



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0020265
(43) 공개일자 2013년02월27일

(51) 국제특허분류(Int. Cl.)
H04L 12/26 (2006.01) H04L 12/22 (2006.01)
H04L 9/00 (2006.01) G06F 21/00 (2006.01)

(21) 출원번호 10-2011-0082787

(22) 출원일자 2011년08월19일

심사청구일자 2011년08월19일

(71) 출원인
고려대학교 산학협력단
서울 성북구 안암동5가 1

(72) 발명자
김희강
서울특별시 서초구 양재동
고폴린
서울특별시 종로구 창경궁로22길 5-7 (원남동)

(74) 대리인
특허법인충현

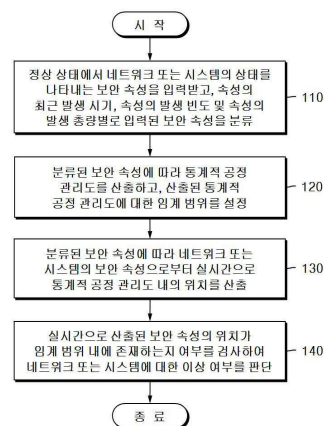
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 통계적 공정 관리도를 이용하여 이상증후를 탐지하는 방법

(57) 요약

본 발명은 통계적 공정 관리도를 이용하여 이상증후를 탐지하는 방법 및 그 방법을 기록한 기록매체에 관한 것으로, 본 발명에 따른 이상증후를 탐지하는 방법은, 정상 상태에서 네트워크 또는 시스템의 상태를 나타내는 보안 속성을 입력받고, 속성의 최근 발생 시기, 속성의 발생 빈도 및 속성의 발생 총량별로 입력된 보안 속성을 분류하고, 분류된 보안 속성에 따라 통계적 공정 관리도를 산출하며, 산출된 통계적 공정 관리도에 대한 임계 범위를 설정하고, 분류된 보안 속성에 따라 네트워크 또는 시스템의 보안 속성으로부터 실시간으로 통계적 공정 관리도를 산출하며, 실시간 통계적 공정 관리도가 임계 범위 내에 존재하는지 여부를 검사하여 네트워크 또는 시스템에 대한 이상 여부를 판단한다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호	ITAC1090111100020001000100100
부처명	정보통신산업진흥원
연구사업명	대학IT연구센터육성지원사업
연구과제명	스마트 그리드 보안 연구센터(스마트 그리드 보안기술 개발)
주관기관	고려대학교 산학협력단
연구기간	2011.01.01 ~ 2011.12.31

특허청구의 범위

청구항 1

적어도 하나의 프로세서(processor)를 구비한 탐지 장치가 이상증후를 탐지하는 방법에 있어서,
정상 상태에서 네트워크 또는 시스템의 상태를 나타내는 보안 속성을 입력받고, 상기 속성의 최근 발생 시기, 상기 속성의 발생 빈도 및 상기 속성의 발생 총량별로 상기 입력된 보안 속성을 분류하는 단계;
상기 분류된 보안 속성에 따라 통계적 공정 관리도(statistical process control)를 산출하고, 산출된 통계적 공정 관리도에 대한 임계 범위를 설정하는 단계;
상기 분류된 보안 속성에 따라 상기 네트워크 또는 시스템의 상기 보안 속성으로부터 실시간으로 통계적 공정 관리도 내의 위치를 산출하는 단계; 및
상기 실시간으로 산출된 보안 속성의 위치가 상기 설정된 임계 범위 내에 존재하는지 여부를 검사하여 상기 네트워크 또는 시스템에 대한 이상 여부를 판단하는 단계를 포함하는 방법.

청구항 2

제 1 항에 있어서,
상기 산출된 통계적 공정 관리도를 시각화하고, 상기 이상증후가 감지된 네트워크 또는 시스템의 보안 속성을 상기 설정된 임계 범위와 비교하여 상기 시각화된 통계적 공정 관리도 상에 표시하는 단계를 더 포함하는 방법.

청구항 3

제 1 항에 있어서,
상기 통계적 공정 관리도는,
상기 보안 속성에 대한 표본의 평균을 중앙선으로 설정하고,
상기 표본의 표준편차의 소정 배율만큼을 상기 임계 범위로 설정하는 것을 특징으로 하는 방법.

청구항 4

제 1 항에 있어서,
상기 통계적 공정 관리도는,
상기 보안 속성의 변동 범위가 상기 임계 범위 내에 있도록 상기 보안 속성에 대한 표본의 평균의 소정 배율만큼을 상기 임계 범위로 설정하는 것을 특징으로 하는 방법.

청구항 5

제 1 항에 있어서,
상기 통계적 공정 관리도는,
상기 보안 속성에 대한 표본의 실패율의 평균을 중앙선으로 설정하고,
상기 실패율의 분포의 표준편차만큼을 상기 임계 범위로 설정하는 것을 특징으로 하는 방법.

청구항 6

제 1 항에 있어서,
상기 보안 속성은 정규분포에 따라 발생하는 것을 특징으로 하는 방법.

청구항 7

제 1 항 내지 제 6 항 중에 어느 한 항의 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을

수 있는 기록매체.

명세서

기술분야

- [0001] 본 발명은 네트워크 또는 시스템의 이상증후를 탐지하는 방법에 관한 것으로, 특히 네트워크 트래픽이나 보안 이벤트가 정규분포에 따르는 환경에서 통계적 공정 관리도를 이용하여 비정상적인 공격 및 이상증후를 실시간으로 탐지할 수 있는 방법 및 그 방법을 기록한 기록매체에 관한 것이다.

