InitialInputRepositories
O = {0}	//FinalOutputRepositories

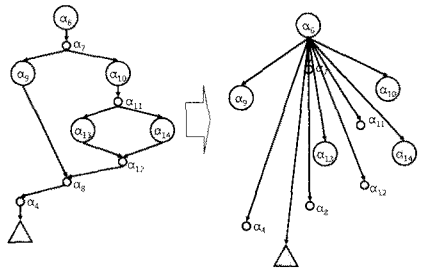


(그림 4) 특정기업의 전자결재(간이출장) 프로세스의 제어흐름 중심 ICN 표현 및 α_6 액티비티와 제어흐름상 연관된 액티비티

$\delta = \delta_i \cup \delta_o$	$\delta_i(\alpha_1) = \{0\}$; $\delta_i(\alpha_1) = \{\alpha_1\}$; $\delta_i(\alpha_2) = \{\alpha_1\}$; $\delta_i(\alpha_3) = \{\alpha_2\}$; $\delta_i(\alpha_4) = \{\alpha_5, \alpha_8\}$; $\delta_i(\alpha_5) = \{\alpha_3\}$; $\delta_i(\alpha_6) = \{\alpha_3\}$; $\delta_i(\alpha_7) = \{\alpha_6\}$; $\delta_i(\alpha_8) = \{\alpha_9, \alpha_{12}\}$; $\delta_i(\alpha_9) = \{\alpha_7\}$; $\delta_i(\alpha_{10}) = \{\alpha_7\}$; $\delta_i(\alpha_{11}) = \{\alpha_{10}\}$; $\delta_i(\alpha_{12}) = \{\alpha_{13}, \alpha_{14}\}$; $\delta_i(\alpha_{13}) = \{\alpha_{11}\}$; $\delta_i(\alpha_{14}) = \{\alpha_{11}\}$; $\delta_i(\alpha_F) = \{\alpha_4\}$	$\delta_o(\alpha_1) = \{\alpha_1\}$; $\delta_o(\alpha_1) = \{\alpha_2\}$; $\delta_o(\alpha_2) = \{\alpha_3\}$; $\delta_o(\alpha_3) = \{\alpha_5, \alpha_6\}$; $\delta_o(\alpha_4) = \{\alpha_F\}$; $\delta_o(\alpha_5) = \{\alpha_4\}$; $\delta_o(\alpha_6) = \{\alpha_7\}$; $\delta_o(\alpha_7) = \{\alpha_9, \alpha_{10}\}$; $\delta_o(\alpha_8) = \{\alpha_4\}$; $\delta_o(\alpha_9) = \{\alpha_8\}$; $\delta_o(\alpha_{10}) = \{\alpha_{11}\}$; $\delta_o(\alpha_{11}) = \{\alpha_{13}, \alpha_{14}\}$; $\delta_o(\alpha_{12}) = \{\alpha_8\}$; $\delta_o(\alpha_{13}) = \{\alpha_{12}\}$; $\delta_o(\alpha_{14}) = \{\alpha_{12}\}$; $\delta_o(\alpha_F) = \{0\}$;
$\gamma = \gamma_i \cup \gamma_o$	$\gamma_i(\alpha_1) = \{0\}$; $\gamma_i(\alpha_1) = \{0\}$; $\gamma_i(\alpha_2) = \{\rho_1, \rho_3\}$; $\gamma_i(\alpha_3) = \{0\}$; $\gamma_i(\alpha_4) = \{0\}$; $\gamma_i(\alpha_5) = \{\rho_3\}$; $\gamma_i(\alpha_6) = \{\rho_3\}$; $\gamma_i(\alpha_7) = \{0\}$; $\gamma_i(\alpha_8) = \{0\}$; $\gamma_i(\alpha_9) = \{\rho_3\}$; $\gamma_i(\alpha_{10}) = \{\rho_3\}$; $\gamma_i(\alpha_{11}) = \{0\}$; $\gamma_i(\alpha_{12}) = \{0\}$; $\gamma_i(\alpha_{13}) = \{\rho_3\}$; $\gamma_i(\alpha_{14}) = \{\rho_3\}$; $\gamma_i(\alpha_F) = \{0\}$;	$\gamma_o(\alpha_1) = \{0\}$; $\gamma_o(\alpha_1) = \{\rho_1, \rho_2, \rho_3\}$; $\gamma_o(\alpha_2) = \{\rho_2\}$; $\gamma_o(\alpha_3) = \{0\}$; $\gamma_o(\alpha_4) = \{0\}$; $\gamma_o(\alpha_5) = \{0\}$; $\gamma_o(\alpha_6) = \{\rho_2\}$; $\gamma_o(\alpha_7) = \{0\}$; $\gamma_o(\alpha_8) = \{0\}$; $\gamma_o(\alpha_9) = \{0\}$; $\gamma_o(\alpha_{10}) = \{0\}$; $\gamma_o(\alpha_{11}) = \{0\}$; $\gamma_o(\alpha_{12}) = \{0\}$; $\gamma_o(\alpha_{13}) = \{0\}$; $\gamma_o(\alpha_{14}) = \{0\}$; $\gamma_o(\alpha_F) = \{0\}$;
