



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0098673
(43) 공개일자 2018년09월04일

- (51) 국제특허분류(Int. Cl.)
G06N 5/04 (2006.01) G06N 99/00 (2010.01)
- (52) CPC특허분류
G06N 5/04 (2013.01)
G06N 99/005 (2013.01)
- (21) 출원번호 10-2018-7023119
- (22) 출원일자(국제) 2017년01월12일
심사청구일자 없음
- (85) 번역문제출일자 2018년08월10일
- (86) 국제출원번호 PCT/US2017/013267
- (87) 국제공개번호 WO 2017/123819
국제공개일자 2017년07월20일
- (30) 우선권주장
14/996,154 2016년01월14일 미국(US)

- (71) 출원인
업테크 테크놀로지스 인코포레이티드
미국 60654 일리노이주 시카고 웨스트 시카고 애비뉴 600 스위트 620
- (72) 발명자
헤르조그 제임스
미국 60654 일리노이주 시카고 웨스트 시카고 애비뉴 600 스위트 620
- (74) 대리인
리엔특허법인

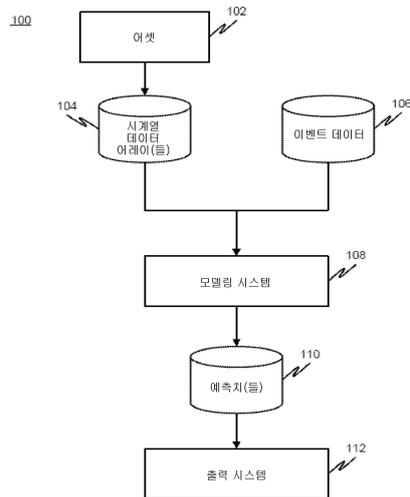
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 로컬화된 시간적 모델 예측

(57) 요약

본원에는 이벤트 데이터로 증첩된 다변량 시계열 데이터 모델링에 관련된 시스템, 컴퓨터 판독가능 매체 및 방법이 개시되어 있다. 특히, 몇몇 예들에는 최근 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 선택하고 이벤트 데이터에 기초하여 상기 유사한 과거 시계열 데이터 어레이를 필터링하는 것이 포함된다. 몇몇 예들에는 또한 상기 필터링된 과거 시계열 데이터 어레이를 사용하여 로컬화된 시간적 예측 모델을 트레이닝하는 것이 포함될 수 있다. 몇몇 예들에는 예측이 필요한 시간에나 또는 그 시간 부근에 로컬화된 시간적 예측 모델을 구축 및/또는 트레이닝하는 것이 포함될 수 있다.

대표도 - 도1



명세서

청구범위

청구항 1

컴퓨팅 시스템에 있어서,

상기 컴퓨팅 시스템은,

적어도 하나의 프로세서;

비-일시적 컴퓨터 판독가능 매체; 및

상기 적어도 하나의 프로세서에 의해 실행 가능한 비-일시적 컴퓨터 판독가능 매체상에 저장된 프로그램 명령어;

를 포함하며,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

- a) 시스템에 연관된 하나 이상의 매개변수의 세트를 나타내는 제1 시계열 데이터 어레이를 획득하게 하고,
- b) 상기 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 식별하게 할 수 있으며 - 상기 하나 이상의 과거 시계열 데이터 어레이는 하나 이상의 매개변수의 세트의 대응하는 과거 값을 나타냄 -,
- c) 상기 제1 시계열 데이터 어레이에 연관된 제1 이벤트 데이터 및 상기 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터를 얻게 하고,
- d) 상기 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터 및 제1 이벤트 데이터의 비교에 기초해 상기 하나 이상의 과거 시계열 데이터 어레이를 필터링하여 필터링된 하나 이상의 과거 시계열 데이터 어레이를 얻게 하며,
- e) 상기 필터링된 하나 이상의 과거 시계열 데이터 어레이를 사용하여 시계열 예측 모델을 트레이닝시키게 하고, 그리고
- f) 상기 트레이닝된 시계열 예측 모델을 사용하여 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 미래 값의 예측치를 생성하게 하도록 상기 적어도 하나의 프로세서에 의해 실행 가능한, 컴퓨팅 시스템.

청구항 2

제1항에 있어서,

a) - f)는 예측치 생성 요구의 수신 또는 기간 만료 중 하나에 기초하여 반복되는, 컴퓨팅 시스템

청구항 3

제1항에 있어서,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

시간적 변화를 검출하게 하고,

상기 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하며, 그리고

상기 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 하도록 상기 적어도 하나의 프로세서에 의해 부가적으로 실행 가능한, 컴퓨팅 시스템.

청구항 4

제1항에 있어서,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 변화를 검출하게 하고,

상기 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하며, 그리고

상기 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 하도록 상기 적어도 하나의 프로세서에 의해 부가적으로 실행 가능한, 컴퓨팅 시스템.

청구항 5

제1항에 있어서,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

예측치를 다른 시스템으로 전송하게 하도록 상기 적어도 하나의 프로세서에 의해 부가적으로 실행 가능한, 컴퓨팅 시스템.

청구항 6

제1항에 있어서,

상기 제1 이벤트 데이터는 상기 제1 시계열 데이터 어레이의 시간에 이루어지고 상기 제1 시계열 데이터 어레이의 시간에 후속하는 시스템의 작동에 영향을 줄 것으로 예상되는 이벤트를 나타내는, 컴퓨팅 시스템.

청구항 7

제1항에 있어서,

상기 필터링된 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 대응하는 과거 이벤트 데이터는 상기 제1 이벤트와 유사한 이벤트를 나타내는, 컴퓨팅 시스템.

청구항 8

프로그램 명령어가 저장된 비-일시적 컴퓨터 판독가능 매체에 있어서,

상기 프로그램 명령어는, 컴퓨팅 시스템으로 하여금,

a) 시스템에 연관된 하나 이상의 매개변수의 세트를 나타내는 제1 시계열 데이터 어레이를 얻게 하고,

b) 상기 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 식별하게 하며 - 상기 하나 이상의 과거 시계열 데이터 어레이는 하나 이상의 매개변수의 세트의 대응하는 과거 값들을 나타냄 -,

c) 상기 제1 시계열 데이터 어레이에 연관된 제1 이벤트 데이터 및 상기 하나 이상의 과거 시계열 데이터 어레이 각각에 관련된 과거 이벤트 데이터를 얻게 하고,

d) 상기 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터 및 상기 제1 이벤트 데이터의 비교에 기초해 상기 하나 이상의 과거 시계열 데이터 어레이를 필터링하여 필터링된 하나 이상의 과거 시계열 데이터 어레이를 얻게 하며,

e) 상기 필터링된 하나 이상의 과거 시계열 데이터 어레이를 사용하여 시계열 예측 모델을 트레이닝시키고, 그리고

f) 상기 트레이닝된 시계열 예측 모델을 사용하여 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 미래 값의 예측치를 생성하게 하도록 상기 컴퓨팅 시스템의 적어도 하나의 프로세서에 의해 실행 가능한, 비-일시적 컴퓨터 판독가능 매체.

청구항 9

제8항에 있어서,

a) - f)는 예측치 생성 요구의 수신 또는 기간 만료 중 하나에 기초하여 반복되는, 비-일시적 컴퓨터 판독가능 매체.

청구항 10

제8항에 있어서,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

시간적 변화를 검출하게 하고,

상기 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하며, 그리고

상기 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 하도록 상기 적어도 하나의 프로세서에 의해 부가적으로 실행 가능한, 비-일시적 컴퓨터 판독가능 매체.

청구항 11

제8항에 있어서,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 변화를 검출하게 하고,

상기 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하며, 그리고

상기 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 하도록 상기 적어도 하나의 프로세서에 의해 부가적으로 실행 가능한, 비-일시적 컴퓨터 판독가능 매체.

청구항 12

제8항에 있어서,

상기 프로그램 명령어는, 상기 컴퓨팅 시스템으로 하여금,

예측치를 다른 시스템으로 전송하게 하도록 상기 적어도 하나의 프로세서에 의해 부가적으로 실행 가능한, 비-일시적 컴퓨터 판독가능 매체.

청구항 13

제8항에 있어서,

상기 제1 이벤트 데이터는 상기 제1 시계열 데이터 어레이의 시간에 이루어지고 상기 제1 시계열 데이터 어레이의 시간에 후속하는 시스템의 작동에 영향을 줄 것으로 예상되는 이벤트를 나타내는, 비-일시적 컴퓨터 판독가능 매체.

청구항 14

제8항에 있어서,

상기 필터링된 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 대응하는 과거 이벤트 데이터는 상기 제1 이벤트와 유사한 이벤트를 나타내는, 비-일시적 컴퓨터 판독가능 매체.