배경기술

- [0002] 전통적으로 생산 활동으로 획득되는 생산 제품에 관한 품질을 관리하고자 하는 다양한 노력의 일환으로 행해진 통계적인 품질 관리(statistical quality control)는 시장성 있는 제품을 가장 경제적으로 생산하기 위해서 생산 활동의 모든 단계에서 통계적인 방법을 응용하고자 하는 시도에도 접근되었다.
- [0003] 이와 관련하여, 통계적 공정 관리(statistical process control, SPC)는 통계적 자료와 분석 기법으로 품질과 공정의 상태를 파악하여 원하는 생산자/관리자가 희망하는 상태 및 품질에 따라 생산이 이루어지도록 관리하는 방법을 의미한다. 즉, 통계적 공정 관리 활동이라 함은, 공정 품질 유지 활동과 개선 활동을 포함한다. 따라서, 통계적 공정 관리 활동은 공정의 이상이 발생한 것을 사전 탐지하여 기피 요인은 찾아 제거하거나, 이를 사전에 예방하는 품질 관리 활동이다.
- [0004] 통계적 공정 관리에 있어서, 과거에는 주로 산포를 줄이고자 하는 노력이 이루어졌으나, 최근에는 목표치의 설정 또는 달성 방법의 변경 정도를 연구하여 공정의 능력이나 공정에 수반되는 다양한 인자들이 적절하게 설정되어 있는지 여부를 분석하는 활동을 포함하여 통계적 방법을 활용하여 공정의 상태를 관리하는 과정 전부를 포괄하는 추세이다. 이하에서 소개되는 선행기술문헌에는 이러한 통계적 공정 관리가 생산 공정 내에서 어떠한 흐름에 따라 활용될 수 있는지, 그 기법들은 어떠한지에 관한 개괄적인 내용을 소개하고 있다.

선행기술문헌

비특허문헌

- [0005] (비특허문헌 0001) 신용백, 통계적 공정관리(SPC)를 위한 통계적 품질관리(SQC)수법, 전기제품 안전21 11 52-59, 한국제품안전협회, 2007.

발명의 내용

해결하려는 과제

- [0006] 본 발명이 해결하고자 하는 기술적 과제는, 종래의 침입 탐지 기술들에 있어서 이상증후 탐지법이 높은 오탐율을 갖는 문제점을 해결하고, 통계적 공정 관리도가 단순한 공정 내지 품질 관리에 국한되어 활용되는 한계를 극복하며, 현재의 네트워크 또는 시스템의 보안 상황에 대한 정보를 직관적이고 시각적으로 제공할 수 있는 수단이 존재하지 않음으로써 발생하는 관리상의 불편함을 해소하고자 한다.

과제의 해결 수단

- [0007] 상기 기술적 과제를 해결하기 위하여, 본 발명의 일 실시예에 따른 이상증후를 탐지하는 방법은, 정상 상태에서 네트워크 또는 시스템의 상태를 나타내는 보안 속성을 입력받고, 상기 속성의 최근 발생 시기, 상기 속성의 발생 빈도 및 상기 속성의 발생 총량별로 상기 입력된 보안 속성을 분류하는 단계; 상기 분류된 보안 속성에 따라 통계적 공정 관리도(statistical process control)를 산출하고, 산출된 통계적 공정 관리도에 대한 임계 범위를 설정하는 단계; 상기 분류된 보안 속성에 따라 상기 네트워크 또는 시스템의 상기 보안 속성으로부터 실시간으로 통계적 공정 관리도 내의 위치를 산출하는 단계; 및 상기 실시간으로 산출된 보안 속성의 위치가 상기 설정된 임계 범위 내에 존재하는지 여부를 검사하여 상기 네트워크 또는 시스템에 대한 이상 여부를 판단하는 단계를 포함한다.

- [0008] 상기된 이상증후를 탐지하는 방법에서, 보안 속성은 정규분포에 따라 발생한다.
- [0009] 상기된 이상증후를 탐지하는 방법에서, 상기 통계적 공정 관리도는, 상기 보안 속성에 대한 표본의 평균을 중앙선으로 설정하고, 상기 표본의 표준편차의 소정 배율만큼을 상기 임계 범위로 설정할 수 있다.
- [0010] 또한, 상기된 이상증후를 탐지하는 방법에서, 상기 통계적 공정 관리도는, 상기 보안 속성의 변동 범위가 상기 임계 범위 내에 있도록 상기 보안 속성에 대한 표본의 평균의 소정 배율만큼을 상기 임계 범위로 설정할 수 있다.
- [0011] 또한, 상기된 이상증후를 탐지하는 방법에서, 상기 통계적 공정 관리도는, 상기 보안 속성에 대한 표본의 실패율의 평균을 중앙선으로 설정하고, 상기 실패율의 분포의 표준편차만큼을 상기 임계 범위로 설정할 수 있다.
- [0012] 나아가, 상기된 네트워크의 이상증후를 탐지하는 방법은, 상기 산출된 통계적 공정 관리도를 시각화하고, 상기 이상증후가 감지된 네트워크의 보안 속성을 상기 설정된 임계 범위와 비교하여 상기 시각화된 통계적 공정 관리도 상에 표시하는 단계를 더 포함한다.
- [0013] 한편, 이하에서는 상기 기재된 이상증후를 탐지하는 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

발명의 효과

- [0014] 본 발명의 다양한 실시예들은 정상 상태에서 상기 네트워크 또는 시스템의 상태를 나타내는 보안 속성을 입력받고, 최근 발생 시기, 발생 빈도 및 발생 총량별로 분류된 보안 속성에 따라 통계적 공정 관리도를 산출하여 실시간 통계적 공정 관리도와 비교함으로써, 이상증후 탐지법의 정확도를 향상시키고, 통계적 공정 관리도를 이상증후 탐지의 관리 수단으로써 활용하여 이를 시각화하여 관리자에게 제공함으로써, 사용자로 하여금 현재의 네트워크 또는 시스템의 보안 상황에 대한 정보를 직관적으로 제공할 수 있다.