$\varepsilon = \varepsilon_a \cup \varepsilon_p$	$\varepsilon_p(\alpha_1) = \{0\}$; $\varepsilon_p(\alpha_1) = \{\omega_1\}$; $\varepsilon_p(\alpha_2) = \{\omega_2\}$; $\varepsilon_p(\alpha_3) = \{0\}$; $\varepsilon_p(\alpha_4) = \{0\}$; $\varepsilon_p(\alpha_5) = \{\omega_3\}$; $\varepsilon_p(\alpha_6) = \{\omega_4\}$; $\varepsilon_p(\alpha_7) = \{0\}$; $\varepsilon_p(\alpha_8) = \{0\}$; $\varepsilon_p(\alpha_9) = \{\omega_3\}$; $\varepsilon_p(\alpha_{10}) = \{0\}$; $\varepsilon_p(\alpha_{11}) = \{0\}$; $\varepsilon_p(\alpha_{12}) = \{0\}$; $\varepsilon_p(\alpha_{13}) = \{\omega_3\}$; $\varepsilon_p(\alpha_{14}) = \{\omega_5\}$; $\varepsilon_p(\alpha_F) = \{0\}$;	$\varepsilon_a(\omega_1) = \{\alpha_1\}$; $\varepsilon_a(\omega_2) = \{\alpha_2\}$; $\varepsilon_a(\omega_3) = \{\alpha_5, \alpha_9, \alpha_{13}\}$; $\varepsilon_a(\omega_4) = \{\alpha_6\}$; $\varepsilon_a(\omega_5) = \{\alpha_{14}\}$
$\pi = \pi_p \cup \pi_c$	$\pi_p(\eta_1) = \{\omega_1\}$; $\pi_p(\eta_2) = \{\omega_2\}$; $\pi_p(\eta_3) = \{\omega_2\}$; $\pi_p(\eta_4) = \{\omega_3\}$; $\pi_p(\eta_5) = \{\omega_3\}$; $\pi_p(\eta_6) = \{\omega_4\}$; $\pi_p(\eta_7) = \{\omega_4\}$; $\pi_p(\eta_8) = \{\omega_4\}$; $\pi_p(\eta_9) = \{\omega_5\}$; $\pi_p(\eta_{10}) = \{\omega_5\}$;	$\pi_c(\omega_1) = \{\eta_1\}$; $\pi_c(\omega_2) = \{\eta_2, \eta_3\}$; $\pi_c(\omega_3) = \{\eta_4, \eta_5\}$; $\pi_c(\omega_4) = \{\eta_6, \eta_7, \eta_8\}$; $\pi_c(\omega_5) = \{\eta_9, \eta_{10}\}$
$\kappa = \kappa_i \cup \kappa_o$	$\kappa_i(\alpha_1) = \{0\}$; $\kappa_i(\alpha_1) = \{d\}$; $\kappa_i(\alpha_2) = \{or_1\}$; $\kappa_i(\alpha_3) = \{or_2\}$; $\kappa_i(\alpha_4) = \{d\}$; $\kappa_i(\alpha_5) = \{d\}$; $\kappa_i(\alpha_6) = \{d\}$; $\kappa_i(\alpha_7) = \{d\}$; $\kappa_i(\alpha_8) = \{d\}$; $\kappa_i(\alpha_9) = \{d\}$; $\kappa_i(\alpha_{10}) = \{d\}$; $\kappa_i(\alpha_{11}) = \{d\}$; $\kappa_i(\alpha_{12}) = \{d\}$; $\kappa_i(\alpha_{13}) = \{d\}$; $\kappa_i(\alpha_{14}) = \{d\}$; $\kappa_i(\alpha_F) = \{d\}$;	$\kappa_o(\alpha_1) = \{d\}$; $\kappa_o(\alpha_1) = \{d\}$; $\kappa_o(\alpha_2) = \{d\}$; $\kappa_o(\alpha_3) = \{d\}$; $\kappa_o(\alpha_4) = \{d\}$; $\kappa_o(\alpha_5) = \{d\}$; $\kappa_o(\alpha_6) = \{or_1, or_2\}$; $\kappa_o(\alpha_7) = \{d\}$; $\kappa_o(\alpha_8) = \{d\}$; $\kappa_o(\alpha_9) = \{d\}$; $\kappa_o(\alpha_{10}) = \{d\}$; $\kappa_o(\alpha_{11}) = \{d\}$; $\kappa_o(\alpha_{12}) = \{d\}$; $\kappa_o(\alpha_{13}) = \{d\}$; $\kappa_o(\alpha_{14}) = \{d\}$; $\kappa_o(\alpha_F) = \{0\}$;



