Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.26, NO.6, Dec. 2016 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2016.26.6.1493

웹 사용자의 실시간 사용 패턴 분석을 이용한 정상 사용자 판별 방법

장 진 구,† 문 종 섭‡
고려대학교 정보보호대학원

A Real-Time User Authenticating Method Using Behavior Pattern Through Web

Jin-gu Jang,† Jong Sub Moon‡
Center for Information Security Technologies, Korea University

요 약

인터넷을 통한 사이버 위협이 증대됨에 따라 개인정보 침해도 지속적으로 발생하고 있다. 악의적인 사용자들은 유출된 개인정보를 도용하여 정상 사용자처럼 해당 웹사이트를 접근하고 불법적인 행동을 할 수 있다. 본 논문에서는 이러한 불법 사용자의 접근을 실시간으로, 효과적으로 탐지하기 위해 정상 사용자의 웹사이트 평시 사용 패턴을 멤버십 분석(membership analysis)과 마르코프 체인 모델(markov chain model)을 기반으로 프로파일링 함으로써, 정상 사용자를 판별하는 방법을 제안한다. 아울러 이러한 프로파일에 시간적인 특성, 즉 시간 가중치(time weight)를 적용하여, 시간적으로 변하는 사용자의 행동을 사용자의 프로파일에 반영한다. 이에 따라 시간에 따른 사용자의 성향을 반영한 결과를 얻을 수 있다. 본 연구를 통해 생성한 사용자별 프로파일을 기반으로 개인정보를 도용한 악의적인 사용자를 적발할 수 있고, 정상적인 사용자이더라도 민감한 정보에 접근하는 것을 방지할 수 있다. 본 연구를 적용한 결과, 정상 사용자에 대해 96%의 높은 판별 정확도를 보여주었다.

ABSTRACT

As cyber threats have been increased over the Internet, the invasions of personal information are constantly occurring. A malicious user can access the Web site as a normal user using leaked personal information and does illegal activities. This paper proposes an effective method which authenticates a genuine user with real-time. The method use the user's profile which is a record of user's behavior created by Membership Analysis(MA) and Markov Chain Model(MCM). In addition to, user's profile is augmented by a Time Weight(TW) which reflects the user's tendency. This method can detect a malicious user who camouflage normal user. Even if it is a genuine user, it can be determined as an abnomal user if the user acts beyond the record profile. The result of experiment showed a high accuracy, 96%, for the correct user.

Keywords: Machine Learning, Markov Chain Model, Membership Analysis, Time Weight

I. 서 론

정보통신기술(ICT)이 발달하면서 개인정보보호 환경도 함께 변화하고 있다. 즉, ICT의 발달은 우리의 생활을 편리하고 스마트하게 변화시켰지만, 수집·활용되는 개인정보의 양이 방대해지고 다양해지면서 개인은 프라이버시 침해의 위험에 노출되기 쉬운 환경에 처하게 되었다. 이에 따라 개인정보 침해도 지속적으로 발생하고 있다. 이렇게 유출된 개인정보들

Received(09. 20. 2016), Modified(11. 07. 2016), Accepted(11. 21. 2016)
† 주저자, mbcmbc@naver.com
‡ 교신저자, jsmoon@korea.ac.kr(Corresponding author)

은 악의적인 사용자들에 의해 도용되어 정상 사용자인 것처럼 해당 웹사이트에 접근하고 불법적인 행동을 할 수 있다. 이러한 문제를 해결하기 위해 웹 관리자는 아이디와 패스워드 외에도 추가 암호, 인증, 권한 관리 등과 같은 추가적인 방법으로 사용자를 확인할 수 있다[1].

본 논문에서는 이러한 문제점들을 해결하기 위해 정상 사용자의 웹사이트 사용 패턴을 멤버십 분석(Membership Analysis)과 기계학습 방법의 일종인 마르코프 체인(markov chain model)을 기반으로 프로파일링 하여, 실시간으로 정상 사용자를 판별하는 방법을 제안한다. 멤버십 분석에 대해서는 3장의 3.3.1에서 자세히 살펴본다. 아울러 분석 프로파일에 시간적인 특성, 즉 시간 가중치(time weight)를 적용하여 시간에 따른 사용자의 성향을 반영할 수 있으며 더욱 정확하게 정상 사용자를 판별할 수 있다.

전체적인 과정은 크게 두 단계 분석 모델로 진행되며 다음과 같다. 1단계 분석 모델에서는 사용자의 웹사이트 접속 정보들을 멤버십 분석을 이용해서 패턴을 분석하고, 2단계 분석 모델에서는 사용자의 웹사이트 이용 패턴을 마르코프 체인 모델(markov chain model)을 이용해서 분석한다. 이때 각 분석 모델마다 각 사용자의 행위에 대한 프로파일 생성 과정과 사용자 진위를 판별하는 사용자 판별 과정으로 나누어진다.

논문의 전체적인 구성은 다음과 같다. 2장에서는 본 논문의 분석 방법과 관련된 유사 연구를 살펴보고, 3장에서는 실험에 사용한 이론과 이론의 적용 방법을 다룬다. 4장에서는 실제 실험 과정과 그 결과를 소개 한다. 마지막 섹션에서는 실험 결과가 가지는 의미와 향후 연구 방안을 제안하는 것으로 논문의 결론을 맺는다.

II. 관련 연구

사용자의 행동 패턴을 기반으로 사용자 진위 판별에 대한 유사 연구는 Peng et al.(2016)이 발표한 논문[2]에서 살펴볼 수 있으며 다음과 같다.

일찍이 Umphress와 Williams(1985)는 키보드 입력 시간에 따라 사용자를 판별할 수 있는 방법을 제안했고[3], 이를 확장하여 높은 정확도와 빠른 분석 시간을 요구하는 연구가 진행되었다[4][5][6].

Alexandre(1997)는 암호와 생체 인식 기법(음성 인식, 지문 인식 등)을 결합하여 키보드 서명의 동작을 인식하는 기법을 제안했고[7], Li et al.(2006)은 프로세스의 CPU 사용량과 응용프로그램 실행 및 윈도우창 수의 패턴을 이용하는 기법을 제안했다[8]. 또한 Vizer et al.(2009)은 사용자의 키보드 입력 패턴을 기반으로 사용자 진위 판별 방법을 제안했다[9].

Bhaskaran et al.(2011)은 눈동자의 움직임을 분석하여 사용자를 판별할 수 있는 방법을 연구했다[10].