도면의 간단한 설명

- [0015] 도 1은 본 발명의 일 실시예에 따른 네트워크의 이상증후를 탐지하는 방법을 도시한 흐름도이다.
- 도 2는 본 발명의 일 실시예에 따른 도 1의 이상증후 탐지 방법에서 네트워크의 상태를 나타내는 보안 속성을 분류하고 정렬하는 방법을 예시한 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 도 1의 이상증후 탐지 방법과 관련하여 통계적 공정 관리도를 설명하기 위한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 도 1의 이상증후 탐지 방법에서 이상증후가 감지된 네트워크의 보안 속성을 임계 범위와 비교하여 시각화된 통계적 공정 관리도 상에 표시하는 방법을 예시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 본 발명의 실시예들을 설명하기에 앞서, 우선 본 발명의 실시예들이 구현되는 기술 분야, 즉 침입 탐지 기술의 개괄에 대해 소개하고, 이러한 실시예들이 구현되는 환경의 특징으로부터 안출될 수 있는 본 발명의 기본적인 아이디어를 제시하고자 한다.
- [0017] 침입 탐지(intrusion detection)라 함은 정보 시스템의 보안을 위협하는 침입 행위가 발생할 경우 이를 탐지하는 기술로서, 침입 탐지 시스템(intrusion detection system, IDS)은 일반적으로 시스템에 대한 위협이 되는 내부 및 외부로부터의 조작을 탐지하여 관리자에게 알려준다. 이를 위해 침입 탐지 시스템은 전통적인 방화벽이 탐지할 수 없는 모든 종류의 악의적인 네트워크 트래픽 및 컴퓨터 사용을 탐지할 수 있어야 한다. 따라서, 침입 탐지 시스템이 탐지하는 대상은, 취약한 서비스에 대한 네트워크 공격과 애플리케이션에서의 데이터 처리 공격(data driven attack), 권한 상승(privilege escalation) 및 침입자 로그인, 침입자에 의한 주요 파일 접근, 멀웨어(malware)(컴퓨터 바이러스, 트로이 목마, 웜 등)와 같은 호스트 기반 공격을 포함한다.
- [0018] 이러한 침입 탐지 기술은 크게 이상증후 탐지법(anomaly based intrusion detection)과 오용 탐지법(misuse detection)의 두 가지로 분류될 수 있다. 이상증후 탐지법은 네트워크 또는 시스템의 상태가 과거의 통계적 정상 행위와 다른 비정상 행위를 보일 경우 이를 침입으로 간주하는 방법이며, 오용 탐지법은 네트워크 또는 시스템의 상태가 미리 설정된 공격 패턴과 일치할 경우 이를 침입으로 간주하는 방법이다. 특히, 이상증후 탐지법은 통계에 기반하는 접근법 또는 예측 모델링을 활용하며, 알려지지 않은 공격이나 보안 장치를 우회하는 공격 등

기존의 보안 시스템에서 정의하고 있지 않은 예상치 못한 공격을 탐지할 때 유용하다고 알려져 있다.

- [0019] 이와 관련하여, 이상증후 탐지를 위해 활용될 수 있는 재료 및 기술적 수단으로는 통계(statistics), 전문가 시스템(expert system), 신경망(neural networks), 컴퓨터 면역학(computer immunology), 데이터 마이닝(data mining) 및 HMM(hidden Markov models) 등이 존재한다. 이 중, 특히 통계 기반 탐지 방법론은 침입 탐지 시스템에서 가장 많이 적용되는 방법론으로서, 통계적인 특징 및 관련 공식을 사용하여 침입 가능성을 추정하게 된다. 즉, 보안과 관련된 다양한 상태 변수의 통계적 수치(평균, 분산, 표준편차 등이 될 수 있다.)를 산출하여 침입 판정을 위한 판단의 근거로서 활용한다.
- [0020] 다만, 이러한 이상증후 탐지법은 그 속성상 오탐율이 높게 나타날 수 있는데, 본 발명의 실시예들은 이러한 오탐율(false-positive)을 낮추기 위해 RFM(recency, frequency, monetary value) 분석 기법과 통계적 공정 관리 기법을 도입하여 새로운 탐지 방법을 제안하고자 한다. 각각의 개별 기술 요소들은 이후 도 1 내지 도 3을 통해 구체적으로 설명한다.
- [0021] 한편, 수집되고 분석된 데이터를 사용자 내지 관리자에게 효과적으로 보고할 수 있는 다양한 기법들이 존재한다. 이 중, 시각화 보고 기술은 보고하고자 하는 데이터를 가공하여 사용자에게 직관적인 영상으로 제공함으로써, 보다 신속하고 정확한 판단을 돕는다. 네트워크 또는 시스템 보안과 관련하여, 기존의 보안 관제 기술들은 현재의 네트워크 상황에 대한 많은 양의 정보를 제공하는데 집중할 뿐, 이러한 보고 내용에 대한 이해와 판단은 전적으로 사용자에게 의존하였다. 이로 인해, 상당한 보안 전문 지식이 없을 경우, 사용자는 현재의 네트워크 또는 시스템의 상황을 정확하게 이해하거나, 앞으로 변화하는 보안 상태를 예측하는 것이 용이하지 않았다. 결과적으로, 고도의 전문가가 아닌 이상 이러한 보안 관제 시스템을 효과적으로 활용할 수 없을 뿐만 아니라, 해당 분야의 전문가 역시 이러한 보안 관련 정보를 분석하고 예측하는데 많은 시간이 소요되었다. 따라서, 각종 침입 탐지와 관련하여 실시간으로 네트워크 또는 시스템을 감시, 분석 및 대응하기 위한 효과적인 보안 관제 서비스가 필요하며, 신속하고 정확한 판단에 도움을 줄 수 있는 시각화 도구가 요청된다.
- [0022] 이하에서 제시될 본 발명의 실시예들은 침입 탐지 기술과 RFM 분석, 통계적 공정 관리 기법을 효과적으로 융합하되, 그 결과를 사용자에게 쉽고 직관적으로 제시할 수 있는 시각화 보고 기술을 추가적으로 포함한다. 즉, 본 발명의 실시예들은 통계적 방법에 기반하여 이상증후를 탐지한 후, 이를 시각화하여 사용자에게 보고한다. 특히, RFM 분석 기법과 통계적 공정 관리 기법을 결합함으로써 이상증후 탐지 기법에서 발생하는 오탐율을 크게 낮출 수 있으며, 이로 인해 신뢰도 높은 분석이 가능하다. 또한, 통계적 공정 관리도에 기초한 시각화 기술을 통해 단순한 보안 상황을 인지시키는 정도에 그치는 것이 아니라, 이상증후 탐지의 상황과 그 결과를 사용자에게 제시할 수 있으며, 전문 지식이 부족한 사용자도 현재의 네트워크 내지 시스템의 상태를 쉽게 이해하도록 돕는다.
- [0023] 이하에서는 도면을 참조하여 본 발명의 실시예들을 구체적으로 설명한다.
- [0024] 도 1은 본 발명의 일 실시예에 따른 적어도 하나의 프로세서(processor)를 구비한 탐지 장치가 네트워크의 이상증후를 탐지하는 방법을 도시한 흐름도로서, 다음과 같은 단계들을 포함한다.
- [0025] 110 단계에서, 탐지 장치는 정상 상태에서 네트워크 또는 시스템의 상태를 나타내는 보안 속성을 입력받고, 보안 속성의 최근 발생 시기, 보안 속성의 발생 빈도 및 보안 속성의 발생 총량별로 이상에서 입력된 보안 속성을 분류한다. 이러한 110 단계는 탐지 장치가 관측하고자 하는 대상 데이터를 수집하는 단계로서, 이러한 대상 데이터에는 네트워크의 패킷(packet) 정보 또는 시스템 로그(system log) 정보가 포함될 수 있다.
- [0026] 이하에서 기술되는 실시예들은 110 단계에서 입력 값으로 사용되는 보안 속성을 네트워크의 보안 상태를 나타내는 이벤트에 집중하여 예시하고 있으나, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명의 다양한 실시예들에서 이러한 입력 값이 단지 네트워크의 보안 이벤트에 제약되지 않으며, 다양한 시스템 속성 중 어느 하나가 될 수 있음을 알 수 있다. 예를 들어, 이러한 보안 속성이란 시스템의 이벤트 로그를 활용하여 파악이 가능하며, 다양한 운영체제에서 각각의 시스템으로부터 적합하게 선택될 수 있다. 예를 들어 윈도우즈 운영체제의 경우, 어플리케이션 로그(application log), 시스템 로그(system log), 시큐리티 로그(security log) 등이 활용 가능한 보안 속성이며, UNIX 시스템의 경우, wtmpx, utmpx, last log, history log, syslog, messages, secure log, authlog, pacct, ftp/http log 등의 활용 가능한 보안 속성이 될 것이다. 나아가, 사용자 행태 정보와 같은 정보 역시 시스템, 사용자, 데이터베이스 또는 네트워크의 이상 증후를 탐지할 수 있는 보안 속성으로 활용 가능하다.
- [0027] 데이터를 분류함에 있어서, 본 실시예는 RFM 분석 기법을 활용하여, 보안 속성들로부터 특징을 추출하고, 추출