청구항 15

컴퓨터 구현 방법에 있어서,

상기 컴퓨터 구현 방법은,

a) 시스템에 연관된 하나 이상의 매개변수의 세트를 나타내는 제1 시계열 데이터 어레이를 얻는 단계;

b) 상기 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터를 식별하는 단계 - 상기 하나 이

상의 과거 시계열 데이터 어레이는 하나 이상의 매개변수의 세트의 대응하는 과거 값들을 나타냄 -;

c) 상기 제1 시계열 데이터 어레이에 연관된 제1 이벤트 데이터 및 상기 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터를 얻는 단계;

d) 상기 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터 및 상기 제1 이벤트 데이터의 비교에 기초해 상기 하나 이상의 과거 시계열 데이터 어레이를 필터링하여 필터링된 하나 이상의 과거 시계열 데이터 어레이를 얻는 단계;

e) 상기 필터링된 하나 이상의 과거 시계열 데이터 어레이를 사용하여 시계열 예측 모델을 트레이닝하는 단계; 및

f) 상기 트레이닝된 시계열 예측 모델을 사용하여 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 미래 값의 예측치를 생성하는 단계;

를 포함하는, 컴퓨터 구현 방법.

청구항 16

제15항에 있어서,

상기 컴퓨터 구현 방법은,

시간적 변화를 검출하는 단계;

상기 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻는 단계; 및

상기 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하는 단계;

를 추가적으로 포함하는, 컴퓨터 구현 방법.

청구항 17

제15항에 있어서,

상기 컴퓨터 구현 방법은,

상기 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 변화를 검출하는 단계;

상기 검출된 변화에 기초해, a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻는 단계; 및

상기 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하는 단계;

를 추가적으로 포함하는, 컴퓨터 구현 방법.

청구항 18

제15항에 있어서,

상기 컴퓨터 구현 방법은,

예측치 생성 요구의 수신 또는 기간 만료 중 하나에 기초하여 a) - f)를 반복하는 단계;

를 추가적으로 포함하는, 컴퓨터 구현 방법.

청구항 19

제15항에 있어서,

상기 제1 이벤트 데이터는 상기 제1 시계열 데이터 어레이의 시간에서 이루어지고 상기 제1 시계열 데이터 어레이의 시간에 후속하는 시스템의 작동에 영향을 줄 것으로 예상되는 이벤트를 나타내는, 컴퓨터 구현 방법.

청구항 20

제15항에 있어서,

상기 필터링된 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 대응하는 과거 이벤트 데이터는 상기 제1 이벤트와 유사한 이벤트를 나타내는, 컴퓨터 구현 방법.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본원은 발명의 명칭이 "로컬화된 시간적 모델 예측(Localized Temporal Model Forecasting)"이고 2016년 1월 14일자로 출원된 미국 정식 특허출원 제14/996,154호에 대한 우선권을 주장하며, 상기 특허출원의 전체내용이 인용에 의해 본원에 보완된다.

배경 기술

[0003] 시간적 예측은 모든(또는 대부분의) 이용 가능한 데이터를 사용하여 트레이닝된 예측 모델들을 활용하는 것이 전형적이다. 모든 이용 가능한 데이터를 사용하여 예측하기 위해 트레이닝된 모델들은 주어진 최근 데이터 패턴에 대한 예측들을 정확하게 생성하지 못할 수 있다. 또한, 일부 시간적 예측 모델들은 이벤트 데이터(예컨대, 간헐적, 비동기적 그리고/또는 일시적 이벤트를 나타내는 데이터)를 고려하지 못할 수 있고, 결과적으로는 이벤트가 포함되어 있을 때 단기간 시스템 매개변수 값 및/또는 응답을 정확하게 예측하지 못할 수 있다.

발명의 내용

[0004] 일반적으로, 일부 구현 예들에는 로컬화된 시간적 예측 모델을 트레이닝하기 위해 이벤트 데이터로 중첩된 다변량 시계열 데이터를 사용하는 모델 예측에 관련된 시스템, 컴퓨터 판독가능 매체 및 방법이 포함될 수 있다. 특히, 일부 예들에는 최근 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 선택하고 이벤트 데이터에 기초하여 상기 유사한 과거 시계열 데이터 어레이를 필터링하는 것이 포함된다. 일부 예들에는 또한 상기 필터링된 과거 시계열 데이터 어레이를 사용하여 로컬화된 시간적 예측 모델을 트레이닝하는 것이 포함될 수 있다. 일부 예들에는 예측이 필요한 시간이나 또는 그 시간 부근에(예를 들어, 그리디 학습(greedy learning) 기법을 통해) 상기 로컬화된 시간적 예측 모델을 구축 및/또는 트레이닝하는 것이 포함될 수 있다.

[0005] 위에서 검토한 바와 같이, 본원 명세서에서 제공되는 일부 예들은 모델링에 관련된 것이다. 일 실시형태에서, 컴퓨팅 시스템이 제공된다. 컴퓨팅 시스템은 적어도 하나의 프로세서, 비-일시적 컴퓨터 판독가능 매체, 및 적어도 하나의 프로세서에 의해 실행 가능한 비-일시적 컴퓨터 판독가능 매체상에 저장된 프로그램 명령어를 포함한다. 실행될 때, 프로그램 명령어는 컴퓨팅 시스템으로 하여금 a) 시스템에 연관된 하나 이상의 매개변수의 세트를 나타내는 제1 시계열 데이터 어레이를 획득하게 하고, 그리고 b) 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 식별하게 할 수 있으며, 상기 하나 이상의 과거 시계열 데이터 어레이는 하나 이상의 매개변수의 세트의 대응하는 과거 값을 나타낸다. 실행될 때, 프로그램 명령어는 또한 컴퓨팅 시스템으로 하여금 c) 제1 시계열 데이터 어레이에 연관된 제1 이벤트 데이터 및 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터를 얻게 하고, 그리고 d) 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터 및 제1 이벤트 데이터의 비교에 기초해 하나 이상의 과거 시계열 데이터 어레이를 필터링하여 필터링된 하나 이상의 과거 시계열 데이터 어레이를 얻게 할 수 있다.

[0006] 실행될 때, 프로그램 명령어는 부가적으로 상기 컴퓨팅 시스템으로 하여금 e) 필터링된 하나 이상의 과거 시계열 데이터 어레이를 사용하여 시계열 예측 모델을 트레이닝시키게 하고, 그리고 f) 트레이닝된 시계열 예측 모델을 사용하여 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 미래 값의 예측치를 생성하게 할 수 있다. 일부 예들에서, a) - f)는 예측치 생성 요구의 수신 또는 기간 만료 중 하나에 기초하여 반복될 수 있다. 프로그램 명령어는, 컴퓨팅 시스템으로 하여금 시간적 변화를 검출하게 하고, 그리고 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하도록 적어도 하나의 프로세서에 의해 부가적으로 실행될 수 있다. 실행될 때, 프로그램 명령어는 또한 컴퓨팅 시스템으로 하여금 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 할 수 있다.

[0007] 프로그램 명령어는 컴퓨팅 시스템으로 하여금 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 변화를

검출하게 하고, 그리고 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하도록 적어도 하나의 프로세서에 의해 부가적으로 실행될 수 있다. 실행될 때, 프로그램 명령어는 또한, 컴퓨팅 시스템으로 하여금 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 할 수 있다.

[0008] 프로그램 명령어는 컴퓨팅 시스템으로 하여금 예측치를 다른 시스템으로 전송하게 하도록 적어도 하나의 프로세서에 의해 부가적으로 실행될 수 있다. 일부 예들에서, 제1 이벤트 데이터는 제1 시계열 데이터 어레이의 시간에 이루어지고 제1 시계열 데이터 어레이의 시간에 후속하는 시스템의 작동에 영향을 줄 것으로 예상되는 이벤트를 나타낸다. 필터링된 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 대응하는 과거 이벤트 데이터는 제1 이벤트와 유사한 이벤트를 나타낼 수 있다.

[0009] 또 다른 실시형태에서, 본원 명세서에는, 컴퓨팅 시스템으로 하여금 a) 시스템에 연관된 하나 이상의 매개변수의 세트를 나타내는 제1 시계열 데이터 어레이를 얻게 하고, 그리고 b) 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 식별하게 하도록 상기 컴퓨팅 시스템의 적어도 하나의 프로세서에 의해 실행 가능한 프로그램 명령어가 저장된 비-일시적 컴퓨터 판독가능 매체가 개시되어 있으며, 하나 이상의 과거 시계열 데이터 어레이는 하나 이상의 매개변수의 세트의 대응하는 과거 값들을 나타낸다.

