부정행위 탐색을 위한 시간 논리 기반의 패턴 유효성 검사 방법[1)]

이건수° 김민구 이형수*
아주대학교 정보통신 전문 대학원°, 유비쿼터스 컴퓨팅 연구센터*
{lks7256°, minkoo}@ajou.ac.kr, hslee@keti.re.kr*

Pattern Validation using Temporal Logic for Fraud Detection

Keonsoo Lee° Minkoo Kim
Graduate School of Information and Communication, Ajou University°
Hyung-Su Lee*
Ubiquitous Computing Research Center*

요 약

부정행위 탐지는 개별 사용자의 행동 기록과 그 사용자와 유사한 프로필을 갖고 있는 사용자들의 행동 기록을 바탕으로 행동 패턴 혹은 행동 규칙을 찾아내, 이 패턴/규칙과의 비교를 통해 현재 행위가 부정한 것인지를 결정하는 방법을 주로 사용한다. 그러나, 특정 사용자의 행위 패턴이 급격하게 바뀌는 경우, 과거의 기록을 바탕으로 생성된 패턴의 유효성은 보장받을 수 없다. 더구나 기존 기록과 상이한 행위에 대한 새로운 패턴이 생성되기 위해서는 계속해서 그런 행위가 쌓여야만 하고, 그 쌓이는 양은 기존 패턴의 견고성에 비례된다. 또한, 동일 사용자에게 여러 패턴을 적용시키는 방법 역시 패턴간의 충돌이 일어나는 등의 한계가 존재한다. 본 논문에서는 시간 논리(Temporal Logic)를 적용하여, 과거의 패턴의 유효성을 검증하고, 신규 패턴을 빠르게 찾아내는 방법을 제안하고자 한다.

1. 서 론

하루가 다르게 발전하는 컴퓨터의 성능과 보편화된 인터넷은 보다 쉽게 정보를 공유하기 위한 환경을 제공해 줬다. 이처럼 보다 쉽게 타인의 정보에 접근할 수 있게 됨에 따라 보안의 중요성이 강조되고 있다. 부정행위 탐지는 특정인에게 주어진 자원에 대한 권한이 올바르게 사용되고 있는가를 탐색하는 보안 기법으로 주로 전화통신, 신용카드, 은행 등의 도메인에서 특정인에게 주어진 권한을 타인이 부정한 방법으로 획책하여 사용하는 것을 탐지하여 그 피해를 막는 기법이다. 일반적으로 부정행위를 탐색하기 위해 선택 트리(Decision tree), 분류(Classification), 군집화(Clustering), 신경망(Neural Network), 연결 분석(link analysis)의 어느 방법을 사용하는 것에 상관없이 패턴 혹은 규칙을 찾아내기 위한 데이터 마이닝 과정에서는 사용자의 처리 데이터를 기반으로 사용한다. 카드 회사에서는 특정 고객의 이전 구매 기록을 바탕으로 사용자의 구매 패턴을 찾아낸 뒤, 그 패턴과 일치하지 않는 구매 활동을 부정행위로 탐지한다. 혹은, 전체 구매 기록 사이에서 아이템간의 연결 관계를 찾아내어, 그 관계에 적합하지 않는 구매 활동을 부정행위로 의심한다. 이렇게 사용자들의 구매 기록을 바탕으로 패턴 혹은 규칙을 찾아내는 방법은 부정행위를 탐색해 내는데 효율적은 성능을 보여주지만, 사용자의 구매 패턴이 바뀌는 경우, 패턴이 변화하였다는 것을 감지하기보다는, 부정행위로 판단하게 된다. 더욱이 변화된 패턴을 새롭게 감지하기 위해서는, 기존 패턴에 비례하여 다수의 새로운 패턴에 속하는 행위가 발생해야 한다.

이에 본 논문에서는 사용자의 패턴이 급격하게 변하는 경우, 사용자의 지난 패턴이 신규패턴의 생성에 영향을 미치지 않도록 하기 위해 사용자의 행위 기록에 시간 요소를 도입해 사용자의 구매 습성의 변화에 보다 능동적으로 대처할 수 있는 패턴 생성 및 지난 패턴의 유효성을 검증할 수 있는 방법을 제안하고자 한다. 2장에서는 기존의 부정행위 탐색 방법과 본 논문에서 이용하고 있는 시간 논리에 대하여 알아보고, 3장에서는 시간 논리를 사용한 패턴의 유효성 검사 방법을 제안한다. 제안된 방법의 효율성은 4장의 시뮬레이션을 통해 검증하고, 5장에서 그 결론을 내리도록 한다.