된 특징으로부터 분석에 필요한 정보를 수집한다. 통상적으로 RFM 은 최근 발생 시기(recency), 발생 빈도(frequency) 및 발생 금액(monetary value)의 3가지 측정 기준들에 따라 고객의 행동과 가치를 평가하는 기법으로서, 마케팅 또는 CRM(customer relationship management) 등에서 활용되는 개념이다. 그러나, 본 실시예에서는 이러한 통상적인 활용에서 벗어나, 고객의 행위를 보안과 관련된 이벤트로 치환하여 RFM 분석 기법을 적용하였다. 따라서, RFM 분석의 각각의 분류/측정 기준은 보안 속성의 최근 발생 시기, 보안 속성의 발생 빈도 및 보안 속성의 발생 총량이 된다.

[0028] 이렇게 분류된 보안 속성은 이후 120 단계에서 통계적 공정 관리도를 산출하는데 활용되는데, 이러한 보안 속성은 정규분포에 따라 발생한다. 즉, 이하에서 기술되는 본 발명의 실시예들에서 통계적 공정 관리도는 정규 분포를 따르는 데이터에 대해 적용된다고 전제한다. 다시 말해, 네트워크 트래픽이나 보안 이벤트는 정규 분포에 따른다(중심극한정리).

[0029] 120 단계에서, 탐지 장치는 110 단계를 통해 분류된 보안 속성에 따라 통계적 공정 관리도(statistical process control)를 산출하고, 산출된 통계적 공정 관리도에 대한 임계 범위를 설정한다. 앞서 설명한 바와 같이 통계적 공정 관리는, 품질 관리(quality control, QC) 분야에서 품질 변동에 대한 다양한 원인으로 인해 발생할 수 있는 공정의 패턴을 발견하고, 이를 관리하기 위한 수단으로서 제시된 통계 기법이다. 그러나, 본 실시예에서는 네트워크 트래픽을 하나의 관리 공정으로 치환시켜, 품질 변동에 대한 원인(다양한 보안 이벤트의 보안 속성을 의미한다.)에 따라 발생할 수 있는 이상증후를 탐지하는 관리도구로서 활용하고자 한다.

[0030] 이를 위해, 탐지 장치는 RFM 분석 기법에 따라 수집/분류된 보안 속성들로부터 통계적 공정 관리도를 산출하고, 산출된 공정 관리도 내에서 정상 행위라고 간주될 수 있는 수준의 범위를 임계 범위로서 설정한다. 이러한 임계 범위는 정상 상태에서 측정된 네트워크 트래픽의 다양한 보안 속성에 기초하여 실험적으로 설정될 수 있으며, 보안 유지의 수준 및 사용자의 요구에 따라 적절히 변형하여 적용될 수 있을 것이다.

[0031] 따라서, 일단 통계적 공정 관리도가 산출되고, 이에 대한 임계 범위가 설정되면, 새롭게 측정된 네트워크 또는 시스템의 보안 속성이 산출된 공정 관리도의 어느 위치에 해당하는지를 파악함으로써 그 적절한 관리 수준(이상증후가 존재하는지 여부를 의미한다.)을 파악할 수 있다. 상기된 110 단계 및 120 단계를 통해 본 발명의 실시예는 정상 상태에서 네트워크의 트래픽이나 보안 이벤트를 수집하여 분류한 후, 이로부터 통계적 공정 관리도와 임계 범위를 산출한다. 앞서 전제한 바와 같이 네트워크 트래픽이나 보안 이벤트는 정규 분포에 따르므로, 통계적 공정 관리도는 정규 분포를 따르는 데이터에 대해 적용된다. 따라서, 네트워크 내에 별다른 침입이 없을 경우, 여전히 새롭게 측정된 보안 속성 역시 정규 분포에 따른 범위 내에 존재할 것이다. 즉, 설정된 임계 범위를 넘지 않을 것이다.