또한 사용자의 행동 패턴을 마르코프 체인 모델로 분석하여 사용자 진위 판별을 하는 유사 연구를 살펴보면 다음과 같다.

마르코프 체인 모델을 이용해 컴퓨터에서 비정상 행동 패턴을 탐지하려는 본격적인 시도는 Wen-Hua el at.(2001)와 Ye el at.(2004)가 발표한 논문[11][12]에서 엿볼 수 있다. 컴퓨터의 비정상 행동 패턴 탐지에 마르코프 체인 모델을 적용하는 효율적인 계산 방법을 제안한다. 실험 결과의 정확도는 높았지만 분석 대상이 유닉스를 대상으로 분석했다는 점에서 인터넷을 이용하는 일반적인 사용자들을 대상으로 적용하여 분석하기에는 제한적인 한계점을 갖고 있다.

Jongho Choy el at(2001)도 침입탐지 시스템을 위한 은닉 마르코프 모델(hidden markov model)을 모델링(단일, 사용자별, 그룹) 방식에 따라 다양하게 분석하였지만, 발생하는 시스템 이벤트에 대해서만 분석이 가능하다는 제한사항을 갖고 있다[13].

또한 Juan el at.(2004)는 비정상 사용자 탐지를 위해 네트워크의 HTTP 트래픽을 대상으로 마르코프 체인 모델을 적용하여 사용자의 행동 패턴을 분석하였지만, HTTP 트래픽 전체에 대한 하나의 프로파일을 기반으로 분석을 진행하여 각 사용자에 대해 판별이 제한적인 한계점을 갖고 있다[14].

이 밖에도 마르코프 체인 모델을 이용한 여러 연구가 선행되어 왔지만, 대부분 공통적으로 각 사용자마다 프로파일을 생성하여 테스트하는 것이 아닌, 전체적인 하나의 프로파일을 기반으로 분석을 진행한다는 한계점을 갖고 있다.

III. 제안하는 기술

3.1 프로파일 생성 과정 - 데이터 취득

웹서버에서 생성된 로그데이터는 사용자의 행동에

대해 시간 순서에 따라 자동적으로 생성하고 유지된다. 즉, 사용자의 패턴이 담겨 있는 데이터이며, 그 속에서 사용자의 패턴을 나타낼 수 있는 항목을 선택하여 사용해야 한다.

사용자가 웹사이트를 접속할 때 accept, browser, version, connection, clientIP 등 여러 가지 데이터가 발생한다. 이 중에서 해당 사용자임을 확인할 수 있는 항목은 사용자 IP 주소와 사용자 ID가 있다. 또한 시간 순서를 나타내는 접속 시간, 접속 요일 항목이 필요하다. 또한 사용자가 웹사이트 접속 후 이동하는 경로를 나타내는 URL 주소 항목이 필요하다.

따라서 Table 1.의 항목을 기반으로 시간 순서에 따라 발생하는 사용자의 행동 패턴이 담긴 프로파일을 구성할 수 있다.

Table 1. Log Data Category of Web Server

Category	Contents
ProcessTime	Log Creation Time
ClientIP	User IP Address
requestURL	User Access URL
userID	User Access ID
sessionID	User Session ID

3.2 프로파일 생성 과정 - 시간 가중치 적용

사용자의 프로파일은 일정 기간 동안의 웹사이트 사용 이력이 반영된다. 이때, 수집 기간 혹은 데이터의 중요도에 따라 사용자의 시간적인 성향을 반영하기 위해 시간 가중치를 적용한다. 기본적으로 가장 최근 데이터에 제일 큰 비중의 가중치를 적용하고, 오래된 데이터일수록 가중치 비중은 낮아진다.

예를 들어, Fig.1을 참조하면, 지난 4주간의 사용 이력은 〔1주 전 데이터 : x_t, 2주 전 데이터: x_{t-1}, 3주 전 데이터 : x_{t-2}, 4주 전 데이터: x_{t-3}〕이고, 시간적인 가중치(w_t)는 각각 〔1주 전 데이터 : $w_t(0.5)$, 2주 전 데이터 : $w_{t-1}(0.3)$, 3주 전 데이터 : $w_{t-2}(0.15)$, 4주 전 데이터 : $w_{t-3}(0.05)$, 총합 : 1〕을 적용한다. 수식으로 나타내면 다음과 같다.

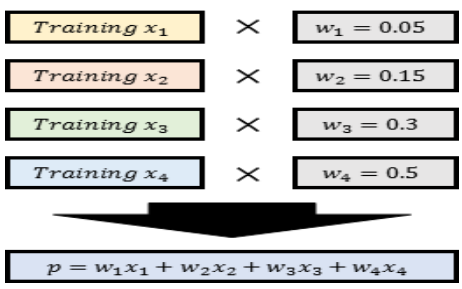

Fig. 1. The Weighted Profile

$$p = w_t x_t + w_{t-1}x_{t-1} + w_{t-2}x_{t-2} + w_{t-3}x_{t-3}$$
$$= 0.5x_t + 0.3x_{t-1} + 0.15x_{t-2} + 0.1x_{t-3} \quad (1)$$

이때 p는 프로파일을 나타내고, x_t는 프로파일에 사용된 주차별 데이터이다. 또한 프로파일에 적용하는 전체 기간을 T라고 할 수 있으며 다음과 같이 정의할 수 있다.

$$p = \sum_{i=0}^{T-1} w_{t-i} \cdot x_{t-i}, where \begin{cases} \sum_{i=0}^{T-1} w_{t-i} = 1 \\ w_i > w_j, \forall i,j \\ T = 1 \sim N \end{cases} \quad (2)$$

누적된 데이터에 대한 가중치 적용 대상은 각 단계별 프로파일 형성 값이 아닌 일정기간 동안의 누적 횟수에 대해 가중치를 부여한다. 각 단계별 프로파일 형성 값은 시간구간 내의 상대적인 값이기 때문에 기준이 동일하지 않다. 따라서 동일한 기준을 적용하기 위해서 기간별 누적 횟수에 대해 가중치를 부여하여 적용한 뒤 각 단계별 프로파일 형성 값을 계산한다. 예를 들어, Fig.1.을 참조하면, Training 데이터는 사용자의 일정기간(1주) 동안의 누적 횟수 데이터를 사용한다. 따라서 시간 가중치 평균은 아래 각 단계별 분석에서 사용되는 데이터 항목들의 누적 횟수에 대해 가중치를 부여할 때 사용된다.