(그림 4) 특정기업의 전자결재(간이출장) 프로세스의 제어흐름 중심 ICN 표현 및 α_6 액티비티와 제어흐름상 연관된 액티비티



(그림 5) AA-알고리즘 적용 및 AA-Nets의 시각적 표현법

본 연구를 통해 제안하고 있는 액티비티 분석 방법을 적용하고자 위의 관련연구를 통해 보여주었던 어느 특정 조직의 전자결재(출장승인) 프로세스를 예시로 들고자 한다. ICN 모델을 통해 표현된 전자결재 프로세스를 액티비티와 제어흐름을 중심으로 만들어 내고, 임의의 액티비티를 'α_6'이라고, 가정하였을 경우, 'α_6' 액티비티와 제어흐름상 관련이 있는 액티비티는 그림 4와 같이 파악할 수 있다. 이는 본 장의 내용 중 AA-알고리즘을 적용하면 좀더 쉽게 파악해 낼 수 있으며, 마지막으로 그림 5에서 표현한 것과 같이 AA-Nets 모델로 구성 할 수 있다. 그리고 본 연구에서 예시로 쓰였던, 전자결재 프로세스를 AA-알고리즘을 통해 분석, AA-Nets을 통해 정형적 표기법으로 표현한다면, 다음 표 2과 같이 표현해 낼 수 있다.

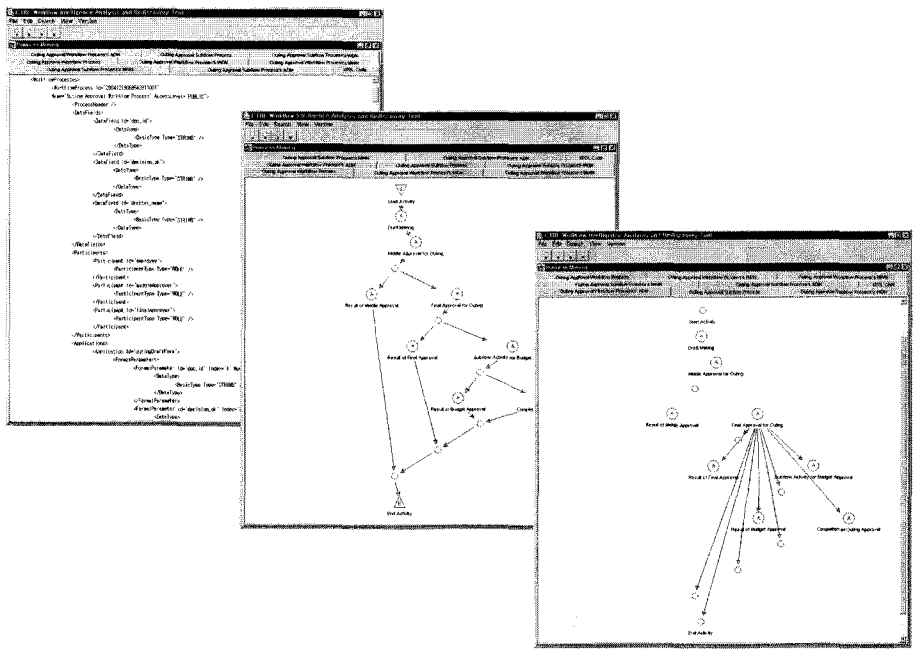
(표 2) α_6에서의 특정기업의 전자결재 프로세스의 정형적 AA-Nets 표현