[0010] 실행될 때, 프로그램 명령어는 또한, 컴퓨팅 시스템으로 하여금 c) 제1 시계열 데이터 어레이에 연관된 제1 이벤트 데이터 및 하나 이상의 과거 시계열 데이터 어레이 각각에 관련된 과거 이벤트 데이터를 얻게 하고, 그리고 d) 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터 및 제1 이벤트 데이터의 비교에 기초해 상기 하나 이상의 과거 시계열 데이터 어레이를 필터링하여 필터링된 하나 이상의 과거 시계열 데이터 어레이를 얻게 할 수 있다.

[0011] 실행될 때, 프로그램 명령어는 부가적으로 컴퓨팅 시스템으로 하여금 e) 필터링된 하나 이상의 과거 시계열 데이터 어레이를 사용하여 시계열 예측 모델을 트레이닝시키게 하고, 그리고 f) 트레이닝된 시계열 예측 모델을 사용하여 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 미래 값의 예측치를 생성하게 할 수 있다. 일부 예에서, a) - f)는 예측치 생성 요구의 수신 또는 기간 만료 중 하나에 기초하여 반복될 수 있다. 프로그램 명령어는 부가적으로 컴퓨팅 시스템으로 하여금 시간적 변화를 검출하게 하고, 그리고 검출된 변화에 기초하여 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하도록 적어도 하나의 프로세서에 의해 실행될 수 있다. 실행될 때, 프로그램 명령어는 또한, 컴퓨팅 시스템으로 하여금 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 할 수 있다.

[0012] 프로그램 명령어는 부가적으로 컴퓨팅 시스템으로 하여금 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 변화를 검출하게 하고, 그리고 검출된 변화에 기초해 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻게 하도록 적어도 하나의 프로세서에 의해 실행될 수 있다. 실행될 때, 프로그램 명령어는 또한, 컴퓨팅 시스템으로 하여금 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하게 할 수 있다.

[0013] 프로그램 명령어는 부가적으로 컴퓨팅 시스템으로 하여금 예측치를 다른 시스템으로 전송하게 하도록 적어도 하나의 프로세서에 의해 실행될 수 있다. 일부 예들에서, 제1 이벤트 데이터는 제1 시계열 데이터 어레이의 시간에 이루어지고 제1 시계열 데이터 어레이의 시간에 후속하는 시스템의 작동에 영향을 줄 것으로 예상되는 이벤트를 나타낸다. 필터링된 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 대응하는 과거 이벤트 데이터는 제1 이벤트와 유사한 이벤트를 나타낼 수 있다.

[0014] 또 다른 실시형태에서, 본원 명세서에는 a) 시스템에 연관된 하나 이상의 매개변수의 세트를 나타내는 제1 시계열 데이터 어레이를 얻는 단계, 및 b) 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터를 식별하는 단계 - 하나 이상의 과거 시계열 데이터 어레이는 하나 이상의 매개변수의 세트의 대응하는 과거 값들을 나타냄 - 를 포함하는 컴퓨터 구현 방법이 개시되어 있다.

[0015] 상기 방법은 또한, c) 상기 제1 시계열 데이터 어레이에 연관된 제1 이벤트 데이터 및 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터를 얻는 단계, 및 d) 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 과거 이벤트 데이터 및 제1 이벤트 데이터의 비교에 기초해 하나 이상의 과거 시계열 데이터 어레이를 필터링하여 필터링된 하나 이상의 과거 시계열 데이터 어레이를 얻는 단계를 포함할 수 있다. 상기 방법은 e) 필터링된 하나 이상의 과거 시계열 데이터 어레이를 사용하여 시계열 예측 모델을 트레이닝하는 단계, 및

f) 트레이닝된 시계열 예측 모델을 사용하여 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 미래 값의 예측치를 생성하는 단계를 추가적으로 포함할 수 있다.

[0016] 일부 예들에서, 상기 방법은 또한, 시간적 변화를 검출하는 단계, 및 검출된 변화에 기초하여 a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻는 단계를 포함할 수 있다. 상기 방법은 업데이트된 제1 시계열 데이터 어레이에 기초하여 b) - f)를 반복하는 단계를 추가적으로 포함할 수 있다.

[0017] 일부 예들에서, 본 방법은 또한, 하나 이상의 매개변수의 세트의 적어도 하나의 매개변수의 변화를 검출하는 단계, 및 검출된 변화에 기초해, a)를 반복하여 이전의 제1 시계열 데이터 어레이의 시간에 후속하는 시간에 하나 이상의 매개변수의 세트를 나타내는 업데이트된 제1 시계열 데이터 어레이를 얻는 단계를 포함할 수 있다.

[0018] 상기 방법은 또한, 예측치 생성 요구의 수신 또는 기간 만료 중 하나에 기초하여 a) - f)를 반복하는 단계를 포함할 수 있다. 제1 이벤트 데이터는 제1 시계열 데이터 어레이의 시간에서 이루어지고 제1 시계열 데이터 어레이의 시간에 후속하는 시스템의 작동에 영향을 줄 것으로 예상되는 이벤트를 나타낼 수 있다. 필터링된 하나 이상의 과거 시계열 데이터 어레이 각각에 연관된 대응하는 과거 이벤트 데이터는 제1 이벤트와 유사한 이벤트를 나타낼 수 있다.

[0019] 통상의 기술자는 이하의 개시 내용을 읽음으로써 이들뿐만 아니라 다수의 다른 실시형태를 이해할 것이다.

도면의 간단한 설명

[0020] 도 1은 적어도 하나의 구현 예에 따른 대표적인 로컬화된 시간적 모델링 환경의 다이어그램이다.
 도 2는 적어도 하나의 구현 예에 따른 로컬화된 시간적 모델링에 기초하여 예측을 생성하도록 구성된 대표적인 컴퓨팅 시스템의 다이어그램이다.
 도 3은 적어도 하나의 구현 예에 따른 로컬화된 시간적 모델링을 위한 대표적인 방법의 플로차트이다.
 도 4는 적어도 하나의 구현 예에 따른 그리디 학습(greedy learning) 기법의 세부사항을 보여주는 로컬화된 시간적 모델링을 위한 대표적인 방법의 플로차트이다.

발명을 실시하기 위한 구체적인 내용

[0021] 도 1은 적어도 하나의 구현 예에 따른 대표적인 모델링 환경(100)의 다이어그램이다. 상기 환경(100)은 하나 이상의 어셋(asset)(102), 모델링 시스템(108), 및 출력 시스템(112)을 포함한다. 대표적인 어셋은 다른 예들 중에서 운송 기계(예컨대, 기관차, 항공기, 승용차, 세미-트레일러 트럭, 선박 등), 산업 기계(예컨대, 광산 장비, 건설 장비, 공장 자동화 등), 의학 기계(예컨대, 의료 이미징 장비, 수술 장비, 의료 모니터링 시스템, 의료 실험 장비 등), 유틸리티 기계(예컨대, 터빈, 솔라 팜(solar farm) 등)를 포함할 수 있다. 또한, 어셋에는 의료 시설 또는 직원을 고용하는 다른 타입의 시설이 포함될 수 있다. 대표적인 구현 예에서, 상기 어셋들은 각각 동일한 타입(예컨대, 다른 예들 중에서 기관차 또는 항공기 집합체, 풍력 터빈 그룹 또는 MRI 기계 세트)을 지닐 수 있으며, 아마도 동일한 클래스(예컨대, 동일한 브랜드 및/또는 모델)를 지닐 수 있다. 다른 예들에서, 어셋은 타입, 브랜드, 모델 등에 따라 다를 수 있다. 통상의 기술자라면 이들은 단지 어셋의 몇 가지 예일 뿐이며 많은 다른 것들이 본 명세서에서 가능하고 고려될 수 있음을 이해할 것이다.

[0022] 동작시, 어셋(102)은 모델링 시스템(108)에 제공될 수 있는 시계열 데이터(104)를 생성한다. 시계열 데이터(104)는 최근 또는 과거 시점에서나 또는 최근 또는 과거 시간 범위에 걸쳐 어셋의 시스템 또는 시스템들에 연관된 하나 이상의 변수, 신호, 매개변수 및/또는 공변수(covariant)를 나타내는 데이터를 포함할 수 있다. 시계열 데이터(104)로부터, 모델링 시스템(108)은 제1 시계열 데이터 어레이(예컨대, 최근 시간으로부터의 시계열 데이터 어레이) 및/또는 하나 이상의 과거 시계열 데이터 어레이를 추출할 수 있다. 제1 시계열 데이터 어레이는 최근 기간(또는 윈도우)으로부터의 시스템의 하나 이상의 매개변수의 세트를 나타내는 시계열 데이터(예컨대, 1분 전부터의 주어진 데이터 샘플 수 등)를 포함할 수 있다. 시계열 데이터(104)는 모델링 시스템(108)에 의해 저장될 수 있다.