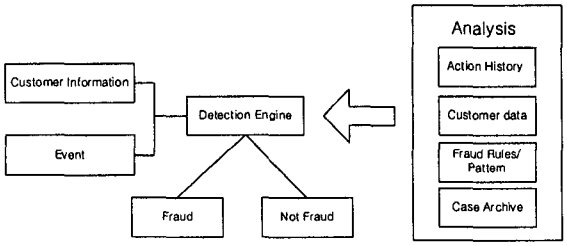
2. 관련연구

2.1 부정행위 탐색

부정행위 탐색은 특정인에게 주어진 권한을 타인이 획책하여 사용하는 것을 탐지하여 그 피해를 막기 위한 방법이다. 타인의 부정한 사용을 찾아내기 위하여 원 사용자의 행위 기록을 바탕으로 데이터 마이닝 과정을 거쳐 그 사용자의 행동 패턴 혹은 행동 규칙을 찾아낸다. 이

1) 본 논문은 유비쿼터스 컴퓨팅 및 네트워크 원천기반 기술 개발사업 내 uT-DB 기반의 지능성장 학습 엔진 사업의 일환으로 지원 받아 수행되었음 (과제번호 M103KT010007-04K2001-00713)

렇게 찾아낸 패턴 혹은 규칙을 바탕으로 새로이 발생하는 사용자의 행위를 비교해 그 행위가 부정행위인가 아닌가를 탐지할 수 있다. <그림 1>은 일반적인 부정행위 탐지 시스템의 구조도를 보여주고 있다.



<그림 1> 부정행위 탐지 시스템 구조도

이때, 마이닝 과정을 통해 생성되는 패턴 혹은 규칙은 사용자의 개별 행위 기록과 전체 사용자의 일반적인 행위 기록을 통해 생성된다. 즉, 특정 사용자의 행위 기록을 통해 그 사용자의 과거 행위 패턴을 찾아내고, 그 행위 패턴에 맞지 않는 행동을 부정행위로 탐지하는 방법이 있고, 사용자의 프로파일을 바탕으로 사용자별 클러스터링을 한 뒤, 그 사용자 집합에 속하는 사용자들이 행위를 조사해 일반적인 행위 패턴을 생성한다. 이 경우 사용자 A가 "20대 초반, 미혼, 월수입 30미만, 대도시"의 그룹에 속해있다면, 그런 사람들의 일반적인 행위 패턴을 따를 것이라는 가정을 바탕으로 동일 조건의 일반적인 행위 패턴과 비교해 부정행위를 검사한다.

2.2 시간 논리

시간 논리는 1960년대 발표된 시제논리(Tense Logic)의 확장으로 논리체계 안에서 시간 개념을 갖고 있는 정보를 표현하기 위한 방법으로 "항상(Always)", "때때로(Sometimes)" 그리고 "절대로(Never)'의 3가지 개념을 기본 연산자로 사용하고, 그 이외의 "~부터(Since)", "~까지(Until)", "~후에(Next)" 또는 "~전에(Before)" 등의 확장 연산자들을 통해 표현된다.

<표 1> 시간 논리의 연산자

연산자	의미
Always	존재하는 모든 시간
Sometimes	특정 시간
Never	존재하지 않는 시간
Since	특정 시간 이후 계속
Until	특정 시간 이전 계속
Next	특정 시간 이후
Before	특정 시간 이전

이상의 연산자들을 통해 각 현상들은 $t \leq_T t'$인 $(T, \leq_T)$의 집합으로 표현될 수 있고 이는 곧 $t \in T$ 인 사건이 일어난 뒤 $t' \in T$인 사건이 일어났음을 의미한다. 이처럼 시간 요소를 논리에 적용함으로써, 특정 현상이 참 또는 거짓이 되기 위해 다른 현상과의 연속적인 발생과정을 명시할 수 있고 이는 곧 각 현상들이 서로 어떤 연결 관계를 갖고 있는가를 표현할 수 있다.

