



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0078605  
(43) 공개일자 2016년07월05일

(51) 국제특허분류(Int. Cl.)

G06Q 50/06 (2012.01)

(21) 출원번호 10-2014-0188075

(22) 출원일자 2014년12월24일

심사청구일자 2014년12월24일

(71) 출원인

충남대학교산학협력단

대전광역시 유성구 대학로 99 (궁동, 충남대학교)

(72) 발명자

이규철

대전 유성구 배울2로 133, 205동 203호 (용산동, 경남아너스빌2단지)

조우승

경남 창원시 마산회원구 합성서11길 148-1, 409호 (합성동, 태우아파트)

백문기

충북 제천시 용두대로7길 25, 107동 602호 (하소동, 그린코아루아파트)

(74) 대리인

특허법인태동

전체 청구항 수 : 총 9 항

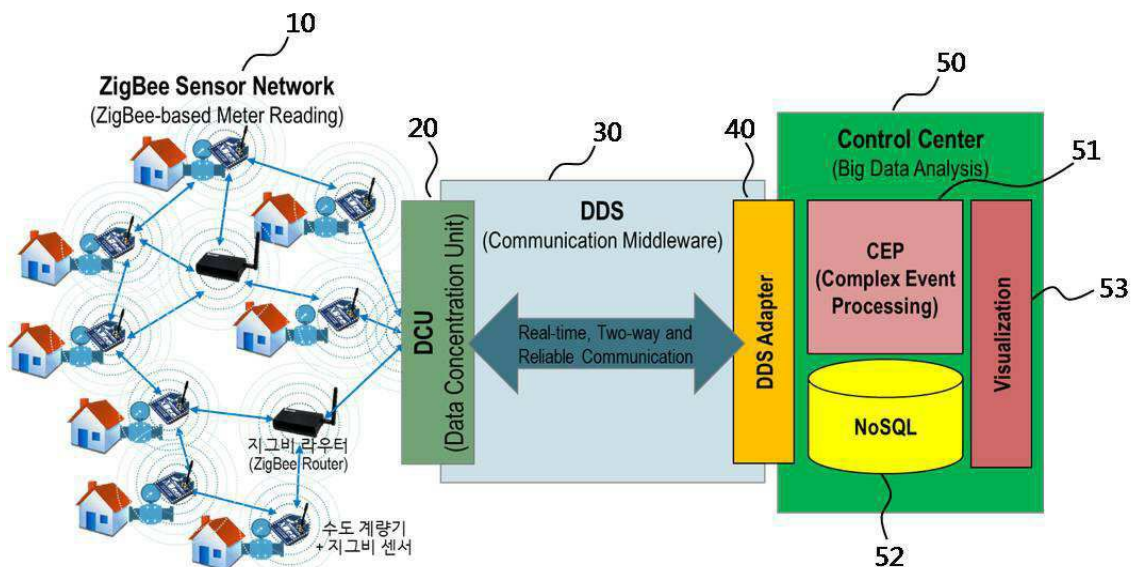
(54) 발명의 명칭 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 및 누수 탐지방법

### (57) 요약

본 발명은 본 발명은 관제센터에서 유량을 측정하는 무선 센서 네트워크로부터 데이터 분배 시스템을 통해 수신되는 데이터에 기초하여 가정별 물 수요량을 예측하는 방법에 있어서, 데이터 분배 시스템을 통해 수신되는 유량 데이터 토픽과 이벤트 토픽을 데이터베이스에 저장하는 제 1 단계, 상기 데이터베이스에 저장된 데이터 셋을 학

(뒷면에 계속)

대표도 - 도1



습 데이터 셋으로 변환한 후, 기계 학습 모듈을 통해 예측 모델을 생성하는 제 2 단계, 데이터 분배 시스템을 통해 실시간으로 수신되는 유량 데이터 토픽과 이벤트 토픽을 CEP 이벤트 타입 객체로 변환하는 제 3 단계, CEP 엔진에서 상기 예측 모델을 호출하여 설정된 예측 주기에 대한 각 가정별 유량을 예측하는 제 4 단계를 포함하는 것을 특징으로 한다.

본 발명에 따르면, 가정별 물 수요량을 실시간 예측하여 물 공급량을 빠르게 산정하여 안정적인 물의 공급이 가능할 뿐 아니라, 기계학습을 사용하여 미수집된 유량데이터를 예측하여 누수량 추정 정확도를 높임으로써 보다 정확하게 누수량을 예측할 수 있는 효과가 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호 2014B7243010116

부처명 교육부

연구관리전문기관 충남대학교 산학협력단

연구사업명 LINC사업

연구과제명 복합 이벤트처리와 기계학습 빅데이터 처리 기술을 이용한 물 수요 예측 기법 개발

기 여 율 1/1

주관기관 한국연구재단

연구기간 2014.07.01 ~ 2014.12.31

## 명세서

### 청구범위

#### 청구항 1

관제센터에서 유량을 측정하는 무선 센서 네트워크로부터 데이터 분배 시스템을 통해 수신되는 데이터에 기초하여 가정별 물 수요량을 예측하는 방법에 있어서,

데이터 분배 시스템을 통해 수신되는 유량 데이터 토픽과 이벤트 토픽을 데이터베이스에 저장하는 제 1 단계;

상기 데이터베이스에 저장된 데이터 셋을 학습 데이터 셋으로 변환한 후, 기계 학습 모듈을 통해 예측 모델을 생성하는 제 2 단계;

데이터 분배 시스템을 통해 실시간으로 수신되는 유량 데이터 토픽과 이벤트 토픽을 CEP 이벤트 타입 객체로 변환하는 제 3 단계;

CEP 엔진에서 상기 예측 모델을 호출하여 설정된 예측 주기에 대한 각 가정별 유량을 예측하는 제 4 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 방법.

#### 청구항 2

제 1 항에 있어서,

상기 데이터 토픽은 수도 계량기 식별자, 유량 데이터 및 데이터 값을 포함하고, 상기 데이터 값은 블록번호, 계량기 유형 정보를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 방법.

#### 청구항 3

제 2 항에 있어서,

상기 이벤트 토픽은 수도 계량기 식별자, 이벤트 이름 및 이벤트 값을 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 방법.

#### 청구항 4

제 1 항에 있어서,

상기 제 2 단계는

유량 데이터를 수집하여 저장하여 초기 데이터 셋을 생성하는 단계;

초기 데이터 셋에서 각 가정 및 학습을 위한 연도를 선택하여 데이터를 추출한 후 이를 시계열 데이터 셋으로 변환하는 단계;

시계열 데이터 셋으로부터 최종적으로 날짜가 가지는 속성을 추출하여 학습 데이터 셋으로 변환하는 단계; 및

상기 학습 데이터 셋을 기계 학습 모듈에 입력하여 예측 모델을 생성하는 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 방법.