<table>
<tr><td colspan="3">$\Gamma = (\sigma, \tau, I, O)$ over A, C
$AA = \{\alpha_6\}$
$DA = \{\alpha_4, \alpha_7, \alpha_8, \alpha_9, \alpha_{10}, \alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{14}, \alpha_F\}$
$C = \{d(default)\}$</td><td>//AAN
//Actual Activities
//Dependent Activities
//Connector</td></tr>
<tr><td>$\delta = \delta_i \cup \delta_o$</td><td>$\delta_i(\alpha_6) = \{0\}$; $\delta_i(\alpha_4) = \{\alpha_6\}$; $\delta_i(\alpha_6) = \{\alpha_6\}$;
$\delta_i(\alpha_7) = \{\alpha_6\}$; $\delta_i(\alpha_8) = \{\alpha_6\}$; $\delta_i(\alpha_9) = \{\alpha_6\}$;
$\delta_i(\alpha_{10}) = \{\alpha_6\}$; $\delta_i(\alpha_{11}) = \{\alpha_6\}$; $\delta_i(\alpha_{12}) = \{\alpha_6\}$;
$\delta_i(\alpha_{13}) = \{\alpha_6\}$; $\delta_i(\alpha_{14}) = \{\alpha_6\}$; $\delta_i(\alpha_F) = \{\alpha_6\}$;</td><td>$\delta_o(\alpha_6) = \{\alpha_4, \alpha_7, \alpha_8, \alpha_9, \alpha_{10}, \alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{14}, \alpha_F\}$
$\delta_o(\alpha_4) = \{0\}$; $\delta_o(\alpha_6) = \{0\}$; $\delta_o(\alpha_7) = \{0\}$;
$\delta_o(\alpha_8) = \{0\}$; $\delta_o(\alpha_9) = \{0\}$; $\delta_o(\alpha_{10}) = \{0\}$;
$\delta_o(\alpha_{11}) = \{0\}$; $\delta_o(\alpha_{12}) = \{0\}$; $\delta_o(\alpha_{13}) = \{0\}$;
$\delta_o(\alpha_{14}) = \{0\}$; $\delta_o(\alpha_F) = \{0\}$;</td></tr>
</table>

$K = K_i \cup K_o$	$K_i(\alpha_1) = \{0\}$; $K_i(\alpha_1) = \{0\}$; $K_i(\alpha_2) = \{0\}$; $K_i(\alpha_3) = \{0\}$; $K_i(\alpha_4) = \{d\}$; $K_i(\alpha_5) = \{0\}$; $K_i(\alpha_6) = \{d\}$; $K_i(\alpha_7) = \{d\}$; $K_i(\alpha_8) = \{d\}$; $K_i(\alpha_9) = \{d\}$; $K_i(\alpha_{10}) = \{d\}$; $K_i(\alpha_{11}) = \{d\}$; $K_i(\alpha_{12}) = \{d\}$; $K_i(\alpha_{13}) = \{d\}$; $K_i(\alpha_{14}) = \{d\}$; $K_i(\alpha_F) = \{d\}$;	$K_o(\alpha_1) = \{0\}$; $K_o(\alpha_1) = \{0\}$; $K_o(\alpha_2) = \{0\}$; $K_o(\alpha_3) = \{0\}$; $K_o(\alpha_4) = \{0\}$; $K_o(\alpha_5) = \{0\}$; $K_o(\alpha_6) = \{d[10]\}$; $K_o(\alpha_7) = \{0\}$; $K_o(\alpha_8) = \{0\}$; $K_o(\alpha_9) = \{0\}$; $K_o(\alpha_{10}) = \{0\}$; $K_o(\alpha_{11}) = \{0\}$; $K_o(\alpha_{12}) = \{0\}$; $K_o(\alpha_{13}) = \{0\}$; $K_o(\alpha_{14}) = \{0\}$; $K_o(\alpha_F) = \{0\}$;