3. 유효성 검사 방법

일반적인 부정행위 탐지 시스템의 동작 예는 다음과 같다. 가령, 카드를 만든 뒤, 주로 값싼 술집에서 한달에 2~3번 술 마시는 경우에만 카드를 사용한 20대의 한 남성이 있다고 가정하자. 이 남성의 구매 패턴은 대략, "저녁 시간대 저렴한 술집, 사용빈도는 낮음"이 될 것이다. 어느 날, 이 남성이 어떤 여성을 소개받아 서로 사귀기 시작한다면, 이 남성의 구매 패턴은 급격하게 변하게 된다. 카드는 술집보다는 고급 레스토랑에서 결제가 이루어지고, 남성은 카드로 평생 사본적도 없는 여성용 화장품이나, 옷가지를 구매하는 된다. 이런 행위는 기존의 패턴과 비교해 분명 부정행위로 분류된다. 부정행위로 분류된 경우, 시스템은 부정행위가 발생했음을 알려준다. 만약, 이 행위가 부정행위가 아닌, 사용자의 새로운 행동 패턴인 경우라면, 새로운 행동 타입에 맞는 패턴을 새로 만들어야 한다. 이렇게 급격히 변하는 경우, 기존의 사용자의 행위 기록은 사용될 수 없기 때문에 결국, 사용자의 프로필과 유사한 조건의 다른 사용자들의 패턴을 차용하는 방법이 사용된다. 이 같은 방법으로 사용자의 행동 패턴이 할당되는 경우 두 가지 문제점이 발생할 수 있는데, 할당된 패턴들 사이에 충돌이 일어날 수 있다는 것과, 할당된 패턴들 중에서 더 이상 유효하지 않은 패턴을 찾아내 사용자에게서 제거해야 한다는 것이다.

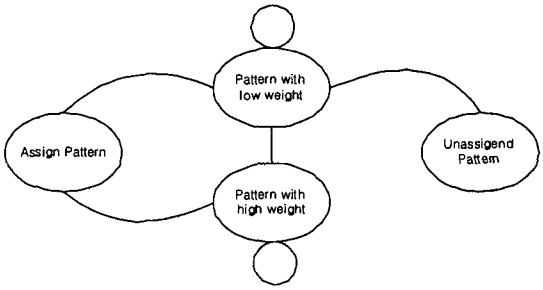
이를 위해 본 논문에서는 사용자에게 할당된 패턴들의 유효성을 검사함으로써, 현재 사용자의 행동 패턴을 가장 잘 표현하는 패턴 집합을 찾아내는 방법을 제안한다. 이를 위해 각 패턴들에 가중치를 적용해 충돌이 일어나는 패턴들 사이의 우선순위를 기반으로 부정행위의 확률을 찾아내고 그 값이 임계 값을 넘는 경우 부정행위로 간주한다. 이 때 각 패턴들의 가중치 값은 현재 각 패턴의 유효성을 표현하고 이 값은 시간 논리를 기반으로 결정된다. 각 패턴의 가중치를 결정하는 시간 논리 규칙은 다음과 같다.

<표 2> 제안된 유효성 검사 규칙

```
1. Always ({assigned(Pattern) }
     -> {hasWeight(Pattern)} )
2. Always ({assign(Pattern)}
     -> (next{ decreasedWeight(assigned(Pattern))}) )
3. Always ({occur(Behavior)}
     -> (next{ increasedWeight(matched(Pattern))}) )
4. Always ({occur(Behavior)}
     -> (next{decreasedWeight(Not{matched(Pattern)})}) )
5. Always ({assign(Pattern) and Not{next(occur{Behavior
   })} -> (next{ decreasedMoreWeight(Pattern)}) )
6. Always ({occur(Behavior) and {next(occur{Behavior})}
     -> (next{ increasedMoreWeight(matched{Pattern})}) )
7. {validate(Pattern)} Since {{assigned(Pattern)}
8. {validate(Pattern)} Until {over(Weight, threshold)}
```

<표 2>의 유효성 검사 규칙을 통해, 사용자에게 할당된

패턴은 시간의 흐름과 새로운 사용자 행위에 따라 그 유효성을 검증 받는다. <그림 2>는 유효성 검증 결과에 의한 패턴의 생성에서부터 소멸에 이르는 상태변환 단계를 보여주고 있다.



<그림 2> 상태 변환 관계

각 패턴은 지지 행위의 발생, 그 행위의 시간 간격, 신규 패턴의 생성 및 소멸현상에 따라 중요도가 변하고 결국, 현재 사용자의 행위를 가장 적절히 표현할 수 있는 패턴과 유효도 쌍의 집합으로 표현된다.