#### 청구항 5

제 1 항에 있어서,

상기 제 4 단계는

CEP 엔진으로 입력된 유량데이터를 수도 계량기 식별자 속성에 기초하여 가정별로 분할하는 단계;

가정별로 분할된 유량 데이터를 예측주기마다 합산하는 단계; 및

예측주기의 유량 합계를 사용하여 CEP 엔진에 등록된 학습 모델 함수의 결과를 반환하여 예측값을 생성하는 단

계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 방법.

## 청구항 6

관제센터에서 유량을 측정하는 무선 센서 네트워크로부터 데이터 분배 시스템을 통해 수신되는 데이터에 기초하여 수용가의 누수를 탐지하는 방법에 있어서,

데이터 분배 시스템을 통해 수신되는 유량 데이터 토픽과 이벤트 토픽을 데이터베이스에 저장하고, 저장된 데이터 셋을 학습 데이터 셋으로 변환한 후, 기계 학습 모듈을 통해 예측 모델을 생성하는 제 1 단계;

데이터 분배 시스템을 통해 실시간으로 수신되는 유량 데이터 토픽과 이벤트 토픽을 CEP 이벤트 타입 객체로 변환하는 제 2 단계;

블록 유량계에서 수집되는 유량데이터를 가정용 계량기별로 분할하는 제 3 단계;

가정용 계량기별로 분할된 데이터를 집계하여 수용가별 예측시간 단위의 유량 데이터로 변환하는 제 4 단계;

상기 변환된 유량 데이터를 예측 모델로 입력하여 각 수용가별 유량을 예측하는 제 5 단계;

가정용 계량기별로 분할된 데이터를 블록별로 그룹화한 후 합산하되, 상기 이벤트 토픽에 유량 데이터 미수집 이벤트가 포함된 경우 상기 5단계에서 산출된 예측 유량을 해당 주기의 예측 유량으로 반영하는 제 6 단계;

블록 유량계의 측정유량과 상기 6단계에서 합산된 유량의 차를 계산하여 그 차가 0 이상인 블록을 찾아내어 누수를 탐지하는 제 7 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 누수 탐지방법.

## 청구항 7

제 6 항에 있어서,

상기 제 3 단계에서 CEP 이벤트 타입 객체의 계량기 유형정보를 통해 데이터가 분할되고, 상기 제 6 단계에서는 CEP 이벤트 타입 객체의 블록 번호 속성을 통해 그룹화가 이루어지는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 누수 탐지방법.

## 청구항 8

제 6 항에 있어서,

상기 제 1 단계는

유량 데이터를 수집하여 저장하여 초기 데이터 셋을 생성하는 단계;

초기 데이터 셋에서 각 가정 및 학습을 위한 연도를 선택하여 데이터를 추출한 후 이를 시계열 데이터 셋으로 변환하는 단계;

시계열 데이터 셋으로부터 최종적으로 날짜가 가지는 속성을 추출하여 학습 데이터 셋으로 변환하는 단계; 및

상기 학습 데이터 셋을 기계 학습 모듈에 입력하여 예측 모델을 생성하는 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 누수 탐지방법.

## 청구항 9

제 6 항에 있어서,

상기 제 5 단계는

CEP 엔진으로 입력된 유량데이터를 수도 계량기 식별자 속성에 기초하여 가정별로 분할하는 단계;

가정별로 분할된 유량 데이터를 예측주기마다 합산하는 단계; 및

예측주기의 유량 합계를 사용하여 CEP 엔진에 등록된 학습 모델 함수의 결과를 반환하여 예측값을 생성하는 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 누수 탐지방법.

## 발명의 설명

## 기술분야

[0001] 본 발명은 가정별 물 수용량을 예측하고 누수를 탐지하는 방법에 관한 것으로, 보다 상세하게는 실시간으로 가정 수준의 물 수요 예측을 하여 빠르게 공급량을 산정하여 안정적인 물 공급이 가능하고, 가정별 물 수요량 예측기법을 통해 누수 탐지시에 미 수집된 유량 데이터가 발생하는 경우에도 가정별 수요예측을 사용하여 누수 탐지 정확도를 높일 수 있는 기술에 관한 것이다.

## 배경기술

[0002] 기존의 수자원 관리는 노후화된 기반구조로 인한 유량과 유압 등에 대한 실시간 데이터 수집의 제한, 상수도관의 누수 발생 문제, 도시의 거대화, 지구 기후변화 등의 문제점을 가지고 있으며, 이로 인해 제한된 수자원의 효율적인 공급 및 관리방법의 필요성이 대두되고 있다.

[0003] 이러한 환경적 변화의 문제점에 능동적으로 대처하고자 스마트 워터 그리드가 제시되고 있다. 스마트 워터 그리드는 기존의 수자원 관리 시스템의 한계를 극복하기 위해 첨단정보통신기술(ICT; Information Communication Technologies)을 융합하는 차세대 물관리 인프라 시스템으로, 물이나 지하수, 해수, 하천수 등 다양한 수자원을 활용하고 물을 효율적으로 관리하여 수자원의 지역적, 시간적 불균형을 해소하는 기술로 정의되어 있다.

[0004] 스마트 워터 그리드의 핵심기술로는 지능형 검침 인프라(AMI; Advanced Metering Infrastructure), 스마트 그리드를 이용한 물/에너지 최적화, 수자원 및 수질 관리 센서네트워크를 통한 수자원 정보의 지식 네트워크 구축, 국가 단위의 효율적 광역 수자원 관리 시스템의 개발 및 적용을 예로 들 수 있으며, 제조기술, 정보통신기술 그리고 통합운영 관리기술이 접목되어 다양한 형태의 융복합 기술이 개발/적용되고 있다. 이중에 지능형 검침 인프라는 스마트 워터 그리드의 핵심 기술이다.

[0005] 지능형 검침 인프라는 물 소비량 및 계량기의 상태와 관련된 데이터를 측정/전송/분석에 활용되는 다양한 스마트 미터, 통신방식 및 인프라, 데이터 통합 관리시스템을 총칭한다. 지능형 검침 인프라 중 상수도 원격 검침 시스템(Automatic Meter Reading)은 수용가의 지속적 증가에 따른 검침 업무의 난해함을 해소하고, 검침 누락 예방 및 고지업무의 투명성을 확보할 수 있는 장점을 가진다. 또한 용수 사용의 실시간 분석을 통해 누수탐지가 가능하게 되면서 빠른 누수처리가 가능하다.

[0006] 원격 검침 시스템을 통해서 수집된 데이터를 통해서 누수를 탐지하는 연구들이 있었다. 이들 연구들에서는 매년 총공급량의 10% 이상의 누수가 발생하고 있으며, 2012년의 누수량은 6억2,600만 톤으로 약 5,100억 원의 손실이 발생하였다고 한다.

[0007] 국내의 누수 탐지에 대한 연구들은 블록 시스템을 대상으로 누수를 탐지하고 있다. 블록 시스템은 배수관망을 대/중/소의 단위구역으로 나눠 적정유량과 유압이 유지되도록 하며, 수도물 공급량/소비량/누수량 등을 분석가능하도록 관망을 구성하여 관리하는 체계이다.

