



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0060657
(43) 공개일자 2025년05월07일

(51) 국제특허분류(Int. Cl.)
H04L 67/1095 (2022.01) G06N 3/098 (2023.01)
H04L 67/50 (2022.01)
(52) CPC특허분류
H04L 67/1095 (2022.05)
G06N 3/098 (2023.01)
(21) 출원번호 10-2023-0144912
(22) 출원일자 2023년10월26일
심사청구일자 2023년10월26일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
이현석
서울특별시 성동구 상원길 63, 107동 602호(성수동1가, 쌍용아파트)
김지완
경기도 용인시 기흥구 동백죽전대로527번길 81, 103동 1102호(중동, 신동백 서해그랑블 1차)
(74) 대리인
수안특허법인, 민영준

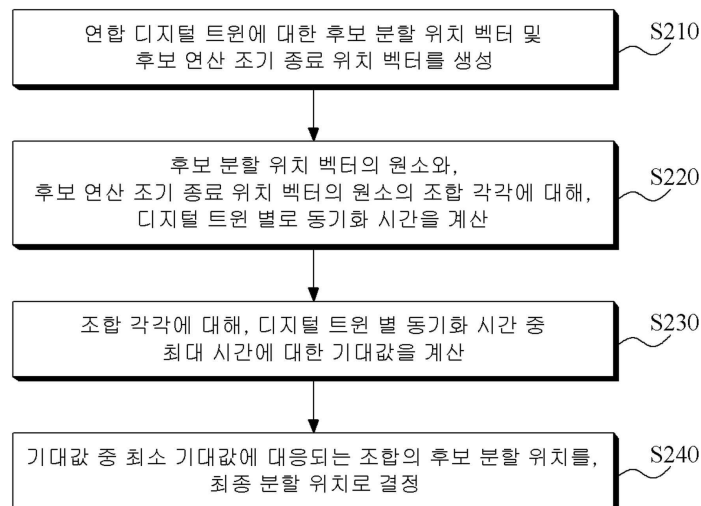
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 연합 디지털 트윈 동기화 방법

(57) 요약

엣지 서버 및 클라우드 서버가 이용되는 분할 컴퓨팅 환경과 조기 종료 모델이 이용되는 환경에서의 연합 디지털 트윈에 대한 동기화 방법이 개시된다. 개시된 연합 디지털 트윈 동기화 방법은, 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 상기 연합 디지털 트윈에 대한 후보 분할 위치 벡터를 생성하고, 상기 신경망에 대한 후보 연산 조기 종료 위치로부터, 상기 연합 디지털 트윈에 대한 후보 연산 조기 종료 위치 벡터를 생성하는 단계; 상기 후보 분할 위치 벡터의 원소와, 상기 후보 연산 조기 종료 위치 벡터의 원소의 조합 각각에 대해, 상기 디지털 트윈 별로 동기화 시간을 계산하는 단계; 상기 조합 각각에 대해, 상기 디지털 트윈 별 동기화 시간 중 최대 시간에 대한 기대값을 계산하는 단계; 및 상기 기대값 중 최소 기대값에 대응되는 조합의 후보 분할 위치를 최종 분할 위치로 결정하는 단계를 포함한다.

대표도 - 도2



(52) CPC특허분류

H04L 67/50 (2022.05)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193284
과제번호	2022-0-00545-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	디지털트윈연합핵심기술개발사업
연구과제명	(2세부) 지능형 디지털 트윈 연합 객체 구성 및 데이터 프로세싱 기술 개발
기 여 율	1/2
과제수행기관명	한국전자기술연구원
연구기간	2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193909
과제번호	2021-0-01816-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	메타버스 자율트윈 핵심기술 연구
기 여 율	1/2
과제수행기관명	세종대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 상기 연합 디지털 트윈에 대한 후보 분할 위치 벡터를 생성하고, 상기 신경망에 대한 후보 연산 조기 종료 위치로부터, 상기 연합 디지털 트윈에 대한 후보 연산 조기 종료 위치 벡터를 생성하는 단계;

상기 후보 분할 위치 벡터의 원소와, 상기 후보 연산 조기 종료 위치 벡터의 원소의 조합 각각에 대해, 상기 디지털 트윈 별로 동기화 시간을 계산하는 단계;

상기 조합 각각에 대해, 상기 디지털 트윈 별 동기화 시간 중 최대 시간에 대한 기대값을 계산하는 단계; 및

상기 기대값 중 최소 기대값에 대응되는 조합의 후보 분할 위치를 최종 분할 위치로 결정하는 단계

를 포함하는 연합 디지털 트윈 동기화 방법.

청구항 2

제 1항에 있어서,

상기 후보 분할 위치 벡터의 원소 각각은

상기 신경망 각각의 후보 분할 위치로부터, 상기 신경망 별로 선택된 하나의 후보 분할 위치를 포함하며,

상기 후보 연산 조기 종료 위치 벡터의 원소 각각은

상기 신경망 각각의 후보 연산 조기 종료 위치로부터, 상기 신경망 별로 선택된 하나의 후보 연산 조기 종료 위치를 포함하는,

연합 디지털 트윈 동기화 방법.

청구항 3

제 2항에 있어서,

상기 기대값을 계산하는 단계는

상기 후보 연산 조기 종료 위치 벡터의 원소에 포함된 후보 연산 조기 종료 위치에서의 조기 종료 확률값을 이용하여, 상기 기대값을 계산하는

연합 디지털 트윈 동기화 방법.

청구항 4

제 3항에 있어서,

상기 기대값을 계산하는 단계는

하기 [수학식]을 이용하여, 상기 기대값을 계산하는

연합 디지털 트윈 동기화 방법.