3.3 단계별 사용자 프로파일 형성 방법과 사용자 판별 방법

본 논문에서 제안하는 단계별 사용자 프로파일 형성 과정은 Fig.2.와 같다. 단계별 분석 모델은 사용자의 로우데이터(raw data)로부터 일정기간 동안의 누적 횟수 데이터를 얻게 되고 시간 가중치를 적용하여 각각의 프로파일을 형성한다.

또한 상기 데이터 취득 과정에서 일정 기간 동안 취득한 데이터 중에서 ClientIP, ProcessTime(시

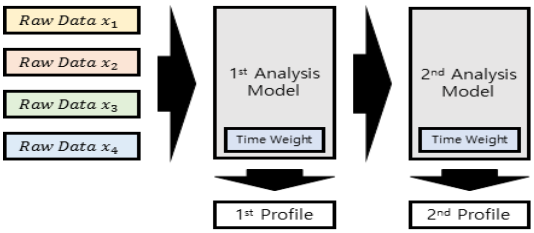

Fig. 2. Step-by-step Create Process of Profile

간, 요일)은 1단계 접속 정보 프로파일 생성과정에 사용되고, requestURL, userID, sessionID는 2단계 프로파일 생성과정에 사용된다.

3.3.1 1단계 접속 정보 분석 - 프로파일 형성 방법

본 1단계에서는 사용자의 접속 형태에 대한 프로파일을 생성한다. 멤버십 분석을 이용해 프로파일을 형성하고, 추후 진위 판별을 위한 경계면(threshold)과 경계 범위(threshold range)를 생성하여 진위를 판별한다. 여기서 말하는 멤버십 분석이란, 퍼지 멤버십 함수의 이론을 준용하여 본 논문에서 새롭게 정의하여 제안하는 분석 방법이다[15][16]. 즉, 일반적인 집합 원소들의 소속 정도를 구하는 방법을 응용한 분석 방법이다. 이때 소속 정도는 0과 1사이의 실수로 표현되고, 집합에 가장 많이 존재하는 원소는 1, 전혀 속하지 않는 경우에는 아주 작은 $\alpha_{1^{st}}$값으로 나타내며, 그 사이에 존재하는 원소 중에서 가장 적게 존재하는 원소는 최소값을 가지고, 나머지 원소들은 최소값과 1 사이의 값을 균등하게 분배한다. 여기서 집합에 속하지 않는 원소의 경우에 $\alpha_{1^{st}}$를 할당하는 이유는 멤버십 분석을 계산하는 과정이 곱의 형태이기 때문에 0 대신에 아주 작은 $\alpha_{1^{st}}$값을 할당하여 계산한다. 이때 $\alpha_{1^{st}}$값은 1단계 분석 모델에서 사용하는 접속 정보의 개수에 따라 다르게 사용된다. 이때 멤버십 분석에 적용하여 나온 각 요소별 결과 값을 멤버십 값(membership value)이라 정의하며, 전체 결과값을 통합 값(total value)라 하면, 다음과 같다.

$$\text{total value} : \prod_{i=1}^{I_{\max}} m_{ij} \tag{3}$$

membership value : m_{ij}

i : 접속 정보 인덱스 ($1 \leqq i \leqq I_{\max}$, IP 주소, 접속 시간, 접속 요일 등), $I_{\max}$는 최대 접속 정보의 개수

j : 해당 접속 정보의 누적 횟수를 내림차순으로 정렬했을 때, 누적 횟수 전체 카테고리($N^i_{\max}$)의 j번째 카테고리 ($1 \leqq j \leqq N^i_{\max}$)

$$m_{ij} = \begin{cases} \alpha_i & (\forall j \notin Category, 1 \leq i \leq I_{\max}) \\ \beta^i_{\min} + \frac{j-1}{N^i_{\max}-1} \times (\beta^i_{\max} - \beta^i_{\min}) & (1 \leq i \leq I_{\max}, 1 \leq j \leq N^i_{\max}) \end{cases} \tag{4}$$

$\alpha_{1^{st}i}$: i번째 접속 정보를 사용할 때, 누적 횟수 카테고리가 존재하지 않는 경우에 대해서 부여 하는 예외상수

$\beta^i_{\min}$: i번째 접속 정보의 최소 누적 횟수 할당 값(누적 횟수가 가장 최소일 때 부여받는 값)

$\beta^i_{\max}$: i번째 접속 정보의 최대 누적횟수 할당 값(누적 횟수가 가장 최대일 때 부여받는 값)

위 정의를 그림으로 표현하면 Fig.3.과 같다.

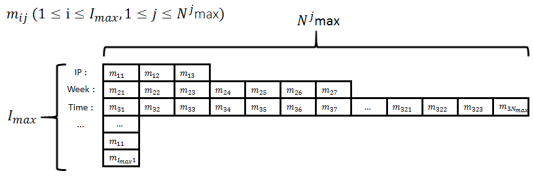

Fig. 3. Access Information Type in Step 1

본 연구에서는 사용자의 접속 정보에 대해 누적된 횟수를 원소로 사용한다. 기본적으로 사용되는 접속 정보($I_{\max}$)는 3가지(접속 IP 주소, 접속 시간, 접속 요일)이며, 각 접속 정보($I_{\max}$)마다 누적 횟수의 카테고리($N^i_{\max}$)가 정해진다. 여기서 α_i, $\beta^i_{\min}$, $\beta^i_{\max}$은 다음과 같은 조건을 만족하는 상수로 지정하여 사용한다.

정상 사용자의 접속 패턴이 존재하여 멤버십 값을 형성할 때, 악의적인 사용자가 정상 사용자를 흉내내어 접속을 한다고 가정한다. 이때, 악의적인 사용자는 정상 사용자의 모든 접속 패턴 중에서 가장 높은 멤버십 값을 흉내 내지만, 한 가지는 패턴에 존재하지 않는 경우라고 가정하자. 즉, 하나의 잘못된 패턴만으로도 악의적인 사용자를 판별할 수 있어야 한다는 것이다. 또한, 악의적인 사용자의 접속 패턴 판별 값은 정상 사용자의 접속 패턴 중 가장 작은 확률

로 발생하는 행동들의 판별값보다 작아야한다. 즉, 정상 사용자를 판별하는 임계치(threshold)는 악의적인 사용자의 판별값보다 크고 정상 사용자의 판별값보다 작아야 한다.