[0032] 130 단계에서, 탐지 장치는 이상과 같은 원리에 기초하여 110 단계를 통해 분류된 보안 속성에 따라 네트워크 또는 시스템의 보안 속성으로부터 실시간으로 통계적 공정 관리도 내의 위치를 산출한다.

[0033] 140 단계에서, 탐지 장치는 130 단계를 통해 실시간으로 산출된 보안 속성의 위치가 120 단계를 통해 설정된 임계 범위 내에 존재하는지 여부를 검사하여 네트워크 또는 시스템에 대한 이상 여부를 판단한다.

[0034] 즉, 130 단계 및 140 단계에서는, 앞서 산출된 통계적 공정 관리도를 통해 과거 학습된 정상 행위와 현재의 측정된 보안 속성을 비교/분석함으로써 이상증후 판단의 근거로 삼는다. 이제, 이상증후가 판단되었다면, 이를 사용자에게 보고하게 된다.

[0035] 본 실시예에서 탐지 장치는 이상에서 기술된 일련의 연산을 수행하기 위한 적어도 하나의 프로세서를 구비한다. 이러한 프로세서는 네트워크 또는 시스템의 다양한 보안 속성을 특정 기준에 따라 분류하고, 통계적 공정 관리도와 임계 범위를 산출 및 설정한다. 또한, 프로세서는 실시간으로 네트워크의 보안 속성으로부터 통계적 공정 관리도 내의 위치를 산출하여 임계 범위와 비교/검사함으로써 네트워크의 이상 여부를 판단하는 역할을 수행한다. 나아가 탐지 장치는 상기된 프로세서가 일련의 연산을 수행하는 과정에서 연산 그 자체 또는 임시 저장을 위한 공간으로서 사용되는 메모리(memory)를 더 포함할 수 있다. 물론, 이들 연산들을 상기된 프로세서와 메모리를 통해 수행하기 위해 부가적인 소프트웨어 코드(software code)가 활용될 수 있음은 당연하다.

[0036] 도 2는 본 발명의 일 실시예에 따른 도 1의 이상증후 탐지 방법에서 네트워크의 상태를 나타내는 보안 속성을 분류하고 정렬하는 방법을 예시한 도면이다. 앞서 간단히 소개한 바와 같이 RFM 분석 기법이란 R(Recency), F(Frequency), M(Monetary)의 3가지 판단 요소(factor)를 이용하여 사용자의 패턴을 진단해 내는 방법을 말한다. 이러한 RFM 분석 기법을 본 발명에 적용하는 구현의 측면에서, R(Recency)은 보안과 관련된 이벤트가 가장 최근에 언제 발생했는지를 의미하고, F(Frequency)는 보안과 관련된 이벤트가 어떤 주기/빈도로 발생했는

지를 의미하며, M(Monetary)은 보안과 관련된 이벤트의 발생 총량으로서, 로그인(login)시 로그인 시간 간격(duration) 값 또는 로그인 시도에 따른 CPU 평균 사용량과 같은 정량적인 값을 의미한다.

[0037] 이러한 보안 속성은, 네트워크의 패킷(packet) 정보 내지 네트워크 내의 시스템의 보안 상태에 대한 속성 정보 중 적어도 하나로 선택될 수 있다. 따라서, 보안 속성은 네트워크 내의 패킷 또는 네트워크 내의 시스템의 이벤트 로그(event log) 중 적어도 하나로부터 획득될 수 있다. 이제 이렇게 수집된 보안 속성을 각각의 기준에 따라 분류하고, 필요에 따라서는 정렬한다. 이 때, R(Recency), F(Frequency), M(Monetary)의 3가지 기준은 각각 독립적인 변수로서 사용될 수도 있으며, 각각의 필요한 순서에 따라 종속적인 관계로 연결하여 사용할 수도 있다. 도 2에서는 이러한 3가지 기준을 일련의 순서에 따라 연결하여 활용하는 방법의 일례를 소개하고 있으나, 본 발명의 실시예들에서는 도 2와 같은 일례에 제한되지 않으며, 상황에 따라 각각의 분류 기준을 독립적으로 활용하는 방법과 이들 기준을 정렬하여 사용하는 방법 등을 유연하게 변형하여 채택하는 것이 가능하다.

[0038] 첫째, 보안 속성의 최근 발생 시기와 관련하여, 본 실시예에 따른 탐지 장치는 대상 데이터를 최근 날짜 및 시간을 기준으로 내림차순 정렬한 다음, 정렬된 데이터를 복수 개의 동일한 크기의 그룹으로 분류할 수 있다. 예를 들어, 5개의 그룹으로 분류할 경우, 개별 그룹당 20%의 데이터가 포함될 것이다. 이렇게 분류된 데이터에 대해 각 개별 그룹의 결과 값을 산출한다. 이 때의 결과 값이란 개별 데이터로부터 파악하고자 하는 보안에 관련된 의미있는 데이터를 의미한다. 예를 들어, 로그인 실패 횟수 또는 네트워크 접근에 따른 질의 부하(query load)의 정도 등이 이러한 결과 값이 될 수 있다. 물론 이러한 결과 값은 파악하고자 하는 보안 관련 이슈에 따라 적절히 선택되고 달라질 수 있다.

[0039] 둘째, 보안 속성의 발생 빈도와 관련하여, 본 실시예에 따른 탐지 장치는 대상 데이터를 기간별 평균 발생 빈도를 기준으로 내림차순 정렬할 수 있다. 이후의 절차는 상기된 보안 속성의 최근 발생 시기에 대한 일련의 처리 절차와 같다.

