(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년11월18일
(11) 등록번호 10-2179290
(24) 등록일자 2020년11월10일

(51) 국제특허분류(Int. Cl.)
G06F 11/30 (2006.01) *G06F 17/18* (2006.01)
G06N 20/00 (2019.01)
(52) CPC특허분류
G06F 11/3034 (2013.01)
G06F 11/3055 (2013.01)
(21) 출원번호 10-2019-0141937
(22) 출원일자 2019년11월07일
심사청구일자 2019년11월07일
(56) 선행기술조사문헌
KR1020170084445 A*
KR1020160042616 A*
KR101965598 B1*
KR100840129 B1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
윤호영
서울특별시 서대문구 연세로 50 4공학관 909호 (신촌동,연세대학교)
김경섭
서울특별시 서대문구 연세로 50 4공학관 919호 (신촌동,연세대학교)
(74) 대리인
특허법인우인

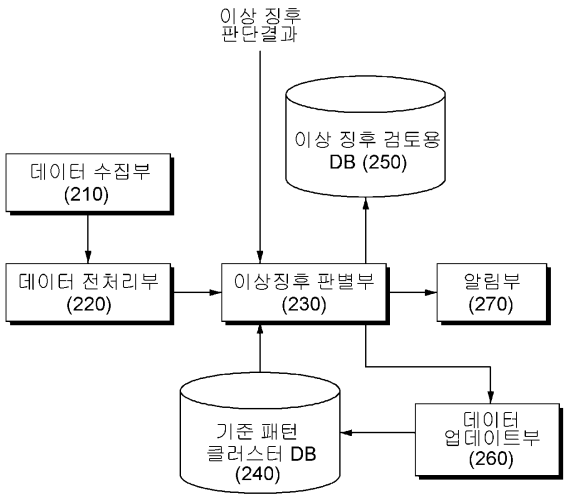
전체 청구항 수 : 총 3 항 심사관 : 김계준

(54) 발명의 명칭 **워크로드 데이터에 대한 이상징후 판별 방법**

(57) 요 약

프로세서 및 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 이상징후 판별 방법에 있어서, 상기 프로세서가 워크로드 데이터를 수집하는 단계, 상기 수집된 워크로드 데이터의 패턴이 미리 정해진 적어도 하나의 기준 패턴 클러스터에 속하는지 여부를 판단하는 단계, 상기 수집된 워크로드 데이터가 상기 기준 패턴 클러스터에 속하지 않는 것으로 판단되면, 기 설정된 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터에 대한 이상징후를 판별하는 단계 및 상기 이상징후가 판별된 워크로드 데이터의 패턴을 기초로, 상기 기준 패턴 클러스터에 대한 정보를 업데이트 하는 단계를 포함할 수 있다.

대 표 도 - 도2



(52) CPC특허분류
G06F 17/18 (2013.01)
G06N 20/00 (2019.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	2019R1F1A1059276
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	일반연구자지원사업
연구과제명	클라우드 데이터센터의 최적 운영을 위한 의사결정지원 모형 연구(1/3)
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2019.06.01 ~ 2020.02.29

명 세 서

청구범위

청구항 1

프로세서 및 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 이상징후 판별 방법으로서,

상기 프로세서가 워크로드 데이터를 수집하는 단계;

상기 수집된 워크로드 데이터의 패턴이 미리 정해진 적어도 하나의 기준 패턴 클러스터에 속하는지 여부를 판단하는 단계;

상기 수집된 워크로드 데이터가 상기 기준 패턴 클러스터에 속하지 않는 것으로 판단되면, 기 설정된 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터에 대한 이상징후를 판별하는 단계; 및

상기 이상징후가 판별된 워크로드 데이터의 패턴을 기초로, 상기 기준 패턴 클러스터에 대한 정보를 업데이트 하는 단계;를 포함하되,

상기 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터에 대한 이상징후를 판별한 결과, 상기 워크로드 데이터에 이상징후가 있는 것으로 판별하면, 상기 워크로드 데이터를 검토용 데이터베이스에 저장하는 단계; 상기 검토용 데이터베이스에 저장된 워크로드 데이터를 화면에 표시하는 단계; 및 사용자로부터 상기 화면에 표시된 워크로드 데이터에 대한 이상징후 판단 결과를 입력 받는 단계;를 더 포함하며,

상기 기준 패턴 클러스터에 대한 정보를 업데이트 하는 단계는, 상기 사용자로부터 입력 받은 이상징후 판단 결과에 따라 워크로드 패턴과 관련된 이상징후 패턴을 학습하여 상기 기준 패턴 클러스터에 대한 정보를 업데이트 하는 것을 특징으로 하는 이상징후 판별 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 사용자로부터 상기 이상징후 판단 결과를 입력 받는 단계는,

상기 워크로드 데이터에 대한 복수의 이상징후 판단 결과값들을 입력 받는 단계; 및

수집된 복수의 이상징후 판단 결과값들을 고려하여 상기 워크로드 데이터에 대한 이상징후 판단 결과를 도출하되, 상기 복수의 이상징후 판단 결과값들별 로 적용되는 미리 정해진 적어도 일부 서로 다른 가중치를 고려하여 상기 이상징후 판단 결과를 도출하는 단계;를 더 포함하는 것을 특징으로 하는 이상징후 판별 방법.

청구항 4

삭제

청구항 5

제1항에 있어서,

상기 프로세서는, 상기 워크로드 데이터의 패턴과 상기 기준 패턴 클러스터를 비교함에 따라 상기 워크로드 데이터에 대한 이상징후 여부를 1차적으로 판단하고,

상기 기 설정된 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터의 패턴에 대한 통계적 분석을 수행함에 따라 상기 워크로드 데이터에 대한 이상징후 여부를 2차적으로 판별하는 것을 특징으로 하는 이상징후 판별 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 워크로드 데이터에 대한 이상징후 판별 방법에 관한 것이다. 보다 상세하게는, 본 발명은 반지도 학습을 기반으로 하는 이상징후 판별 방법에 관한 것이다.