따라서 정상 사용자를 판별하는 임계치는 항상 "악의적인 사용자의 판별값 ≦ Threshold 〈 정상 사용자의 판별값" 이 성립되어야 한다. 멤버십 분석의 결과 값은 $\prod_{i=1}^{I_{\max}=3} m_{ij}$으로 표현되며, 다음과 같은 조건의 함수가 된다.

$$\alpha_{1^{st}i}\prod_{i=1}^{I_{\max}-1}(\beta^i_{\max}) \leqq \theta_{1^{st}}(threshold) < \prod_{i=1}^{I_{\max}}(\beta^i_{\min}) \quad (5)$$

여기서 각 요소마다 같은 $\alpha_{1^{st}i}$, $\beta^i_{\min}$, $\beta^i_{\max}$을 사용한다. 즉, 접속 정보로 사용되는 속성의 개수가 정해지면 각 속성마다 사용하는 $\alpha_{1^{st}i}$, $\beta^i_{\min}$, $\beta^i_{\max}$을 동일하게 사용한다. 따라서 이후 표기를 $\alpha_{1^{st}i}=\alpha_{1^{st}}$, $\beta^i_{\min}=\beta_{\min}, \beta^i_{\max}=\beta_{\max}, \forall i \in I_{\max}$로 표기한다. 이때 앞서 언급한 멤버십 분석 기준에 따라 $\beta^i_{\max}=\beta_{\max}=1$ 을 갖는다. 수식 5을 표기법을 다시 적용하면 다음과 같다.

$$\alpha_{1^{st}}\prod_{i=1}^{I_{\max}-1}(1) \leqq \theta_{1^{st}}(threshold) < \prod_{i=1}^{I_{\max}}(\beta_{\min}) \quad (6)$$

수식 6에서 $\prod_{i=1}^{I_{\max}}(\beta_{\min})$ 을 더 간단하게 $\gamma_{1^{st}}$로 표기할 수 있으며 다음과 같다.

$$\alpha_{1^{st}} \leqq \theta_{1^{st}}(threshold) < \gamma_{1^{st}} \quad (7)$$

$\theta_{1^{st}}$은 1번째 요소의 판별 경계면으로써, 정상 사용자의 행동 중 최소값을 만족해야하므로 $\theta_{1^{st}}=\gamma_{1^{st}}$로 사용한다.

이때, 수식 6을 살펴보면 접속 정보($I_{\max}$)의 개수가 늘어날수록 정상 사용자의 판별값이 점진적으로 낮아진다는 것을 알 수 있다. 따라서 본 연구에서 사용하는 접속 정보의 개수(3개)에 적정 최소값($\beta_{\min}$)을 부여하여 나머지 값을 계산한다. 즉, $I_{\max}$가 3개일 때 $\beta_{\min}$에 0.5를 부여하게 되면, $\gamma_{1^{st}}$는 0.125를 갖게 되고, $\alpha_{1^{st}}$는 $\gamma_{1^{st}}$보다 작은 값이 되어야하기 때문에 0.1을 만족하는 것이 가장 효과적이다.

따라서, 상기 방법을 적용하면 다음과 같이 접속 정보의 개수가 2 ~ 5개로 바뀔 때, 각 $\alpha_{1^{st}}$, $\beta_{\min}$, $\theta_{1^{st}}$에 Table 2.과 같은 상수로 할당할 수 있다.

Table 2. Constant Value depend on The $I_{\max}$

$I_{\max}$	2	3	4	5
$\alpha_{1^{st}}$	0.1	0.1	0.1	0.1
$\beta_{\min}$	0.4	0.5	0.6	0.65
$\theta_{1^{st}}=\gamma_{1^{st}}$	0.16	0.125	0.1296	0.116

즉, 사용되는 $I_{\max}$의 개수에 따라 Table 2.의 상수를 이용하여 멤버십 값을 계산하고 프로파일을 형성한다.

$I_{\max}$의 개수에 따라 형성된 Table 2.의 상수들이 공격자에 의해 유출되더라도 사용자별 행동 패턴의 확률이 아니기 때문에 공격자는 정상 사용자의 프로파일에 대한 정보를 유추할 수 없다.

또한 1단계 분석 즉, 웹사이트 접속 정보를 판별할 때만 사용되어, 사용자의 웹사이트 이용 패턴 분석에는 독립적이기 때문에 사용자 행위 패턴 분석에는 영향을 주지 않는다.

3.3.2 1단계 접속 정보 분석 - 사용자 판별 방법

1단계 모델에서 사용자에 대한 판별은 학습기간 동안 형성된 정상 사용자의 프로파일을 기반으로 사용자의 접속 정보를 분석한다. 즉, 사용자가 웹사이트를 접속했을 때 발생하는 접속 정보가 정상 사용자의 프로파일 범주에 속하는지 판단한다. 프로파일 형성 과정에서 할당된 상수 $\alpha_{1^{st}}$, $\beta_{\min}$, $\theta_{1^{st}}$와 판별 수식 7을 이용하며 세부 판별 방법은 다음과 같다.

가. 우선 판별 대상의 로그데이터에서 접속 정보(IP 주소, 접속 시간, 접속 요일 데이터)를 추출한다.

나. 추출한 데이터를 1단계 분석 모델에 적용하면, 정상 사용자의 프로파일에서 추출 데이터에 해당하는 멤버십 값을 찾아 계산한다.

다. 계산된 결과값이 판별 경계면($\theta_{1^{st}}$)을 만족하게 되면 정상 사용자로 판별하고 그렇지 않은 경우는 악의적인 사용자로 판별한다.

3.3.3 2단계 이용 패턴 분석 - 프로파일 형성 방법

본 논문에서는 사용자의 행동과 시간 순서에 따라 발생하는 사건을 모델링하는 방법으로 마르코프 프로세스(markov process) 분석 방법을 사용한다. 마르코프 프로세스란 현재의 사건에 대한 조건 확률이 가장 최근의 사건에 대해서만 영향을 받는 확률적인 특징을 갖는 프로세스로, 실생활에서 확률적인 특징을 갖는 상황에 대해 예측 또는 모델링을 하는데 가장 적합한 확률적 프로세스이다. 특히 시점과 상태를 이산적인 값으로 취할 때는 마르코프 체인(markov chain)이라 한다[17].

본 2단계에서는 이러한 마르코프 체인을 이용하여 웹 사용자 이용 패턴을 실시간으로 분석하며 다음과 같이 정의할 수 있다.