[0008] 종래에는 시스템을 통해 먼 단위 누수탐지방식으로 소블록의 누수가 탐지되면, 인력에 의해 선/점 단위의 누수를 탐지하고 있다. 이처럼 소블록 누수탐지는 해당 블록의 정확한 누수 발생여부를 발견하는 것이 아니라, 원격 검침 시스템으로 수집된 유량 데이터를 계산하여 높은 누수량을 산출하여 인력에 의한 누수탐지의 우선순위를 부여하는데 사용된다.

[0009] 이와 같은 방식으로 누수를 탐지하는 기존시스템들은 실시간 누수탐지를 위한 시스템을 구축하지 않고 있다. 대규모의 일괄 처리 시스템을 구축하여 일 단위의 누수 탐지를 수행하기 때문에, 빠르게 누수에 대한 처리를 수행할 수 없다. 이를 위해 실시간으로 대량의 유량 데이터를 실시간으로 처리하기 위한 기술이 필요하다.

[0010] 이를 위해 본 출원인은 복합 이벤트 처리 기법(CEP : Complex Event Processing)을 사용하여 대량의 유량 데이터를 실시간으로 처리하여 빠른 누수탐지를 하는 기술을 개발한 바 있다.

[0011] 한편, 원격 검침 시스템으로부터 수집되는 유량 데이터가 주파수 간섭이나 하드웨어 문제로 최종 서버로 전송되지 못하게 된다면 수집되지 못한 유량만큼을 누수량으로 추정하게 되는데, 이에 의해 잘못된 누수량 산출의 원인이 되고 있다. 관련 연구에 따르면 상수도 원격 검침 시스템의 정상 수신율은 평균 85 ~ 90% 정도로서, 10 ~ 15%의 데이터가 정상적으로 수집되지 못하고 있는 실정이다.

[0012] 따라서, 미수집된 유량 데이터가 발생하는 경우에도 정확하게 누수량을 산출할 수 있는 방법이 요구되고 있다.

[0013] 다른 한편으로, 국내에서의 용수에 대한 문제는 주로 새로운 용수 공급원을 개발하여 용수 수요를 충족시키는데 초점을 두었다. 따라서 수요량을 조절하기보다는 공급량을 늘려 낮은 가격 수준에서 수요량을 충족시키고자 하였다. 그 결과 공급량이 한정된 현실에서 용수 수요량을 충족시킬 수 없는 현상이 예상될 뿐만 아니라 공급원가에도 못 미치는 가격으로 인하여 공급 비용을 충당하기 위해 정부예산의 비효율적 사용을 야기하였다. 또한 국내의 물 수요 예측은 배수지 수준의 물 수요량을 예측하기 때문에, 가정의 세부적인 물 수요량과 큰 오차가 유발되고 있다. 배수지는 물을 많이 사용하는 시간대에는 많은 물을 공급하고, 물의 사용이 적은 새벽에는 물의 공급량을 줄이는 등, 급수량을 조절하면서 물을 공급하는 시설이다. 배수지별로 예측을 수행하게 된다면, 배수지의 급수 계통 내에 존재하는 수용가에서의 급격한 물 수요량 변동을 정확하게 예측할 수 없는 문제점이 있다.

[0014] 이를 위해 원격검침시스템(AMR; Automatic Meter Reading)을 통해 가정별 실측 데이터를 기반으로 물 수요량 예측 모델을 생성하는 연구가 진행된 바 있으나, 이 연구는 실측자료를 토대로 물 수요를 예측하기 위한 모델을 만드는 수준에 그쳤다. 때문에 실시간으로 물 수요예측이 불가능하기 때문에 빠르게 용수 공급량을 예상할 수 없었다.

## 선행기술문헌

### 비특허문헌

[0015] (비특허문헌 0001) IT 융합을 통한 스마트 워터 그리드 추진 방안에 대한 연구, The Journal of Digital Policy & Management, 2013

(비특허문헌 0002) 국외 상수도 원격 검침 시스템의 개발 동향 및 현장 적용 사례 고찰, 대한환경공학회, 2012

## 발명의 내용

### 해결하려는 과제

[0016] 본 발명의 목적은 스마트 그리드에서 사용되는 실시간 가정별 수요량 예측 기법을 적용하여 누수 탐지시에 미 수집된 유량 데이터가 발생하는 경우 가정별 수요 예측을 사용하여 누수탐지시에 정확도를 높일 수 있도록 하는 것이다.

[0017] 본 발명의 다른 목적은 실시간으로 가정 수준의 물 수요 예측을 하여 빠르게 공급량을 산정하여 안정적인 물 공급이 가능하도록 하는 것이다.

### 과제의 해결 수단

[0018] 상기와 같은 목적을 달성하기 위한 본 발명의 일 측면에 따르면, 관제센터에서 유량을 측정하는 무선 센서 네트워크로부터 데이터 분배 시스템을 통해 수신되는 데이터에 기초하여 가정별 물 수요량을 예측하는 방법에 있어서, 데이터 분배 시스템을 통해 수신되는 유량 데이터 토픽과 이벤트 토픽을 데이터베이스에 저장하는 제 1 단계, 상기 데이터베이스에 저장된 데이터 셋을 학습 데이터 셋으로 변환한 후, 기계 학습 모듈을 통해 예측 모델을 생성하는 제 2 단계, 데이터 분배 시스템을 통해 실시간으로 수신되는 유량 데이터 토픽과 이벤트 토픽을 CEP 이벤트 타입 객체로 변환하는 제 3 단계, CEP 엔진에서 상기 예측 모델을 호출하여 설정된 예측 주기에 대한 각 가정별 유량을 예측하는 제 4 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 가정별 물 수요량 예측 방법이 제공된다.

[0019] 상기와 같은 목적을 달성하기 위한 본 발명의 다른 일 측면에 따르면, 관제센터에서 유량을 측정하는 무선 센서 네트워크로부터 데이터 분배 시스템을 통해 수신되는 데이터에 기초하여 수용가의 누수를 탐지하는 방법에 있어서, 데이터 분배 시스템을 통해 수신되는 유량 데이터 토픽과 이벤트 토픽을 데이터베이스에 저장하고, 저장된 데이터 셋을 학습 데이터 셋으로 변환한 후, 기계 학습 모듈을 통해 예측 모델을 생성하는 제 1 단계, 데이터 분배 시스템을 통해 실시간으로 수신되는 유량 데이터 토픽과 이벤트 토픽을 CEP 이벤트 타입 객체로 변환하는 제 2 단계, 블록 유량계에서 수집되는 유량데이터를 가정용 계량기별로 분할하는 제 3 단계, 가정용 계량기별로 분할된 데이터를 집계하여 수용가별 예측시간 단위의 유량 데이터로 변환하는 제 4 단계, 상기 변환된 유량 데

이터를 예측 모델로 입력하여 각 수용가별 유량을 예측하는 제 5 단계, 가정용 계량기별로 분할된 데이터를 블록별로 그룹화한 후 합산하되, 상기 이벤트 토픽에 유량 데이터 미수집 이벤트가 포함된 경우 상기 5단계에서 산출된 예측 유량을 해당 주기의 예측 유량으로 반영하는 제 6 단계, 블록 유량계의 측정유량과 상기 6단계에서 합산된 유량의 차를 계산하여 그 차가 0 이상인 블록을 찾아내어 누수를 탐지하는 제 7 단계를 포함하는 것을 특징으로 하는 복합 이벤트 처리와 기계학습 기반의 누수 탐지방법이 제공된다.