[수학식]

$$\sum_{e \in E} x^e \cdot \max_{u \in U} \{t_u(p_u, e_u)\}$$

여기서, E는 상기 후보 연산 조기 종료 위치 벡터, e는 상기 후보 연산 조기 종료 위치 벡터의 원소,

$\max_{u \in U} \{t_u(p_u, e_u)\}$ 는 상기 최대 시간, x^e 는 상기 후보 연산 조기 종료 위치 벡터의 원소에 포함된 후보 연산 조기 종료 위치에서의 조기 종료 확률값을 곱한 값을 나타냄.

청구항 5

제 1항에 있어서,

상기 최종 분할 위치에 의해 분할된 신경망 중 입력 노드를 포함하는 제1분할 신경망은 엣지 서버에 할당되고, 출력 노드를 포함하는 제2분할 신경망은 클라우드 서버에 할당되는

연합 디지털 트윈 동기화 방법.

청구항 6

연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 상기 디지털 트윈 각각에 대한 후보 분할 위치 벡터를 생성하고, 상기 신경망에 대한 후보 연산 조기 종료 위치로부터, 상기 디지털 트윈 각각에 대한 후보 연산 조기 종료 위치 벡터를 생성하는 단계;

상기 디지털 트윈 별로, 상기 후보 분할 위치 벡터의 원소와, 상기 후보 연산 조기 종료 위치 벡터의 원소의 조합 각각에 대한 동기화 시간을 계산하는 단계;

상기 디지털 트윈 각각에 대해, 상기 동기화 시간에 대한 기대값을 계산하는 단계; 및

상기 기대값 중 최소 기대값에 대응되는 후보 분할 위치를 상기 디지털 트윈 각각의 신경망에 대한 최종 분할 위치로 결정하는 단계

를 포함하는 연합 디지털 트윈 동기화 방법.

청구항 7

제 6항에 있어서,

상기 기대값을 계산하는 단계는

상기 신경망 각각에 대한 상기 후보 연산 조기 종료 위치 각각에서의 조기 종료 확률값을 이용하여, 상기 기대값을 계산하는

연합 디지털 트윈 동기화 방법.

청구항 8

제 7항에 있어서,

상기 최종 분할 위치에 의해 분할된 신경망 중 입력 노드를 포함하는 제1분할 신경망은 엣지 서버에 할당되고, 출력 노드를 포함하는 제2분할 신경망은 클라우드 서버에 할당되는

연합 디지털 트윈 동기화 방법.

청구항 9

제 8항에 있어서,

상기 기대값을 계산하는 단계는

하기 [수학식]을 이용하여, 상기 기대값을 계산하는

연합 디지털 트윈 동기화 방법.

[수학식]

$$\sum_{e_u \in E_u} x_u^{e_u} \cdot t_u^E(p_u, e_u) + \sum_{e_u \in E_u} x_u^{e_u} \cdot t_u^{tr}(p_u, e_u) + \sum_{e_u \in E_u} x_u^{e_u} \cdot t_u^C(p_u, e_u)$$

여기서, u 는 상기 디지털 트윈의 인덱스, p 는 상기 후보 분할 위치, e 는 상기 후보 연산 초기 종료 위치, E 는

상기 후보 연산 초기 종료 위치 벡터, $t_u^E(p_u, e_u)$ 는 상기 엣지 서버가 상기 제1분할 신경망의 연산을

수행하는데 소요되는 시간, $t_u^{tr}(p_u, e_u)$ 는 상기 제1분할 신경망의 출력값이 상기 클라우드 서버로 전송되

는 시간, $t_u^C(p_u, e_u)$ 는 상기 클라우드 서버가 상기 제2분할 신경망의 연산을 수행하는데 소요되는 시

간, $x_u^{e_u}$ 는 상기 후보 연산 초기 종료 위치에서의 초기 종료 확률값을 나타냄.

청구항 10

연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치를 입력받는 단계;

상기 후보 분할 위치별로 상기 신경망 각각의 동기화 시간을 계산하는 단계; 및

상기 신경망 각각의 동기화 시간 중 최소 시간을 탐색하고, 상기 최소 시간 중 최대 시간에 대응되는 후보 분할 위치를 상기 디지털 트윈의 최종 분할 위치로 결정하는 단계

를 포함하는 연합 디지털 트윈 동기화 방법.

발명의 설명

기술 분야

[0001] 본 발명은 연합 디지털 트윈 동기화 방법에 관한 것으로서, 더욱 상세하게는 엣지 서버 및 클라우드 서버가 이용되는 분할 컴퓨팅 환경과 초기 종료 모델이 이용되는 환경에서의 연합 디지털 트윈에 대한 동기화 방법에 관한 것이다.

배경 기술

[0003] 디지털 트윈(digital twin)이란 현실 세계에서 실존하는 사람, 자동차, 제조설비 등을, 디지털 공간에 실물과 똑같은 쌍둥이를 만들고, 현실 객체의 동작과 행위를 가상 세계에서 실현시켜서, 현실 세계와 가상세계의 쌍

등이 개체가 서로의 변화를 동기화시키는 것이다. 그리고 연합 디지털 트윈(federated digital twin)이란, 하나 이상의 디지털 트윈들로 구성된 집합을 의미한다.

[0004] 물리적인 대상들을 인터넷에 연결하고 데이터를 교환할 수 있는 IoT 환경에서 연합 디지털 트윈 동기화는 현실 세계에 존재하는 물리적인 객체들을 IoT 센서들을 통해 지속적으로 탐지하고 감지하여 현실 객체들의 형태와 동작 등을 가상세계에서도 실현시켜 현실 세계와 가상세계의 쌍둥이 개체들이 서로의 변화를 동기화시키는 것이다.