시간 t는 이산적인 값 : $0 \leqq t \leqq T$ (8)

상태 공간(State Space) X : $X=(1,2,....,N)$ (9)

시간 t에서의 상태가 x_t일 때, 상태 천이 벡터 : $(x_0,...,x_t)$ (10)

시간 t+1에서의 상태가 x_{t+1}이 될 확률 : $P(x_{t+1}|x_0,x_1,...,x_t)=P(x_{t+1}|x_t)$ (11)

상태 천이 확률 :

$$P(x_{t+1}\,|\,x_t),\ \sum_{t=0}^{N} P(x_{t+1}\,|\,x_t)=1 \quad (12)$$

초기 상태 확률 : $\pi_u = P(x_0), \sum_{u=0}^{N}\pi_u = 1$ (13)

정상 사용자의 연속적인 어떤 사건들$(x_1,x_2,...,x_t)$이 관측되었을 때, x_{t+1}가 관측될 확률은 바로 이전에 관측된 x_t만 관련 있고, 나머지 과거 관측$(x_1,x_2,...,x_{t-1})$과는 독립이라고 가정한다. 따라서 시간 $T=1,2,...,t$에 대하여 각각 $X=(x_1,x_2,...,x_t)$라는 사건이 순차적으로 관측될 확률은 다음과 같이 구할 수 있다.

$$\begin{aligned} P&=(X)\\ &=P(x_1,x_2,...,x_t)\\ &=P(x_1)P(x_2|x_1)P(x_3|x_2)...P(x_t|x_{t-1})\\ &=P(x_1)\prod_{i=1}^{T-1}P(x_{i+1}|x_i) \end{aligned} \quad (14)$$

정상 사용자의 연속적인 모든 사건들에 대해 수식 14를 이용하면 연속된 사건이 발생할 확률을 구할 수 있다. 이때 관측된 각 사건의 확률들을 이용하여 상태 천이 확률에 대한 행렬을 생성한다. 즉, 정상 사용자의 웹사이트 이용 패턴에 대한 프로파일을 형성한다. 여기서 상태(state)란 정상 사용자가 이동한 웹사이트의 requestURL을 의미한다. 예를 들어, Fig.4.와 같이 사용자가 이동한 경로가 존재할 때, 각 상태에서 다음 상태로 이동한 횟수에 따라, 즉, 발생한 천이에 대한 확률을 구할 수 있으며, 이때 상태 천이 확률을 테이블로 나타내면 Table 3.와 같다.

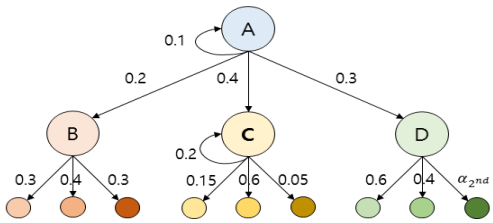

Fig. 4. State Transition Probabilities And Route

Table 3. State Transition Probabilities Tables

From / To	A	B	C	D
A	0.1	0.2	0.4	0.3

From / To	E	F	G	-
B	0.3	0.4	0.3	-

From / To	C	H	I	J
C	0.2	0.15	0.6	0.05

From / To	K	L	M	-
D	0.6	0.4	α_{2nd}	-

이때, 상태에는 존재하나 정상 사용자가 사용하지 않은 상태에 대한 확률은 Fig. 8.의 상태 D처럼 0이 아닌 $\alpha_{2^{nd}}$라는 아주 작은 값을 대체한다. 마르코프 체인은 곱으로 계산되는 수식이기에 사용하지 않은 상태를 선택하게 될 경우, 확률 값이 0이 되기 때문에 $\alpha_{2^{nd}}$를 지정하여 사용한다. $\alpha_{2^{nd}}$를 지정하는 방법은 다음과 같이 사용자마다 가장 최적의 $\alpha_{2^{nd}}$를 계산 및 적용하여 사용한다.

악의적인 사용자가 정상 사용자의 이용 패턴 중에서 가장 높은 천이 확률을 갖는 패턴을 따라할 때, 한 가지는 천이가 존재하지 않는 경우라고 가정하자. 이때의 확률은 정상 사용자의 행동 패턴에서 발생할 수 있는 가장 작은 천이 확률보다 작아야 한다.

$\alpha_{2^{nd}}$: 천이가 존재하지 않는 경우의 확률

$\pi^{i}_{\min}$: 각 상태에서 존재하는 천이 확률 중 가장 작은 값

$\pi^i_{\max}$: 각 상태에서 존재하는 천이 확률 중 가장 큰 값
N : 사용자의 URL 천이 횟수

$$\alpha_{2^{nd}}\prod_{i=1}^{N-1}\pi^i_{\max} \leqq \theta_{2^{nd}}(Threshold) < \prod_{i=1}^{N}\pi^i_{\min} \quad (15)$$

수식 15를 $\alpha_{2^{nd}}$에 대해 정리하면 $\alpha_{2^{nd}} \leqq \frac{\theta_{2^{nd}}(Threshold)}{\prod_{i=1}^{N-1}\pi^i_{\max}} < \frac{\prod_{i=1}^{N}\pi^i_{\min}}{\prod_{i=1}^{N-1}\pi^i_{\max}}$이 된다. 즉, $\alpha_{2^{nd}}$는 $\frac{\prod_{i=1}^{N}\pi^i_{\min}}{\prod_{i=1}^{N-1}\pi^i_{\max}}$보다 충분히 작은 값을 가져야하기 때문에 $0<k<1$의 실수인 임의의 k값을 곱해주면 효과적인 $\alpha_{2^{nd}} = k\frac{\prod_{i=1}^{N}\pi^i_{\min}}{\prod_{i=1}^{N-1}\pi^i_{\max}}$를 가질 수 있고, 수식 15에 대입하면 결과적으로 다음과 같은 수식 16을 갖는다.

$$k\prod_{i=1}^{N}\pi^i_{\min} \leqq \theta_{2^{nd}}(Threshold) < \prod_{i=1}^{N}\pi^i_{\min} \quad (16)$$

수식 16에서 $\prod_{i=1}^{N}\pi^i_{\min}$를 더 간단하게 $\gamma_{2^{nd}}$로 표기하며 다음과 같다.

$$k\gamma_{2^{nd}} \leqq \theta_{2^{nd}}(Threshold) < \gamma_{2^{nd}} \quad (17)$$

이렇게 계산된 $\alpha_{2^{nd}}$는 다음 단계인 사용자 판별에 사용된다.

