



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0081555
(43) 공개일자 2024년06월10일

- | | |
|---|---|
| <p>(51) 국제특허분류(Int. Cl.)
G06N 3/098 (2023.01) G06N 3/0464 (2023.01)</p> <p>(52) CPC특허분류
G06N 3/098 (2023.01)
G06N 3/0464 (2023.01)</p> <p>(21) 출원번호 10-2022-0163804</p> <p>(22) 출원일자 2022년11월30일
심사청구일자 2022년11월30일</p> | <p>(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)</p> <p>(72) 발명자
박대영
인천광역시 연수구 송도문화로84번길 24, 208동 3104호(송도동, 송도 에듀포레 푸르지오)</p> <p>김민식
경기도 수원시 영통구 동수원로 448, 105동 402호(매탄동, 현대아파트)</p> <p>(74) 대리인
양성보</p> |
|---|---|

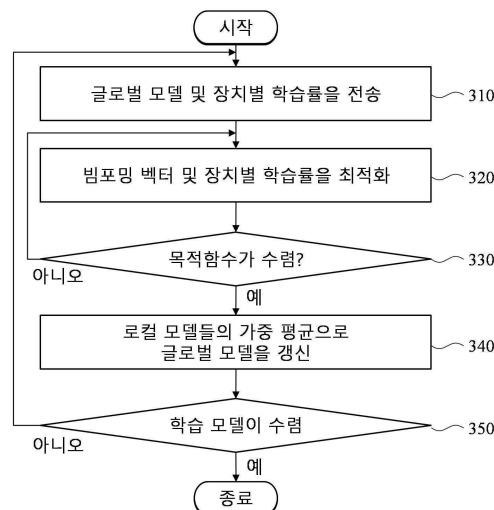
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **연합 학습을 위한 무선 계산에서 빔포밍 벡터와 학습률을 동시에 최적화하는 방법 및 장치**

(57) 요약

연합 학습을 위한 무선 계산에서 빔포밍 벡터와 학습률을 동시에 최적화하는 방법 및 장치를 개시한다. 일실시예에 따른 최적화 방법은 글로벌 모델 및 장치별 학습률을 복수의 로컬 장치들로 전송하는 제1 단계, 빔포밍 벡터 및 상기 장치별 학습률을 최적화하는 제2 단계, 목적함수가 수렴할 때까지 상기 제2 단계를 반복하는 제3 단계, 상기 복수의 로컬 장치들로부터 수신된 로컬 모델들의 가중 평균으로 상기 글로벌 모델을 갱신하는 제4 단계 및 현재 빔포밍 벡터와 이전 빔포밍 벡터간의 차이가 기설정된 값 이하로 수렴할 때까지 상기 제1 단계 내지 상기 제4 단계를 반복 수행하는 제5 단계를 포함할 수 있다.

대 표 도 - 도3



이 발명을 지원한 국가연구개발사업

과제고유번호	1711174241
과제번호	2022R1H1A2092541
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	제도약연구(중견연구)
연구과제명	[Ezbaro] 대규모 연결을 위한 사물인터넷 시스템 최적화 연구
기 여 율	80/100
과제수행기관명	인하대학교
연구기간	2022.09.01 ~ 2023.08.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152759
과제번호	2020-0-01389-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합연구센터지원사업(국고)
연구과제명	[Ezbaro][정부] 인공지능융합연구센터지원(3차년도)
기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711174118
과제번호	RS-2022-00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신인재양성(교육운영수익)
연구과제명	[Ezbaro][정부+민간] 인공지능융합혁신인재양성(1차년도)
기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.07.01 ~ 2022.12.31

명세서

청구범위

청구항 1

적어도 하나의 프로세서를 포함하는 컴퓨터 장치의 최적화 방법에 있어서,

상기 적어도 하나의 프로세서에 의해, 글로벌 모델 및 장치별 학습물을 복수의 로컬 장치들로 전송하는 제1 단계;

상기 적어도 하나의 프로세서에 의해, 빔포밍 벡터 및 상기 장치별 학습물을 최적화하는 제2 단계;

상기 적어도 하나의 프로세서에 의해, 목적함수가 수렴할 때까지 상기 제2 단계를 반복하는 제3 단계;

상기 적어도 하나의 프로세서에 의해, 상기 복수의 로컬 장치들로부터 수신된 로컬 모델들의 가중 평균으로 상기 글로벌 모델을 갱신하는 제4 단계; 및

상기 적어도 하나의 프로세서에 의해, 현재 빔포밍 벡터와 이전 빔포밍 벡터간의 차이가 기설정된 값 이하로 수렴할 때까지 상기 제1 단계 내지 상기 제4 단계를 반복 수행하는 제5 단계

를 포함하는 최적화 방법.

청구항 2

제1항에 있어서,

상기 복수의 로컬 장치 각각은, 상기 장치별 학습물을 반영하여 상기 글로벌 모델을 업데이트하는 과정을 주어진 에포크만큼 반복 수행하여 로컬 모델을 생성하도록 구현되는 것을 특징으로 하는 최적화 방법.