[0040] 셋째, 보안 속성의 발생 총량과 관련하여, 본 실시예에 따른 탐지 장치는 대상 데이터를 기간별 발생 총량의 평균을 기준으로 내림차순 정렬할 수 있다. 이후의 절차는 상기된 보안 속성의 최근 발생 시기에 대한 일련의 처리 절차와 같다.

[0041] 이상과 같은 과정을 통해 보안 속성에 대해 분류된 R, F, M 값을 도출할 수 있으며, 이들 분류된 그룹에 대해서는 필요에 따라 특정 그룹 코드를 부여하여 전자적 처리를 용이하게 할 수 있다. 도 2에는 이러한 기준에 따라 각각 5개의 그룹으로 분류한 결과를 예시하고 있으며, 분류된 R, F, M의 그룹들을 재차 순차적으로 연결하여 정렬한 결과를 도시하였다. 따라서, 총 분류된 그룹은 $125(=5*5*5)$ 개이며, 각각의 그룹들에는 편의상 그룹 코드를 부여하였다. 본 실시예에서는 이들 그룹들 각각을 분석의 대상으로 활용할 수도 있으나, 필요에 따라서는 이들 그룹을 연결하여 정렬함으로써 분석하고자 하는 목적에 따라 활용할 수 있다. 정렬의 기준이 되는 R, F, M의 순서 역시 필요에 따라 달라질 수 있을 뿐만 아니라, 각 그룹에 가중치를 달리하여 부여할 수 있음은 물론이다. 즉, 이러한 RFM 분석 기법을 통해 본 실시예에 따른 보안 속성을 각각의 세부 그룹 특성에 따라 분류하고, 개별 그룹의 결과 값을 통해 해당 그룹의 속성을 확인할 수 있다.

[0042] 도 3은 본 발명의 일 실시예에 따른 도 1의 이상증후 탐지 방법과 관련하여 통계적 공정 관리도를 설명하기 위한 도면이다. 도 3에 도시된 바와 같이 통계적 공정 관리도는 시간을 의미하는 가로축과 중심선(center line, CL)을 중심으로 상/하로 연장되는 세로축을 포함한다. 이 때, 중심으로부터 세로축의 양(+)의 방향으로 일정 위치만큼 떨어진 곳에 해당 공정에서 허용되는 관리 상한선(upper control limit, UCL)이 설정되고, 중심으로부터 세로축의 음(-)의 방향으로 역시 마찬가지로 관리 하한선(lower control limit, LCL)이 설정된다. 이러한 관리 상한선과 관리 하한선은 공정을 통해 생산되는 산출물 변인(속성)의 변동 하에서 품질의 변동이 허용되는 한도를 의미하는 것으로, 양자에 의해 임계 영역이 결정되게 된다. 즉, 도 3에 도시된 바와 같이 시간의 흐름에 따라 각각 발생하는 지점(point)들은 해당 공정이 허용 범위 내에 있으며, 현재의 공정의 품질이 적절하게 관리되고 있음을 나타낸다. 만약 이러한 통계적 공정 관리도에서 이상 요인이 발생할 경우, 품질을 나타내는 지점은 관리 범위(임계 범위) 외에서 나타나게 되며, 이러한 이상 요인들에 의해 산포가 악화되게 된다.

[0043] 이상과 같은 원리를 차용하여 본 발명의 실시예들에서 네트워크 트래픽을 하나의 관리 공정으로 간주하면, 네트워크 또는 시스템의 보안 속성을 관찰하고, 이로부터 통계적 공정 관리도를 산출하며, 적절한 임계 범위를 설정한다. 앞서 전제된 바와 같이 통계적 공정 관리도는 정규 분포를 따르는 데이터에 대해 적용되며, 본 발명의 실시예들이 대상으로 하는 네트워크 트래픽이나 보안 이벤트는 정규 분포를 따른다고 가정하였으므로, 만약 측정된 보안 속성이 설정된 관리 상한 또는 관리 하한의 영역을 벗어난 경우가 감지된다면 현재의 네트워크 또는 시스템 내에 이상원인(Anomaly)이 존재한다고 판단할 수 있다.

[0044] 통계적 공정 관리도의 활용 및 구현의 측면에서, 본 발명의 실시예들은 X-bar chart, R-chart 또는 P-chart 등의 다양한 유형의 관리도를 활용할 수 있다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 탐지 장치를 통해 분석하고자 하는 데이터의 특성에 따라 적절한 관리도를 선택하여 효과적으로 이상증후를 탐지할 수 있다. 이하에서는 각각의 유형을 간략하게 소개하도록 한다.

[0045] 첫째, $\overline{X}chart$ (X-bar chart)를 활용할 수 있다. $\overline{X}chart$ 는 길이, 무게, 시간, 강도 성분, 수확물 등과 같이 데이터가 연속적인 계량치로 나타나는 공정을 관리할 때 사용하며 정규분포를 따르는 자료의 평균치 관리도이다. 특히, $\overline{X}chart$ 를 이용하게 되면 품질 특성치 x 가 정규분포가 아닌 다른 분포를 하는 경우에도 중심극한정리에 따라 $\overline{x}$ 는 정규분포에 가까운 분포를 하게 되므로 정규분포의 성질을 그대로 이용할 수 있는 장점이 있다. 이러한, $\overline{X}chart$ 의 관리 상한선과 관리 하한선은 다음의 수학식 1과 같이 정의된다.

수학식 1

$$UCL_{\overline{X}} = \overline{X} + A \cdot \overline{R}$$

$$LCL_{\overline{X}} = \overline{X} - A \cdot \overline{R}$$

[0046]

[0047] 수학식 1에서 $\overline{X}$ 는 관리도의 중앙선, 즉 이전 표본의 평균 또는 목표치를 의미하고, A 는 표본의 3σ (σ : 표준편차)를 의미한다.

[0048] 요약하건대, 본 실시예에서 통계적 공정 관리도는, 보안 속성에 대한 표본의 평균을 중앙선으로 설정하고, 표본의 표준편차의 일정 배율(예를 들어, 중앙선의 상/하로 표본의 표준편차의 3배가 될 것이다.)만큼을 임계 범위로 설정함으로써 네트워크 또는 시스템의 보안 상태의 이상증후를 관리할 수 있다.