[0005] 이때 연합 디지털 트윈의 각 디지털 트윈 동기화 과정은 현실 객체 감지, 객체 형태/동작 예측, 계산된 데이터 업로드등의 작업들을 모두 완료함으로써 이루어질 수 있다. 디지털 트윈 동기화 과정을 수행하기 위해서, 복잡한 특징이나 패턴을 입력 데이터에서 학습할 수 있는 능력으로, 이미지 분류, 객체 감지와 같은 다양한 응용 분야에서 뛰어난 성능을 보여왔던 강력한 기계 학습 모델인 Deep Neural Network(DNN) 모델이 일반적으로 사용된다. 특히 DNN 파이프라인 모델은 end-to-end로 훈련될 수 있어, 전체 객체를 감지하고 객체의 형태나 동작을 예측하는 과정을 통합된 프레임워크에서 최적화할 수 있다. 그리고 DNN 모델은 여러 객체를 동시에 감지하도록 훈련될 수 있으며 전체 동작 과정을 엣지 서버(edge server)와 클라우드 서버(cloud server)에서 모두 활용할 수 있다.

[0006] 그러나 이러한 DNN 모델을 디지털 트윈 동기화 과정에서 사용하는 것은, 일반적으로 수행 시간이 많이 소요되고 계산 비용이 많이 든다. 특히 연합 디지털 트윈과 같이, 여러 디지털 트윈들을 동시에 다룰 때 더욱 DNN 모델을 사용하기 어려울 수 있다.

[0007] 이때 연합 디지털 트윈 동기화 과정을 효율적으로 수행하기 위해, 엣지와 클라우드 서버를 계층적인 구조로 두고, DNN 모델 수행 과정을 엣지 서버와 클라우드 서버로 분할하여 계층적으로 작동하도록 하는 분할 컴퓨팅 구조와 조기 종료(Early-Exit) 구조가 활용될 수 있다.

[0008] 분할 컴퓨팅 구조는 DNN 모델을 미리 정해진 레이어(layer)를 기준으로 헤드 모델과 테일 모델로 분할하고, 헤드 모델과 테일 모델을 각각 엣지 서버와 클라우드 서버에서 실행하는 구조이다. 헤드 모델의 출력값이 엣지 서버에서 클라우드 서버로 전송되면, 클라우드 서버에서 헤드 모델의 출력값을 테일 모델에 입력하여, 테일 모델에 대한 연산을 수행한다. 이러한 구조를 활용하면 컴퓨팅 부하를 엣지 서버와 클라우드 서버로 분산시키고 원시적(raw) 입력 데이터를 전부 하나의 서버로 전송하여 수행하는 것을 방지함으로써, 동기화에 필요한 수행 시간을 최소화하고 시스템의 전체적인 효율성을 향상할 수 있다.

[0009] 그리고 조기 종료 구조는 DNN 모델 내부의 여러 “조기 종료 위치”를 두고 해당 위치에서 예측 결과를 조기에 출력할 수 있도록 하며, 이미 충분히 예측 결과가 정확하다면 DNN 모델의 모든 레이어를 수행하지 않더라도 동기화 과정 수행을 조기 종료하는 구조이다. 여기서 각 조기 종료 위치에서는 fully connected layer를 통해 결과값을 만들며, 미리 설정한 신뢰 수준과 결과값을 비교하여 신뢰 수준에 도달하면 DNN 모델을 종료하고 도달하지 못한다면 마지막 조기 종료 위치까지 같은 과정을 반복하여 진행해 수행한다. 이러한 구조는 동기화 작업 수행에 필요한 계산 리소스 및 시간을 크게 줄일 수 있어 시스템의 효율성을 더욱 향상할 수 있다.

[0010] 종래기술 및 연구들은 디지털 트윈 동기화를 수행하기 위해 다양한 시도를 진행하였으나, 각 디지털 트윈의 동기화 수행 시간을 최소화하기 위해 상기에 설명한 엣지 서버와 클라우드 서버로 나누어진 계층적인 구조 및 조기 종료 모델을 활용한 방법들은 연구가 부족하다. 그리고 하나의 DNN 모델에 대해서 시간 모델을 고려하여 분할 위치를 찾는 기존의 연구들은 존재하지만, 다수의 디지털 트윈들의 시간 모델을 모두 고려하면서 조기 종료 위치별 확률을 공동으로 고려하는 구조는 존재하지 않는다.

[0011] 관련 선행문헌으로 대한민국 공개특허 제2020-0113744호, 제2023-0072395호, 제2023-0008599호, 대한민국 등록특허 제10-2577436호가 있다.

발명의 내용

해결하려는 과제

[0013] 본 발명은 복수의 디지털 트윈이 이용되는 연합 디지털 트윈에 대한 동기화 방법을 제공하기 위한 것이다.

[0014] 특히 본 발명은, 연합 디지털 트윈에 대한 동기화 시간을 감소시킬 수 있는 동기화 방법을 제공하기 위한 것이다

다.

과제의 해결 수단

[0016] 상기한 목적을 달성하기 위한 본 발명의 일 실시예에 따르면, 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 상기 연합 디지털 트윈에 대한 후보 분할 위치 벡터를 생성하고, 상기 신경망에 대한 후보 연산 초기 종료 위치로부터, 상기 연합 디지털 트윈에 대한 후보 연산 초기 종료 위치 벡터를 생성하는 단계; 상기 후보 분할 위치 벡터의 원소와, 상기 후보 연산 초기 종료 위치 벡터의 원소의 조합 각각에 대해, 상기 디지털 트윈 별로 동기화 시간을 계산하는 단계; 상기 조합 각각에 대해, 상기 디지털 트윈 별 동기화 시간 중 최대 시간에 대한 기대값을 계산하는 단계; 및 상기 기대값 중 최소 기대값에 대응되는 조합의 후보 분할 위치를 최종 분할 위치로 결정하는 단계를 포함하는 연합 디지털 트윈 동기화 방법이 제공된다.