4. 시뮬레이션 및 결과

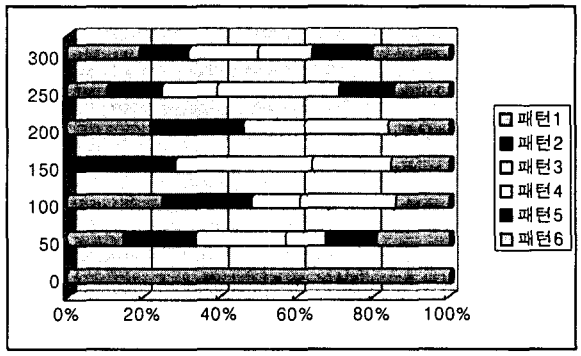
본 논문에서 제안한 방법을 검증하기 위해 가정한 실험 환경은 <표 3>에 나타나 있다.

<표 3> 실험 환경

실험 환경	
행위 시퀀스	300
유효 패턴 수	6
초기 패턴 가중치	100
가중치 감소폭(min/max)	5/10
가중치 증가폭(min/max)	5/10
최저 유효 가중치	25
부정행위 검출 조건	
선행 처리	패턴별 가중치 정규화
부정행위 검출 값	패턴 비교 값 25% 이하

실험 조건에서는 사용자의 행위 기록을 바탕으로 생성된 특성화된 행동 패턴 하나와 전체 행위 데이터를 기반으로 생성된 5개의 일반 패턴이 존재한다. 사용자는 6개의 패턴 중 하나에 속하는 행위를 무작위로 선택하여 300수행하고 모든 행위는 부정하지 않다고 가정한다. 이 경우 사용자가 행하는 여러 행동에 부합하는 패턴을 사용자에게 적용시키고, 그 행위 패턴의 유효성을 시간논리를 기반으로 추출함으로써, 현재 사용자의 행동 패턴을 다양한 패턴들과 각 패턴들의 가중치 집합으로 표현하였다.

이상의 조건으로 실험한 결과는 <그림 3>에 나타나 있다. 무작위로 변하는 사용자의 300번의 행위 변화에 따라 총 6개의 패턴들이 사용자의 현재 행동 패턴을 설명하기 위해 다양한 유효치를 갖고 적용되었다. 150번째 행동 결과에서 나타나는 것처럼, 기존에 적용된 패턴도, 계속해서 그 패턴의 행동이 나타나지 않는다면, 유효성을 상실하게 되고, 200번째 행동 결과는 과거에 적용됐던 패턴의 행동이 다시금 나타난다면, 그 패턴은 새로운 패턴이 생기는 것보다 높은 유효성을 갖고 적용됨을 보여준다.



<그림 3> 실험 결과

5. 결 론

본 논문에서는 시간 논리를 기반으로 사용자에게 할당된 패턴들의 유효성을 계산하여, 새로운 행동에 대한 부정행위를 탐지하는 방법을 제안했다. 본 방법은 신규 발생하는 사용자의 행위에 따라 기존에 할당된 패턴들의 가중치를 조절함으로써, 현재 사용자의 행동 패턴을 가장 잘 표현할 수 있는 패턴 집합과 그 집합의 패턴들 사이의 중요도 관계를 표현할 수 있다. 이 방법을 통해 각 사용자에게 할당된 여러 패턴들 사이의 관계를 기반으로 효율적인 부정행위를 탐지가 가능하다.

참고 자료

[1] Richard J. Roiger, Michael W. Geatz "Data Mining: a tutorial-based primer" Addison Wesley, ISBN 0-201-74128-8

[2] Jose R. Dorronsoro, Fransicso Ginel, Carmen Sanchez and Carlos Sana Cruz "Neural Fraud Detectoin in Credit Card Operations" IEEE Transactions on Neural Networks. Vol 8. no 4 July 1997

[3] Mads Dam, "Temporal Logic, Automata, and Classical Theories An Introduction" , Notes for the Sixth European Summer School in logic, language, and information, Copenhagen, 1994

[4] Ph. Schnoebelen, "The Complexity of Temporal Logic Model Checking", Proc. 4th Int. Workshop, AiML '2002, Toulouse, France

[5] Peter Cabena, Pablo Handjinian, Rolf Stadler, Jaap Verhees, Alessandro Zanasi, "Discovering Data mining-From Concept to Implementation", Prentice Hall, ISBN 0-13-743980-6