배 경 기 술

[0002] 이상징후 탐지(Anomaly Detection)는 자료에서 예상과는 다른 패턴을 보이는 개체 또는 자료를 찾는 것을 일컫는다. 비슷한 개념으로 아웃라이어 감지(outlier detection)가 있는데, 아웃라이어는 시간과 관련이 없이 대상을 표현하는 숫자들의 위치를 보고 보편적인 대상과 벗어나는 것을 찾아내는 것이고, 이상감지는 시간 또는 순서가 있는 흐름에 따른 패턴이 보편적인 상황 또는 보편적인 패턴들과 다른 것들을 찾아내는 것이다. 즉, 이상징후 탐지는 시계열 데이터에서 아웃라이어를 찾는 것이라고 할 수 있다.

[0003] 이상징후를 탐지하기 위해서는 "이상하지 않은 것" 즉 정상적인 데이터가 정의되어야 한다. 정상적인 데이터가 정의되면, 정상 데이터에서 일정 범위를 벗어나는 데이터를 이상징후라고 정의하는 것이 가장 기본적인 방법(예를 들어, 3 sigma rule, Z-score)이다. 하지만, 정상 데이터를 정의하는 것조차 현실적으로 쉽지 않다. CPU의 시계열 패턴에 대하여 이상징후로 검출된 패턴들은 지속적으로 반복되기 때문에 상황에 따라 일반적인 패턴으로도 볼 수도 있다. 이는 상황에 따라 환경에 따라 다르게 적용된다.

[0004] 과거에 이상징후라고 판별한 것이 최근에는 이상징후가 아닌 것으로 판별될 수도 있으며, 그 반대도 성립될 수 있다.

[0005] 따라서, 이상징후 탐지 시스템은 환경에 따라 다른 방식으로 적용되는 것이 바람직하다. 개발사의 관점에서는 다양한 고객사에 납품을 하게 되는데, 다양한 고객사에 맞게 커스터마이징(customizing)하는 것은 현실적으로 어려운 일이다. 또한, 자체망으로 운영되는 환경의 고객사 같은 경우는 직접 현장으로 가서 패치를 수행해야 하는데, 이는 비용과 시간이 많이 소모되는 단점이 존재한다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국 등록 특허 제10-1965598호 (등록)

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기 전술한 종래의 문제점을 해결하기 위해 적용되는 현장의 환경에 맞는 이상징후를 보다 효과적이고 효율적으로 판별할 수 있도록, 관리자 기반의 반지도 학습을 통해 워크로드 데이터의 패턴에 대한 이상징후를 판별하는 이상징후 탐지 방법을 제공하는 것을 목적으로 한다.

[0008] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다

과제의 해결 수단

[0009] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 프로세서 및 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 이상징후 판별 방법은, 상기 프로세서가 워크로드 데이터를 수집하는 단계, 상기 수집된 워크로드 데이터의 패턴이 미리 정해진 적어도 하나의 기준 패턴 클러스터에 속하는지 여부를 판단하는 단계, 상기 수집된 워크로드 데이터가 상기 기준 패턴 클러스터에 속하지 않는 것으로 판단되면, 기 설정된 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터에 대한 이상징후를 판별하는 단계 및 상기 이상징후가 판별된 워크로드 데이터의 패턴을 기초로, 상기 기준 패턴 클러스터에 대한 정보를 업데이트 하는 단계를 포함할 수 있다.

[0010] 또한, 상기 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터에 대한 이상징후를 판별한 결과, 상기 워크로드 데이터에 이상징후가 있는 것으로 판별하면, 상기 워크로드 데이터를 검토용 데이터베이스에 저장하는 단계, 상기 검토용 데이터베이스에 저장된 워크로드 데이터를 화면에 표시하는 단계 및 사용자로부터 상기 화면에 표

시된 워크로드 데이터에 대한 이상징후 판단 결과를 입력 받는 단계를 더 포함할 수 있다.

[0011] 또한, 상기 사용자로부터 상기 이상징후 판단 결과를 입력 받는 단계는, 상기 워크로드 데이터에 대한 복수의 이상징후 판단 결과값들을 입력 받는 단계 및 수집된 복수의 이상징후 판단 결과값들을 고려하여 상기 워크로드 데이터에 대한 이상징후 판단 결과를 도출하되, 상기 복수의 이상징후 판단 결과값들별 로 적용되는 미리 정해진 적어도 일부 서로 다른 가중치를 고려하여 상기 이상징후 판단 결과를 도출하는 단계를 더 포함할 수 있다.

[0012] 또한, 상기 기준 패턴 클러스터에 대한 정보를 업데이트 하는 단계는, 상기 사용자로부터 입력 받은 이상징후 판단 결과에 따라 워크로드 패턴과 관련된 이상징후 패턴을 학습하여 상기 기준 패턴 클러스터에 대한 정보를 업데이트할 수 있다.

[0013] 또한, 상기 프로세서는, 상기 워크로드 데이터의 패턴과 상기 기준 패턴 클러스터를 비교함에 따라 상기 워크로드 데이터에 대한 이상징후 여부를 1차적으로 판단하고, 상기 기 설정된 이상탐지 알고리즘을 이용하여 상기 워크로드 데이터의 패턴에 대한 통계적 분석을 수행함에 따라 상기 워크로드 데이터에 대한 이상징후 여부를 2차적으로 판별할 수 있다.

발명의 효과

[0014] 본 발명의 실시예에 따른 이상징후 판별 방법은 관리자 기반의 반지도 학습을 통해 워크로드 데이터의 패턴에 대한 이상징후를 판별함에 따라, 적용되는 현장의 환경에 맞는 이상징후를 보다 효과적이고 효율적으로 판별할 수 있다.

도면의 간단한 설명

[0015] 도1은 본 발명의 일 실시예에 따른 이상징후 판별 장치의 구성을 개략적으로 도시한 블록도이다.

도2는 본 발명의 일 실시예에 따른 프로세서가 이상징후 판별을 위해 수행하는 동작 개념을 설명하기 위해 도시한 블록도이다.

도3은 본 발명의 일 실시예에 따른 입력 모듈에 표시되는 화면을 예시한 예시도이다.

도4는 본 발명의 일 실시예에 따라 사용자로부터 입력 받은 이상징후 판단 결과를 고려하여 프로세서가 워크로드 데이터에 대한 최종적인 이상징후를 판별하는 과정을 설명하기 위해 도시한 참고도이다.

도5는 본 발명의 일 실시예에 따라 기준 패턴 클러스터를 업데이트하는 과정에 대하여 설명하기 위해 도시한 참고도이다.