[0017] 또한 상기한 목적을 달성하기 위한 본 발명의 다른 실시예에 따르면, 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 상기 디지털 트윈 각각에 대한 후보 분할 위치 벡터를 생성하고, 상기 신경망에 대한 후보 연산 초기 종료 위치로부터, 상기 디지털 트윈 각각에 대한 후보 연산 초기 종료 위치 벡터를 생성하는 단계; 상기 디지털 트윈 별로, 상기 후보 분할 위치 벡터의 원소와, 상기 후보 연산 초기 종료 위치 벡터의 원소의 조합 각각에 대한 동기화 시간을 계산하는 단계; 상기 디지털 트윈 각각에 대해, 상기 동기화 시간에 대한 기대값을 계산하는 단계; 및 상기 기대값 중 최소 기대값에 대응되는 후보 분할 위치를 상기 디지털 트윈 각각의 신경망에 대한 최종 분할 위치로 결정하는 단계를 포함하는 연합 디지털 트윈 동기화 방법이 제공된다.

[0018] 또한 상기한 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따르면, 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치를 입력받는 단계; 상기 후보 분할 위치별로 상기 신경망 각각의 동기화 시간을 계산하는 단계; 및 상기 신경망 각각의 동기화 시간 중 최소 시간을 탐색하고, 상기 최소 시간 중 최대 시간에 대응되는 후보 분할 위치를 상기 디지털 트윈의 최종 분할 위치로 결정하는 단계를 포함하는 연합 디지털 트윈 동기화 방법이 제공된다.

발명의 효과

[0020] 본 발명의 일 실시예에 따르면, 연합 디지털 트윈에 대한 동기화 시간이 감소하여, 효율적으로 연합 디지털 트윈의 동기화가 수행될 수 있다.

도면의 간단한 설명

[0022] 도 1은 본 발명의 일 실시예에 따른 연합 디지털 트윈 동기화 방법의 개념을 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시예에 따른 연합 디지털 트윈 동기화 방법을 설명하기 위한 도면이다.

도 3은 본 발명의 다른 실시예에 따른 연합 디지털 트윈 동기화 방법을 설명하기 위한 도면이다.

도 4 및 도 5는 본 발명의 일 실시예에 따른 연합 디지털 트윈 동기화 방법의 수행 결과를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0023] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0024] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.

- [0026] 도 1은 본 발명의 일실시예에 따른 연합 디지털 트윈 동기화 방법의 개념을 설명하기 위한 도면이다.
- [0027] 본 발명의 일실시예는, 엣지 서버와 클라우드 서버가 이용되는 분산 컴퓨팅 환경에서, 신경망의 분할 위치를 탐색하여, 연합 디지털 트윈의 효율적인 동기화를 지원한다. 본 발명의 일실시예는 메모리 및 메모리와 전기적으로 연결된 프로세서를 포함하는 컴퓨팅 장치에서 수행될 수 있다.
- [0028] 본 발명의 일실시예에 따른 컴퓨팅 장치는 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치를 입력받아, 후보 분할 위치별로 신경망의 동기화 시간을 계산한다. 여기서 신경망은, 디지털 트윈 각각에 대응되는 실물 객체에 대한 센싱 데이터 등의 정보를 입력받아, 실물 객체의 위치나 동작을 예측하는 등의 태스크를 수행하는 신경망일 수 있다. 신경망은, 디지털 트윈 각각에 대해 할당되며, 신경망에서 예측된 위치나 동작은 디지털 트윈에 반영된다. 그리고 후보 분할 위치는 신경망의 레이어를 나누는 위치로서, 실시예에 따라서는 다양하게 결정될 수 있으며, 신경망의 레이어의 개수를 초과하지 않는다.
- [0029] 컴퓨팅 장치는 후보 분할 위치에 따라서 신경망을 분할하여, 동기화 시간을 계산할 수 있다. 동기화 시간은 분할된 신경망 중 입력 노드를 포함하는 제1분할 신경망에 대한 연산이 엣지 서버에서 수행되는 시간, 제1분할 신경망의 출력값이 엣지 서버에서 클라우드 서버로 전송되는 시간, 클라우드 서버가 분할된 신경망 중 출력 노드를 포함하는 제2분할 신경망에 제1분할 신경망의 출력값을 입력하여, 제2분할 신경망에 대한 연산을 수행하는 시간을 포함할 수 있다.
- [0030] 그리고 컴퓨팅 장치는 신경망 각각의 동기화 시간 중 최소 시간을 탐색하고, 최소 시간 중 최대 시간에 대응되는 후보 분할 위치를 디지털 트윈의 최종 분할 위치로 결정한다. 연합 디지털 트윈의 동기화 시간을 줄이기 위해, 동기화 시간 중 최소 시간을 고려하되, 병렬적으로 처리되는 연합 디지털 트윈에 대한 동기화가 완료되는 시간은, 각각의 동기화 시간 중 최대 시간에 대응되므로, 컴퓨팅 장치는 최소 시간 중 최대 시간에 대응되는 후보 분할 위치를 디지털 트윈의 최종 분할 위치로 결정할 수 있다. 실시예에 따라서, 신경망의 최종 분할 위치는 동일하거나, 전부 또는 일부가 다르도록 결정될 수 있다.
- [0031] 컴퓨팅 장치는 실시예에 따라서, 조기 종료 모델이 추가로 이용되는 환경에서, 신경망의 분할 위치를 탐색할 수 있다. 이 경우 컴퓨팅 장치는 컴퓨팅 장치는 후보 분할 위치 뿐만 아니라, 후보 연산 조기 종료 위치까지 고려하여, 디지털 트윈별 동기화 시간을 계산할 수 있다.
- [0032] 도 1과 같이, 디지털 트윈의 동기화를 위해, 5개의 레이어(111 내지 115)로 구성된 신경망이 이용되는 경우, 제1 내지 제4지점(121 내지 124)이 후보 분할 위치와 후보 연산 조기 종료 위치로 이용될 수 있다. 제3지점(123)이 최종 분할 위치로 결정되는 경우, 제1 내지 제3레이어(121 내지 123)에 대한 연산은 엣지 서버에서 수행되고, 제4 내지 제5레이어(124, 125)에 대한 연산은 클라우드 서버에서 수행될 수 있다. 이 때, 연산 조기 종료 위치가 제3지점(123)이라면, 클라우드 서버에서의 연산은 수행되지 않고, 연산 조기 종료 위치가 제4지점(124)이라면, 클라우드 서버에서는 제4레이어(124)에 대한 연산만이 수행될 수 있다.
- [0033] 컴퓨팅 장치는 일실시예로서, [수학식 1]을 이용해, 동기화 시간을 계산할 수 있다.