### 발명의 효과

- [0020] 본 발명에 따르면, 가정별 물 수요량을 실시간 예측하여 물 공급량을 빠르게 산정하여 안정적인 물의 공급이 가능한 효과가 있다.
- [0021] 또한, 기계학습을 사용하여 미수집된 유량데이터를 예측하여 누수량 추정 정확도를 높임으로써 보다 정확하게 누수량을 예측할 수 있는 효과도 있다.

### 도면의 간단한 설명

- [0022] 도 1은 본 발명에 따른 원격 검침 시스템의 전체 구성도이다.
- 도 2는 유량 데이터 토픽 명세를 나타낸 것이다.
- 도 3은 이벤트 토픽 명세를 나타낸 것이다.
- 도 4는 데이터 토픽을 CEP 이벤트 타입으로 변환하는 과정을 도시한 것이다.
- 도 5는 데이터 토픽을 데이터베이스에 저장하는 과정을 도시한 것이다.
- 도 6은 본 발명에 따른 유량 데이터 예측 구조를 도시한 것이다.
- 도 7은 기계 학습 과정을 설명하기 위한 구성도이다.
- 도 8은 데이터베이스에 저장된 초기 데이터 셋을 예시한 것이다.
- 도 9는 중간 데이터 셋을 예시한 것이다.
- 도 10은 학습 데이터 셋을 예시한 것이다.
- 도 11은 CEP 엔진에서의 스트림 처리과정을 도시한 것이다.
- 도 12는 미수집 유량이 발생한 경우 예측 유량을 반영한 누수 탐지 성능 평가 결과 그래프이다.
- 도 13은 미수집 유량이 발생하지 않은 경우 누수 탐지 성능 평가 결과를 나타낸 그래프이다.
- 도 14는 예측 정확도 평가 결과 그래프이다.

### 발명을 실시하기 위한 구체적인 내용

- [0023] 본 발명에 관한 설명은 구조적 내지 기능적 설명을 위한 실시예에 불과하므로, 본 발명의 권리범위는 본문에 설명된 실시예에 의하여 제한되는 것으로 해석되어서는 아니 된다. 즉, 실시예는 다양한 변경이 가능하고 여러 가지 형태를 가질 수 있으므로 본 발명의 권리범위는 기술적 사상을 실현할 수 있는 균등물들을 포함하는 것으로 이해되어야 한다. 또한, 본 발명에서 제시된 목적 또는 효과는 특정 실시예가 이를 전부 포함하여야 한다거나 그러한 효과만을 포함하여야 한다는 의미는 아니므로, 본 발명의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.

- [0024] 도 1은 본 발명에 따른 원격 검침 시스템의 전체 구성도이다.

- [0025] 도 1에 도시된 바와 같이, 본 발명에 따른 원격 검침 시스템은 무선 센서 네트워크(10)와, 데이터 집중 유닛(20, DCU : Data Concentration Unit), 데이터 분배 시스템(30, DDS : Data Distribution System), DDS 어댑터(40) 및 관제 센터(50)를 포함하여 구성된다.

- [0026] Zigbee 기반의 무선 센서 네트워크(10)를 통해 각 가정의 수도 계량기로부터 측정되는 유량 데이터는 데이터 집중 유닛(20)에 수집되고, 데이터 집중 유닛(20)은 통신 미들웨어인 데이터 분배 시스템(30)을 사용하여 실시간으로 유량 데이터/이벤트 데이터를 관제센터(50)로 전송한다.
- [0027] 데이터 분배 시스템(DDS : Data Distribution System)은 OMG(Object Management Group)에서 국제 표준으로 정한 실시간 통신 미들웨어로서, 분산 환경을 위한 발간/구독 모델을 가진다. DDS의 특징은 많은 사용자들이 물리적인 네트워크의 특성에 상관없이 동적인 네트워크를 형성하는 것이다. 또한 발간자와 구독자간 정보 교환 시 실시간성과 신뢰성을 만족한다. DDS는 글로벌 데이터 공간에서 토픽이라는 가상 채널을 통해 데이터 작성자(Data Writer)와 데이터 독자(Data Reader)간의 데이터를 송수신하게 된다.
- [0028] DDS 어댑터(40)는 DDS 상에 발간된 유량 데이터 및 이벤트 데이터를 구독하여 복합 이벤트 처리부(51) 및 데이터베이스(52)에 전송하기 위한 것으로서, 구독된 토픽을 복합 이벤트 처리부(51) 및 데이터베이스(52)에 적합한 객체로 변환하는 기능을 수행한다.
- [0029] 도 2는 유량 데이터 토픽 명세를 나타낸 것이고, 도 3은 이벤트 토픽 명세를 나타낸 것이다.
- [0030] 수도계량기로부터 데이터와 이벤트가 생성되는데, 데이터 및 이벤트 특징에 맞는 토픽을 명세하여 데이터 집중 유닛(20)과 DDS 어댑터(40) 간의 통신을 수행한다. 토픽명세는IDL(Interface Description Language)이 사용될 수 있다.
- [0031] 도 2의 데이터 토픽은 지속적으로 발생하는 데이터를 전달하기 위한 토픽으로 유량은 데이터이므로 데이터 토픽으로 발간된다.
- [0032] 데이터 토픽 명세에서 WaterMeterID는 수도 계량기의 식별자로서, 계량기 제조업체에 의해 부여되는 번호이고, DataName은 발간하는 데이터의 이름으로 유량은 WaterFlow라는 이름을 가진다. Value는 데이터의 값을 가지며 블록번호(BN), 계량기타입(MS), 유량(flow)를 가진다.
- [0033] 이벤트는 특정상태가 변하거나 특정 조건에 해당하여 처리된 결과정보를 의미한다. 수도계량기로부터 발생하는 역류, 단선, 배터리 부족은 이벤트 토픽을 통해 전송된다.
- [0034] 도 3의 이벤트 토픽명세에서, WaterMeterID는 수도 계량기의 식별자로서, 계량기 제조업체에 의해 부여되는 번호이고, EventName은 발간하는 이벤트의 이름이다. EventValue는 이벤트의 상태값을 가진다. EventName은 CableCut(단선), LowBattery(배터리 부족), Non-nonnected(비연결) 등일 수 있고, EventValue는 state값이 0인 경우 이상이 발생한 것으로 볼 수 있다.
- [0035] 도 4는 데이터 토픽을 CEP 이벤트 타입으로 변환하는 과정을 도시한 것으로서, DDS 상에 발간되는 유량 데이터/이벤트를 CEP 엔진에서 사용하기 위해서는 토픽 인스턴스를 CEP 엔진에서 처리할 수 있는 이벤트로 정의해야 한다. CEP 엔진은 POJO 객체 형태를 사용하며, getter 메소드를 제공해야 한다.
- [0036] 도 4를 참조하면, DDS 어댑터(40)는 데이터 분배 시스템(30)의 데이터/이벤트 토픽으로부터 토픽 인스턴스를 구독하면, 데이터 토픽 인스턴스는 데이터 이벤트 객체로 변환하고, 이벤트 토픽 인스턴스는 이벤트 객체로 변환하여 복합 이벤트 처리부(51) 즉, CEP 엔진으로 전송한다.
- [0037] 도 5는 데이터 토픽을 데이터베이스에 저장하는 과정을 도시한 것으로서, DDS 어댑터(40)는 데이터 분배 시스템(30)의 토픽 인스턴스를 구독하면, 데이터베이스(52)에 저장한다. 기계학습을 통해서 예측모델을 만들기 위해서는 장기간의 유량 데이터 셋을 저장해야 하므로 데이터베이스로서 NoSQL기반의 Cassandra를 사용하는 것이 바람직하다. Cassandra JDBC는 CQL(Cassandra Query Language)문을 통해서 Cassandra에 접근하여 데이터를 저장한다.
- [0038] 관제센터(50)는 복합 이벤트 처리부(51)와, 데이터베이스(52) 및 시각화부(53)를 포함하여 구성된다.
- [0039] 복합 이벤트 처리부(51)는 유량 데이터를 실시간 처리하여 누수를 발견한다. 복합 이벤트 처리(CEP : Complex Event Processing)는 다수의 이벤트를 실시간으로 처리하기 위한 기술로서, 이벤트 간의 복합처리가 가능하다.
- [0040] CEP 엔진은 여러 시스템에서 발생하는 이벤트를 지정하고, 추출하고자 하는 이벤트 패턴을 등록하면, 여러 가지의 이벤트 스트림을 정제, 합산, 수집, 결합한 후 패턴 매칭을 통해 원하는 이벤트 패턴의 감지하는 기능을 제공한다. 즉, CEP 엔진은 같이 여러 이벤트 소스로부터 발생하는 데이터를 수집하여, 그 중에서 필요한 이벤트를 발견하는 역할을 수행한다.