도6은 본 발명의 일 실시예에 따른 이상징후 판별 방법을 시간의 흐름에 따라 도시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

[0017] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계 없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0018] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함" 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록"등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로구현될 수 있다.

[0019] 이하, 본 발명의 일 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 또한, 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략할 수 있다.

[0020] 이하에서는 본 발명의 실시예에 따른 이상징후 판별 장치 및 방법 구성을 관련된 도면을 참조하여 상세히 설명한다.

[0021] 도1은 본 발명의 일 실시예에 따른 이상징후 판별 장치의 구성을 개략적으로 도시한 블록도이다. 본 발명의 이상징후 판별 장치(10)는 도1에 도시된 바와 같이 입력 모듈(100), 프로세서(200), 및 메모리(300)를 포함하는 컴퓨팅 디바이스로 구현될 수 있다.

[0022] 본 발명의 이상징후 판별 장치는 컴퓨터, 휴대폰, 단말기 등과 같은 컴퓨팅 장치에 탑재되어 상기 컴퓨팅 장치에서 수행하는 데이터 처리 동작에 따른 작업 량, 및 작업의 성격 등과 같은 데이터 처리 작업과 관련된 워크로드 데이터를 상기 컴퓨팅 장치로부터 전달 받을 수 있다.

[0023] 입력 모듈(100)은 위와 같은 컴퓨팅 장치의 동작에 따라 생성된 워크로드 데이터를 입력 받을 수 있다. 그리고, 입력 모듈(100)은 프로세서(200)에서 워크로드 데이터의 이상징후를 판별하기 위해 사용자로부터의 명령 메시지를 입력 받을 수 있다. 본 발명의 입력 모듈(100)은 본 발명의 이상징후 판별 장치가 탑재되는 컴퓨팅 장치와의 통신을 통해 워크로드 데이터를 수신하는 통신 모듈을 포함하는 개념일 수 있다. 또한, 본 발명의 입력 모듈(100)은 디스플레이 장치로 이루어진 GUI(Graphical user interface)로 구현될 수 있다.

[0024] 프로세서(200)는 예를 들면, 소프트웨어(예: 프로그램)를 실행하여 프로세서(200)에 연결된 전자 장치의 적어도 하나의 다른 구성요소(예: 하드웨어 또는 소프트웨어 구성요소)을 제어할 수 있고, 다양한 데이터 처리 또는 연산을 수행할 수 있다. 일 실시 예에 따르면, 데이터 처리 또는 연산의 적어도 일부로서, 프로세서(200)는 다른 구성요소(예: 센서 모듈 또는 통신 모듈)로부터 수신된 명령 또는 데이터를 휘발성 메모리에 로드하고, 휘발성 메모리에 저장된 명령 또는 데이터를 처리하고, 결과 데이터를 비휘발성 메모리에 저장할 수 있다. 일 실시 예에 따르면, 프로세서(200)는 메인 프로세서 (예: 중앙 처리 장치 또는 어플리케이션 프로세서), 및 이와는 독립적으로 또는 함께 운영 가능한 보조 프로세서(123)(예: 그래픽 처리 장치, 이미지 시그널 프로세서, 센서 허브 프로세서, 또는 커뮤니케이션 프로세서)를 포함할 수 있다. 추가적으로 또는 대체적으로, 보조 프로세서는 메인 프로세서보다 저전력을 사용하거나, 또는 지정된 기능에 특화되도록 설정될 수 있다.

[0025] 메모리(300)는, 컴퓨팅 장치(10)인 이상징후 판별 장치의 적어도 하나의 구성요소(예: 프로세서(200) 또는 센서 모듈)에 의해 사용되는 다양한 데이터를 저장할 수 있다. 데이터는, 예를 들어, 소프트웨어(예: 프로그램) 및, 이와 관련된 명령에 대한 입력 데이터 또는 출력 데이터를 포함할 수 있다. 메모리(300)는, 휘발성 메모리 또는 비휘발성 메모리를 포함할 수 있다

[0026] 도2는 본 발명의 일 실시예에 따른 프로세서가 이상징후 판별을 위해 수행하는 동작 개념을 설명하기 위해 도시한 블록도이다. 도2의 블록도에는 데이터 수집부(210), 데이터 전처리부(220), 이상징후 판별부(230), 기준 패턴 클러스터 DB(240), 이상 징후 검토용 DB(250), 업데이트부(260), 및 알림부(270)가 도시되어 있다. 본 발명의 일 실시예에 따르면, 도2에 개시된 구성요소들은 소프트웨어적으로 마련되어 메모리에 저장되고, 프로세서에 의하여 일련의 절차가 실행되도록 구현될 수 있다. 또한, 위 구성요소들 중 적어도 일부는 하드웨어적으로도 구현될 수 있음은 자명하다.

[0027] 데이터 수집부(210)는 입력 모듈(100)로부터 기 설정된 주기로 입력되거나 비주기적으로 입력되는 워크로드 데이터를 수집할 수 있다.

[0028] 데이터 전처리부(220)는 수집된 워크로드 데이터의 차원을 축소하고, 차원이 축소된 워크로드 데이터의 패턴을 좌표 데이터로 변환함으로써, 좌표 상에 상기 워크로드 데이터의 패턴을 나타낼 수 있도록 한다. 보다 구체적으로는, 데이터 전처리부(220)는 3차원으로 수집되는 워크로드 데이터를 2차원으로 차원 축소하고, 2차원으로 차원 축소된 워크로드 데이터를 2차원 좌표로 표시 가능한 좌표 데이터로 변환할 수 있다.

[0029] 이상징후 판별부(230)는 좌표 데이터로 변환된 워크로드 데이터의 패턴이 미리 정해진 적어도 하나의 기준 패턴 클러스터에 속하는지 여부를 판단할 수 있다.

[0030] 여기서, 상기 기준 패턴 클러스터는 워크로드 데이터에 있어서 이상징후가 있는 패턴들이 이상징후의 유형별로 적어도 하나의 패턴들이 클러스터링되어 있는 클러스터들을 의미한다. 이와 같은 기준 패턴 클러스터들은 도2에 도시된 바와 같은 기준 패턴 클러스터 DB(240)에 저장되어 있을 수 있다.