수학식 1

$$t_u(p_u, e_u) = t_u^E(p_u, e_u) + t_u^{tr}(p_u, e_u) + t_u^C(p_u, e_u)$$

- [0035] 여기서 u 는 신경망, 즉 디지털 트윈에 대한 인덱스, p_u 는 후보 분할 위치, e_u 는 후보 연산 조기 종료 위치,

$t_u^E(p_u, e_u)$ 는 제1분할 신경망의 연산이 엣지 서버에서 수행되는 시간, $t_u^{tr}(p_u, e_u)$ 는 제1분할 신경망의 출력값이 엣지 서버에서 클라우드 서버로 전송되는 시간, $t_u^C(p_u, e_u)$ 는 제2분할 신경망의 연산이 클라우드 서버에서 수행하는 시간을 나타낸다.

- [0036] 그리고 동기화 시간 중 최대 시간, 즉 연합 디지털 트윈의 동기화가 모두 완료되는 시간은, [수학식 2]와 같이

표현될 수 있다.

수학식 2

$$t_u(p, e) = \max_{u \in U} \{t_u(p_u, e_u)\}$$

[0037]

[0038] 초기 종료 모델이 이용되는 경우, 신경망 분할 위치는 연산 초기 종료 지점에 영향을 받기 때문에, 컴퓨팅 장치는 동기화 시간 중 최대 시간의 기대값 중 최소 기대값에 대응되는 후보 분할 위치를 최종 분할 위치로 결정할 수 있다.

[0039] 본 발명의 일실시예에 따르면, 연합 디지털 트윈에 대한 동기화 시간이 감소하여, 효율적으로 연합 디지털 트윈의 동기화가 수행될 수 있다.

[0041] 도 2는 본 발명의 일실시예에 따른 연합 디지털 트윈 동기화 방법을 설명하기 위한 도면이다.

[0042] 도 2를 참조하면, 본 발명의 일실시예에 따른 컴퓨팅 장치는 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 연합 디지털 트윈에 대한 후보 분할 위치 벡터를 생성하고, 신경망에 대한 후보 연산 초기 종료 위치로부터, 연합 디지털 트윈에 대한 후보 연산 초기 종료 위치 벡터를 생성(S210)한다. 후보 분할 위치 벡터의 원소 각각은, 신경망 각각의 후보 분할 위치로부터, 신경망 별로 선택된 하나의 후보 분할 위치를 포함하며, 후보 연산 초기 종료 위치 벡터의 원소 각각은 신경망 각각의 후보 연산 초기 종료 위치로부터, 신경망 별로 선택된 하나의 후보 연산 초기 종료 위치를 포함한다.

[0043] 예컨대, 연합 디지털 트윈에 포함된 디지털 트윈이 2개이고, 디지털 트윈 각각에 할당된 제1 및 제2신경망의 레이어가 도 1과 같이 5개라면, 제1 및 제2신경망의 후보 분할 위치와 후보 연산 초기 종료 위치는 모두 4개일 수 있다. 따라서, 후보 분할 위치 벡터 및 후보 연산 초기 종료 위치 벡터 각각은, 총 16개의 원소를 포함한다. 그리고 후보 분할 위치 벡터의 각 원소는 2개의 후보 분할 위치를 포함하고, 후보 연산 초기 종료 위치 벡터의 각 원소 역시 2개의 후보 연산 초기 종료 위치를 포함한다.

[0044] 컴퓨팅 장치는 후보 분할 위치 벡터의 원소와, 후보 연산 초기 종료 위치 벡터의 원소의 조합 각각에 대해, 디지털 트윈 별로 동기화 시간을 계산(S220)한다. 전술된 예시에서, 후보 분할 위치 벡터의 원소와, 후보 연산 초기 종료 위치 벡터의 원소의 조합은 총 256개가 결정될 수 있으며, 2개의 디지털 트윈이 이용되므로, 총 512개의 동기화 시간이 계산될 수 있다.

[0045] 컴퓨팅 장치는 조합 각각에 대해, 디지털 트윈 별 동기화 시간 중 최대 시간에 대한 기대값을 계산(S230)하며, [수학식 1]과 같이 각 조합에 대해 동기화 시간을 계산하여, 최대 시간을 탐색할 수 있다. 컴퓨팅 장치는 전술된 예시에서, 디지털 트윈 별로 동기화 시간 2개 중 최대 시간을 탐색하되, 총 256개의 조합 각각에 대해 최대 시간을 탐색하여, 총 256개의 기대값을 계산할 수 있다.