3.3 워크플로우 가시적 검증 및 마이닝을 위한 액티비티 분석 도구 구현

본 연구를 통해 분석할 워크플로우 모델은 그래픽적 표현 방법 및 정형적 표현방법과 관련하여 ICN 기법 등으로 표현될 수 있다. 하지만, 일반적으로 대부분의 기업이나, 연구단체 등에서는 WfMC(Workflow Management Coalition)에서 제정한 프로세스 정의 표준 언어인 XPDL(XML-based Process Definition Language)[15]로 변환하여 사용한다[7, 9]. 따라서, 본 연구에서는 워크플로우 모델을 ICN 기법을 통해 모델링하고, XPDL 표준 언어로 변환하고, XPDL 형태의 언어를 ICN으로 표현, 분석해 내는 방법으로 구현하였다. 예제로 쓰인 모델은 위의 관련연구 및 방법을 소개할 때와 마찬가지로 앞서 소개한 같은 조직의 전자결재 프로세스를 사용하였다.



(그림 6) 구현예제 - 워크플로우 프로세스의 XPDL, ICN, AA-Nets 표현

4. 결론 및 향후연구

본 논문에서는 최근의 연구 개발 이슈인 워크플로우 마이닝과 관련하여, 워크플로우내 존재하

는 워크플로우 프로세스 제어흐름의 재발견을 위한 워크플로우 액티비티 분석을 하고자 하였다.

프로세스내 존재하는 임의의 액티비티가 변경될 때, 워크플로우 프로세스의 제어흐름에 어떠한 영향을 미칠 수 있는지를 파악고자, 워크플로우 마이닝을 위한 액티비티 분석 방법을 제안하였다. 워크플로우 마이닝을 위한 액티비티 분석 방법은 본 논문에서 제시한 AA-알고리즘과 AA-Net을 통해, 처리 및 표현 할 수 있도록 구체화 시켰으며, '특정기업의 차세대 정보기술 인프라를 위한 e-비즈니스 프로세스 기반 차세대 전자결시스템 적용 및 발전 방안 연구' 보고서[17]내에 정의된 바 있는 '전자결재(간이 결재) 프로세스'의 예제 프로세스로서 ICN 모델로 정의하였다. 또한, '특정기업의 전자결재(간이 결재) 프로세스'에 AA-알고리즘을 적용하여, AA-Nets으로 표현할 수 있도록 하였다.

또한, 워크플로우 마이닝을 위한 액티비티 분석 도구의 구현을 통해, 워크플로우 프로세스 정의 표준 언어인 XML 기반 프로세스 정의 언어(XPDL)로 표현된 '특정기업의 전자결재(간이 결재) 프로세스'에 AA-알고리즘을 적용하여, 워크플로우 재발견의 가능성을 확인 하였다.

본 연구를 통해 얻어진 액티비티 분석 방법과 도구는 워크플로우 관리 시스템이 현재보다 점점 발전되어, 기본적인 워크플로우 관리 시스템의 기능뿐 아니라, 진보된 형태의 워크플로우 관리 시스템의 형태를 구축하게 될 경우 적용 가능한 방법이 될 수 있을 것이다.

앞으로 진보된 워크플로우 관리 시스템의 출현을 위해, 본 연구를 통해 제안한 워크플로우 재발견 방법이 포함된 워크플로우 마이닝 또는 워크플로우 지능 시스템이 개발 되어야 할 것이다. 또한 AA-알고리즘 및 AA-Nets 모델도 좀 더 확장할 수 있는 형태로 구성하여, 향후 개발될 워크플로우 마이닝과 관련된 여러 가지 알고리즘 등도 다양한 방법으로 표현해 낼 수 있도록 할 것이다.