[0031] 이상징후 판별부(230)는 수집된 워크로드 데이터와 기준 패턴 클러스터 DB(240)에 저장되어 있는 적어도 하나의 기준 패턴 클러스터와 비교하여 상기 수집된 워크로드 데이터에 대한 이상징후를 1차적으로 판별할 수 있다.

[0032] 이때, 이상징후 판별부(230)는 판별 대상인 상기 수집된 워크로드 데이터가 상기 기준 패턴 클러스터들 중 적어도 하나에 속한다면, 알림부(270)에 상기 수집된 워크로드 데이터에 이상징후가 있음을 전달할 수 있다.

[0033] 반면, 이상징후 판별부(230)가 상기 수집된 워크로드 데이터와 기준 패턴 클러스터들을 비교한 결과, 상기 수집

된 워크로드 데이터가 기준 패턴 클러스터들에 속하지 않는 것으로 판단되면, 이상징후 통계적 분석을 통해 워크로드 데이터에 대한 이상징후를 2차적으로 판별할 수 있다.

[0034] 2차 판별 방법으로, 이상징후 판별부(230)는 먼저 기 설정된 이상탐지 알고리즘을 이용하여 상기 수집된 워크로드 데이터의 이상징후 여부를 판별한다. 여기서, 이상탐지 알고리즘이란 통계적인 분석을 이용하는 S-H-ESD 알고리즘일 수 있다. 이상징후 판별부(230)가 이용하는 이상탐지 알고리즘은 상기 S-H-ESD 알고리즘에 한정되지 않고, 다양한 이상탐지 알고리즘을 이용할 수도 있다.

[0035] 그리고, 이상징후 판별부(230)가 상기 워크로드 데이터와 이상탐지 알고리즘을 이용하여 판별한 결과를 이상 징후 검토용 DB(250)로 전달한다.

[0036] 이상 징후 검토용 DB(250)는 상기 워크로드 데이터와 이상탐지 알고리즘을 기반으로 상기 이상징후 판별부로부터 판별된 결과를 저장한다.

[0037] 본 실시예에 따른 이상 징후 검토용 DB(250)에 저장된 워크로드 데이터의 양이 임계치 이상으로 저장되어 있거나, 저장된 워크로드 데이터가 임계시간 이상 저장되어 있는 경우, 이상징후 판별부(230)는 상기 입력 모듈(100)로 이상 징후 검토용 DB(250)에 저장되어 있는 워크로드 데이터를 전달함으로써, 입력 모듈(100)의 GUI(Graphical user interface)는 전달 받은 워크로드 데이터를 화면(디스플레이) 상에 표시할 수 있다.

[0038] 입력 모듈(100)은 이상징후 판별부(230)로부터 전달 받은 워크로드 데이터를 화면 상에 표시함으로써 사용자로부터 이상 징후가 있는 워크로드 패턴에 대하여 확인받을 수 있도록 한다. 이상징후 패턴은 주어진 환경마다 다르게 정의되어야 하기 때문에, 적용되는 환경별로 이상징후의 패턴이 상이하여 현장 관리자가 해당 환경에 따른 이상징후 패턴을 정의하는 것이 이상징후 판별에 있어서 높은 정확도를 도출할 수 있다.

[0039] 도3은 본 발명의 일 실시예에 따른 입력 모듈에 표시되는 화면을 예시한 예시도이다. 도3을 참조하면, 입력 모듈의 화면에는 시간의 흐름에 따라 데이터 처리 동작과 관련된 작업 량, 및 작업의 성격을 나타내는 패턴을 나타내는 화면(30)과, 이상징후 목록들(31) 및 각 이상징후 목록별 사용자의 의견을 기록할 수 있는 화면(32)이 표시될 수 있다.

[0040] 입력 모듈의 화면 상에 표시되는 이상징후 목록은 워크로드 패턴에서 시간구간별 구분된 패턴별로 나열될 수 있다. 사용자가 이상징후 목록에서 일 이상징후 항목([1])을 선택하면, 해당 항목에 대응하는 패턴([1])이 화면 상에 표시되고, 사용자는 이에 대한 사용자 의견을 GUI를 통해 입력하여 기록할 수 있다.

[0041] 일 실시예에 따른 도3에서 복수의 이상징후 항목들([1]~[8]) 중 이상징후 패턴이라고 판단한 항목(예를 들어, 항목 [1])에는 이상징후 패턴임을 나타내는 표시(33)가 화면 상에 표시될 수 있는데, 이는 이상징후 판별부(230)가 기 설정된 이상탐지 알고리즘을 통해 2차적으로 이상징후를 판별한 결과를 나타낸 것이며, 일 예인 도3의 항목 [1]의 경우, 이상징후 판별부(230)가 해당 항목([1])은 이상징후가 있다고 판별한 것이다.

[0042] 사용자는 이상징후 판별부(230)로부터 판별된 결과를 참고하여, 해당 워크로드 데이터에 대한 패턴을 심층적으로 분석하여 이상징후 판단 결과값을 입력할 수 있고, 상기 해당 워크로드 데이터에 대한 패턴에 대한 판별 결과가 잘못 판별된 것으로 보고, 이에 대한 판별 결과를 정정하기 위한 이상징후 판단 결과값을 입력할 수도 있다.

[0043] 도4는 본 발명의 일 실시예에 따라 사용자로부터 입력 받은 이상징후 판단 결과를 고려하여 프로세서가 워크로드 데이터에 대한 최종적인 이상징후를 판별하는 과정을 설명하기 위해 도시한 참고도이다.

[0044] 사용자는 도3에 도시된 바와 같은 화면 상의 워크로드 데이터에 대한 이상징후 패턴을 확인하여 각 이상징후 패턴에 대한 이상징후 판단 결과값을 입력할 수 있다. 예컨대, 사용자는 검토한 워크로드 데이터에 대한 패턴에 이상징후가 있는 것으로 판단하면 '1'로 입력하고, 패턴에 이상징후가 없는 것으로 판단하면 '0'으로 입력할 수 있다.

[0045] 일 실시예에 따르면, 이상징후 판별부(230)는 한 명의 사용자로부터 상술한 바와 같이 입력된 이상징후 판단 결과값에 따라서 상기 워크로드 데이터에 대한 패턴의 이상징후 여부를 결정할 수도 있지만, 다른 실시예로 이상징후 판별부(230)는 다수의 사용자들로부터 입력되는 이상징후 판단 결과값들을 고려하여 워크로드 데이터에 대한 패턴의 이상징후 여부를 결정할 수도 있다.