[0046] 그리고 전술된 바와 같이, 신경망 분할 위치는 연산 초기 종료 지점에 영향을 받기 때문에, 컴퓨팅 장치는 후보 연산 초기 종료 위치 벡터의 원소에 포함된 후보 연산 초기 종료 위치에서의 초기 종료 확률값을 이용하여, 기대값을 계산한다. 후보 연산 초기 종료 위치에서의 초기 종료 확률값은, 신경망에 대한 시뮬레이션을 통해 획득될 수 있다.

[0047] 컴퓨팅 장치는 일실시예로서, [수학식 3]과 같이, 기대값을 계산할 수 있다.

수학식 3

$$\sum_{e \in E} x^e \cdot \max_{u \in U} \{t_u(p_u, e_u)\}$$

[0048]

[0049] E는 후보 연산 초기 종료 위치 벡터, e는 후보 연산 초기 종료 위치 벡터의 원소,

$\max_{u \in U} \{t_u(p_u, e_u)\}$ 는 최대 시간을 나타낸다. 그리고 x^e 는 후보 연산 초기 종료 위치 벡터의 원소에 포함된 후보 연산 초기 종료 위치에서의 초기 종료 확률값을 곱한 값을 나타낸다. 전술된 예시에서, 2개의 후보 연산 초기 종료 위치에 대한 초기 종료 확률값이 곱해져, 총 16개의 x^e 가 계산될 수 있다. 컴퓨팅 장치는 16개의 x^e 와 최대 시간을 곱한 값의 합을, 기대값으로 산출할 수 있다.

[0050] 그리고 컴퓨팅 장치는 기대값 중 최소 기대값에 대응되는 조합의 후보 분할 위치를 최종 분할 위치로 결정(S240)한다. 최종 분할 위치에 의해 분할된 신경망 중 입력 노드를 포함하는 제1분할 신경망은 엣지 서버에 할당되고, 출력 노드를 포함하는 제2분할 신경망은 클라우드 서버에 할당될 수 있다.

[0052] 도 3은 본 발명의 다른 실시예에 따른 연합 디지털 트윈 동기화 방법을 설명하기 위한 도면이다.

[0053] 도 3에서 설명되는 연합 디지털 트윈 동기화 방법은, 도 2에서 설명된 실시예보다, 최종 분할 위치를 탐색하는데 소요되는 비용을 줄이기 위한 방법이다.

[0054] 도 3을 참조하면, 본 발명의 일 실시예에 따른 컴퓨팅 장치는 연합 디지털 트윈에 포함된 디지털 트윈 별 신경망에 대한 후보 분할 위치로부터, 디지털 트윈 각각에 대한 후보 분할 위치 벡터를 생성하고, 신경망에 대한 후보 연산 초기 종료 위치로부터, 디지털 트윈 각각에 대한 후보 연산 초기 종료 위치 벡터를 생성(S310)한다. 전술된 실시예에서는 후보 분할 위치 벡터 및 후보 연산 초기 종료 위치 벡터가 연합 디지털 트윈에 대해서 생성되는데 반해, 도 3에서는 디지털 트윈 별로 후보 분할 위치와 후보 연산 초기 종료 위치 벡터가 생성된다.

[0055] 예컨대 연합 디지털 트윈에 포함된 디지털 트윈이 2개이고, 디지털 트윈 각각에 할당된 제1 및 제2신경망의 레이어가 도 1과 같이 5개라면, 제1 및 제2신경망의 후보 분할 위치와 후보 연산 초기 종료 위치는 모두 4개일 수 있다. 따라서, 제1신경망에 대한 후보 분할 위치 벡터 및 후보 연산 초기 종료 위치 벡터 각각은, 총 4개의 원소를 포함하며, 제1신경망에 대한 후보 분할 위치 벡터 및 후보 연산 초기 종료 위치 벡터 각각 역시 4개의 원소를 포함할 수 있다.

[0056] 컴퓨팅 장치는 디지털 트윈 별로, 후보 분할 위치 벡터의 원소와, 후보 연산 초기 종료 위치 벡터의 원소의 조합 각각에 대한 동기화 시간을 계산(S320)한다. 전술된 예시에서, 후보 분할 위치 벡터의 원소와, 후보 연산 초기 종료 위치 벡터의 원소의 조합은 총 16개로 결정될 수 있으며, 디지털 트윈 별로 총 16개의 동기화 시간이 계산될 수 있다.

[0057] 그리고 컴퓨팅 장치는 디지털 트윈 각각에 대해, 동기화 시간에 대한 기대값을 계산(S330)한다. 전술된 예시에서 디지털 트윈 별로 16개의 기대값이 계산되어 총 32개의 기대값이 계산될 수 있다.

[0058] 컴퓨팅 장치는 신경망 각각에 대한 후보 연산 초기 종료 위치 각각에서의 초기 종료 확률값을 이용하여, 기대값을 계산하며, 일 실시예로서, [수학식 4]를 이용하여 기대값을 계산할 수 있다.

수학식 4

$$\sum_{e_u \in E_u} x_u^{e_u} \cdot t_u^E(p_u, e_u) + \sum_{e_u \in E_u} x_u^{e_u} \cdot t_u^{tr}(p_u, e_u) + \sum_{e_u \in E_u} x_u^{e_u} \cdot t_u^C(p_u, e_u)$$