- [0041] CEP 엔진은 EPL(Event processing Language)을 사용한다. EPL은 원하는 이벤트 스트림 중에서 원하는 조건에 충족되는 상황을 추출하는 구문을 정의하는데, SQL처럼 Select/From/Where 절을 기본적으로 사용한다. 또한 CEP 엔진은 디스크를 사용하지 않고, 메모리상에서 이벤트를 처리하기 위해 슬라이딩 윈도우(Sliding Window)를 사용한다. 슬라이딩 윈도우는 일정 시간 또는 일정 개수의 이벤트만을 메모리상에 보관하고, EPL을 사용하여 슬라이딩 윈도우 상에 보관된 이벤트를 대상으로 처리한다.
- [0042] 복합 이벤트 처리부(51)는 이러한 특징을 통해 다수의 수도 계량기로 구성된 블록시스템에 적합하며, 여러 수도 계량기와 블록 유량계를 복합적으로 처리하여 누수를 발견할 수 있다. 또한 복합 이벤트 처리부(51)는 미수집 유량 이벤트가 발생한 경우도 복합처리하여 미수집 유량에 대한 대처를 수행할 수 있는 것이 특징적인 구성이다.
- [0043] 데이터베이스(52)는 NoSQL 기반의 유량 데이터베이스로서, 다수의 수도 계량기에서 발생하는 유량 데이터를 저장한 후, 기계학습을 위한 학습 데이터로 활용된다.
- [0044] 기계학습은 관측된 과거의 데이터로부터 학습을 통하여 예측모델을 구축하고, 이를 바탕으로 발생하는 여러 문제들에 대하여 최적의 의사결정을 이끌어내는 방법론이다. 기계학습 과정은 과거의 저장된 많은 데이터로부터 특징 벡터를 추출하고, 추출된 특징벡터를 토대로 학습 데이터 셋을 만들고, 학습 데이터 셋은 기계학습 알고리즘에 기반하여 예측모델을 생성하게 된다. 이후 새로운 데이터는 예측모델을 토대로 예측결과를 생성하게 된다.
- [0045] 시각화부(53)는 최종 누수결과를 시각화 도구를 통해 지도상에서 누수가 발생한 블록의 위치를 표시한다.
- [0046] 도 6은 본 발명에 따른 유량 데이터 예측 구조를 도시한 것이다.
- [0047] 미수집된 유량 데이터를 예측하기 위해서 기계학습과 CEP 기술이 사용된다. 유량 데이터의 예측과정은 많은 데이터들을 기반으로 기계학습을 통한 예측 모델을 생성하는 일괄 처리과정과 복합 이벤트 처리기술을 이용하여 유량을 예측하는 스트림 처리과정으로 구성된다.
- [0048] 도 7은 기계 학습 과정을 설명하기 위한 구성도이고, 도 8은 데이터베이스에 저장된 초기 데이터 셋을 예시한 것이고, 도 9는 중간 데이터 셋을 예시한 것이고, 도 10은 학습 데이터 셋을 예시한 것이다.
- [0049] 본 발명에서 예측 모델을 생성하기 위해서는 장기간의 데이터 셋을 필요로 하기 때문에, Cassandra에 저장된 데이터 셋을 사용한다.
- [0050] 도 8은 Casandra에 저장된 초기 데이터의 예시로서, DDS 어댑터(40)를 통해 제공되는 유량 데이터가 자동적으로 수집되어 저장된다.
- [0051] 도 9는 중간 데이터 셋을 예시한 것으로서, 배치 집합기(60 : Batch Aggregator)는 초기 데이터 셋에서 CQL 언어를 사용하여 각 가정 및 학습을 위한 연도를 선택하여 데이터를 추출한 후 이를 시계열 데이터 셋으로 변환한다.
- [0052] 하기 문장은 배치 집합기(60)에서 수행하는 CQL 문장의 예시로서 WaterFlow 컬럼 패밀리에 저장된 "SN0011" 수용가의 2013년도에 저장된 모든 유량 데이터를 추출한다.
- ```
select * from WaterFlow where SN="SN0011" AND Year = "2013"
```
- [0053]
- [0054] 추출된 데이터는 중간 데이터 셋(시계열 데이터 셋)으로 변환을 수행하게 된다. 도 9의 시계열 데이터 셋은 15분 주기를 가지도록 변환된 사례가 예시되어 있다.
- [0055] 배치 집합기(60)는 시계열 데이터 셋으로부터 최종적으로 날짜가 가지는 속성을 추출하여 학습 데이터 셋으로 변환한다. 본 실시예에서 추출한 속성으로는 월, 요일, 오전/오후, 최근 유량을 사용하였다. 도 10은 학습 데이터 셋의 예시로서, 시계열 데이터 셋의 날짜 속성들을 사용하여 요일, 오전/오후, 최근 3개의 유량을 속성을 추출하였다.
- [0056] 학습 데이터 셋은 학습 모듈(70)로 전송된다. 학습모듈(70)은 기계 학습도구인 Weka의 입력 데이터로 사용된다. 이때 학습 알고리즘으로는 SVM(Support Vector Machine)를 사용하였다. SVM은 기존 수요량 예측연구에서 우수한 성능을 가진다는 것이 검증되었다. 생성된 모델은 각 수용가별로 파일로 저장되어 스트림 과정에서 모델이 호출

되어 사용된다.