[0023] 시계열 데이터(104)에 추가하여, 이벤트 데이터(106)는 모델링 시스템(108)에 제공될 수 있다. 이벤트 데이터(106)는 어셋(102) 또는 다른 시스템에 의해 제공될 수 있다. 이벤트 데이터(106)는 (예컨대, 제1 시계열 데이터 어레이의 시간으로부터의 이벤트를 나타내는) 제1 이벤트 데이터 및 과거 이벤트 데이터를 포함할 수 있다. 이벤트 데이터(106)는 하나 이상의 시간 의존 이산 간헐 이벤트, 비동기 이벤트, 랜덤 이벤트 및/또는 반복 이

벤트(예컨대, 시간, 요일, 일자, 휴일 등)에 대응하는 데이터를 포함할 수 있다.

[0024] 일부 예들에서, 모델을 트레이닝하기 전에, 예측 모델링 기법은 시스템의 어떤 매개변수(또는 신호)가 모델링에 사용되어야 하는지와 더불어 선택될 수 있다. 이전에 관측된 값에 기초하여 미래 값을 예측하기 위해 모델을 이용하는 모든 예측 기법은 개시된 주제로부터 이익을 얻을 수 있다. 일반적으로, 미래의 결과를 예측하기 위해 과거 데이터를 이용하는 타입의 모델링 기술은 회귀 모델링 기법이다. 이 기술은 커널 회귀, 유사도 기반 모델링, 가우스 프로세스, 반복 신경 네트워크, 지원 벡터 회귀, 및 ARMA(Autoregressive Moving Average), ARIMA(Autoregressive Integrated Moving Average) 및 VAR(Vector Autoregression)과 같은 자동 회귀 기법들을 포함하지만, 이들에 국한되지 않는다. 일부 예들에서, 모델링에 사용된 시간 윈도우의 크기(예컨대, 샘플의 수)가 선택된다. 일부 예들에서, 미래에 예측치가 생성될 시간이 선택된다. 모델링 시스템(108)은 이벤트 데이터(106)와 함께 시계열 데이터(104)를 사용하여 예측 모델을 트레이닝시킬 수 있다. 예를 들면, 모델링 시스템(108)은 제1 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이를 식별할 수 있다. 식별된 유사한 과거 시계열 데이터 어레이는 이벤트 데이터에 기초하여 필터링될 수 있다. 예를 들면, 필터링된 과거 시계열 데이터 어레이는 제1 이벤트와 유사한 연관된 과거 이벤트를 갖는 그러한 과거 시계열 데이터 어레이만을 포함할 수 있다. 필터링된 과거 시계열 데이터 어레이는 예측 모델을 트레이닝하는데 사용될 수 있다. 일부 구현 예들은 이벤트 데이터를 사용하여 필터링된 최근 시계열 데이터와 유사한 과거 이력 시계열 데이터로 트레이닝된 모델에 기초하여 더 정확한 예측을 생성하는 이점을 제공할 수 있다.

[0025] 모델링 시스템(108)은 트레이닝된 예측 모델을 사용하여 예측치(110)를 생성할 수 있다. 예측치(110)는 모델링되는 시스템의 하나 이상의 매개변수의 미래 값을 나타낼 수 있고, 미래의 시점에서나 또는 미래의 시간 범위에 걸쳐 모델링되는 시스템의 하나 이상의 변수, 신호, 매개변수의 하나 이상의 값을 포함할 수 있다. 예측의 변수, 신호 및/또는 매개변수는 시계열 데이터의 변수, 신호, 매개변수 및/또는 공변수와 동일할 수도 있고 다를 수도 있다.

[0026] 예측치(110)는 출력 시스템(112)에 제공될 수 있다. 출력 시스템(112)은 어셋(102)과 동일하거나 다를 수 있다. 또한, 어셋(102), 모델링 시스템(108) 및 출력 시스템(112)은 동일한 시스템 또는 하나 이상의 다른 시스템에 포함될 수 있다. 출력 시스템(112)은 예측치(110)에 기초하여 출력을 생성할 수 있다.

[0027] 예를 들면, 모델은 디젤 기관차와 같은 원격 운송 장치상에서 중요한 시스템의 향후 작동을 예측하도록 구성될 수 있다. 이러한 중요한 시스템은 디젤 엔진, 엔진 냉각수, 트랙션 모터 및 AC 발전기 시스템을 포함할 수 있지만 이들에 국한되지는 않는다. 시계열 데이터는 원격 측정 장치를 통해 일정한 간격으로 전송되는 기관차 상의 센서로부터의 데이터를 포함할 수 있다. 이벤트 데이터는 기관차 상의 제어 시스템에 의해 생성되고 동일하거나 상이한 원격 측정 장치에 의해 전송되는 이벤트 코드로 이루어질 수 있다. 이벤트 코드는 제어 시스템에 의해 검출된 비정상 조건의 특정 세트를 식별하는 고유한 정수로 이루어질 수 있다. 이벤트 코드는 비정상 조건의 검출과 동시에 수집된 센서 측정과 같은 추가 데이터를 포함할 수 있다. 모델링 시스템(108)은 기관차 상의 적어도 하나의 중요한 시스템의 향후 작동의 예측을 생성할 수 있다. 시계열 데이터로 이벤트 데이터를 중첩함으로써, 모델링 시스템(108)은 시계열 데이터만을 사용하여 생성된 예측보다 더 정확할 수 있는 중요 시스템의 예측치를 생성할 수 있다.

[0028] 다른 예에서, 모델은 풍력 터빈의 전력 출력을 예측하도록 구성될 수 있다. 시계열 데이터는 풍력 터빈의 매개변수를 포함할 수 있다. 상기 이벤트 데이터는 기상(meteorological) 예측 데이터를 포함할 수 있다. 모델링 시스템(108)은 풍력 터빈의 향후 전력 출력의 예측치를 생성할 수 있다. 시계열 데이터로 이벤트 데이터를 중첩함으로써, 모델링 시스템(108)은 종래 기법을 사용하여 생성된 예측보다 더 정확할 수 있는 전체 전력 출력의 예측치를 생성할 수 있다. 전체 전력 출력의 예측치는 에너지 시장에서 터빈에 의해 생성된 전력을 거래하는 것을 돕기 위해 에너지 트레이더(energy trader)에 의해 사용될 수 있다. 예를 들어, 더 정확한 예측치는 에너지 트레이더가 터빈에 의해 생성된 전력에 대해 더 높은 가격을 얻는 것을 허용할 수 있다.

[0029] 다른 예에서, 어셋은 의료 시설일 수 있다. 모델링 시스템(108)은 의료 시설의 예상 인력 수요를 예측하도록 구성될 수 있다. 예상 인력 수요는 의료 시설에서 일할 인력을 스케줄링하는 데 사용될 수 있다.

[0030] 도 2는 적어도 하나의 구현 예에 따라 모델링(예컨대, 로컬화된 시간적 모델 예측)을 위해 구성된 대표적인 컴퓨팅 시스템(200)의 다이어그램이다. 특히, 대표적인 컴퓨팅 시스템(200)은 도 1에 도시된 모델링 시스템(108)을 포함할 수 있다. 컴퓨팅 시스템(200)은 프로세서(202), 운영 시스템(204), 컴퓨터 판독가능 매체(206) 및 I/O 인터페이스(208)를 포함한다. 컴퓨터 판독가능 매체(206)는 모델링 애플리케이션(210)(예컨대, 로컬화된 시간적 모델 예측 애플리케이션) 및 데이터베이스(212)(예컨대, 시계열 데이터, 이벤트 데이터, 예측치 등을 저장

하도록 구성됨)을 포함할 수 있다.