[0060] 여기서, u는 디지털 트윈의 인덱스, p는 후보 분할 위치, e는 후보 연산 초기 종료 위치, E는 후보 연산 초기

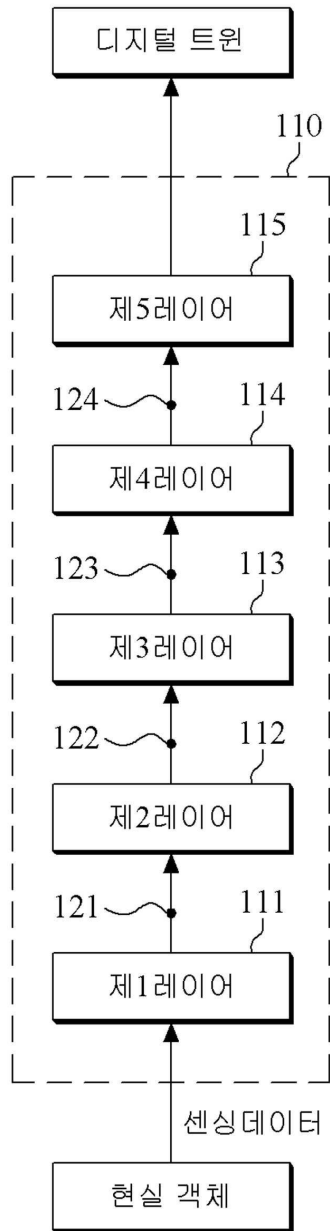
종료 위치 벡터, $x_u^{e_u}$ 는 후보 연산 초기 종료 위치에서의 초기 종료 확률값을 나타낸다. 전술된 예시에서, 디지털 트윈 별로 4개의 후보 연산 초기 종료 위치가 이용되므로, 후보 분할 위치 별로 4개의 후보 연산 초기 종

료 위치에 대한 조기 종료 확률값이 이용된다.

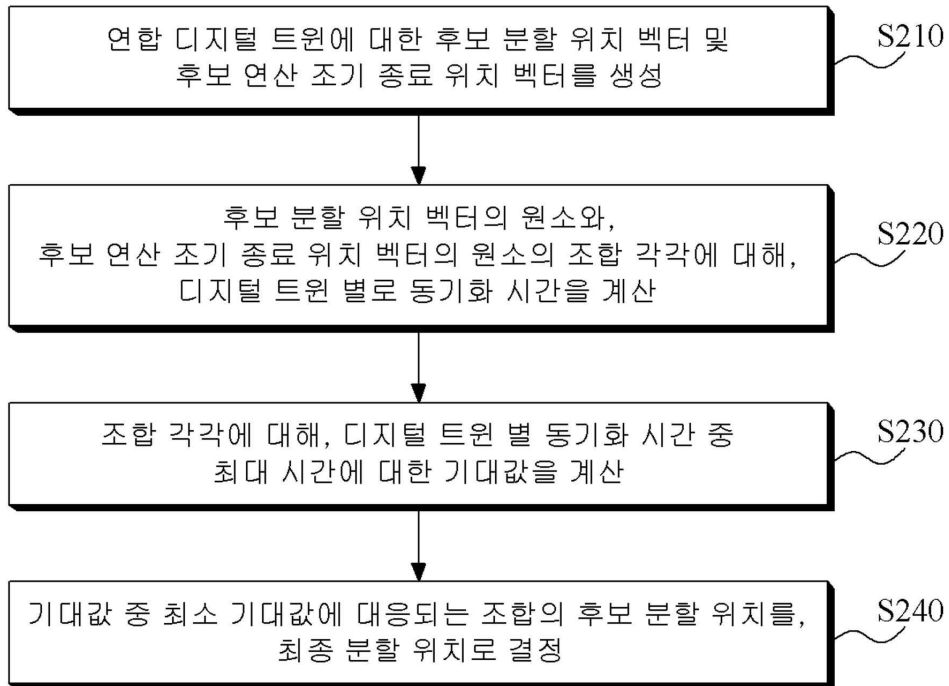
- [0061] 컴퓨팅 장치는 기대값 중 최소 기대값에 대응되는 후보 분할 위치를 디지털 트윈 각각의 신경망에 대한 최종 분할 위치로 결정(S340)한다. 그리고 최종 분할 위치에 의해 분할된 신경망 중 입력 노드를 포함하는 제1분할 신경망은 엣지 서버에 할당되고, 출력 노드를 포함하는 제2분할 신경망은 클라우드 서버에 할당된다.
- [0062] 도 3의 실시예에 따르면, 계산되는 기대값의 개수가 감소하여, 도 2의 실시예와 비교하여, 동기화 소요 비용이 감소함을 알 수 있다.
- [0064] 도 4 및 도 5는 본 발명의 일실시예에 따른 연합 디지털 트윈 동기화 방법의 수행 결과를 설명하기 위한 도면이다.
- [0065] 도 4에서는, 엣지-클라우드 서버 분할 컴퓨팅 구조 및 조기 종료 모델이 이용되고, 다섯 개의 후보 연산 조기 종료 위치와 여섯 개의 후보 분할 위치가 존재하는 신경망이 이용되는 환경에서, 세개의 디지털 트윈을 포함하는 연합 디지털 트윈에 대한 동기화 수행 결과가 설명된다. 디지털 트윈의 동기화 작업으로는, CNN 모델을 통한 CIFA-10 이미지 세트에 대해 이미지 분류 작업이 사용되었고, 이미지 분류 작업에 일반적으로 사용되는 신경망 모델 중 ResNet 모델이 사용되었다.
- [0066] 그리고 도 4에서, OA는 도 2에서 설명된 실시예, HPA는 도 3에서 설명된 실시예, Random은 디지털 트윈의 분할 위치가 랜덤으로 결정되어 동기화가 수행되는 방법, Edge-only는 신경망의 분할없이 레이어에 대한 모든 연산이 엣지 서버에서만 진행되어 동기화가 수행되는 방법, Cloud-Only는 신경망의 분할없이 레이어에 대한 모든 연산이 클라우드 서버에서만 진행되는 동기화가 수행되는 방법을 나타낸다. 그리고 Edge는 엣지 서버에서 소요되는 시간, Tr은 데이터 전송에 소요되는 시간, Cloud는 클라우드 서버에서 소요되는 시간을 나타낸다.
- [0067] 도 4 및 도 5의 수행 결과는 총 1000번의 전체 동기화 과정을 각 알고리즘을 통해 진행한 결과로서, 전체 1000번의 동기화 과정에 대해 일반화한 평균 동기화 수행 시간 값(NATC)이 도시된다. 일반화한 평균 동기화 수행 시간 값(NATC)은 매 동기화 과정마다 각 알고리즘의 결과값을 다섯 개의 알고리즘들 중 가장 큰 수행 시간 값을 뽑아 나누어 준 값을 의미한다. 이 때, 도 5는 도 4와 비교하여, 엣지 서버와 클라우드 서버 사이의 데이터 전송 시간이 상대적으로 긴 환경에서의 수행 결과를 나타낸다.
- [0068] 도 4 및 도 5에 도시된 바와 같이, 본 발명의 일실시예에 따른 동기화 방법(OA, HPA)에 의할 경우, 가장 빠르게 동기화가 수행됨을 알 수 있다.
- [0070] 앞서 설명한 기술적 내용들은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예들을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 하드웨어 장치는 실시예들의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0072] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