Acknowledgement

본 논문은 2006학년도 경기대학교 학술연구비 지원에 의하여 수행되었음

참고문헌

[1] Clarence A. Ellis, Gary J. Nutt, "ICNs Revisited and Revised", Technical Report, 1996

[2] Daniela Grigori, Fabio Casati, Malu Castellanos, Umeshwar Dayal, Mehmet Sayal and Ming-Chien Shan: Business Process Intelligence. Journal of Computers in Industry, Vol.53, Issue 3, Elsevier, April 2004

[3] Guido Schimm: Mining exact models of concurrent workflows. Journal of Computers in Industry, Vol. 53, Issue 3, Elsevier, April 2004

[4] Joachim Herbsta, and Dimitris Karagiannisb: Workflow mining with InWoLvE. Journal of Computers in Industry, Vol. 53, Issue 3, Elsevier, April 2004

[5] Jonathan E. Cook, Zhidian Du, Chongbing Liu and Alexander L. Wolf: Discovering models of behavior for concurrent workflows. Journal of Computers in Industry, Vol. 53, Issue 3, Elsevier, April 2004

[6] Kwang-Hoon Kim, "Beyond Workflow Mining", Lecture Notes in Computer Science, 4102, pp.49 -64, 2006.09

[7] Kwanghoon Kim, Minjae Park, "A Group Workflow modeling System", CollabTech2005, pp.97 - 102, 2005.07

[8] 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

[9] Min-Jae Park, Kwang-Hoon Kim , "Control-path Oriented Workflow Intelligence Analysis and Mining System", Proceedings of 2007 International Conference on Convergence

Information Technology, pp.951-960, 2007.11
[10] Minjae Park and Kwanghoon Kim: An Efficient Workcase Classification Method and Tool in Workflow Mining. Proceedings of the DMIN2005:The International Conference on Data Mining, Monte Carlo Resort, Las Vegas, Nevada, USA, June, 2005
[11] Shlomit S. Pinter, and Mati Golani: Discovering workflow models from activities' lifespans Journal of Computers in Industry, Vol. 53, Issue 3, Elsevier, April 2004
[12] San-Yih Hwang, Chih-Ping Wei and Wan-Shiou Yang: Discovery of temporal patterns from process Instances. Journal of Computers in Industry, Vol. 53, Issue 3, Elsevier, April 2004
[13] W. M. P. van der Aalst, B. F. van Dongena; J. Herbst, L. Marustera, G. Schimm and A. J. M. M. Weijters: Workflow mining: A survey of issues and approaches. Journal of Data & Knowledge Engineering, Vol. 47, Issue 2, pp. 237-267, Elsevier, November 2003
[14] W. M. P. van der Aalst and A. J. M. M. Weijters: Process mining: a research agenda. Journal of Computers in Industry, Vol. 53, Issue 3, Elsevier, April 2004
[15] Workflow Management Coalition Specification Document, "Workflow Coalition Interface 1: Process Definition Interchange Process Model", Document Number: WFMC TC-1016-P, August 5, 1998.
[16] 박민재, 김광훈 , "워크플로우 마이닝 기술 및 연구개발 동향", 한국인터넷정보학회학회지, 6권 1호, pp.75 - 81, 2005.03
[17] 유혁재, 한승혁, 정재우, 이정훈, 안형진, 박민재, 원재강, 김광훈 , "한국도로공사의 차세대 정보기술 인프라를 위한 e-비즈니스 프로세스 기반 차세대 전자결재 시스템 적용 및 발전 방안 연구", 최종연구보고서, 2004.12

◐ 저 자 소 개 ◑

김 광 훈 (Kwang-Hoon Kim)
1984년 경기대학교 전자계산학과 졸업(이학사)
1986년 중앙대학교 전자계산학과 졸업(이학석사)
1994년 Univ. of Colorado at Boulder, 컴퓨터과학과 졸업(이학석사)
1998년 Univ. of Colorado at Boulder, 컴퓨터과학과 졸업(이학박사)
1986년 ~ 1991년 : 한국전자통신연구원 연구원
1998년 ~ 현재 : 경기대학교 정보과학부 교수
관심분야 : 워크플로우, 비피엠, 그룹웨어, CSCW, 분산처리기술, 데이터베이스
E-mail : kwang@kgu.ac.kr