3.3.4 2단계 이용 패턴 분석 - 사용자 판별 방법

학습기간 동안 형성된 정상 사용자의 프로파일을 기반으로 사용자의 웹사이트 이용 패턴을 실시간으로 분석한다. 즉, 사용자가 웹사이트를 이용할 때마다 발생하는 정보가 정상 사용자의 프로파일 범주에 속하는지 판단한다. 프로파일 형성 과정에서 계산된 $\alpha_{2^{nd}}$와 판별 수식 15를 이용한다. 또한 $\theta_{2^{nd}}$는 2단계 모델의 판별 경계면으로써, 수식 16와 수식 17에서 정상 사용자의 행동 중 최소값을 만족해야하므로 $\theta_{2^{nd}} = \gamma_{2^{nd}}$로 사용한다. 세부 판별 방법은 다음과 같다.

가. 우선 판별 대상의 로그데이터에서 웹사이트 URL을 추출한다.

나. 추출된 데이터를 2단계 분석 모델에 적용하면, 정상 사용자의 프로파일에서 추출 데이터에 해당하는 상태 천이 확률 값을 찾아 마르코프 체인 모델 수식 14로 계산한다.

다. 계산된 결과값이 판별 경계면($\theta_{2^{nd}}$)을 만족하게 되면 정상 사용자로 판별하고 그렇지 않은 경우는 악의적인 사용자로 판별한다.

IV. 실험 및 결과

본 연구는 일정기간 동안 정상 사용자의 로그데이터를 기반으로 단계별 분석모델을 적용하여 프로파일을 형성하고, 사용자의 실시간 사용 패턴을 판별한다. 세부 실험 과정은 다음과 같다.

4.1 실험 환경

연구에 사용한 로그데이터는 충청도 기관 웹사이트의 일부 기능을 구현하여 수집했다. 구현된 기능은 로그인, 회원가입, 메인 메뉴, 하위 메뉴, 각종 게시판 등이며 메뉴를 세부적으로 살펴보면 Fig.5.와 같이 구성되어 있다.

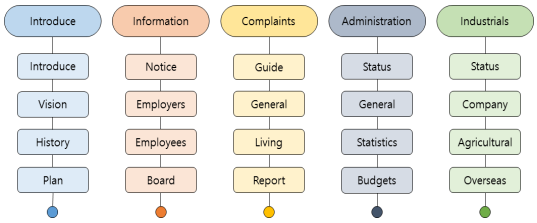

Fig. 5. Menu of Web Site

4.2 데이터 수집

데이터 수집은 3명의 사용자(사용자 A, B, C)를 대상으로 진행했으며, 4주 동안 발생한 로그데이터를 사용했다. 각 사용자마다 자기 자신만의 패턴으로 웹사이트를 접속 및 이용했다. 이러한 로그데이터를 기반으로 사용자별 프로파일을 형성했으며, 사용자

1, 3은 약 2000개, 사용자 2는 약 3000개의 로그 데이터를 수집했다.

4.3 데이터 분석

4주간 수집된 데이터 중에서 ProcessTime, ClientIP, requestURL, userID, sessionID를 분석항목으로 결정하고, 이를 기반으로 1단계 분석 모델에서 접속 시간, 접속 요일, 접속 IP 주소 정보를 분석하고, 2단계 분석 모델에서는 접속 URL 정보를 분석한다. 분석된 주간 데이터에 시간 가중치를 적용하며, 각각 〔1주 전 데이터 : $w_t(0.5)$, 2주 전 데이터 : $w_{t-1}(0.3)$, 3주 전 데이터 : $w_{t-2}(0.15)$, 4주 전 데이터 : $w_{t-3}(0.05)$, 총합 : 1〕를 사용한다.

상기의 조건으로 분석 모델을 실행하면, 수집된 데이터에서 각 단계에서 사용자별 프로파일이 생성되며, 이를 기반으로 실시간 사용자 판별은 다음과 같은 순서로 실험을 진행한다.

1) 사용자 로그인 시, 접속한 userID를 기반으로 1단계 분석(접속 시간, 접속 요일, 접속 IP 주소)을 진행한다.

2) 분석된 정보에 해당하는 멤버십 값을 해당 사용자의 1단계 프로파일에서 불러와 전체 결과값을 계산한다.

3) 계산된 결과가 1단계 판별 경계면을 만족한다면 정상, 만족하지 않는다면 악의적인 사용자로 판별한다.

4) 사용자가 실시간으로 이동하는 URL에 따라 2단계 분석(접속 URL)을 진행한다.

5) 분석된 정보에 해당하는 천이 확률 값을 해당 사용자의 2단계 프로파일에서 불러와 마르코프 확률 값을 계산한다.

6) 계산된 결과가 2단계 판별 경계면을 만족한다면 정상, 만족하지 않는다면 악의적인 사용자로 판별한다.

이러한 판별 결과를 정상 판정(True Positive / True Negative)이라고 하며, 더욱 정확한 결과를 얻기 위해 부정 판정(False Positive / False Negative)에 대한 판별도 반복적으로 수행한다. 즉, 정상 사용자 계정을 사용하는 악의적인 사용자를 탐지하기 위함이며, 위의 실험 과정과 동일하게 진행된다. 사용자별 정상 판정과 부정 판정 각각 100번씩 판별 테스트를 진행한다.

4.4 실험 결과

실험 과정에 따라 분석 모델을 실행하여 사용자별 프로파일을 생성하면, 결과적으로 다음과 같은 프로파일 정보를 얻을 수 있다.

가. 사용자 A의 프로파일 정보

Table 4. Profile of User A in Step 1

Object	Contents
$I_{\max}$	3
$\alpha_{1^{st}}$	0.1
$\beta_{\min}$	0.5
$\theta_{1^{st}} = \gamma_{1^{st}}$	0.125
Threshold	$0.1 \leqq \theta_{1^{st}}(threshold) < 0.125$

Table 5. Profile of User A in Step 2

Object	Contents
k	0.5
N	5
$\alpha_{2^{nd}}$	2.09×10^{-7}
$\theta_{2^{nd}} = \gamma_{2^{nd}}$	1.90×10^{-6}
Threshold	$9.49 \times 10^{-7} \leqq \theta_{2^{nd}}(Threshold) < 1.90 \times 10^{-6}$

나. 사용자 B의 프로파일 정보

Table 6. Profile of User B in Step 1

Object	Contents
$I_{\max}$	3
$\alpha_{1^{st}}$	0.1
$\beta_{\min}$	0.5
$\theta_{1^{st}} = \gamma_{1^{st}}$	0.125
Threshold	$0.1 \leqq \theta_{1^{st}}(threshold) < 0.125$

