



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0030898
(43) 공개일자 2019년03월25일

(51) 국제특허분류(Int. Cl.)
H04L 29/06 (2006.01) G06N 99/00 (2019.01)
H04L 29/08 (2006.01)
(52) CPC특허분류
H04L 63/1425 (2013.01)
G06N 20/00 (2019.01)
(21) 출원번호 10-2017-0118469
(22) 출원일자 2017년09월15일
심사청구일자 없음

(71) 출원인
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
이규철
대전광역시 유성구 배울2로 133, 205동 203호
백문기
대전광역시 유성구 문화원로46번길 24, 302호
오길탁
대전광역시 유성구 상대남로 26, 916동 1104호
(74) 대리인
홍성욱, 심경식

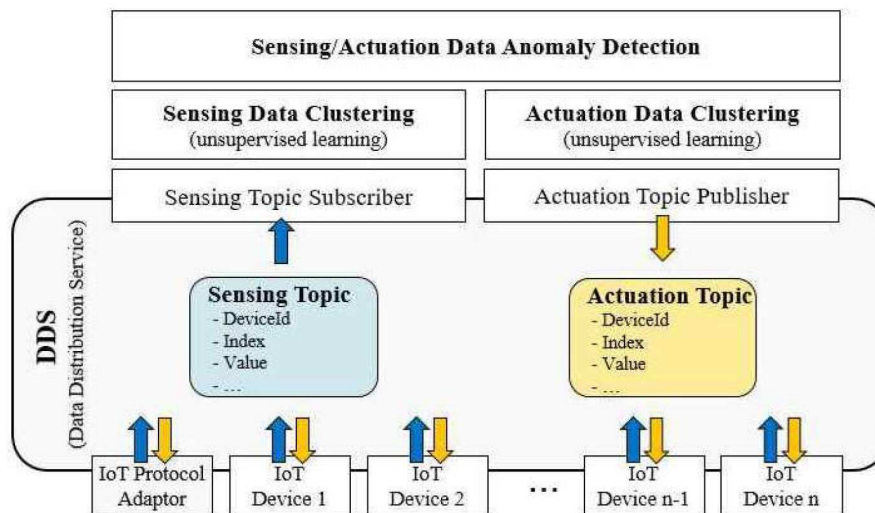
전체 청구항 수 : 총 1 항

(54) 발명의 명칭 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치

(57) 요약

본 발명의 일 실시예에 따른 사물인터넷 환경에서 이질적인 객체의 상호운용을 위한 가상 객체 어댑터가 적용된 데이터 분산 서비스 장치는 이질적인 객체 간 상호운용을 위한 데이터 분산 서비스(Data Distribution Service, DDS) 기반의 가상 객체 어댑터가 적용된 데이터 분산 서비스 장치에 있어서, 적어도 하나의 DDS 노드로부터 수신된 DDS 기반의 복수의 데이터 토픽을 하나의 데이터 토픽으로 통합하여 DDS 엔티티를 생성 및 관리하는 DDS 엔티티 관리부, 상기 통합된 데이터 토픽에 대한 유효성 검사 및 데이터 필터링을 수행하는 토픽 관리부, 사용자가 원하는 가상 객체에 대하여 모델링하도록, 상기 통합된 데이터 토픽에 관한 데이터를 모니터링하는 사용자 인터페이스부, 및 상기 모델링된 가상 객체에 DDS의 서비스 품질(Quality of Service, QoS) 정책을 반영하여 가상 객체 인스턴스를 생성 및 관리하는 가상 객체 인스턴스 관리부를 포함한다.

대표도 - 도1



(52) CPC특허분류

H04L 67/1097 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1345256571

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 지역혁신창의인력양성사업

연구과제명 Beacon을 이용한 영역기반(zone-based) 서비스 핵심기술 개발

기 여 율 1/1

주관기관 충남대학교

연구기간 2016.10.01 ~ 2017.09.30

명세서

청구범위

청구항 1

복수의 사물인터넷 디바이스와 연결된 데이터 분산 서비스(Data Distribution Service, DDS) 기반 사물인터넷 디바이스의 이상 탐지 장치에 있어서,

상기 디바이스 내 센서로 측정된 센싱 데이터에 관한 센싱 토픽 및 상기 디바이스가 수행해야 하는 동작 데이터에 관한 액추에이션 토픽을 상기 디바이스로부터 수신하거나 상기 디바이스로 전송하는 토픽 관리부; 및

마할라노비스 거리 알고리즘, 격리 포레스트 알고리즘 및 k-평균 알고리즘 중 하나인 기계학습 알고리즘을 적용하여 상기 센싱 토픽 및 상기 액추에이션 토픽에 대한 이상여부를 감지하는 토픽 이상 감지부

를 포함하는 것을 특징으로 하는 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치.