청구항 3

제1항에 있어서,

상기 제2 단계는,

상기 복수의 로컬 장치로부터 전송된 신호가 집계된 정보 및 상기 복수의 로컬 장치로부터 수신된 로컬 모델들이 집계된 정보간의 평균 제곱 오차(mean squared error, MSE)를 최소화하는 빔포밍 벡터 및 동적 학습물을 통해 상기 빔포밍 벡터 및 상기 장치별 학습물을 최적화하는 것을 특징으로 하는 최적화 방법.

청구항 4

제3항에 있어서,

상기 신호가 집계된 정보는 로컬 장치와 상기 컴퓨터 장치간의 채널 정보, 간섭제거를 위한 스칼라 변수, 로컬 모델을 위한 송신 신호 및 가우시안 잡음에 의해 계산되는 수신신호와 빔포밍 벡터간의 곱에 기반하여 얻어지는 것을 특징으로 하는 최적화 방법.

청구항 5

제3항에 있어서,

상기 로컬 모델들이 집계된 정보는 상기 복수의 로컬 장치 각각에 대한 동적 학습물의 역과 로컬 모델의 곱에 대한 평균에 기반하여 얻어지는 것을 특징으로 하는 최적화 방법.

청구항 6

제3항에 있어서,

상기 제2 단계는,

상기 빔포밍 벡터와 상기 동적 학습물로 라그랑지안을 최소화한 후 쌍대 변수로 최대화하는 과정을 반복하여 상

기 빔포밍 벡터 및 상기 장치별 학습물을 최적화하는 것을 특징으로 하는 최적화 방법.

청구항 7

컴퓨터 장치와 결합되어 제1항 내지 제6항 중 어느 한 항의 방법을 상기 컴퓨터 장치에 실행시키기 위해 컴퓨터 판독 가능한 기록매체에 저장되는 컴퓨터 프로그램.

청구항 8

제1항 내지 제6항 중 어느 한 항의 방법을 컴퓨터 장치에 실행시키기 위한 컴퓨터 프로그램이 기록되어 있는 컴퓨터 판독 가능한 기록매체.

청구항 9

컴퓨터 장치에 있어서,

상기 컴퓨터 장치에서 판독 가능한 명령을 실행하도록 구현되는 적어도 하나의 프로세서
를 포함하고,

상기 적어도 하나의 프로세서에 의해,

글로벌 모델 및 장치별 학습물을 복수의 로컬 장치들로 전송하는 제1 과정;

빔포밍 벡터 및 상기 장치별 학습물을 최적화하는 제2 과정;

목적함수가 수렴할 때까지 상기 제2 과정을 반복하는 제3 과정;

상기 복수의 로컬 장치들로부터 수신된 로컬 모델들의 가중 평균으로 상기 글로벌 모델을 갱신하는 제4 과정;
및

현재 빔포밍 벡터와 이전 빔포밍 벡터간의 차이가 기설정된 값 이하로 수렴할 때까지 상기 제1 과정 내지 상기 제4 과정을 반복 수행하는 제5 과정

을 처리하는 것을 특징으로 하는 컴퓨터 장치.

청구항 10

제9항에 있어서,

상기 제2 과정에서,

상기 복수의 로컬 장치로부터 전송된 신호가 집계된 정보 및 상기 복수의 로컬 장치로부터 수신된 로컬 모델들이 집계된 정보간의 평균 제곱 오차(mean squared error, MSE)를 최소화하는 빔포밍 벡터 및 동적 학습물을 통해 상기 빔포밍 벡터 및 상기 장치별 학습물을 최적화하는 것

을 특징으로 하는 컴퓨터 장치.

발명의 설명

기술 분야

[0001] 아래의 설명은 연합 학습을 위한 무선 계산에서 빔포밍 벡터와 학습물을 동시에 최적화하는 방법 및 장치에 관한 것이다.

배경 기술

[0002]연합 학습(federated learning, FL)은 원시 데이터(raw data)를 서버와 공유하지 않고 각 장치에서 학습된 모델 정보를 사용하여 서버에서 인공지능 모델을 학습하는 분산형 방식이다. 연합 학습의 과정은 다음과 같이 진행된다. 먼저 서버는 글로벌 모델을 사용 가능한 수많은 장치에 전송한다. 각각의 장치들은 보유하고 있는 자신의 로컬 데이터로 인공지능 모델을 업데이트하고 업데이트된 로컬 모델을 서버로 전송한다. 마지막으로 서버에서는 로컬 모델들의 가중 평균으로 글로벌 모델을 업데이트하며 일련의 과정을 모델이 수렴할 때까지 반복한다. 따라서 연합 학습은 기존의 중앙집중식 방법이 가지는 개인정보, 서버구축 비용, 제한된 무선자원 등의 문제를

해결할 수 있다. 그러나 연합 학습에서 인공지능 모델의 성능을 높이기 위해서는 수많은 장치가 학습에 참여해야 하므로 무선 자원에 대한 문제가 여전히 남아있다.