Table 7. Profile of User B in Step 2

Object	Contents
k	0.5
N	5
$\alpha_{2^{nd}}$	5.67×10^{-7}
$\theta_{2^{nd}} = \gamma_{2^{nd}}$	2.45×10^{-7}
Threshold	$1.22 \times 10^{-7} \leqq \theta_{2^{nd}}(Threshold) < 2.45 \times 10^{-7}$

다. 사용자 C의 프로파일 정보

Table 8. Profile of User C in Step 1

Object	Contents
$I_{\max}$	3
$\alpha_{1^{st}}$	0.1
$\beta_{\min}$	0.5
$\theta_{1^{st}} = \gamma_{1^{st}}$	0.125
Threshold	$0.1 \leqq \theta_{1^{st}}(threshold) < 0.125$

Table 9. Profile of User C in Step 2

Object	Contents
k	0.5
N	5
$\alpha_{2^{nd}}$	7.98×10^{-7}
$\theta_{2^{nd}} = \gamma_{2^{nd}}$	2.77×10^{-3}
Threshold	$1.38 \times 10^{-3} \leqq \theta_{2^{nd}}(Threshold) < 2.77 \times 10^{-3}$

이러한 프로파일의 내용은 단순히 2차원적인 그림과 표만으로 나타내기에는 방대한 데이터이다. 따라서 생성된 프로파일을 기반으로 사용자별 이동 경로 및 확률의 일부만 살펴보면 아래의 Fig, Table과 같다. 파란색 점으로 표시한 부분은 추가적인 경로가 존재하는 것임을 나타낸다.

가. 사용자 A의 이동 경로 및 확률

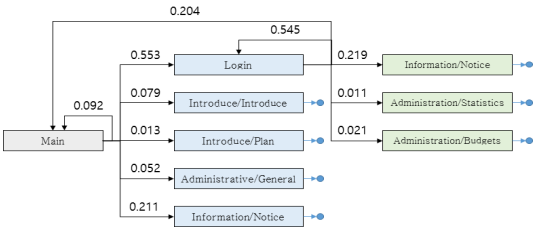

Fig. 6. Route of User A

Table 10. State Transition Probabilities Tables of User A

From / To	Main	Login	Introduce/Introduce	Introduce/Plan	Administration/General	Information/Notice	Total
Main	0.092	0.553	0.079	0.013	0.052	0.211	1

From / To	Login	Main	Information/Notice	Administration/Statistics	Administration/Budgets	Total
Login	0.092	0.204	0.553	0.079	0.013	1

나. 사용자 B의 이동 경로 및 확률

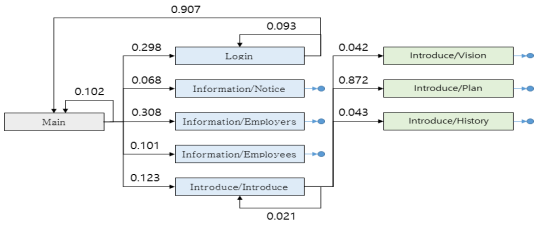

Fig. 7. Route of User B

Table 11. State Transition Probabilities Tables of User B

From / To	Main	Login	Information/Notice	Information/Employers	Information/Employees	Introduce/Introduce	Total
Main	0.102	0.298	0.068	0.308	0.101	0.123	1

From / To	Introduce/Introduce	Introduce/Vision	Introduce/Plan	Introduce/History	Total
Introduce/Introduce	0.092	0.553	0.079	0.013	1

다. 사용자 C의 이동 경로 및 확률

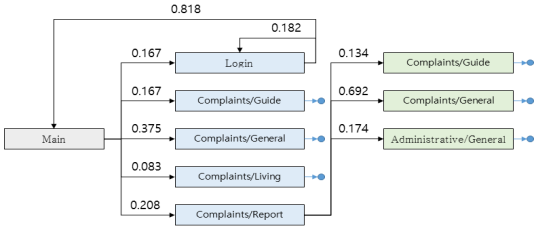

Fig. 8. Route of User C

Table 12. State Transition Probabilities Tables of User C

From / To	Login	Complaints/Guide	Complaints/General	Complaints/Living	Complaints/Report	Total
Main	0.092	0.553	0.079	0.013	0.052	1

From / To	Complaints/Guide	Complaints/General	Administrative/General	Total
Complaints/Report	0.134	0.692	0.174	1

이러한 프로파일을 기반으로, 사용자를 판별한 결과는 다음과 같다.

사용자별 판정 결과를 통계적 분류 기준에 따라 정밀도(precision), 특이도(specificity), 정확도(accuracy), TPR(True Positive Rate), FPR(False Positive Rate) 등을 아래의 수식 18, 19, 20을 이용해서 계산할 수 있으며, Table

15.와 같은 결과를 얻을 수 있다[18].

Table 13. Results of Users (%)

Object		Result	
		True	False
A	True	93	7
	False	6	94
B	True	96	4
	False	3	97
C	True	97	3
	False	1	99

Table 14. Classification Criteria

Objects / Results	True	False
True	True Positive(TP)	False Negative(FN)
False	False Positive(FP)	True Negative(TN)

$$\text{정밀도} : \frac{TP}{TP+FP}\times 100 \quad (18)$$

$$\text{특이도} : \frac{TN}{FP+TN}\times 100 \quad (19)$$

$$\text{정확도} : \frac{TP+TN}{TP+FN+FP+TN}\times 100 \quad (20)$$

Table 15. Result Analysis of Users

Objects	Precision	Specificity	TPR	FPR	Accuracy
A	93.94%	94%	93%	6%	93.5%
B	96.97%	97%	96%	3%	96.5%
C	98.98%	99%	97%	1%	98%

사용자 3명의 데이터를 종합해본 결과, 평균적으로 96%의 높은 판별 정확도를 확인할 수 있었다.

V. 결론 및 향후 연구

본 연구에서는 정상 사용자의 프로파일을 이용해 실시간으로 사용자의 이용 패턴을 분석하여 정상 사용자를 판별하는 방법을 제안하였다. 제안된 모델은 사용자별 확률적 모델링을 위해 멤버십 분석과 마르코프 체인 모델을 사용하여 분석하였다. 이러한 방법을 사용하여 개인정보를 도용한 악의적인 사용자를 적발할 수 있고, 정상적인 사용자이더라도 민감한 정보에 접근하는 것을 방지할 수 있다.