[0046] 일 실시예인 도4를 참고하면, 이상징후 판별부(230)는 본 발명의 이상징후 판별 장치에 미리 등록되어 있는 A~F까지의 사용자들 각각으로부터 평가되어 입력되는 이상징후 판단 결과값을 모두 고려하여 워크로드 데이터에 대

한 패턴(A)의 이상징후 판별을 수행할 수 있는데, 이때 이상징후 판별부(230)는 등록된 사용자들 각각에 대하여 도4의 (a)에 도시된 바와 같이 직급 및 근속연수를 고려하여 미리 정해진 가중치와 상기 각 이상징후 판단 결과값을 고려하여 워크로드 데이터에 대한 패턴(A)의 이상징후 판단 결과를 도출할 수 있다. 이렇게, A~F까지의 사용자들이 각각 입력한 이상징후 판단 결과값과 각 사용자에 상응하는 가중치를 곱하여 산출된 값들을 합산한 값이 중간치(예, 0.5)를 초과할 경우 해당 패턴(A)에 이상징후가 있는 것으로 판별하고, 상기 합산한 값이 중간치 이하인 경우 해당 패턴(A)에 이상징후가 없는 것으로 판별할 수 있다.

[0047] 이에 따라, 데이터 업데이트부(260)는 상기와 같이 판별한 결과를 기반으로 이상징후 패턴을 학습하여 기준 패턴 클러스터 DB(240)에 저장되어 있는 기준 패턴 클러스터에 대한 정보를 업데이트 할 수 있다.

[0048] 도5는 본 발명의 일 실시예에 따라 기준 패턴 클러스터를 업데이트하는 과정에 대하여 설명하기 위해 도시한 참고도이다. 도5의 (a)는 각 워크로드 데이터에 따른 패턴별로 해당 좌표값, 이상징후 판별부로부터 계산된 합산값, 및 사용자로부터 입력된 사용자 의견을 표로 나타낸 것이다.

[0049] 보다 구체적으로는, 데이터 업데이트부(260)는 각 패턴별로 입력된 사용자 의견을 기초로 자연어 처리를 하여 단어 분석을 통해 이상징후 원인 데이터로 활용할 수 있다.

[0050] 일 실시예에 따른 도5의 (b)는 기준 패턴 클러스터 DB(240)에 저장되어 있는 기준 패턴 클러스터들을 좌표 상에 표시한 것을 나타낸 것이다. 본 실시예에 따른 기준 패턴 클러스터들은 키워드(자연어), 및 좌표값을 기준으로 클러스터링 된 그룹들일 수 있다.

[0051] 도5를 참고하면, 예컨대, 데이터 업데이트부(260)는 패턴 A를 참고로 기준 패턴 클러스터를 업데이트 하기 위해, 패턴 A에 대하여 사용자가 입력한 사용자 의견("2013년 12월 15일경 스토리지 교체 작업으로 이상패턴이 발생하였으며, 교체 작업 후에 DISK I/O의 워크로드는 정상 처리 되었음")을 기초로 자연어 처리하여, 상기 사용자 의견의 자연어들 중 기준 패턴 클러스터 DB(240)에 이미 저장되어 있는 기준 패턴 클러스터들에 따른 키워드들 중 매칭되는 것이 있다면, 패턴 A를 매칭되는 키워드에 대응되는 클러스터로 클러스터링 하고, 매칭되는 것이 없다면, 패턴 A를 포함하는 새로운 클러스터를 생성할 수 있다.

[0052] 상술한 바와 같은 동작이 모두 완료되면, 이상징후 판별부(230)는 수집된 워크로드 데이터에 따른 패턴에 대하여 최종적으로 판별된 이상징후 여부와, 이상징후 원인 데이터로 정의될 수 있는 클러스터링된 자연어에 대한 정보를 알림부(270)로 전달할 수 있다.

[0053] 그리고, 알림부(270)는 전달받은 이상징후 여부에 대한 정보와, 클러스터링된 자연어에 대한 정보를 입력 모듈(100)로 전달함으로써, 입력 모듈(100)은 화면 상에 상기 전달 받은 정보들을 표시하여 사용자에게 정보를 제공할 수 있다.

[0054] 이때, 일 실시예에 따른 데이터 업데이트부(260)는 랜덤 또는 사전지식을 기반으로 하여 도5의 (b)와 같은 k개의 기준 패턴 클러스터들에 대하여, 거리 기반의 함수를 통해 각 워크로드 데이터의 패턴을 각 기준 패턴 클러스터와의 거리 차이의 분산을 최소화하도록 클러스터에 할당하여, 이후의 생성된 클러스터들 각각의 내부에서의 유사도는 최대로 하고, 서로 다른 클러스터들 간의 유사도는 최소화할 수 있도록 할 수 있다. 보다 구체적으로, 데이터 업데이트부(260)는 기 설정된 k개의 각 기준 패턴 클러스터의 중심점들에 따라, 상기 각 워크로드 데이터의 패턴을 상기 k개의 중심점들 중 가장 가까운 중심점에 할당한다. 그리고, 데이터 업데이트부(260)는 각 중심점들을 각 기준 패턴 클러스터들의 무게중심으로 이동시키고, 무게중심이 이동된 중심점들과 상기 중심점들 각각과 인접한 위치에 있는 워크로드 데이터에 따른 패턴의 좌표값과의 거리를 다시 계산하여, 상기 패턴의 좌표값과의 거리가 가장 가까운 기준 패턴 클러스터에 할당하도록 한다. 이때, 데이터 업데이트부는 이와 같은 동작을 한번만 수행하는 것이 아니고, 클러스터의 형태가 변하지 않을 때까지 반복하여 수행될 수 있다. 일 예로, 데이터 업데이트부는 K-Means, Agglomerative Clustering 등의 알고리즘을 이용하여 워크로드 데이터에 따른 패턴에 대한 클러스터링을 수행할 수 있다.

[0055] 도6은 본 발명의 일 실시예에 따른 이상징후 판별 방법을 시간의 흐름에 따라 도시한 흐름도이다.