- [0057] 도 11은 CEP 엔진에서의 스트림 처리과정을 도시한 것이다.
- [0058] 스트림 처리 과정은 예측된 모델을 사용하여 유량 데이터가 생성되었을 때, 다음 유량을 예측하는 과정이다. 스트림 처리 과정은 CEP의 Statement를 통해 실시간으로 예측된다.
- [0059] CEP 엔진은 총 3 단계를 거쳐서 유량을 예측하게 된다. 우선, 분할단계로서 유량데이터(Waterflow)가 CEP 엔진으로 들어오면 자동으로 SN 속성을 사용하여 가정별로 분할한다.
- [0060] 그 다음은 집계단계로서, 가정별로 분할된 유량 데이터는 예측주기에 따라 합계가 계산된다. 본 실시예에서의 statement에서는 15분을 예측주기로 설정하였으나, 사용자가 원하는 예측주기에 따라 시간을 설정할 수 있다.
- [0061] 마지막은 예측 단계로서, 예측주기로 설정된 15분간의 유량 합계를 사용하여 CEP 엔진에 등록된  $svm(time, consumption)$  함수의 결과를 반환하여 얻어지는 값이 예측값이 된다. 즉, 예측 statement를 통해 현재 측정된 값이 도착되면 가정별로 예측값이 생성된다.
- [0062] CEP 엔진을 이용한 블록 누수 탐지방법을 보다 상술하면 다음과 같다.
- [0063] 본 발명은 블록 시스템을 대상으로 하기 때문에 블록 시스템에 적합한 누수탐지 패턴에 대해서 정의한 후 정의된 패턴을 CEP 엔진에서 찾아내기 위한 Statement를 정의한다. 정의된 Statement는 CEP 엔진에서 사용하기 위해 등록된다.
- [0064] 배수본관으로부터 들어오는 모든 물의 양은 소블록 유량계를 통해 측정된다. 그리고 소블록 내에 원격검침이 가능한 스마트 수도 계량기들이 설치되어 있어 소블록 내의 가정에서 사용한 물의 양이 측정된다. 소블록 내의 각 수용가 개체 사용량(sq)의 합계는 유량(Q)과 동일 수량이어야 하나 동일값이 나타나지 않을 경우 누수가 제일 큰 요인이므로 차이를 대부분 누수량으로 추정한다. 따라서, 누수량  $L_q = Q - \sum sq$ 의 공식이 성립한다.
- [0065] 실시간으로 추정 누수량을 계산하기 위해 분할, 집계, 조인, 필터링 및 삭제의 총 5단계의 과정이 정의된다.
- [0066] 분할단계는 각 블록 내의 수용가의 합계를 구하기 위해서 블록 유량계에서 수집되는 유량데이터를 분할하는 단계이다. 본 단계에서, CEP 이벤트 타입 객체의 MS 속성(계량기 유형정보)을 통해 데이터가 분할된다.
- [0067] 집계 단계에서는 가정용 계량기에서 분할된 데이터를 블록별로 그룹화한 후 합계가 계산된다. 본 단계에서 CEP 이벤트 타입 객체의 BN 속성(블록 번호)을 통해 그룹화가 이루어진다.
- [0068] 조인단계에서는 분할단계에서 분할되었던 블록 유량과 집계단계에서 계산된 블록별 합계의 차이가 계산된다. 본 단계에서 블록 유량과 블록 유량의 합계에서 BN 속성이 같다면 두 값의 차이를 계산하며, 블록 유량계의 유량 데이터가 수집되는 즉시 조인이 수행된다.
- [0069] 필터링 단계에서는 최종적으로 블록 유량계와 각 블록 유량의 합계 차가 0 이상인 블록을 찾아내고, 필터링 단계를 통과한 이벤트는 리스너를 통해서 결과를 받게 되고 시각화 도구를 통해 출력된다.
- [0070] 삭제단계에서는 불합 단계에 남아 있는 데이터를 삭제하는 단계로서, KeepAll() 함수로 인해 유량 데이터는 윈도우 상에 남아있기 때문에 처리가 완료된 데이터는 삭제해야 한다.
- [0071] 본 발명에서 미 수집 이벤트가 도착하게 되면 가정용 계량기의 유량 데이터가 저장된 윈도우 상에 예측 유량을 반영하여 누수를 탐지하도록 한다.
- [0072] 성능 실험 결과
- [0073] 본 발명자들은 본 발명에 따른 누수탐지 방법을 사용하여 실험 및 성능평가를 하였다. 성능평가는 소블록의 유량계로부터 유량 데이터가 발생하였을 때의 지연시간(Latency)과 예측 정확도를 평가하였다.
- [0074] 학습 데이터는 1년간의 데이터를 사용하여 예측 모델 생성하였으며, 기간은 2013년 1월 ~ 2013년 12월까지 사용하였다. 그리고, 유량 데이터는 4시간의 수집주기를 가지기 때문에, 예측 모델은 4시간 단위로 예측이 가능하다.
- [0075] 유량 데이터는 2014년 1월 1일 20시에 200개의 가구의 유량데이터를 사용하였고, 블록 유량계의 유량 데이터가

없기 때문에, 블록 유량계는 4시간 동안 가정에서 사용한 총량을 블록 유량계로 사용하였다. 그리고, 누수 발생을 위해 누수량은 총 공급량의 10%로 하였다.

[0076] 데모상황으로는 누수 발생 여부와 모든 유량 데이터가 정상적으로 수집되었는지 아닌지를 비교하였다. 누수 발생상황의 누수량은 2013년 평균인 10% 정도를 발생시켰다. 또한 미수집율은 기존 조사결과를 적용하여 15% 정도 발생시켰다.

[0077] 도 12는 미수집 유량이 발생한 경우 예측 유량을 반영한 누수 탐지 성능 평가 결과 그래프로서, 측정결과, 단일 소블록일 경우 99.45%의 이벤트가 22  $\mu$ s의 지연시간을 가진다. 200개의 블록으로 구성하였을 경우에는 99.12%의 이벤트가 12593  $\mu$ s의 지연시간을 가졌다. 본 시스템을 통한 누수탐지는 200개의 블록을 대상으로 12ms의 실시간 처리성능을 보장함을 확인할 수 있다.

[0078] 도 13은 미수집 유량이 발생하지 않은 경우 누수 탐지 성능 평가 결과를 나타낸 그래프로서, 단일 소블록을 대상으로 하였을 경우, 99.13%의 이벤트가 14  $\mu$ s의 지연시간을 보임을 확인하였다. 블록의 개수가 200개가 될 경우 99.43%의 이벤트가 2ms의 성능을 보장함을 보였다. 본 시스템을 통한 누수탐지는 200 개의 블록을 대상으로 2ms의 실시간 처리성능을 보장함을 확인할 수 있다.