Juan el at.(2004) 연구처럼 기존 대부분의 연구에서 마르코프 체인 모델을 활용해 하나의 프로파일을 생성하여 사용자의 이상 탐지 모델을 제안하였지만, 각 사용자를 구별할 수 없다는 한계점을 갖고 있다[13]. 본 연구는 이러한 한계점을 보완하여 각 사용자별 프로파일을 기반으로 정상 사용자를 판별할 수 있으며, 더 나아가 정상 사용자 임에도 불구하고 민감한 데이터에 악의적인 접근하는 것 또한 탐지가 가능하다.

본 연구의 탐지 방법의 가장 큰 핵심 중 하나는 데이터 정보량이라고 할 수도 있다. 즉, 정상 사용자별 분석 데이터가 충분히 클수록 더욱 정확하게 판별할 수 있다는 것이다. 하지만 본 연구에서는 전체 데이터 크기가 변화하는 상황에 대한 정확도를 측정하지 못하여 추후 연구가 필요하다.

또한, 단계별 사용자 판별 방법에서 단순히 정상 / 비정상의 결과를 확인하는 것이 아니라, 데이터의 민감한 정도와 분석 크기에 따라 결과를 세분화 할 수 있다. 예를 들어, 정상 경계면(θ) 값 미만의 경우, 경계 범위를 지정하여 여러 위험 레벨로 세분화 할 수 있다. 이러한 방법은 정상을 비정상으로 오탐할 가능성을 낮춰줄 수 있는 장점이 있다.

마지막으로, 웹을 통해 개인정보시스템에 접근해서 정보를 탈취하려는 위협에 대해, 마찬가지로 접근 패턴 분석을 이용해 실시간으로 정상 사용자를 판별하는 추가적인 연구가 필요하며, 이는 SQL Query 정보와 관련된 로그데이터 분석이 수행되어야 할 것이다. 차후 이러한 점을 개선하고 증명할 수 있다면, 판별 성능을 더욱 향상 시킬 수 있을 것으로 보인다.

References

[1] Huang, Xinyi, et al. "Robust multi-factor authentication for fragile communications." IEEE Transactions on Dependable and Secure Computing, vol. 11, no. 6, pp. 568-581. Nov. 2014.

[2] Peng, Jian, Kim-Kwang Raymond Choo, and Helen Ashman. "User profiling in intrusion detection: A review." Journal of Network and Computer Applications, vol. 72, pp. 14-27, Jul. 2016.

[3] Umphress, David, and Glen Williams.

"Identity verification through keyboard characteristics." International journal of man-machine studies, vol. 23, no. 3, pp. 263-273, Apr. 1985.

[4] Bergadano, Francesco, Daniele Gunetti, and Claudia Picardi. "User authentication through keystroke dynamics." ACM Transactions on Information and System Security (TISSEC), vol. 5, no. 4, pp. 367-397 Nov. 2002.

[5] Revett, Kenneth. "A bioinformatics based approach to user authentication via keystroke dynamics." International Journal of Control, Automation and Systems, vol. 7, no. 1, pp. 7-15, Mar. 2009.

[6] Pannell, Grant, and Helen Ashman. "User modelling for exclusion and anomaly detection: a behavioural intrusion detection system." International Conference on User Modeling, Adaptation, and Personalization. Springer Berlin Heidelberg, LNCS 6075, pp. 207-218, 2010.

[7] Alexandre, Thomas J. "Biometrics on smart cards: An approach to keyboard behavioral signature." Future Generation Computer Systems, vol. 13, no. 1, pp. 19-26, Jul. 1997.

[8] Ling, Li, Sui Song, and C. N. Manikopoulos. "Windows nt user profiling for masquerader detection." 2006 IEEE International Conference on Networking, Sensing and Control. IEEE, pp. 386-391, Apr. 2006.

[9] Vizer, Lisa M., Lina Zhou, and Andrew Sears. "Automated stress detection using keystroke and linguistic features: An exploratory study." International Journal of Human-Computer Studies, vol. 67, no. 10, pp. 870-886, Aug. 2009.

[10] Bhaskaran, Nisha, et al. "Deceit detection via online behavioral learning." Proceedings of the 2011 ACM Symposium on Applied Computing. ACM, pp 29-30, Mar. 2011.

[11] Ju, Wen-Hua, and Yehuda Vardi. "A hybrid high-order Markov chain model for computer intrusion detection." Journal of Computational and Graphical Statistics, vol. 10. no. 2, pp. 277-295, Jan. 2001.

[12] Ye, Nong, Yebin Zhang, and Connie M. Borror. "Robustness of the Markov-chain model for cyber-attack detection." IEEE Transactions on Reliability, 53(1), pp. 116-123, Mar 2004.

[13] Jongho Choy et al "Application of Hidden Markov Model to Intrusion Detection System." Journal of KISS : Software and Applications, vol. 2, no. 6, pp. 429-438, Jun. 2001.

[14] J.M. Estévez-Tapiador, P. García-Teodoro, J.E. DíazVerdejo, "Measuring Normality in HTTP Traffic for Anomaly-Based Intrusion Detection." in. Computer Networks, vol. 45, no. 2, pp. 175-193, Jun 2004.

[15] Zadeh, Lotfi A. "Fuzzy sets." Informationand control, vol. 8, no. 3, pp. 338-353, Jun 1965.

[16] Monthly Electronic Technology. "Electronics Dictionaries." Seongandang, pp. 643, 2005.

[17] Diaconis, Persi, and David Freedman. "de Finetti's theorem for Markov chains." The Annals of Probability, vol. 8, no, 1, pp. 115-130, Feb. 1980.

[18] Sokolova, Marina, and Guy Lapalme. "A systematic analysis of performance measures for classification tasks." Information Processing & Management, vol. 45, no. 4, pp. 427-437, Jul. 2009.

〈저자소개〉

장 진 구 (Jin Gu Jang) 정회원
2008년 2월: 고려대학교 전자 및 정보공학과 졸업
2015년 3월~현재: 고려대학교 정보보호대학원 석사과정
〈관심분야〉 정보보호, 패턴인식, 네트워크 보안, 클라우드 보안

문 종 섭 (Jong Sub Moon) 종신회원
1981년 1월: 서울대학교 계산통계학과 졸업
1983년 1월: 서울대학교 대학원 계산통계학과 석사
1991년 5월: Illinois Institute of Technology 전산학 박사
2002년 3월~현재: 고려대학교 전자 및 정보공학과 교수
〈관심분야〉 정보보호, 전자공학, 통신공학