발명의 설명

기술 분야

[0001] 본 발명은 데이터 사물인터넷 디바이스의 이상 탐지 장치에 관한 것으로, 보다 상세하게는 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치에 관한 것이다.

배경 기술

[0002] 4차 산업혁명은 인공지능이 주도하는 차세대 산업혁명으로 사물인터넷(Internet of Things)을 통해 수집된 원시 데이터를 가공하여 의미 있는 지식과 정보를 추출해 인간의 삶을 풍요롭게 하는 개념으로 볼 수 있다. 사물인터넷 시스템은 운용 목적에 따라 다양한 구조가 존재한다. 비교적 단순한 3계층 구조의 사물인터넷 시스템은 수집된 원시 데이터를 응용으로 전달하는데 초점이 맞추어져 있고, 비교적 복잡한 5계층 구조의 사물인터넷 시스템은 대규모로 연결된 디바이스를 통해 수집한 데이터로부터 의사결정을 위한 지식을 추출하는데 초점이 맞추어져 있다.

[0003] 사물인터넷 등장 이후 저전력, 저비용의 소형 사물인터넷 디바이스가 폭발적으로 증가하였다. 이 디바이스는 MCU, 센서, 무선통신 모듈, 전원장치로 구성되며, 센서를 통해 원시 데이터를 생산하거나 무선통신을 통해 원격으로 장치를 제어하는 서비스를 다룰 수 있다. 하지만, 일반적인 PC와 비교해 리소스에 제약이 많기 때문에 센서를 통해 생산되는 원시 데이터에 누락과 오류가 빈번하게 발생할 수 있다. 예를 들면, 센서의 잘못된 측정, 주파수 간섭 노이즈, 낮은 배터리로 인한 슬립모드 등이 있다.

[0005] 관련 선행기술로는 대한민국 공개특허공보 제10-2014-0051513호(발명의 명칭: 데이터 분산 서비스 모니터링 장치 및 그 방법, 공개일자: 2014년 05월 02일)가 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명의 일 실시예는 대규모의 사물인터넷 디바이스와 연결 가능한 데이터 분산 서비스 환경에서, 기계학습 알고리즘을 이용하여 디바이스 내 원시 데이터의 이상여부를 감지하기 위한 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치를 제공한다.

[0008] 본 발명이 해결하고자 하는 과제는 이상에서 언급한 과제(들)로 제한되지 않으며, 언급되지 않은 또 다른 과제(들)은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0009] 본 발명의 일 실시예에 따른 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치는 복수의 사물인터넷 디바이스와 연결된 데이터 분산 서비스(Data Distribution Service, DDS) 기반 사물인터넷 디바이스의 이상 탐지 장치에 있어서, 상기 디바이스 내 센서로 측정된 센싱 데이터에 관한 센싱 토픽 및 상기 디바이스가 수행해야 하는 동작 데이터에 관한 액추에이션 토픽을 상기 디바이스로부터 수신하거나 상기 디바이스로 전송하는 토픽 관리부, 및 마할라노비스 거리 알고리즘, 격리 포레스트 알고리즘 및 k-평균 알고리즘 중 하나인 기계학습 알고리즘을 적용하여 상기 센싱 토픽 및 상기 액추에이션 토픽에 대한 이상여부를 감지하는 토픽 이상 감지부를 포함한다.

[0011] 기타 실시예들의 구체적인 사항들은 상세한 설명 및 첨부 도면들에 포함되어 있다.

발명의 효과

[0012] 본 발명의 일 실시예에 따르면, 대규모의 사물인터넷 디바이스와 연결 가능한 데이터 분산 서비스 환경에서, 기계학습 알고리즘을 이용하여 디바이스 내 원시 데이터의 이상여부를 감지할 수 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 일 실시예에 따른 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치의 개념을 설명하기 위해 도시한 도면이다.

도 2는 본 발명의 일 실시예에 있어서, 기계학습 알고리즘을 이용한 센싱 데이터 및 액추에이션 데이터의 이상 감지 과정을 설명하기 위해 도시한 도면이다.