[0003] 다중 접속 채널(multiple access channel, MAC)에서 중첩의 원리를 활용한 무선 계산(over-the-air computation, 이하 'AirComp')은 연합 학습의 특성을 활용할 수 있는 새로운 통신 기법으로 각광받고 있다. 연합 학습의 서버는 로컬 모델들의 가중 평균으로 글로벌 모델을 업데이트 하기 때문에 각각의 로컬 모델의 매개변수 또는 기울기가 아닌 집계된 값에만 관심이 있다. 따라서 송신 신호를 공기와 같은 무선 매체에서 집계하여 수신하는 AirComp는 연합 학습에서 높은 스펙트럼 효율과 낮은 대기 시간을 달성할 수 있다.

[0004] 머신 러닝 관점에서 학습률(learning rate)은 학습 성능과 수렴성에 큰 영향을 미칠 수 있는 중요한 요소이다. 연합 학습을 위한 AirComp에서 빔포밍 벡터와 학습률을 동시에 최적화하면 통신에서 발생하는 집계 오류를 최소화하고 학습 성능을 높일 수 있다. 하지만 집계 오류를 최소화하는 빔포밍 벡터와 학습률을 동시에 찾는 문제는 변수가 결합되어 있고, 비볼록 제한조건을 가지기 때문에 해결하기 어렵다. 기존의 방식은 SDR(Semi-Definite Relaxation) 기반의 알고리즘으로 문제를 해결하지만 준 최적의 해를 찾아 성능이 좋지 않으며 높은 계산 복잡도로 인해 통신과 학습 사이에 병목 현상이 발생한다.

[0005] [선행문헌번호]

[0006] 한국공개특허 제10-2022-0025663호

발명의 내용

해결하려는 과제

[0007] 연합 학습을 위한 무선 계산에서 빔포밍 벡터와 학습률을 동시에 최적화하는 방법 및 장치를 제공한다.

과제의 해결 수단

[0008] 적어도 하나의 프로세서를 포함하는 컴퓨터 장치의 최적화 방법에 있어서, 상기 적어도 하나의 프로세서에 의해, 글로벌 모델 및 장치별 학습률을 복수의 로컬 장치들로 전송하는 제1 단계; 상기 적어도 하나의 프로세서에 의해, 빔포밍 벡터 및 상기 장치별 학습률을 최적화하는 제2 단계; 상기 적어도 하나의 프로세서에 의해, 목적함수가 수렴할 때까지 상기 제2 단계를 반복하는 제3 단계; 상기 적어도 하나의 프로세서에 의해, 상기 복수의 로컬 장치들로부터 수신된 로컬 모델들의 가중 평균으로 상기 글로벌 모델을 갱신하는 제4 단계; 및 상기 적어도 하나의 프로세서에 의해, 현재 빔포밍 벡터와 이전 빔포밍 벡터간의 차이가 기설정된 값 이하로 수렴할 때까지 상기 제1 단계 내지 상기 제4 단계를 반복 수행하는 제5 단계를 포함하는 최적화 방법을 제공한다.

[0009] 일측에 따르면, 상기 복수의 로컬 장치 각각은, 상기 장치별 학습률을 반영하여 상기 글로벌 모델을 업데이트하는 과정을 주어진 에포크만큼 반복 수행하여 로컬 모델을 생성하도록 구현되는 것을 특징으로 할 수 있다.

[0010] 다른 측면에 따르면, 상기 제2 단계는, 상기 복수의 로컬 장치로부터 전송된 신호가 집계된 정보 및 상기 복수의 로컬 장치로부터 수신된 로컬 모델들이 집계된 정보간의 평균 제곱 오차(mean squared error, MSE)를 최소화하는 빔포밍 벡터 및 동적 학습률을 통해 상기 빔포밍 벡터 및 상기 장치별 학습률을 최적화하는 것을 특징으로 할 수 있다.

[0011] 또 다른 측면에 따르면, 상기 신호가 집계된 정보는 로컬 장치와 상기 컴퓨터 장치간의 채널 정보, 간섭제거를 위한 스칼라 변수, 로컬 모델을 위한 송신 신호 및 가우시안 잡음에 의해 계산되는 수신신호와 빔포밍 벡터간의 곱에 기반하여 얻어지는 것을 특징으로 할 수 있다.

[0012] 또 다른 측면에 따르면, 상기 로컬 모델들이 집계된 정보는 상기 복수의 로컬 장치 각각에 대한 동적 학습률의 역과 로컬 모델의 곱에 대한 평균에 기반하여 얻어지는 것을 특징으로 할 수 있다.