[0056] 먼저, S110 단계에서 프로세서(200)는 입력 모듈(100)로부터 기 설정된 주기로 입력되거나 비주기적으로 입력되는 워크로드 데이터를 수집할 수 있다.

[0057] 그리고, S120 단계에서 프로세서(200)는 수집된 워크로드 데이터의 차원을 축소하고, 차원이 축소된 워크로드 데이터의 패턴을 좌표 데이터로 변환함으로써, 좌표 상에 상기 워크로드 데이터의 패턴을 나타낼 수 있도록 한다.

[0058] 그리고, S130 단계에서 프로세서(200)는 좌표 데이터로 변환된 워크로드 데이터의 패턴이 미리 정해진 적어도 하나의 기준 패턴 클러스터에 속하는지 여부를 분석함으로써, S140 단계에서 워크로드 데이터의 패턴에 대하여 1차적인 이상징후 판별을 수행한다.

[0059] 이때, S140 단계에서 프로세서(200)가 워크로드 데이터의 패턴과 기준 패턴 클러스터를 비교한 결과, 상기 워크로드 데이터의 패턴이 기준 패턴 클러스터들 중 적어도 하나에 속한다면, S210 단계로 진행되어 상기 수집된 워크로드 데이터에 이상징후가 있음을 사용자에게 알린다.

[0060] 반면, S140 단계에서 프로세서(200)가 상기 수집된 워크로드 데이터와 기준 패턴 클러스터들을 비교한 결과, 상기 수집된 워크로드 데이터가 기준 패턴 클러스터들에 속하지 않는 것으로 판단되면, S150 단계로 진행되어 이상징후 통계적 분석을 통해 워크로드 데이터에 대한 이상징후를 2차적으로 판별할 수 있다.

[0061] S150 단계에서 프로세서(200)는 기 설정된 이상탐지 알고리즘을 이용하여 상기 수집된 워크로드 데이터의 이상징후 여부를 판별한다. 여기서, 이상탐지 알고리즘이란 통계적인 분석을 이용하는 S-H-ESD 알고리즘일 수 있다.

[0062] 그리고 S160 단계에서 프로세서(200)는 상기 워크로드 데이터와 이상탐지 알고리즘을 기반으로 상기 판별한 결과를 저장한다.

[0063] 그리고, S170 단계에서 프로세서(200)는 상기 2차 이상징후 판별한 워크로드 데이터의 패턴과, 판별한 결과를 입력모듈(100)로 전달하고, 입력모듈(100)은 전달 받은 워크로드 데이터의 패턴과, 판별한 결과를 포함하는 정보를 화면(인터페이스) 상에 표시함으로써, 사용자로부터 이상 징후가 있는 워크로드 패턴에 대하여 확인받을 수 있도록 한다.

[0064] S180 단계에서 입력모듈(100)은 사용자가 상기 화면에 표시된 정보를 확인함에 따라 입력하는 입력정보들을 입력 받는다. 이때, 입력모듈(100)이 사용자로부터 입력받는 입력정보는 이상징후 판단 결과값 및 자연어들로 구성된 사용자 의견일 수 있다.

[0065] S190 단계에서, 프로세서(200)는 입력모듈(100)로 입력된 상기 이상징후 판단 결과값 및 사용자의 의견을 분석함으로써, S200 단계에서 1차 판별(S140)을 통해 클러스터링되지 않은 워크로드 데이터의 패턴에 대한 이상징후 여부 및 이상징후 발생 원인을 학습할 수 있다.

[0066] 보다 구체적으로는 S190 단계에서 프로세서(200)는 입력모듈(100)로 입력된 이상징후 판단 결과값을 기초로, 상기 1차 판별(S140)을 통해 클러스터링되지 않은 워크로드 데이터의 패턴에 대한 이상징후 패턴 여부를 결정하고, 상기 워크로드 데이터의 패턴에 대하여 입력된 사용자 의견을 기초로 자연어 처리를 하여 단어 분석을 함으로써, 상기 분석된 단어를 토대로 이상징후 원인 데이터로 활용하여, 상기 워크로드 데이터의 패턴을 기준 패턴 클러스터의 업데이트 정보로 이용할 수 있다.

[0067] 이에 따라, 다시 S130 단계가 진행되고, 프로세서(200)는 기준 패턴 클러스터가 업데이트 되었음으로, S140 단계에서 상기 워크로드 데이터의 패턴에 대한 기준 패턴 클러스트와의 비교를 한 후, S210 단계에서 이에 대한 알림을 실행할 수 있다.

[0068] 이상에서 설명한 본 발명의 실시예를 구성하는 모든 구성요소들이 하나로 결합하거나 결합하여 동작하는 것으로 기재되어 있다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 또한, 이와 같은 컴퓨터 프로그램은 USB 메모리, CD 디스크, 플래쉬 메모리 등과 같은 컴퓨터가 읽을 수 있는 기록매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 기록매체로서는 자기 기록매체, 광 기록매체 등이 포함될 수 있다.

[0069] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구 범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리 범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0070] 10: 이상징후 판별 장치