- [0031] 동작시, 프로세서(202)는 컴퓨터 판독가능 매체(206)에 저장된 애플리케이션(210)을 실행할 수 있다. 애플리케이션(210)은 프로세서(202)에 의해 실행될 때 프로세서(202)로 하여금 본 개시내용에 따른 코컬화된 시간적 모델 예측 동작들(예컨대, 이하에서 설명되는 단계들(302-314 및/또는 402-406) 중 하나 이상의 단계들을 수행하게 하는 소프트웨어 명령어를 포함할 수 있다. 애플리케이션 프로그램(210)은 데이터베이스(212) 및 운영 시스템(204)과 함께 동작할 수 있다. 애플리케이션(210)은 하나 이상의 모델을 포함할 수 있다.
- [0032] 모델링 시스템(예컨대, 108)은 단일 프로세서 시스템, 멀티-프로세서 시스템(동일 위치 또는 분산형), 클라우드 컴퓨팅 시스템, 또는 이들의 조합을 포함할 수 있지만, 이들에 국한되지 않는다. 일부 구현 예들에서, 모델링 시스템(예컨대, 108)은 어셋(예컨대, 102) 및/또는 출력 시스템(예컨대, 112)에 포함될 수 있다.
- [0033] 모델링 시스템(예컨대, 108)은 서버, 데스크톱 컴퓨터, 랩톱 컴퓨터, 휴대용 컴퓨터, 태블릿 컴퓨팅 장치, 또는 차량 등의 정보 시스템을 포함할 수 있지만, 이들에 국한되는 것은 아니다. 일부 구현 예들에서, 모델링 시스템(108) 및 어셋(102) 및/또는 출력 시스템(112)은 동일한 시스템일 수 있다.
- [0034] 일부 예들에서, 모델링 시스템(예컨대, 108)은 네트워크를 통해 어셋(102) 및/또는 출력 시스템(112)에 접속될 수 있다. 네트워크는 유선 또는 무선 네트워크 일 수 있으며, WiFi 네트워크, 근거리 통신 네트워크, 광역 네트워크, 인터넷 또는 상기의 조합을 포함할 수 있지만, 이들에 국한되지 않는다.
- [0035] 모델링 시스템의 스토리지, 메모리 및/또는 컴퓨터 판독가능 매체는 다른 예들 중에서 광학, 자기, 유기 또는 플래시 메모리와 같은 하나 이상의 비-일시적 컴퓨터 판독가능 저장 매체를 포함할 수 있다. 소프트웨어 명령어들은 또한 예를 들면 서버(예컨대, 분산 시스템 및/또는 클라우드 컴퓨팅 시스템)로부터 전달되는 소프트웨어(SaaS)의 형태를 이루는 전자 신호에 포함되어 제공될 수 있다.
- [0036] 도 3은 적어도 하나의 구현 예에 따라 모델링하기 위한 대표적인 방법(300)의 플로차트이다. 프로세싱은 302에서 시작되며, 여기서 모델링 시스템(예컨대, 108)은 최근 시간에서 시스템의 하나 이상의 매개변수의 값을 나타내는 최근 시계열 데이터 어레이를 획득한다. 최근 시계열 데이터 어레이는 최근 과거의 어느 시간(또는 시간 범위)으로부터 이루어진 것일 수 있다. 예를 들면, 몇 시간의 시간 척도(time scale)로 동작하는 시스템의 경우, 최근 시계열 데이터 어레이는 과거 1 시간 또는 그 이상의 시간 내의 시간으로부터 이루어진 것일 수 있다. 다른 예에서, 일부 시스템들은 더 작은 시간 단위 척도(예컨대, 초 또는 초 미만)로 동작할 수 있고, 이들 시스템으로부터의 최근 시계열 데이터 어레이는 과거 1분 또는 그보다 많은 시간 내의 시간으로부터 이루어진 것일 수 있다. 또 다른 예에서, 시스템은 더 긴 시간 단위(예컨대, 일자)로 동작할 수 있고, 그러한 시스템에 대한 최근 시계열 데이터 어레이는 과거 하나 이상의 일자 내의 시간으로부터 이루어진 것일 수 있다. 최근 시계열 데이터 어레이는 특정 시점 또는 시간 범위에 대한 것일 수 있다. 시간 범위의 경우, 시계열 데이터 어레이는 시간 범위 내에서 이루어지는 여러 시점으로부터의 데이터를 포함할 수 있다. 일 예에서, 상기 어레이를 "획득하는" 것은 모델링 시스템에서 착신 데이터로부터 상기 어레이를 생성하는 것을 포함할 수 있다. 다른 예에서, 상기 어레이를 "획득하는" 것은 모델링 시스템 외부의 시스템으로부터 상기 어레이를 수신하는 것을 포함할 수 있다. 프로세싱은 단계 304로 진행된다.
- [0037] 304에서, 모델링 시스템(예컨대, 108)은 최근 시계열 데이터 어레이와 유사한 과거 시계열 데이터 어레이를 식별하기 위해 하나 이상의 과거의 시계열 데이터 어레이를 검색한다. 과거 시계열 데이터 어레이는 과거의 다양한 대응 시간에서 최근 시계열 데이터 어레이의 변수, 신호, 매개변수 및/또는 공변수의 대응 값을 나타내는 데이터를 포함할 수 있다.
- [0038] 최근 시계열 데이터 어레이가 단일 시점으로부터의 변수 값들의 횡단면으로 이루어지면, 시계열 데이터는 벡터 또는 패턴을 형성한다. 최근 시계열 벡터의 내용이 과거 시계열 벡터와 유사한지를 결정하는 데 사용될 수 있는 이용 가능한 여러 방법이 있다. 이러한 방법 중의 가장 일반적인 방법은 벡터 표준에 기초하여 이루어지며 2개의 벡터 간 거리를 측정하는 것이다. 이러한 방법에는 유클리드(Euclidean) 거리, p-표준 거리, 도시-블록(city-block) 거리, 민코프스키(Minkowski) 거리, 마할라노비스(Mahalanobis) 거리, 및 해밍(Hamming) 거리가 포함되지만 이들에 국한되지는 않는다. 원칙적으로, 스칼라 측정치를 생성하는 2개의 벡터의 임의의 함수는 2개의 벡터 간 유사도를 결정하는데 사용될 수 있다. 이러한 벡터 함수는 커널 함수, 공분산 함수, 상관 계수 및 유사도 연산자를 포함한다. 거리 또는 벡터 함수는 최근 시계열 벡터를 모든 과거 시계열 벡터(104)와 비교하는데 사용된다. 결과적으로 얻어진 거리 또는 스칼라 측정치들은 최근 시계열 벡터와 가장 유사한 과거 시계열 벡터의 특정 개수 또는 분수를 식별하도록 정렬(sort) 또는 스케닝된다.