[0013] 또 다른 측면에 따르면, 상기 제2 단계는, 상기 빔포밍 벡터와 상기 동적 학습률로 라그랑지안(Lagrangian)을 최소화한 후 쌍대 변수로 최대화하는 과정을 반복하여 상기 빔포밍 벡터 및 상기 장치별 학습률을 최적화하는 것을 특징으로 할 수 있다.

[0014] 컴퓨터 장치와 결합되어 상기 방법을 컴퓨터 장치에 실행시키기 위해 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램을 제공한다.

[0015] 상기 방법을 컴퓨터 장치에 실행시키기 위한 프로그램이 기록되어 있는 컴퓨터 판독 가능한 기록매체를 제공한

다.

[0016] 컴퓨터 장치에 있어서, 상기 컴퓨터 장치에서 관독 가능한 명령을 실행하도록 구현되는 적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서에 의해, 글로벌 모델 및 장치별 학습률을 복수의 로컬 장치들로 전송하는 제1 과정; 빔포밍 벡터 및 상기 장치별 학습률을 최적화하는 제2 과정; 목적함수가 수렴할 때까지 상기 제2 과정을 반복하는 제3 과정; 상기 복수의 로컬 장치들로부터 수신된 로컬 모델들의 가중 평균으로 상기 글로벌 모델을 갱신하는 제4 과정; 및 현재 빔포밍 벡터와 이전 빔포밍 벡터간의 차이가 기설정된 값 이하로 수렴할 때까지 상기 제1 과정 내지 상기 제4 과정을 반복 수행하는 제5 과정을 처리하는 것을 특징으로 하는 컴퓨터 장치를 제공한다.

발명의 효과

[0017] 연합 학습을 위한 무선 계산에서 빔포밍 벡터와 학습률을 동시에 최적화하는 방법 및 장치를 제공할 수 있다.

도면의 간단한 설명

[0018] 도 1은 본 발명의 일실시예에 따른 네트워크 환경의 예를 도시한 도면이다.

도 2는 본 발명의 일실시예에 따른 컴퓨터 장치의 예를 도시한 블록도이다.

도 3은 본 발명의 일실시예에 있어서, 기지국의 최적화 방법의 예를 도시한 흐름도이다.

도 4는 본 발명의 일실시예에 있어서, 로컬 장치의 최적화 방법의 예를 도시한 흐름도이다.

도 5는 연합 학습에 참여하는 전체 장치 수에 따른 집계 오류에 대한 그래프이다.

도 6은 통신 라운드에 따른 인공지능 모델의 학습 성능에 대한 그래프이다.

도 7은 다양한 데이터 및 인공지능 모델에서의 학습 성능을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0019] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다. 하기 설명 및 첨부 도면에서 본 발명의 요지를 불필요하게 흐릴 수 있는 공지 기능 및 구성에 대한 상세한 설명은 생략한다.

[0020] 본 발명의 실시예들에 따른 최적화 장치는 적어도 하나의 컴퓨터 장치에 의해 구현될 수 있다. 이때, 컴퓨터 장치에는 본 발명의 일실시예에 따른 컴퓨터 프로그램이 설치 및 구동될 수 있고, 컴퓨터 장치는 구동된 컴퓨터 프로그램의 제어에 따라 본 발명의 실시예들에 따른 최적화 방법을 수행할 수 있다. 상술한 컴퓨터 프로그램은 컴퓨터 장치와 결합되어 최적화 방법을 컴퓨터에 실행시키기 위해 컴퓨터 판독 가능한 기록매체에 저장될 수 있다.

[0021] 도 1은 본 발명의 일실시예에 따른 네트워크 환경의 예를 도시한 도면이다. 도 1의 네트워크 환경은 복수의 전자 기기들(110, 120, 130, 140), 복수의 서버들(150, 160) 및 네트워크(170)를 포함하는 예를 나타내고 있다. 이러한 도 1은 발명의 설명을 위한 일례로 전자 기기의 수나 서버의 수가 도 1과 같이 한정되는 것은 아니다. 또한, 도 1의 네트워크 환경은 본 실시예들에 적용 가능한 환경들 중 하나의 예를 설명하는 것일 뿐, 본 실시예들에 적용 가능한 환경이 도 1의 네트워크 환경으로 한정되는 것은 아니다.

[0022] 복수의 전자 기기들(110, 120, 130, 140)은 컴퓨터 장치로 구현되는 고정형 단말이거나 이동형 단말일 수 있다. 복수의 전자 기기들(110, 120, 130, 140)의 예를 들면, 스마트폰(smart phone), 휴대폰, 네비게이션, 컴퓨터, 노트북, 디지털방송용 단말, PDA(Personal Digital Assistants), PMP(Portable Multimedia Player), 태블릿 PC 등이 있다. 일례로 도 1에서는 전자 기기(110)의 예로 스마트폰의 형상을 나타내고 있으나, 본 발명의 실시예들에서 전자 기기(110)는 실질적으로 무선 또는 유선 통신 방식을 이용하여 네트워크(170)를 통해 다른 전자 기기들(120, 130, 140) 및/또는 서버(150, 160)와 통신할 수 있는 다양한 물리적인 컴퓨터 장치들 중 하나를 의미할 수 있다.