도 3a는 본 발명의 일 실시예에 있어서, 기계학습 알고리즘이 적용된 센싱 데이터에 대한 이상 감지 결과를 설명하기 위해 도시한 그래프이다.

도 3b는 본 발명의 일 실시예에 있어서, 기계학습 알고리즘이 적용된 액추에이션 데이터에 대한 이상 감지 결과를 설명하기 위해 도시한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0014] 본 발명의 이점 및/또는 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성요소를 지칭한다.

[0016] 이하에서는 첨부된 도면을 참조하여 본 발명의 실시예들을 상세히 설명하기로 한다.

[0017] 도 1은 본 발명의 일 실시예에 따른 기계학습을 이용한 데이터 분산 서비스 기반 사물인터넷 디바이스의 이상 탐지 장치의 개념을 설명하기 위해 도시한 도면이다.

[0018] 도 2는 본 발명의 일 실시예에 있어서, 기계학습 알고리즘을 이용한 센싱 데이터 및 액추에이션 데이터의 이상 감지 과정을 설명하기 위해 도시한 도면이다.

[0019] 도 3a는 본 발명의 일 실시예에 있어서, 기계학습 알고리즘이 적용된 센싱 데이터에 대한 이상 감지 결과를 설명하기 위해 도시한 그래프이다.

[0020] 도 3b는 본 발명의 일 실시예에 있어서, 기계학습 알고리즘이 적용된 액추에이션 데이터에 대한 이상 감지 결과를 설명하기 위해 도시한 그래프이다.

[0021] 본 발명에서는 기계학습(Machine Learning) 알고리즘을 이용하여 대규모로 연결된 사물인터넷 디바이스로부터 수집된 원시 데이터의 이상을 감지하는 시스템을 제안하고자 한다. 제안하는 시스템 설계는 대규모 사물인터넷

디바이스를 연결하는데 적합한 표준 DDS(Data Distribution Service)를 기반으로 한다.