- [0039] 최근 시계열 데이터 어레이가 여러 시점으로부터의 데이터 벡터 또는 패턴의 그룹으로 이루어지면, 2-단계 계산이 최근 시계열 어레이와 과거 시계열 어레이 간의 유사도를 결정하는데 사용될 수 있다. 제1 단계에서, 최근 시계열 데이터 어레이 내 벡터들 각각은 과거 시계열 어레이 내 대응 벡터와 비교된다. 2개의 어레이 내 벡터들의 대응관계는 어레이들의 편성 방법에 의존한다. 2개의 어레이가 동일한 시간 범위에 걸쳐있는 동일한 개수의 벡터로 시간 순서화된다고 가정하면, 어레이 A의 제1 벡터를 어레이 B의 제1 벡터와 매치(match)시키고 상기 어레이들 내의 벡터들을 통해 마찬가지로 매치시키는 대응관계가 시계열 어레이 내 데이터 벡터들의 상대 순서에 의해 결정될 수 있다. 어레이들이 상이한 개수의 벡터들을 포함하거나 상이한 시간 범위에 걸쳐있는 경우, 벡터들의 타임스탬프(다시 말하면, 데이터가 수집된 시점)가 벡터들 간 최상의 대응관계를 결정하는데 사용된다. 이는 어레이들 내 벡터들 각각의 오프셋, 다시 말하면 각각의 벡터의 타임스탬프 및 어레이 내 첫 번째(가장 최근) 벡터의 타임스탬프 간의 차이를 결정함으로써 수행된다. 최근 시계열 벡터들의 오프셋들은 양자 모두의 어레이들 내 벡터들 간 최상의 매치를 찾아내도록 과거 시계열 벡터들의 오프셋들과 비교된다. 일단 최근 시계열 어레이들 내 벡터들 각각이 과거 시계열 어레이 내 벡터와 매치 되면, 벡터 쌍들 각각 간의 유사도는 위에서 설명한 거리 함수 또는 벡터 함수를 사용하여 계산된다. 제2 단계에서, 평균값, 중앙값 또는 최대값과 같은 통계치를 사용하여 벡터 쌍들 각각에 대한 값들이 조합되고, 그럼으로써 2개의 어레이 간 유사도의 스칼라 측정치를 제공하게 된다. 상기 2-단계 계산을 사용하여, 최근 시계열 어레이는 모든 과거 시계열 어레이(104)와 비교되고, 결과적으로 얻어진 어레이 비교 측정치들이 최근 시계열 어레이와 가장 유사한 과거 시계열 어레이들의 특정 개수 또는 분수를 식별하도록 정렬 및 스케닝된다.
- [0040] 위에서 구체적으로 설명한 2-단계 프로세스에서, 2개의 어레이 내 데이터는 먼저 동일한 상대적인 시점에서의 데이터를 비교함으로써 횡단면으로 분석되고, 그리고 나서 제2 단계에서 데이터가 어레이들에 의해 걸쳐져 있는 시간 범위를 가로질러 길이 방향으로 분석된다. 횡단면 분석 단계 이전에 길이방향 분석 단계를 수행하는 것과 같은 다른 변형 예들은 2개의 어레이 내 데이터가 유사한지를 결정하는데 사용될 수 있다. 추가로, US 8,620,853의 순차적 커널 함수와 같은 행렬 연산은 2개의 어레이 내 데이터가 유사한지를 결정하는데 사용될 수 있다.
- [0041] 306에서, 모델링 시스템(예컨대, 108)은 최근 이벤트 데이터를 획득한다. 이벤트 데이터는 시스템(예컨대, 최근 시계열 데이터 어레이가 대응하게 하는 시스템)의 하나 이상의 매개변수의 단기간 일시적 작동에 영향을 줄 수 있는 이벤트들의 발생을 나타낼 수 있다. 이러한 이벤트들은 특정 시점 또는 기간에 대해 간헐적일 수도 있고 비동기적일 수 있으며 이와는 달리 로컬화될 수도 있으며, 시스템 작동 또는 운영 환경에서 진행중인 패턴을 반영하지 못할 수도 있다. 이러한 이벤트의 효과를 고려하면 더 나은 예측 모델이 제공될 수 있다. 이벤트 데이터는 일시적 이벤트의 성질 또는 타입(예컨대, 기상 이벤트 타입, 사람이 만든 이벤트 타입 등)에 관한 정보를 포함할 수 있다. 프로세싱은 단계 308로 진행된다.
- [0042] 단계 308에서, 하나 이상의 과거 시계열 데이터 어레이가 선택된다. 과거 시계열 데이터 어레이는 최근 시계열 데이터 어레이와 유사한 과거 시계열 데이터 어레이를 식별하는 것과 그리고 나서 과거 시계열 데이터 어레이의 이벤트 데이터를 사용하여 유사한 과거 시계열 데이터 어레이를 필터링하는 것의 조합에 기초하여 선택될 수 있다. 과거 이벤트 데이터는 과거 시계열 데이터와 동일한 소스로부터 또는 다른 소스로부터 획득될 수 있다. 과거 시계열 데이터 어레이 선택 프로세스는 도 4와 관련하여 이하에서 더 구체적으로 설명된다. 프로세싱은 단계 310으로 진행된다.
- [0043] 단계 310에서, 모델링 시스템(예컨대, 108)은 필터링된 과거 시계열 데이터 어레이에 기초하여 모델(예컨대, 로컬화된 시간적 모델)을 트레이닝한다. 프로세싱은 단계 312로 진행된다.
- [0044] 단계 312에서, 모델링 시스템(예컨대, 108)은 트레이닝된 로컬화된 시간적 모델을 사용하여 예측치를 생성한다. 일부 예들에서, 모델은 부가적인 입력 없이 예측치를 생성할 수 있고, 다른 예들에서, 모델은 예측치를 생성하기 위해 최근 시계열 데이터 어레이와 같은 입력을 요구할 수 있다. 예측치는 미래의 시간 또는 시간 범위에서의 최근 시계열 데이터 어레이의 변수, 신호 및/또는 매개변수 중의 하나 이상의 예측된 값을 나타낼 수 있다. 예측치는 하나 이상의 다른 시스템(예컨대, 모델링 시스템 내부 또는 외부의 다른 시스템)에 제공될 수 있다. 예를 들면, 예측치는 기관차 상의 중요한 시스템으로부터의 신호의 예상 작동을 예측하는 것일 수 있다. 다른 예에서, 예측치는 미래의 시간 또는 시간 범위에서 풍력 터빈의 전력 출력을 예측하는 것일 수 있다. 다른 예에서, 예측치는 시설(예컨대, 의료 시설)에서의 인력 수요를 예측하는 것일 수 있다. 프로세싱은 단계 314로 진행된다.
- [0045] 단계 314에서, 모델링 시스템(예컨대, 108)은 새로운 예측치에 대한 요구가 수신되었는지 또는 시간적 변화가

이루어졌는지 또는 로컬화된 시간적 모델의 재-트레이닝 및 새로운 예측치의 생성을 트리거하는 시스템 변경이 이루어졌는지를 결정한다. 예를 들면, 최근 시계열 데이터 어레이에 의해 표현되는 시간 이래로 주어진 시간이 경과(예컨대, 현재 시계 사이클, 현재 매개변수 샘플링 간격 또는 다른 기간이 경과)하면, 모델링 시스템(예컨대, 108)은 로컬화된 시간적 모델을 현재(또는 가장 최근) 시계열 데이터로 재-트레이닝할 수 있다. 다른 예에서, 시간적 변화는 충분한 양의 시간(예컨대, 다음 초, 분, 시간, 일, 주, 월, 계절 등)의 경과를 포함할 수 있다. 다른 예에서, 시간적 변화 또는 시스템 변화는 예측 모델에 특정한 변화(예컨대, 최근 5개의 시계열 데이터 어레이와 비교되는 현재 시계열 데이터 어레이에서의 변화)를 포함할 수 있다. 다른 예에서, 최근 시계열 데이터 어레이가 대응하게 하는 시스템에 대해 시스템 변화(예컨대, 하나 이상의 매개변수 값 또는 공변량의 변화)가 이루어지면, 모델링 시스템(예컨대, 108)은 시스템 변화에 기초하여 로컬화된 시간적 모델을 재-트레이닝할 수 있다. 예를 들면, 차량 교통 스케줄링 시스템 매개변수는 하나 이상의 차량 경로 상의 트래픽 양을 포함할 수 있다. 시스템 매개변수의 변화는 경로 상의 트래픽 레벨 증가(또는 감소)를 포함할 수 있다. 다른 예에서, 매개변수의 변화는 시스템의 동작 모드 변화(예컨대, 정지로부터 실행으로 또는 그 반대로의 동작 모드 변화)를 포함할 수 있다. 시스템 변화는 시스템에 연관된 변수, 신호, 매개변수 및/또는 공변량 중의 하나 이상의 값의 변화에 의해 결정될 수 있다. 로컬화된 시간적 모델의 재-트레이닝을 트리거할 수 있는 또 다른 시스템 변화는 시스템에 연관된 새로운 이벤트의 발생이다. 예를 들면, 새로운 이벤트 코드는 기관차 상의 제어 시스템에 의해 생성되고 이벤트 데이터(106)에 의해 수신될 수 있다. 재-트레이닝은 새로운 최근 시계열 데이터 어레이를 사용하여 블록들(302-312) 중의 하나 이상의 블록들을 반복하는 것을 포함할 수 있다. 예측치에 대한 요구가 수신되지 않았거나 시간적 또는 시스템 변화가 이루어지지 않은 경우, 프로세싱은 단계 314로 복귀된다. 블록(302-312)은 조합될 수도 있고, 스킵될 수도 있으며, 재순서화될 수도 있고, 반복될 수도 있다. 예를 들면, 경우에 따라, 재-트레이닝 기준 중의 하나 이상이 충족될 때까지 트레이닝된 모델을 사용하여 다수의 예측치가 생성될 수 있다. 일부 예들에서, 재-트레이닝은 예측치가 생성되기 직전에 수행될 수 있다.

[0046] 도 4는 적어도 하나의 구현 예에 따라 과거 시계열 데이터 어레이를 선택하는 세부사항을 보여주는 로컬화된 시간 예측을 위한 대표적인 방법의 흐름도이다. 하나 이상의 과거 시계열 데이터 어레이를 선택하는 단계 308에 대한 프로세싱은 단계 402에서 시작하며, 단계 402에서는, 최근 시계열 데이터 어레이와 유사한 하나 이상의 과거 시계열 데이터 어레이가 식별된다. 일부 구현 예들에서, 상기 매칭은 시스템의 과거 이력으로부터의 시계열 데이터 어레이들 및 현재 시계열 데이터 어레이 간 패턴 중첩을 계산하는 함수를 적용하는 것을 포함할 수 있다. 일부 예들에서, 이러한 함수는 데이터의 시간(길이방향) 및 횡단면 비교를 적용하고 패턴 중첩의 스칼라 측정치를 반환할 수 있다. 또한, 상기 함수는 높은 패턴 매치 값들을 갖는 과거로부터의 기간을 식별할 수 있다. 프로세싱은 단계 404로 진행된다.