[0023] 통신 방식은 제한되지 않으며, 네트워크(170)가 포함할 수 있는 통신망(일례로, 이동통신망, 유선 인터넷, 무선 인터넷, 방송망)을 활용하는 통신 방식뿐만 아니라 기기들간의 근거리 무선 통신 역시 포함될 수 있다. 예를 들어, 네트워크(170)는, PAN(personal area network), LAN(local area network), CAN(campus area network), MAN(metropolitan area network), WAN(wide area network), BBN(broadband network), 인터넷 등의 네트워크 중 하나 이상의 임의의 네트워크를 포함할 수 있다. 또한, 네트워크(170)는 버스 네트워크, 스타 네트워크, 링

네트워크, 메시 네트워크, 스타-버스 네트워크, 트리 또는 계층적(hierarchical) 네트워크 등을 포함하는 네트워크 토폴로지 중 임의의 하나 이상을 포함할 수 있으나, 이에 제한되지 않는다.

[0024] 서버(150, 160) 각각은 복수의 전자 기기들(110, 120, 130, 140)과 네트워크(170)를 통해 통신하여 명령, 코드, 파일, 콘텐츠, 서비스 등을 제공하는 컴퓨터 장치 또는 복수의 컴퓨터 장치들로 구현될 수 있다. 예를 들어, 서버(150)는 네트워크(170)를 통해 접속한 복수의 전자 기기들(110, 120, 130, 140)로 서비스를 제공하는 시스템일 수 있다.

[0025] 도 2는 본 발명의 일실시예에 따른 컴퓨터 장치의 예를 도시한 블록도이다. 앞서 설명한 복수의 전자 기기들(110, 120, 130, 140) 각각이나 서버들(150, 160) 각각은 도 2를 통해 도시된 컴퓨터 장치(200)에 의해 구현될 수 있다.

[0026] 이러한 컴퓨터 장치(200)는 도 2에 도시된 바와 같이, 메모리(210), 프로세서(220), 통신 인터페이스(230) 그리고 입출력 인터페이스(240)를 포함할 수 있다. 메모리(210)는 컴퓨터에서 판독 가능한 기록매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 여기서 ROM과 디스크 드라이브와 같은 비소멸성 대용량 기록장치는 메모리(210)와는 구분되는 별도의 영구 저장 장치로서 컴퓨터 장치(200)에 포함될 수도 있다. 또한, 메모리(210)에는 운영체제와 적어도 하나의 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 메모리(210)와는 별도의 컴퓨터에서 판독 가능한 기록매체로부터 메모리(210)로 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록매체를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록매체가 아닌 통신 인터페이스(230)를 통해 메모리(210)에 로딩될 수도 있다. 예를 들어, 소프트웨어 구성요소들은 네트워크(170)를 통해 수신되는 파일들에 의해 설치되는 컴퓨터 프로그램에 기반하여 컴퓨터 장치(200)의 메모리(210)에 로딩될 수 있다.

[0027] 프로세서(220)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(210) 또는 통신 인터페이스(230)에 의해 프로세서(220)로 제공될 수 있다. 예를 들어 프로세서(220)는 메모리(210)와 같은 기록 장치에 저장된 프로그램 코드에 따라 수신되는 명령을 실행하도록 구성될 수 있다.

[0028] 통신 인터페이스(230)는 네트워크(170)를 통해 컴퓨터 장치(200)가 다른 장치(일례로, 앞서 설명한 저장 장치들)와 서로 통신하기 위한 기능을 제공할 수 있다. 일례로, 컴퓨터 장치(200)의 프로세서(220)가 메모리(210)와 같은 기록 장치에 저장된 프로그램 코드에 따라 생성한 요청이나 명령, 데이터, 파일 등이 통신 인터페이스(230)의 제어에 따라 네트워크(170)를 통해 다른 장치들로 전달될 수 있다. 역으로, 다른 장치로부터의 신호나 명령, 데이터, 파일 등이 네트워크(170)를 거쳐 컴퓨터 장치(200)의 통신 인터페이스(230)를 통해 컴퓨터 장치(200)로 수신될 수 있다. 통신 인터페이스(230)를 통해 수신된 신호나 명령, 데이터 등은 프로세서(220)나 메모리(210)로 전달될 수 있고, 파일 등은 컴퓨터 장치(200)가 더 포함할 수 있는 저장 매체(상술한 영구 저장 장치)로 저장될 수 있다.