도면

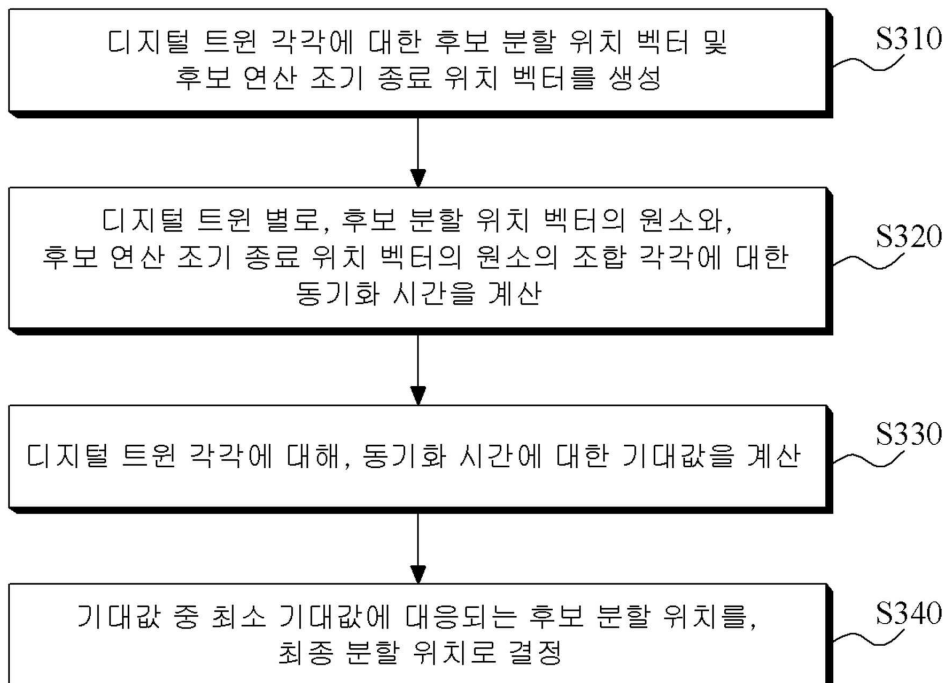
도면1



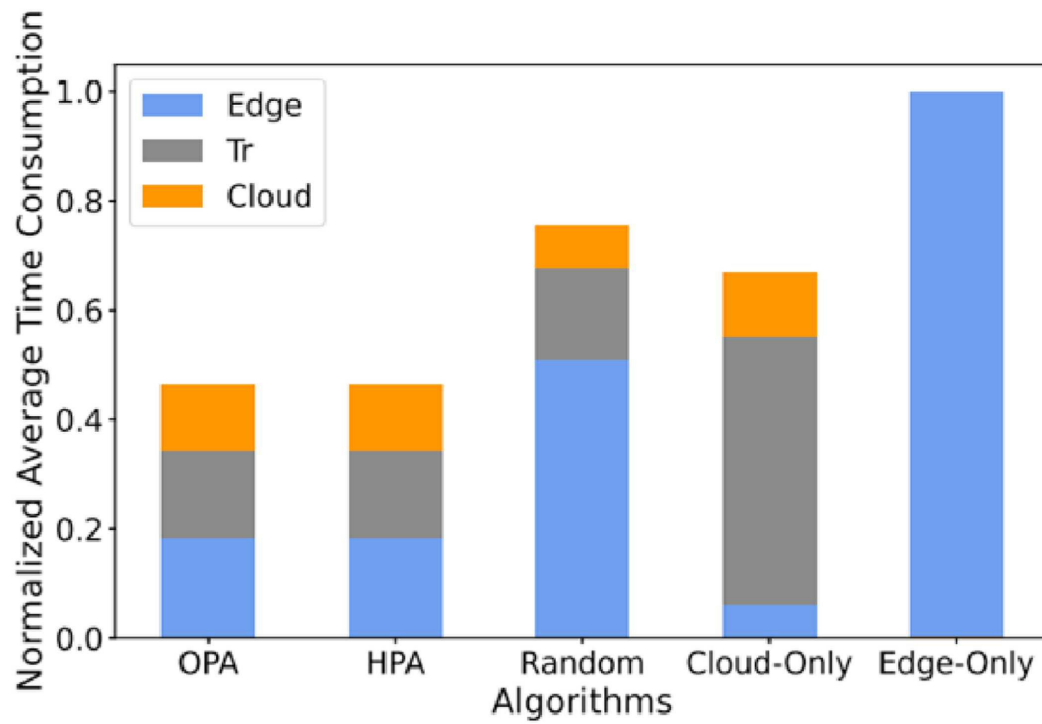
도면2



도면3



도면4



도면5