[0047] 단계 404에서, 최근 시계열 데이터 어레이와 유사한 것으로 단계 402에서 식별된 과거 시계열 데이터 어레이는 이벤트 데이터를 사용하여 필터링될 수 있다. 예를 들면, 최근 시계열 데이터 어레이가 게임이 플레이 되는 스포츠 경기장을 갖는 도시의 게임 일자와 같은 이벤트에 연관되는 경우, 과거 시계열 데이터 어레이와 유사하며 게임 일자 동안 이루어진 그러한 과거 시계열 어레이들을 포함하도록 필터링될 수 있다. 다른 예에서, 상기 이벤트는 비 또는 눈과 같은 기상 이벤트일 수 있다. 이러한 경우에, 과거 시계열 데이터 어레이는 최근 시계열 데이터 어레이에 연관된 이벤트와 유사한 기상 이벤트를 갖는 그러한 과거 시계열 데이터 어레이들만을 포함하도록 필터링될 수 있다. 다른 예에서, 이벤트 데이터는 기관차의 제어 시스템에 의해 생성된 하나 이상의 이벤트 코드를 수신하는 것일 수 있다. 이러한 경우에, 과거 시계열 데이터 어레이는 과거 시계열 데이터가 수집된 기간의 시간과 동일하거나 그러한 기간의 시간 부근에 동일한 이벤트 코드들을 수신한 그러한 과거 시계열 데이터 어레이만을 포함하도록 필터링될 수 있다. 일반적으로 과거 시계열 어레이를 필터링하는 것은 최근 시계열 어레이에 연관된 동일한 이벤트(들)의 발생을 매치시킴으로써 이루어진다. 최근 시계열 어레이에 연관된 동일한 이벤트에 시간적으로 연관된 과거 시계열 어레이를 매치시키는 것 외에도, 이벤트 데이터의 다른 특성은 과거 시계열 어레이를 부가적으로 필터링하는데 사용될 수 있다. 예를 들면, 최근 데이터 어레이에 연관된 이벤트가 지난 24 시간 동안 눈이 내린 것이라면 그 기간에 내린 눈의 양(예컨대, 1 내지 3인치의 눈)은 과거 시계열 어레이에 의해 충족되어야 하는 추가 조건으로 사용될 수 있다. 프로세싱은 단계 406으로 진행된다.

[0048] 이벤트 데이터에 기초한 과거 시계열 데이터 어레이들의 필터링은 최근 기간으로부터의 시계열 데이터(예컨대, 최근 시계열 데이터의 윈도우에 대응하는 시계열 데이터의 윈도우)에 대한 유사도를 지니는 것으로 식별되고 (이벤트의 효과가 있는 경우) 이벤트의 효과를 반영하는 트레이닝 데이터 세트를 제공하는데 중요할 수 있다. 종래의 트레이닝 프로세스는 상기 모델을 트레이닝하도록 모든 과거 시계열 데이터 어레이들을 사용할 수 있다. 그러나 모든 과거 시계열 데이터 어레이들을 사용하는 종래의 트레이닝 세트는 현재 데이터 어레이에 연관된 이

벤트들에 대응하지 않지만 그 외에 매칭되는 과거 데이터 어레이들을 포함할 수 있다(예컨대, 최근 시계열 데이터 어레이는 장마철 동안 "강우량이 적은" 날에 대응되지만, 과거 시계열 데이터 어레이는 최근 시계열 데이터 어레이와 매치하도록 이루어진 가물철 동안 "강우량이 많은" 날에 대응된다). 이러한 매치들을 필터링함으로써, 모델에 사용된 트레이닝 데이터 세트가 개선될 수 있다.

[0049] 단계 406에서, 상기 필터링된 과거 시계열 데이터 어레이들은 모델을 트레이닝하기 위한 로컬화된 시간적 모델에 제공된다. 프로세싱은 위에서 설명한 바와 같이 단계 308로부터 단계 310으로 진행된다.

[0050] 로컬화된 시간적 모델은 필요할 때에만(예컨대, 필요에 따라 또는 요구된 기준으로서 그리고/또는 시간적 또는 시스템 변화에 기초하여) 재-트레이닝되기 때문에 그리디 학습(greedy learning) 알고리즘이다. 시간적 변화 또는 시스템 매개변수 변화에 기초하여 모델을 재-트레이닝함으로써 모델은 시계열 데이터 및/또는 이벤트 데이터의 최근 조건에 지속적으로 적응할 수 있다. 로컬화된 시간적 모델이 현재 시스템 조건과 유사한 데이터로 트레이닝되므로 이용 가능한 데이터를 사용하는 것은 간결하다. 이러한 기법은 빅 데이터 애플리케이션에(예컨대, 어셋, 어셋 그룹, 의료 시설, 기차, 차량, 기상 모니터링 시스템, 과학 컴퓨팅 시스템, 정보 수집 시스템, 분산 시스템 등과 같은 기간당 대량 데이터를 생성하는 시스템들의 작동들에 대한 예측치들을 생성하는 데) 유용할 수 있다.

[0051] 통상의 기술자라면 위에서 설명한 모듈, 프로세스, 시스템 및/또는 섹션이 하드웨어, 소프트웨어에 의해 프로그램된 하드웨어, 비-일시적 컴퓨터 판독가능 매체 상에 저장된 소프트웨어 명령어 또는 상기의 조합으로 구현될 수 있음을 알 수 있을 것이다. 예를 들면, 위에서 설명한 바와 같은 시스템은 비-일시적 컴퓨터 판독가능 매체 상에 저장된 프로그램된 명령들의 시퀀스를 실행하도록 구성된 프로세서를 포함할 수 있다. 예를 들면, 상기 프로세서는 퍼스널 컴퓨터, 모바일 기기 및/또는 워크스테이션, 또는 프로세서, 마이크로프로세서, 마이크로컨트롤러 기기를 포함하거나 예를 들면 ASIC(Application Specific Integrated Circuit)과 같은 집적 회로를 포함하는 제어 로직으로 이루어진 다른 그러한 컴퓨팅 시스템을 포함할 수 있지만, 이들에 국한되지 않는다. 명령어는 Java, C, C ++, C#.net 또는 어셈블리 등과 같은 프로그래밍 언어에 따라 제공되는 소스 코드 명령어로부터 컴파일될 수 있다. 명령어는 또한 예를 들면, Visual Basic™ 언어 또는 다른 구조화된 또는 객체 지향 프로그래밍 언어에 따라 제공되는 코드 및 데이터 객체를 포함할 수 있다. 프로그래밍된 명령어들의 시퀀스 또는 프로그램가능 로직 기기 구성 소프트웨어 및 이에 연관된 데이터는 컴퓨터 메모리와 같은 비-일시적 컴퓨터-판독가능 매체나 또는 ROM, PROM, EEPROM, RAM, 플래시 메모리, 디스크 드라이브 등과 같지만 이들에 국한되지 않는 임의의 적절한 메모리 장치일 수 있는 저장 기기에 저장될 수 있다.

[0052] 또한, 모듈, 프로세스, 시스템 및 섹션은 단일 프로세서로서 또는 분산 프로세서로서 구현될 수 있다. 또한, 여기서 이해해야 할 점은 위에 언급한 단계들이 단일 또는 분산 프로세서(단일 및/또는 다중-코어, 또는 클라우드 컴퓨팅 시스템)상에서 수행될 수 있다는 점이다. 또한, 위의 실시 예들에 대해 그리고 위의 실시 예들의 여러 도면에서 설명한 프로세스, 시스템 컴포넌트, 모듈, 및 서브-모듈은 다수의 컴퓨터 또는 시스템에 걸쳐 분산될 수도 있고 단일 프로세서 또는 시스템에 함께 배치될 수 있다. 본원 명세서에서 설명한 모듈, 섹션, 시스템, 수단 또는 프로세스를 구현하기에 적합한 대표적인 구조적 실시 예가 이하에 제공된다.

[0053] 위에서 설명한 모듈, 프로세서 또는 시스템은 프로그램된 범용 컴퓨터, 마이크로코드로 프로그램된 전자 기기, 하드-와이어드 아날로그 로직 회로, 컴퓨터-판독가능 매체 또는 신호 상에 저장된 소프트웨어, 광학 컴퓨팅 기기, 전자 및/또는 광학 기기들의 네트워크 시스템, 특수 목적 컴퓨팅 기기, 집적 회로 기기, 반도체 칩 및/또는 예를 들면 컴퓨터 판독가능 매체 또는 신호 상에 저장된 소프트웨어 모듈 또는 객체로서 구현될 수 있다.