[0029] 입출력 인터페이스(240)는 입출력 장치(250)와의 인터페이스를 위한 수단일 수 있다. 예를 들어, 입력 장치는 마이크, 키보드 또는 마우스 등의 장치를, 그리고 출력 장치는 디스플레이, 스피커와 같은 장치를 포함할 수 있다. 다른 예로 입출력 인터페이스(240)는 터치스크린과 같이 입력과 출력을 위한 기능이 하나로 통합된 장치와의 인터페이스를 위한 수단일 수도 있다. 입출력 장치(250) 중 적어도 하나는 컴퓨터 장치(200)와 하나의 장치로 구성될 수도 있다. 예를 들어, 스마트폰과 같이 터치스크린, 마이크, 스피커 등이 컴퓨터 장치(200)에 포함된 형태로 구현될 수 있다.

[0030] 또한, 다른 실시예들에서 컴퓨터 장치(200)는 도 2의 구성요소들보다 더 적은 혹은 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 컴퓨터 장치(200)는 상술한 입출력 장치(250) 중 적어도 일부를 포함하도록 구현되거나 또는 트랜시버(transceiver), 데이터베이스 등과 같은 다른 구성요소들을 더 포함할 수도 있다.

[0031] 연합 학습에서 개선된 FedAvg(federated averaging) 방법은 각 장치에서 데이터를 전송하는 대신 k 번째 장치에서 업데이트된 로컬 모델 $\mathbf{v}_k$ 들의 가중 평균으로 기지국으로부터 수신된 글로벌 모델의 파라미터(가중치) $\mathbf{w}$ 를 업데이트 할 수 있다. 이때, 각 장치에서 업데이트된 로컬 모델은 아래 수학적 1과 같이 나타낼 수 있다.

[0032] [수학식 1]

$$\mathbf{v}_k = \mathbf{w} - \mu_k r_k \nabla F_k(\mathbf{w}, \mathbf{B}_k)$$

[0033]

[0034] 여기에서 μ_k 는 고정 학습률, r_k 는 동적 학습률(dynamic learning rate, DLR), $F_k(\mathbf{w}, \mathbf{B}_k)$ 는 모델 파라미터 $\mathbf{w}$ 와 미니 배치(mini batch) $\mathbf{B}_k$ 가 주어졌을 때의 손실 함수(loss function)일 수 있다. 각각의 장치는 업데이트 과정을 정해진 에포크(epoch)만큼 반복할 수 있다. 서버는 업데이트된 로컬 모델로부터 다음 수학식 2와 같은 글로벌 모델을 얻을 수 있다.

[0035] [수학식 2]

$$\mathbf{w} = \frac{1}{K} \sum_{k=1}^K \frac{1}{r_k} \mathbf{v}_k$$

[0036]

[0037] 여기에서 글로벌 모델은 $\frac{1}{K} \sum_{k=1}^K \frac{1}{r_k} = 1$ 일 때 수립하며 $r_k = 1$ 이면 기존의 FedAvg 알고리즘에 대응될 수 있다.

[0038] 무선 계산(over-the-air computation, 이하 'AirComp')에서 각각의 장치들은 제한된 전력 내에서 간섭 제거를 위한 추가 전력 $\mathbf{b}_k$ 을 가지고 신호를 송신할 수 있다. 이때 기지국에서 수신된 신호는 아래 수학식 3과 같이 나타낼 수 있다.

[0039] [수학식 3]

$$\mathbf{y} = \sum_{k=1}^K \mathbf{H}_k \mathbf{b}_k x_k + \mathbf{n}$$

[0040]

[0041] 여기서, $\mathbf{H}_k \in \mathbb{C}^{N \times M}$ 는 k 번째 장치와 기지국 간의 채널 정보, $\mathbf{b}_k \in \mathbb{C}^M$ 는 간섭제거를 위한 스칼라 변수, x_k 는 로컬 모델 $\mathbf{v}_k$ 을 위한 송신 신호, $\mathbf{n} \in \mathbb{C}^N$ 은 평균이 0, 분산이 σ 인 가우시안 잡음을 나타낼 수 있다.

[0042] 기지국에서는 빔포밍 벡터 $\mathbf{m} \in \mathbb{C}^N$ 와 수신신호 $\mathbf{y} \in \mathbb{C}^N$ 를 곱해서 송신 신호 x_k 들이 집계된 정보를 아래 수학식 4와 같이 얻을 수 있다.

[0043] [수학식 4]

$$\sqrt{\eta} \mathbf{m}^H \mathbf{y} = \sqrt{\eta} \mathbf{m}^H \sum_{k=1}^K \mathbf{H}_k \mathbf{b}_k x_k + \sqrt{\eta} \mathbf{m}^H \mathbf{n}$$

[0044]

[0045] 여기에서 η 는 각 장치의 송신 전력이 P_k 보다 작다는 조건을 위한 스칼라 변수일 수 있다.

[0046] AirComp에서 각각의 장치로부터 전송된 신호가 집계된 정보 $\sqrt{\eta} \mathbf{m}^H \mathbf{y}$ 는 연합 학습에서 로컬 모델들이 집계된 정보 $\frac{1}{K} \sum_{k=1}^K \frac{1}{r_k} x_k$ 와 같아야 하므로 아래 수학식 5와 같이 나타나는 평균 제곱 오차(mean squared error, MSE)를 최소화할 수 있다.