[0049] 둘째, R-chart를 활용할 수 있다. R-chart는 패턴 데이터 영역의 변동을 관리하기 위해 사용될 수 있다. 여기서, 특정 패턴의 관찰된 범위는 표본 내의 가장 큰 관찰값에서 가장 작은 관찰값 사이의 차이이다. 이러한 R-chart의 관리 상한선과 관리 하한선은 다음의 수학식 2와 같이 정의된다.

수학식 2

$$UCL_R = D \cdot \overline{R}$$

$$LCL_R = D' \cdot \overline{R}$$

[0050]

[0051] 수학식 2에서 $\overline{R}$ 은 이전에 관측된 R 값의 평균을 의미하고, D , D' 는 관리 한계를 결정하는 상수인 3σ (σ : 표준편차)를 의미한다.

[0052] 요약하건대, 본 실시예에서 통계적 공정 관리도는, 보안 속성의 변동 범위가 임계 범위 내에 있도록 보안 속성

에 대한 표본의 평균의 일정 배율(예를 들어, 표준편차의 3배가 될 것이다.)만큼을 임계 범위로 설정함으로써 네트워크 또는 시스템의 보안 상태의 이상증후를 관리할 수 있다.

[0053] 셋째, P-chart를 활용할 수 있다. P-chart는 공정을 부적합품률에 의해 관리할 경우 사용하며, 이항분포를 따르는 불량률 관리도이다. 계수치로 나타내는 품질특성, 즉 불량률을 이용하게 되면 관리항목이 되는 품질 특성이 여러 개가 있는 제품이라도 불량품/양호품으로 구분하여 P-chart만으로 이들을 동시에 관리할 수 있다. 즉, 본 실시예에서 P-chart는 보안과 관련된 사용자/네트워크의 패턴이 정상인지 아닌지를 측정한다. 이 경우, 보안과 관련된 행동(예를 들어, 로그인 시도가 될 수 있다.)의 실패 또는 성공의 수를 검사하여 표본의 실패 비율을 계산한 것을 확률 변수로 갖는다. 이러한 P-chart의 관리 상한선과 관리 하한선은 다음의 수학적 식 3과 같이 정의된다.

수학적 식 3

$$UCL_p = \bar{p} + z\sigma_p$$

$$LCL_p = \bar{p} - z\sigma_p$$

$$\sigma_p = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

[0054]

[0055] 수학적 식 3에서, σ_p 는 실패율의 분포의 표준편차를 나타내고, n은 표본의 크기를 나타내며, $\bar{p}$ 는 관리도의 중앙선으로서, 과거의 평균 실패율 또는 목표치를 의미한다.

[0056] 요약하건대, 본 실시예에서 통계적 공정 관리도는, 보안 속성에 대한 표본의 실패율의 평균을 중앙선으로 설정하고, 실패율의 분포의 표준편차만큼을 임계 범위로 설정함으로써 네트워크 또는 시스템의 보안 상태의 이상증후를 관리할 수 있다.

[0057] 도 4는 본 발명의 일 실시예에 따른 도 1의 이상증후 탐지 방법에서 이상증후가 감지된 네트워크의 보안 속성을 임계 범위와 비교하여 시각화된 통계적 공정 관리도 상에 표시하는 방법을 예시한 도면이다.

[0058] 본 실시예에서 달성하고자 하는 보안 시각화 기술은 네트워크나 시스템 상에서 발생하는 방대한 양의 이벤트를 실시간으로 시각화함으로써 공격의 탐지, 알려지지 않은 공격 유형 분류, 이상 상태의 발견 등 보안 상황을 관리자가 직관적으로 인지할 수 있도록 하는 것이다. 이를 위해 본 탐지 방법에 채택하고 있는 시각화 기술은 데이터 선정 및 수집, 특성 인자 추출, 연관성 분석, 데이터마이닝, 패턴분석, 이벤트 시각화 등과 같이 현재의 네트워크에서 발생하고 있는 상황(침입 또는 공격이 될 수 있다.)을 알려주는 것뿐만 아니라, 보안 이벤트를 시각화하여 현재의 감지 정보로부터 보안 상황을 알려줄 수 있는 기술적 수단을 제공한다. 따라서, 본 실시예에서 탐지 장치는 앞서 산출된 통계적 공정 관리도를 시각화하고, 이상증후가 감지된 네트워크의 보안 속성을 미리 설정된 임계 범위와 비교하여 시각화된 통계적 공정 관리도 상에 표시한다.

[0059] 도 4를 참조하면, 중앙선(CL)을 기준으로 관리 상한선(UCL)과 관리 하한선(LCL)이 설정되어 있으며, 3개의 기준(1, 2, 3)에 따라 산출된 통계적 공정 관리도가 도시되어 있음을 확인할 수 있다. 이들 통계적 공정 관리도는 각각 정규분포에 따르며, 현재 감지된 보안 속성의 결과 값이 통계적 공정 관리도 상에 표시되었다. 이 때, 도 4에는 임계 범위(도 4에서 보다 정확하게는 관리 상한선을 의미한다.)를 벗어나 이상증후(anomaly)가 감지되었음이 표시되었다. 즉, 본 실시예에서, 탐지 장치는 현재 감지된 네트워크의 보안 속성을 미리 설정된 임계 범위와 비교하여 이상 여부를 판단하고, 그 결과를 시각화된 통계적 공정 관리도 상에 표시한다.

[0060] 상기된 본 발명의 실시예들에 따르면, 정상 상태에서 상기 네트워크의 상태를 나타내는 보안 속성을 입력받고, 특정 기준에 따라 분류된 보안 속성에 따라 통계적 공정 관리도를 산출하여 실시간 통계적 공정 관리도와 비교

함으로써, 이상증후 탐지법의 정확도를 향상시키고, 통계적 공정 관리도를 이상증후 탐지의 관리 수단으로써 활용하여 이를 시각화하여 관리자에게 제공함으로써, 사용자 하여금 현재의 네트워크 또는 시스템의 보안 상황에 대한 정보를 직관적으로 제공할 수 있다. 나아가, 본 발명의 실시예들은 시각적인 데이터 분석을 통해 새로운 관점에서 침입을 탐지하고 연구할 수 있도록 아이디어를 제공하는 역할도 한다.