100: 입력 모듈

200: 프로세서

210: 데이터 수집부

220: 데이터 전처리부

230: 이상징후 판별부

240: 기준 패턴 클러스터 DB

250: 이상 징후 컴토용 DB

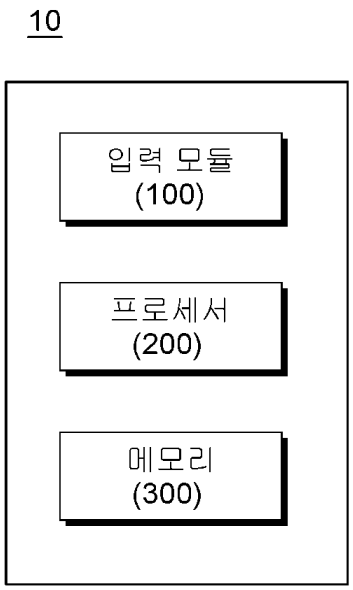
260: 데이터 업데이트부

270: 알림부

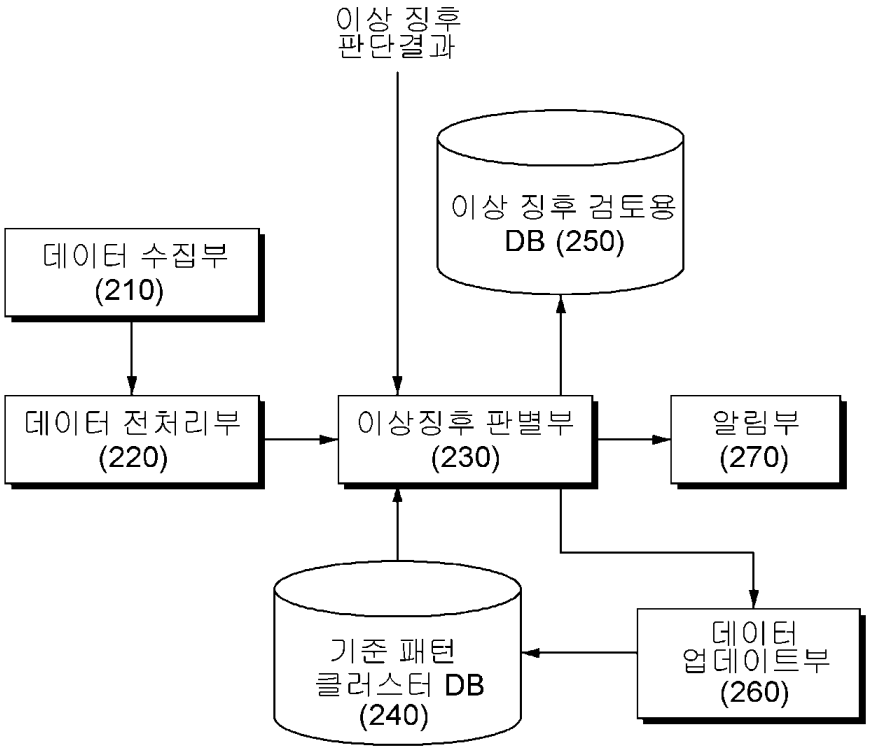
300: 메모리

도면

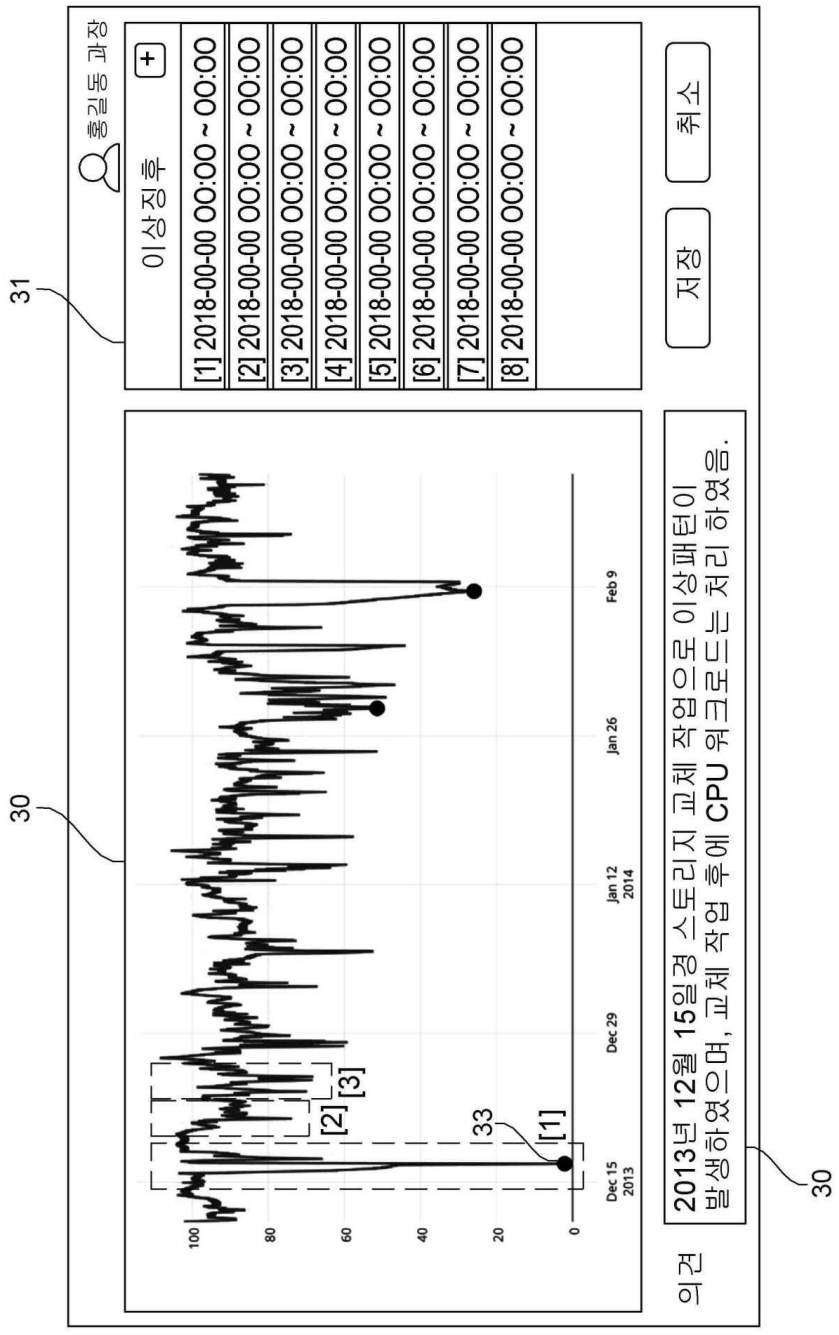
도면1



도면2


이상 징후
판단결과
이상 징후 검토용
DB (250)
데이터 수집부
(210)
데이터 전처리부
(220)
이상징후 판별부
(230)
알림부
(270)
기준 패턴
클러스터 DB
(240)
데이터
업데이트부
(260)

도면3


31
30
홍길동 과장
이상징후
[1] 2018-00-00 OO:OO ~ OO:OO
[2] 2018-00-00 OO:OO ~ OO:OO
[3] 2018-00-00 OO:OO ~ OO:OO
[4] 2018-00-00 OO:OO ~ OO:OO
[5] 2018-00-00 OO:OO ~ OO:OO
[6] 2018-00-00 OO:OO ~ OO:OO
[7] 2018-00-00 OO:OO ~ OO:OO
[8] 2018-00-00 OO:OO ~ OO:OO
저장
취소
100
80
60
40
20
0
Dec 15 2013
Jan 12 2014
Dec 29
Jan 26
Feb 9
[1]
[2]
[3]
33
의견
2013년 12월 15일경 스토리지 교체 작업으로 이상패턴이 발생하였으며, 교체 작업 후에 CPU 워크로드는 처리 하였음.
30