[0047] [수학식 5]

$$\text{MSE} = \mathbb{E} \left[\left| \sum_{k=1}^K \left(\sqrt{\eta} \mathbf{m}^H \mathbf{H}_k \mathbf{b}_k - \frac{1}{K r_k} \right) x_k \right|^2 \right] + \eta \sigma^2 \|\mathbf{m}\|^2$$

[0048]

[0049] 영 강압(zero-forcing)으로 구한 $\mathbf{b}_k = \frac{\mathbf{H}_k^H \mathbf{m}}{\eta K r_k \|\mathbf{H}_k^H \mathbf{m}\|^2}$ 와 각 장치에서의 전력 제한조건 $\|\mathbf{b}_k\|^2 \leq P_{\text{th}}$ 로 구한 $\eta = \max_k \frac{1}{K^2 P_k r_k^2 \|\mathbf{H}_k^H \mathbf{m}\|^2}$ 를 수학식 5에 적용해서 얻은 집계 오류는 다음 수학식 6과 같다.

[0050] [수학식 6]

$$\text{MSE} = \frac{\sigma^2}{K^2} \max_k \frac{\|\mathbf{m}\|^2}{P_k r_k^2 \|\mathbf{H}_k^H \mathbf{m}\|^2}$$

[0051]

[0052] 따라서 집계 오류로서의 MSE를 최소화하는 빔포밍 벡터와 학습률을 찾는 문제는 아래 수학식 7과 같이 나타낼 수 있다.

[0053] [수학식 7]

$$(P1) \min_{\mathbf{m} \in \mathbb{C}^N, \mathbf{s} \in \mathcal{S}} \max_k \frac{s_k^2 \|\mathbf{m}\|^2}{\|\mathbf{G}_k^H \mathbf{m}\|^2}$$

[0054]

[0055] 여기에서 $\mathbf{G}_k = \sqrt{P_k} \mathbf{H}_k$, $\mathbf{s} = [s_1 \dots s_K]^T \in \mathbb{R}^K$, $s_k = \frac{1}{r_k}$ 이고 $\mathcal{S} = \{\mathbf{s} | \sum_{k=1}^K s_k = K, s_{\min} \leq s_k \leq s_{\max}, \forall k\}$ 일 수 있다. 이 문제는 $s_{\min} \leq 1 \leq s_{\max}$ 에서 항상 실현가능(feasible)할 수 있다.

[0056] 수학식 7에서 $\mathbf{m}$ 에 0이 아닌 상수 배를 해도 목적함수 $f_k(\mathbf{m}) = \frac{s_k^2 \|\mathbf{m}\|^2}{\|\mathbf{G}_k^H \mathbf{m}\|^2}$ 의 값은 변하지 않는다는 점을 이용하면 수학식 7에서 최대화에 대한 식을 없애고 다음 수학식 8과 같이 정리할 수 있다.

[0057] [수학식 8]

$$(P2) \min_{\mathbf{m} \in \mathbb{C}^N, \mathbf{s} \in \mathcal{S}} \|\mathbf{m}\|^2$$

$$\text{s.t.} \quad \|\mathbf{G}_k^H \mathbf{m}\|^2 \geq s_k^2, \forall k$$

[0058]

[0059] 수학식 8에서 제한조건은 비볼록으로 해결하기 어렵다. 따라서 MM(majorization minimization) 기법에 기반하여 연속된 볼록 문제들로 제한조건을 아래 수학식 9와 같이 재구성할 수 있다. 제한조건에서 $\|\mathbf{G}_k^H \mathbf{m}\|^2$ 은 $\mathbf{m}$ 에 대해서 볼록 함수이므로 테일러 급수를 1차항까지 전개한 값을 하한으로 가질 수 있다.

[0060] [수학식 9]

$$\|\mathbf{G}_k^H \mathbf{m}\|^2 \geq \|\mathbf{G}_k^H \mathbf{m}^{(l)}\|^2 + \text{Re} \left\{ 2(\mathbf{m}^{(l)})^H \mathbf{G}_k \mathbf{G}_k^H (\mathbf{m} - \mathbf{m}^{(l)}) \right\}$$

$$= 2\text{Re} \left\{ (\mathbf{m}^{(l)})^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{m} \right\} - \|\mathbf{G}_k^H \mathbf{m}^{(l)}\|^2 \equiv \tilde{\varphi}_k(\mathbf{m}; \mathbf{m}^{(l)})$$

[0061]

[0062] 여기에서 $\mathbf{m}^{(l)}$ 은 반복 알고리즘으로부터 구한 l 번째의 값이다. 수학식 9를 수학식 8에 적용하면 아래 수학식 10을 얻을 수 있다.