[0079] 도 14는 예측 정확도 평가 결과 그래프이다.

[0080] 2012년 1월 ~ 2012년 12월까지의 데이터는 학습 데이터 셋으로 사용하여 모델을 생성하였으며, 2013년 1월 ~ 2013년 12월까지의 데이터를 비교하여 평가하였다. 도 14는 1년 간의 평가 결과에서 일부분을 확대한 모습으로서, 파란색 그래프는 테스트 데이터 셋이고, 빨간색 그래프는 모델의 결과이다. 그래프 출력 결과로는 둘 사이에 유사한 패턴을 가진 것을 확인하였다.

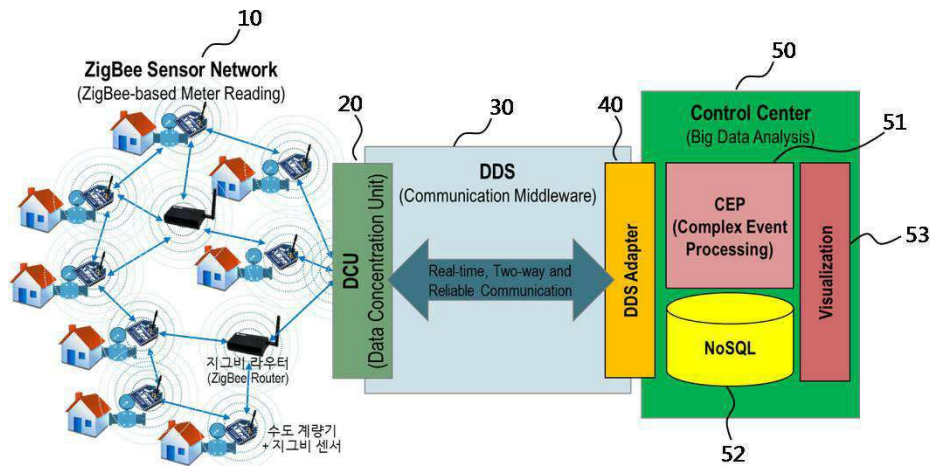
[0081] 상기에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

## 부호의 설명

[0082] 10 : 무선 센서 네트워크 20 : 데이터 집중 유닛  
30 : 데이터 분배 시스템 40 : DDS 어댑터  
50 : 관계데이터 51 : 복합 이벤트 처리부  
52 : 데이터베이스 53 : 시각화부  
60 : 배치 집합기 70 : 학습 모듈

## 도면

### 도면1



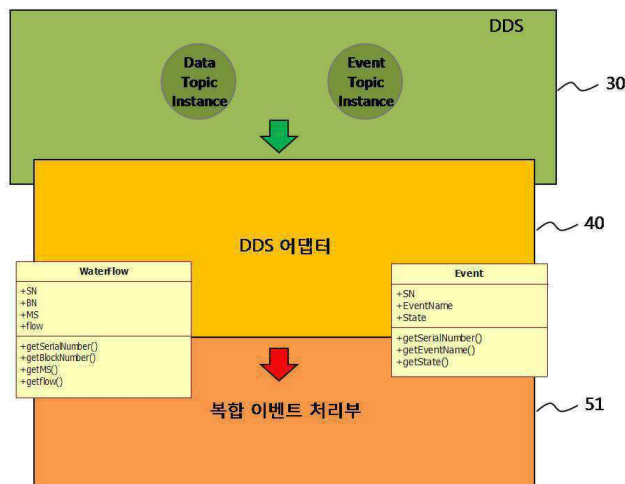
### 도면2

```
struct dataTopic {
    string WaterMeterID;
    string DataName;
    string DataValue;
};
#pragma keylist dataTopic WaterMeterID DataName
```

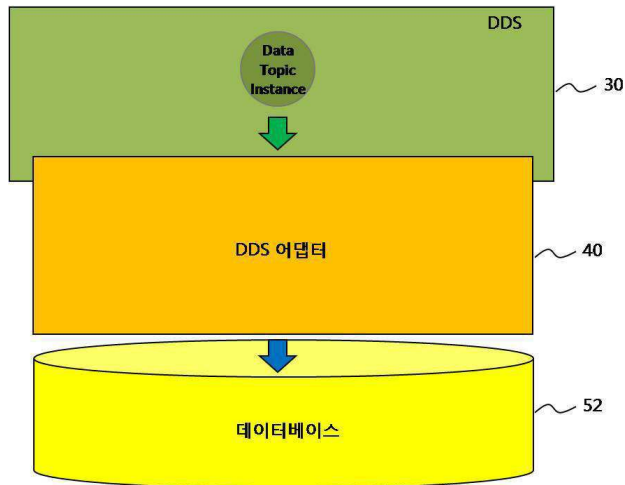
### 도면3

```
struct eventTopic {
    string WaterMeterID;
    string EventName;
    string EventValue;
}
#pragma keylist eventTopic WaterMeterID EventName
```