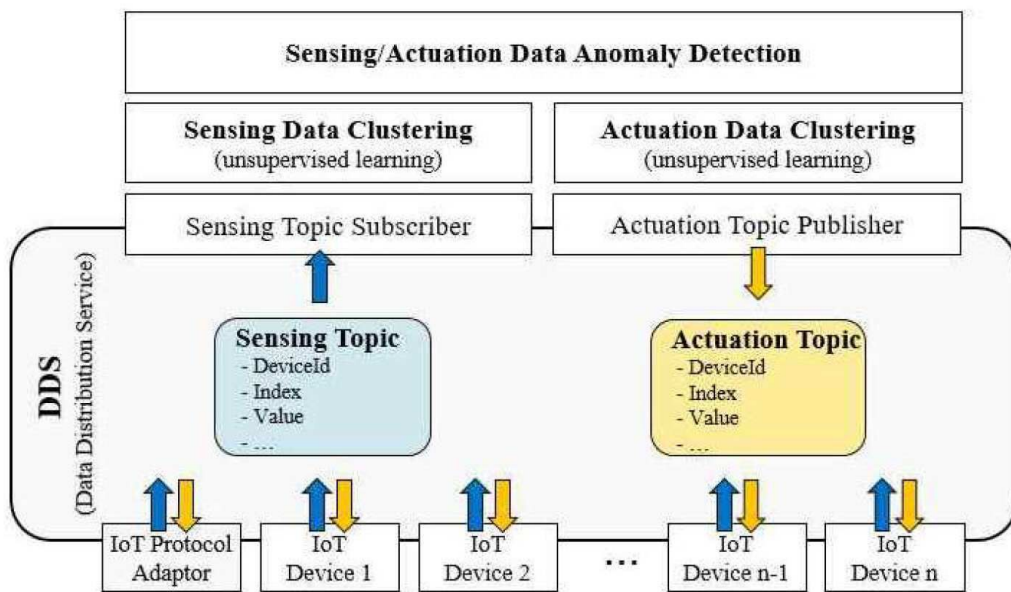
- [0022] DDS는 OMG(Object Management Group)에서 제정한 발간-구독(Publish-Subscribe)방식의 통신 미들웨어 표준이다. 국방과학 분야에서 널리 적용되는 DDS는 발간-구독 통신 방식으로 높은 상호 운용성을 제공하며 같은 도메인 내에서 데이터 중심적인 통신이 가능하다. 최근 사물인터넷의 등장으로, DDS는 점점 경량화 되었으며 대규모의 사물인터넷 디바이스를 연결할 수 있는 통신기술로 분류된다.
- [0023] 본 발명에서 제안하는 DDS 기반 사물인터넷 디바이스의 이상 감지 시스템은 도 1과 같다.
- [0024] n개의 사물인터넷 디바이스가 DDS를 기반으로 연결되어 있으며, 센싱(sensing) 토픽과 액추에이션(actuation) 토픽이라는 단위로 데이터를 전달한다. 센싱 토픽은 사물인터넷 디바이스가 센서로 측정한 온도, 심박수, 가속도 등의 원시 데이터를 포함한다. 액추에이션 토픽은 On/Off, 설정변경 등의 사물인터넷 디바이스가 수행해야 하는 동작을 포함한다.
- [0025] DDS는 UDP를 통해 사물인터넷 디바이스를 연결하는데, 연결된 n개의 디바이스 중에는 지그비(ZigBee) 디바이스와 같이 컴퓨팅 파워의 한계로 TCP/IP를 지원하지 못하는 경우도 많다. 이러한 경우, 사물인터넷 프로토콜 어댑터를 통해 간접적으로 DDS에 연결할 수 있다.
- [0026] 토픽을 통해 수집된 사물인터넷 디바이스의 원천 데이터들은 도 2와 같은 과정을 통해 이상을 감지하게 된다. 수집된 원천 데이터는 전처리 과정을 통해 기계학습에 적합한 훈련 집합(training set)으로 준비되고, 비지도(unsupervised) 학습 알고리즘을 이용하여 모델을 생성한다. 비지도 학습에 사용되는 알고리즘은 마할라노비스 거리(Mahalanobis Distance) 알고리즘, 격리 포레스트(Isolation Forest) 알고리즘 및 k-평균(K-means-) 알고리즘 중 한 가지를 선택하여 사용하는 것이 가장 효율적이다. 실시간성이 중요한 사물인터넷 시스템이라면 학습에 소요되는 시간이 작은 Mahalanobis Distance를 선택하는 것이 유리하고, 신뢰성이 중요한 사물인터넷 시스템이라면 K-means-를 선택하는 것이 좋다. 즉, 학습에 소요되는 시간과 정확도를 고려해 적절한 알고리즘을 선택할 필요가 있다.
- [0027] 도 3a 및 도 3b는 수집된 센싱 데이터 및 액추에이션 데이터의 일부이며, 비지도 학습 알고리즘을 적용하면 (A), (B), (C) 3가지 그룹으로 군집화(clustering)된다.
- [0028] 비지도 학습은 레이블이 없기 때문에 군집화된 데이터들이 무엇을 의미하는지 명시적으로 제시할 수 없다. 하지만, 군집화된 규모를 통해 (A)는 정상적인 원천 데이터들, (B), (C)는 이상치로 볼 수 있다. 참고로, 도 3a는 센싱 데이터가 군집화된 것이다.
- [0029] 보다 정밀한 이상치를 감지하기 위해서는 사물인터넷 디바이스의 특성을 반영할 필요가 있다. 사물인터넷 디바이스는 센싱을 통해 원천 데이터를 생산하고, 이를 기반으로 액추에이션 동작도 수행한다. 온도가 일정 수준 이상이면 에어컨을 작동시키는 스마트 홈 응용을 예로 들 수 있다. 이러한 액추에이션 특성을 반영하여 이상치로 분류된 센싱 데이터 (B), (C) 중 정상적인 데이터가 있는지 검사할 필요가 있다.
- [0030] 도 3b는 액추에이션 데이터가 군집화된 것으로, 군집화된 규모를 통해 (D)는 정상적인 원천 데이터, (E)는 이상치로 볼 수 있다. 액추에이션 데이터는 센싱 데이터와 다르게 지속적으로 생산되지 않고, 명시적인 요청 또는 조건을 만족해야 발생하는 특성이 있기 때문에, (E)는 이상치처럼 보이나 실제로는 정상적인 데이터이다.
- [0031] 센싱-액추에이션 특성을 반영해 도 3a 및 도 3b의 데이터를 비교하면, (C)로 군집화된 데이터가 생산될 때 (E)의 액추에이션이 발생한다는 연관성을 도출할 수 있다. 따라서, (B)로 군집화된 데이터는 사물인터넷 디바이스가 잘못 생산한 원천 데이터일 가능성이 매우 높다.
- [0032] 센싱 데이터 및 액추에이션 데이터를 비지도 학습하여 만들어진 모델은 Value의 범위가 다를 수 있기 때문에, 센싱-액추에이션의 연관성을 도출하기 위해서 일련의 정규화 과정이 추가될 수 있다. 동일 인덱스 범위 안에서 정상으로 군집화된 데이터는 Value를 0으로, 이상치로 군집화된 데이터는 1로 정규화하여 오버랩(overlap)하면 (B)와 같은 센싱 데이터의 이상치를 쉽게 감지할 수 있다.
- [0033] 대규모 사물인터넷 디바이스를 바탕으로 원시 데이터를 수집하여 의사 결정을 지원하는 사물인터넷 시스템은 사물인터넷 디바이스의 특성을 고려해 누락과 오류가 포함된 원시 데이터를 분류할 수 있어야 한다.
- [0034] 본 발명에서는 기계학습 알고리즘을 이용하여 인간의 개입 없이 자동으로 이상 데이터를 감지하는 시스템을 구현하였다. 이 시스템은 사물인터넷 디바이스의 센싱-액추에이션 특성을 고려해 보다 정밀한 이상 감지가 가능하여 eHealth와 같은 중대한 사물인터넷 응용의 신뢰도를 높이는데 기여할 수 있다.