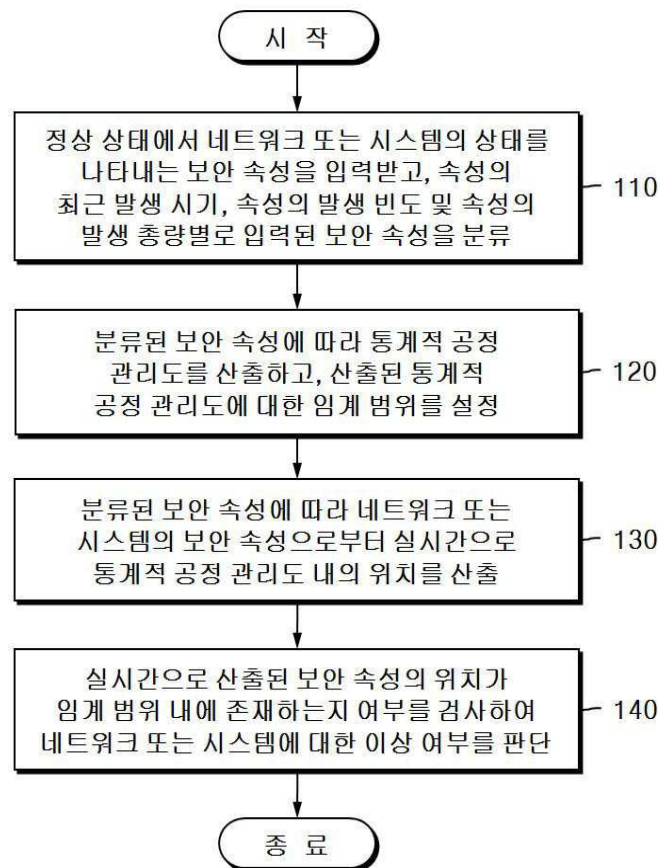
[0061] 한편, 본 발명의 실시예들은 컴퓨터로 읽을 수 있는 기록 매체에 컴퓨터가 읽을 수 있는 코드로 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다.

[0062] 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현하는 것을 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술 분야의 프로그래머들에 의하여 용이하게 추론될 수 있다.

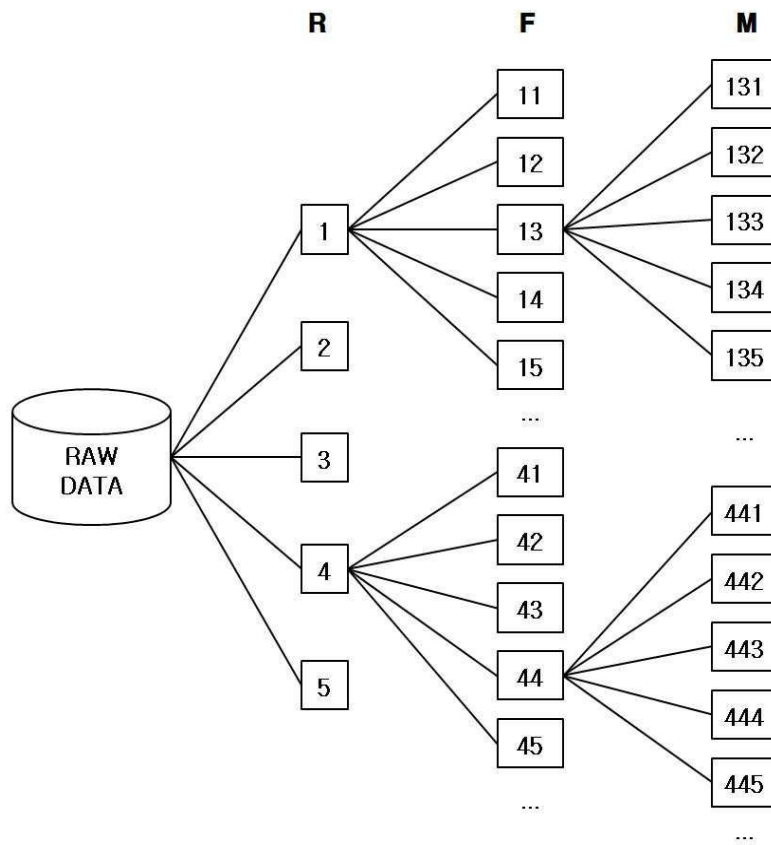
[0063] 이상에서 본 발명에 대하여 그 다양한 실시예들을 중심으로 살펴보았다. 본 발명에 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

도면

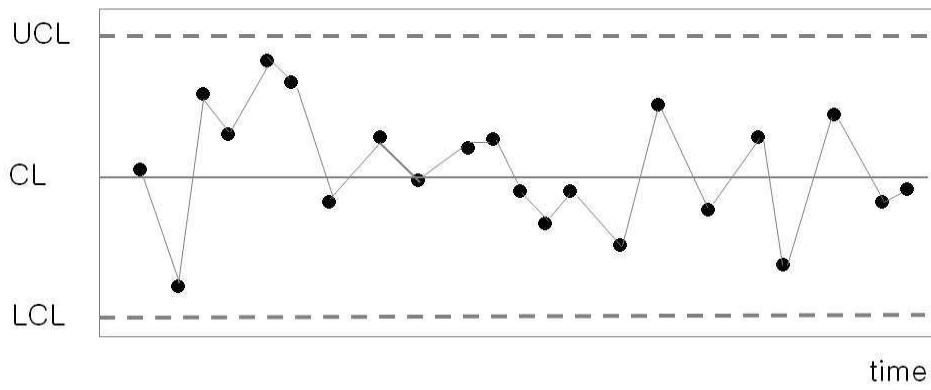
도면1



도면2



도면3



도면4