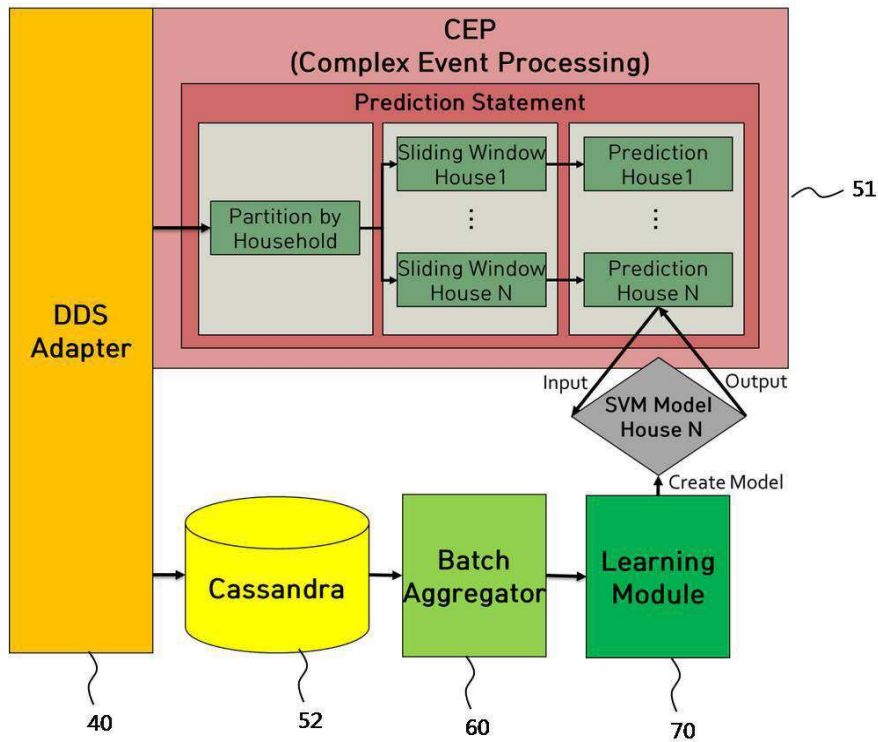
### 도면4



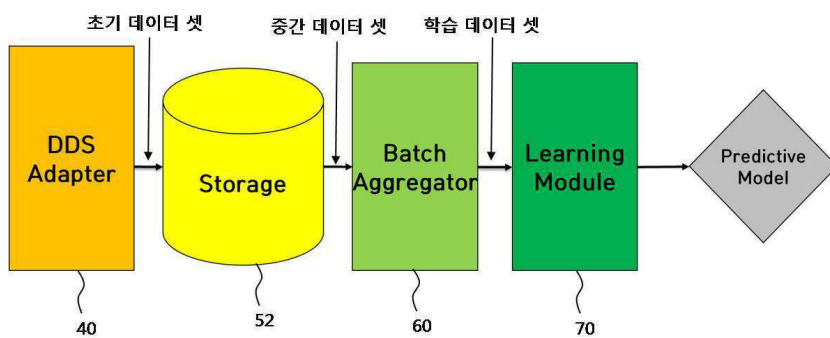
도면5



도면6



도면7



도면8

| SN     | Year | Month | Day | Hour | Min | Sec | flow |
|--------|------|-------|-----|------|-----|-----|------|
| SN0011 | 2013 | 1     | 1   | 0    | 0   | 0   | 1    |
| ...    | ...  | ...   | ... | ...  | ... | ... | ...  |
| SN0011 | 2014 | 11    | 27  | 10   | 20  | 31  | 1    |

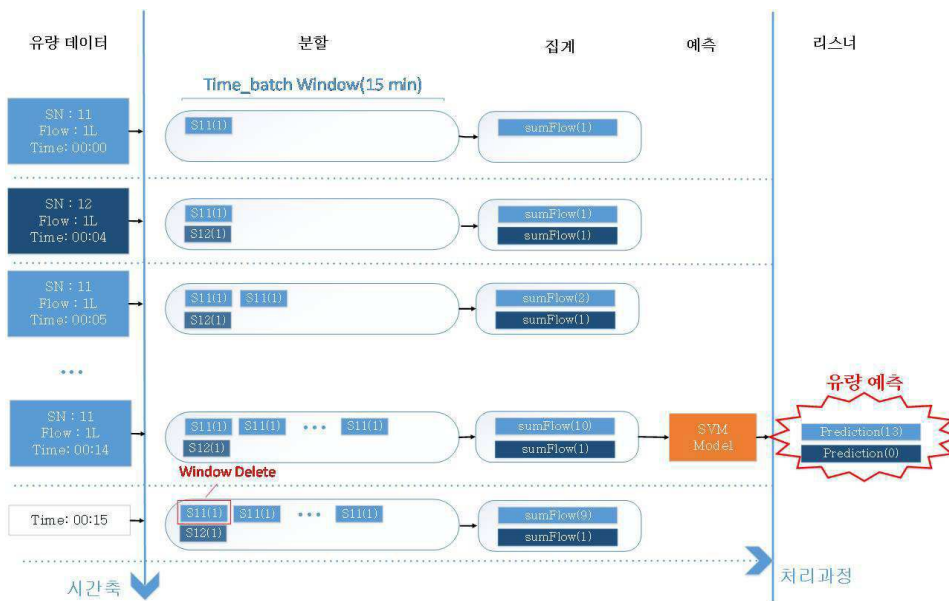
도면9

| SN     | Year | Month | Day | Hour | Min | flow |
|--------|------|-------|-----|------|-----|------|
| SN0011 | 2013 | 11    | 1   | 1    | 0   | 1    |
| SN0011 | 2013 | 11    | 1   | 1    | 15  | 3    |
| SN0011 | 2013 | 11    | 1   | 1    | 30  | 2    |
| SN0011 | 2013 | 11    | 1   | 1    | 45  | 11   |
| SN0011 | 2013 | 11    | 1   | 1    | 16  | 15   |
| ...    | ...  | ...   | ... | ...  | ... | ...  |

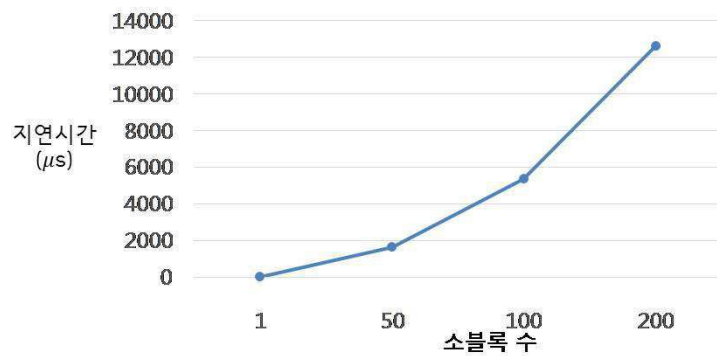
도면10

| SN     | ... | 월   | 요일  | 오전/오후 | Waterflow-1 | ... | Waterflow -3 |
|--------|-----|-----|-----|-------|-------------|-----|--------------|
| SN0011 | ... | 1   | 금   | 오전    | 1           | ... | 12           |
| SN0011 | ... | 1   | 금   | 오후    | 3           | ... | 11           |
| ...    | ... | ... | ... | ...   | ...         | ... | ...          |

도면11

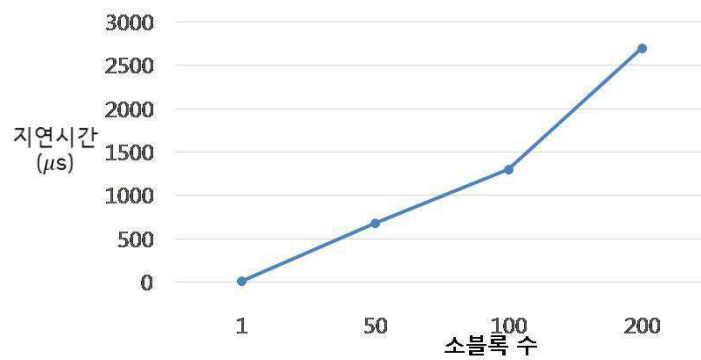


도면12



| 소블록수 | 계량기 수  | 지연시간            |
|------|--------|-----------------|
| 1    | 1500   | 99.45% < 22μs   |
| 50   | 75000  | 99.41 < 1620μs  |
| 100  | 150000 | 99.61 < 5355μs  |
| 200  | 300000 | 99.12 < 12593μs |

도면13



| 소블록수 | 계량기 수  | 지연시간           |
|------|--------|----------------|
| 1    | 1500   | 99.13% < 14μs  |
| 50   | 75000  | 99.56 < 683μs  |
| 100  | 150000 | 99.65 < 1300μs |
| 200  | 300000 | 99.43 < 2758μs |

도면14