[0036] 지금까지 본 발명에 따른 구체적인 실시예에 관하여 설명하였으나, 본 발명의 범위에서 벗어나지 않는 한도 내에서는 여러 가지 변형이 가능함은 물론이다. 그러므로, 본 발명의 범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허 청구의 범위뿐 아니라 이 특허 청구의 범위와 균등한 것들에 의해 정해져야 한다.

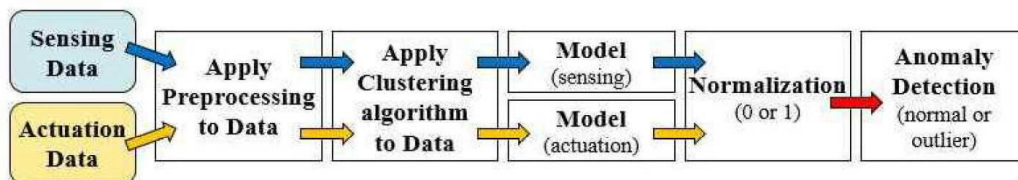
[0037] 이상과 같이 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 이는 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명 사상은 아래에 기재된 특허청구범위에 의해서만 파악되어야 하고, 이의 균등 또는 등가적 변형 모두는 본 발명 사상의 범주에 속한다고 할 것이다.

도면

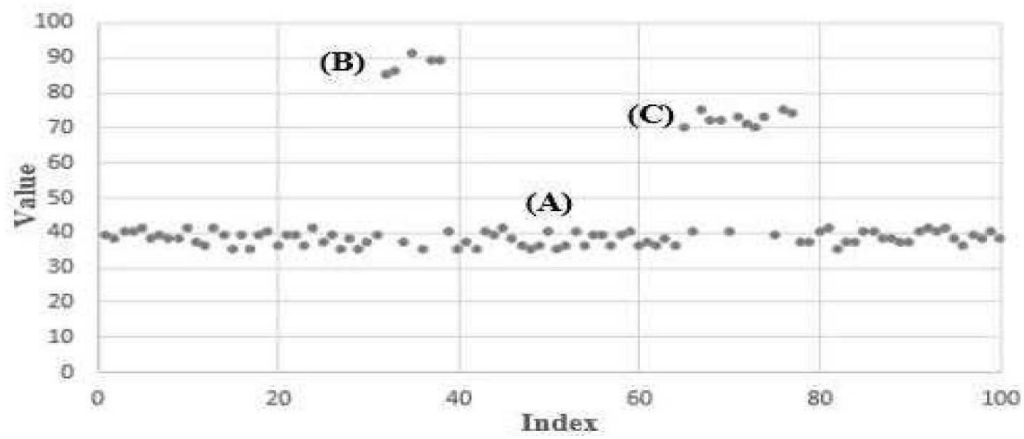
도면1



도면2



도면3a



도면3b