도면4

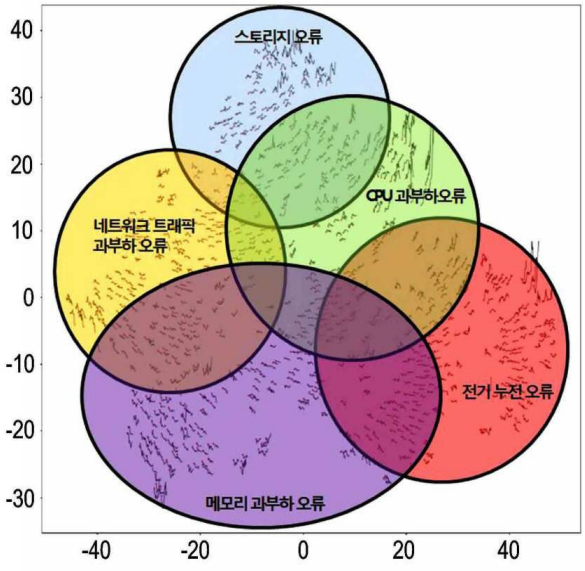
이름	직급	근속연수	가중치
A	차장	15	0.375
B	과장	9	0.225
C	과장	8	0.2
D	대리	3	0.075
E	대리	3	0.075
F	대리	2	0.05
총합		40	1

(a)

도면5

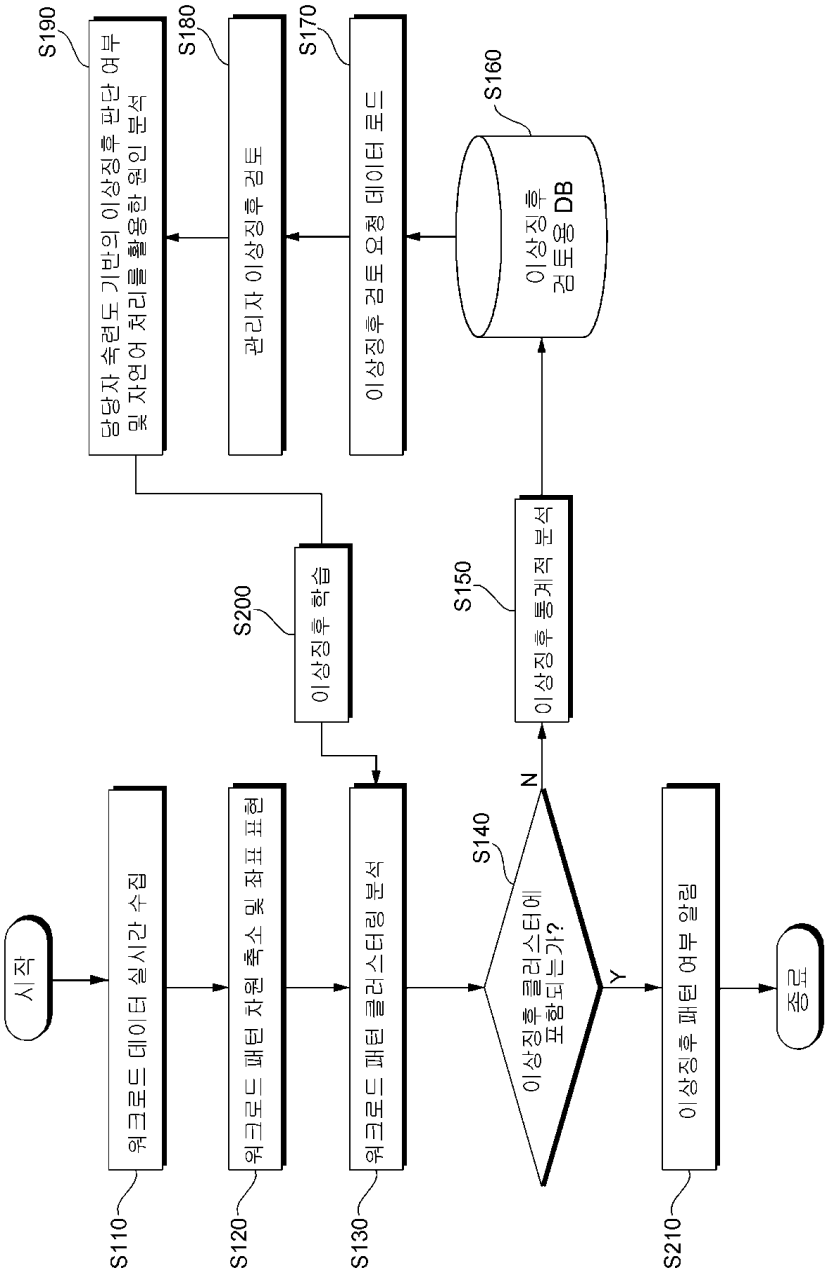
이름	좌표	Anomaly Score	의견
Pattern A	***	0.55	2013년 12월 15일경 스토리지 교체 작업으로 이상패턴이 발생하였으며, 교체 작업 후에 DISK I/O의 워크로드는 정상 처리 되었음
Pattern B	***	0.68	…
Pattern C	***	0.77	…
Pattern D	***	0.64	…
Pattern E	***	0.69	…
Pattern F	***	0.71	…

(a)



(b)

도면6


시작
워크로드 데이터 실시간 수집
S110
워크로드 패턴 차원 축소 및 좌표 표현
S120
워크로드 패턴 클러스터링 분석
S130
이상징후 클러스터에 포함되는가?
S140
N
Y
이상징후 통계적 분석
S150
이상징후 DB
S160
이상징후 학습
S200
이상징후 패턴 여부 알림
S210
종료
이상징후 검토용 데이터 로드
S170
관리자 이상징후 검토
S180
담당자 숙련도 기반의 이상징후 판단 여부 및 자연어 처리를 활용한 원인 분석
S190