[0054] 방법 및 시스템(또는 그 서브-컴포넌트 또는 모듈)의 실시 예들은 범용 컴퓨터, 특수 목적 컴퓨터, 프로그램된 마이크로프로세서 또는 마이크로컨트롤러 및 주변 집적 회로 요소, ASIC 또는 기타 집적 회로, 디지털 신호 프로세서, 개별 요소 회로와 같은 하드와이어드 전자 또는 로직 회로, PLD, PLA, FPGA, PAL 등과 같은 프로그램된 논리 회로 상에서 구현될 수 있다. 일반적으로, 본원 명세서에서 설명한 기능들 또는 단계들을 구현할 수 있는 임의의 프로세서는 방법, 시스템 또는 컴퓨터 프로그램 제품(비-일시적 컴퓨터 판독가능 매체상에 저장된 소프트웨어 프로그램)의 실시 예들을 구현하는데 사용될 수 있다.

[0055] 또한, 개시된 방법, 시스템 및 컴퓨터 프로그램 제품(또는 비-일시적 컴퓨터 판독가능 매체상에 저장된 소프트웨어 명령어)의 실시 예들은, 예를 들면 다양한 컴퓨터 플랫폼상에서 사용 가능한 포터블 소스 코드를 제공하는 객체 또는 객체 지향(object-oriented) 소프트웨어 개발 환경을 사용하여 소프트웨어로 완전히 또는 부분적으로 용이하게 구현될 수 있다. 변형적으로, 개시된 방법, 시스템 및 컴퓨터 프로그램 제품의 실시 예들은 예를 들면 표준 논리 회로 또는 VLSI 설계를 사용하여 하드웨어로 부분적으로 또는 전체적으로 구현될 수 있다. 다른 하드

웨어 또는 소프트웨어는 시스템, 특정 기능, 및/또는 이용되는 특정 소프트웨어 또는 하드웨어 시스템, 마이크로프로세서 또는 마이크로컴퓨터의 속도 및/또는 효율 요건에 따라 실시 예들을 구현하는데 사용될 수 있다. 본 방법, 시스템 및 컴퓨터 프로그램 제품의 실시 예들은 본원 명세서에 제공된 기능 설명으로부터 그리고 소프트웨어 공학 및 어셋 조건 모니터링 기술에 대한 일반적인 기본 지식을 가지고 적용 가능한 통상의 기술자에 의해 임의 공지되었거나 추후 개발될 시스템 또는 구조, 기기 및/또는 소프트웨어를 사용하여 하드웨어 및/또는 소프트웨어로 구현될 수 있다.

[0056] 또한, 개시된 방법, 시스템 및 컴퓨터 판독가능 매체(또는 컴퓨터 프로그램 제품)의 실시 예들은 프로그램된 범용 컴퓨터, 특수 목적 컴퓨터, 마이크로프로세서 등 상에서 실행되는 소프트웨어로 구현될 수 있다.

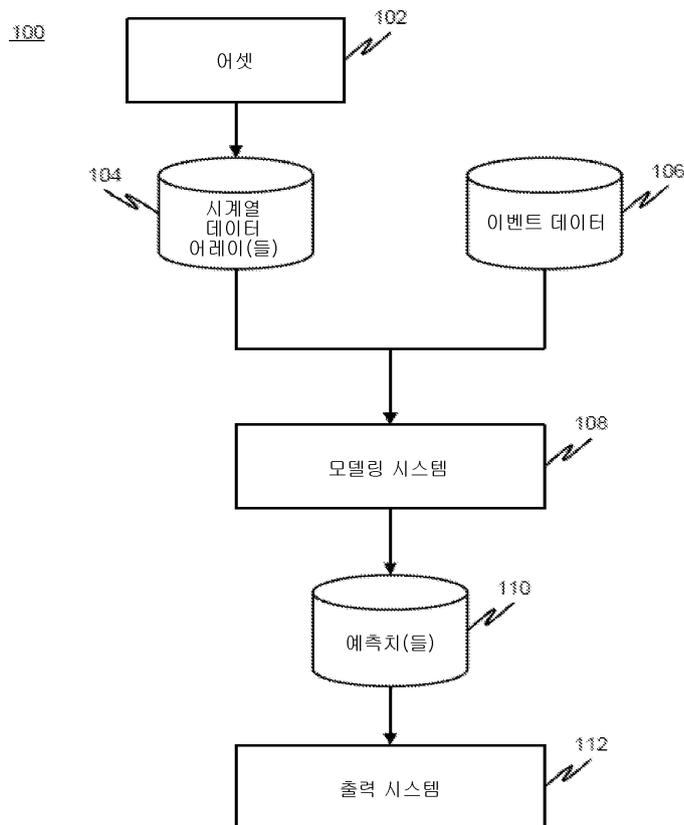
[0057] 그러므로 본원 명세서에 개시된 다양한 실시 예에 따라, 로컬화된 시간적 모델링을 위한 방법, 시스템 및 컴퓨터 판독가능 매체가 제공되는 것이 명백하다.

[0058] 개시된 주제가 다수의 실시 예와 관련하여 설명되었지만, 많은 대안, 수정 및 변형이 적용 가능한 통상의 기술자에게 자명할 것이거나 자명하다는 것이 명백하다. 따라서, 본원 출원인은 개시된 주제의 사상 및 범위 내에 있는 그러한 모든 대안, 수정, 등가물 및 변형을 포함하고자 한 것이다.

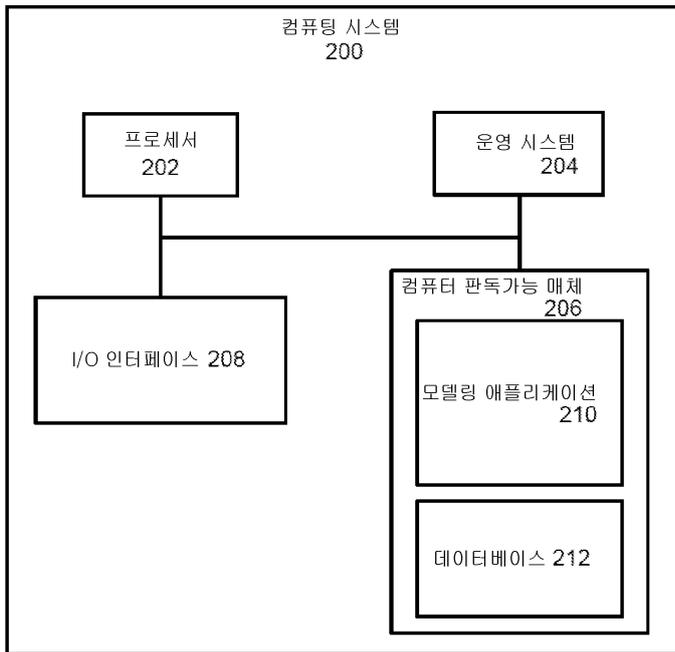
[0059] 또한, 본원 명세서에서 설명한 예가 "인간", "조작자", "사용자" 또는 다른 엔티티와 같은 행위자에 의해 수행되거나 개시되는 동작들을 포함하는 한, 이는 예 및 설명만을 목적으로 한 것이다. 청구항들은 청구항 언어로 명시적으로 기재되어 있지 않는 한 그러한 행위자의 조치를 요하는 것으로 해석되어서는 아니 된다.

도면

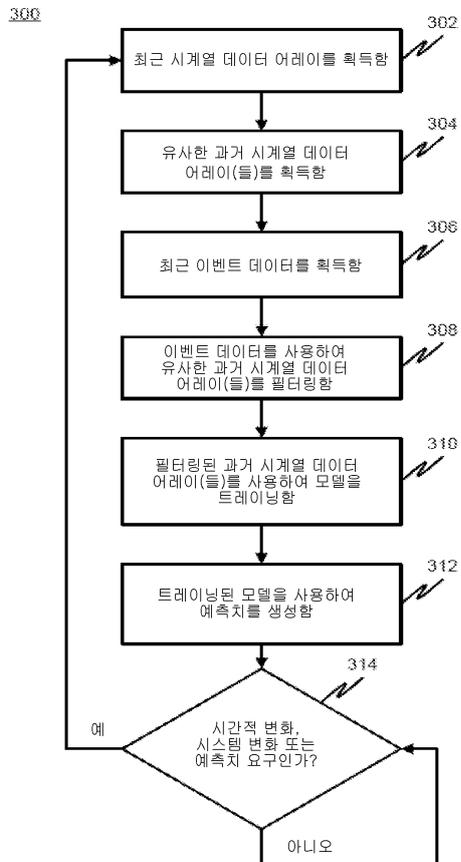
도면1



도면2



도면3



도면4