[0063] [수학식 10]

$$(P3) (\mathbf{m}^{(l+1)}, \mathbf{s}^{(l+1)}) = \arg \max_{\mathbf{m}, \mathbf{s} \in \mathcal{S}} \|\mathbf{m}\|^2$$

$$\text{s.t.} \quad \tilde{\varphi}_k(\mathbf{m}; \mathbf{m}^{(l)}) \geq s_k^2, \forall k$$

[0064]

[0065] 이러한 수학식 10은 볼록 최적화 문제로 CVX 툴박스(toolbox)를 사용해서 내부점법(interior-point method)으로 풀 수 있지만 높은 복잡도를 가진다. 따라서 본 실시예에서는 복잡도를 줄이기 위해 PSM(Projected Subgradient Method)을 적용할 수 있다. 먼저 수학식 10의 라그랑지안(Lagrangian)은 아래 수학식 11과 같이 나타낼 수 있다.

[0066] [수학식 11]

[0067]

[0068]

여기에서 $\boldsymbol{\lambda} = [\lambda_1 \cdots \lambda_K]^T$ 는 제한조건에 대한 쌍대 변수(dual variable)이다. 수학식 11은 원 문제의 변수(primal variable) $\mathbf{m}, \mathbf{s}$ 에 대해서 최소화한 후 쌍대 변수 $\boldsymbol{\lambda}$ 에 대해서 최대화를 하는 것으로 해를 구할 수 있다. 먼저 라그랑지안은 $\mathbf{m}$ 에 대해서 이차 함수(quadratic function)로 최적의 $\mathbf{m}$ 은 기울기의 값이 0이 되도록 하는 값이다. 다음으로 라그랑지안을 최소화하는 $\mathbf{s}$ 는 아래 표 1과 같은 방법으로 구할 수 있다.

[0069] [표 1]

Algorithm 1 Bisection Algorithm for Learning Rate Optimization.

```

1: Input:  $\lambda_k, \forall k$ 
2:  $\mathcal{Z} \leftarrow \{k | \lambda_k > 0\}$ 
3: If  $|\mathcal{Z}|s_{\min} + (K - |\mathcal{Z}|)s_{\max} \geq K$ 
4:    $s_k \leftarrow \begin{cases} s_{\min}, & \text{if } k \in \mathcal{Z} \\ (K - |\mathcal{Z}|s_{\min}) / (K - |\mathcal{Z}|), & \text{otherwise} \end{cases}, \forall k$ 
5: Else
6:    $s_k \leftarrow s_{\max}$  for all  $k \notin \mathcal{Z}$ 
7:   Initialize  $\nu_{\min}, \nu_{\max}$ 
8:   While  $\nu_{\max} - \nu_{\min} > \epsilon$ 
9:      $\nu = (\nu_{\min} + \nu_{\max}) / 2$ 
10:    For  $k \in \mathcal{Z}$ 
11:       $s_k \leftarrow \min \left( \max \left( \frac{\nu}{2\lambda_k}, s_{\min} \right), s_{\max} \right)$ 
12:    End
13:    If  $\sum_{k=1}^K s_k > K$ 
14:       $\nu_{\max} \leftarrow \nu$ 
15:    Else
16:       $\nu_{\min} \leftarrow \nu$ 
17:    End
18:  End
19: End
```

[0070]

[0071]

상술한 표 1에서 ν 는 $\sum_{k=1}^K s_k = K$ 조건을 위한 상수이며 이분법(bisection algorithm)으로 최적의 ν 값과 제한조건을 만족하는 s_k 를 구할 수 있다. 마지막으로 수학식 11에서 $\boldsymbol{\lambda}$ 를 PSM으로 구하면 아래 수학식 12와 같이 나타낼 수 있다.

[0072] [수학식 12]

[0073]

$$\lambda_k = \max(\lambda_k - \delta_j(\tilde{\varphi}_k(\mathbf{m}; \mathbf{m}^{(l)}) - s_k^2), 0), \forall k$$

[0074]

여기에서 δ_j 는 쌍대 변수 갱신을 위한 스텝 사이즈(step size)일 수 있다. 제안하는 저 복잡도 알고리즘(DLR-Subgradient)의 전체 과정은 아래 표 2와 같으며 반복 알고리즘들은 ϵ 이하의 값으로 수렴할 때까지 동작할 수 있다.

[0075] [표 2]

Algorithm 2 Beamforming Vector Design for Federated Learning based on the Subgradient Method.

```

1: Input:  $\mathbf{G}_k, \forall k$ 
2: Initialize  $\mathbf{m}^{(0)}, \mathbf{m}^{(-1)} \leftarrow \mathbf{0}, l \leftarrow 0$ 
3: While  $\|\mathbf{m}^{(l)} - \mathbf{m}^{(l-1)}\| > \epsilon_1$ 
4:    $\mathbf{m} \leftarrow \mathbf{m}^{(l)}, \mathbf{m}_{\text{old}} \leftarrow \mathbf{0}$ 
5:    $\lambda_k \leftarrow 1, \forall k, j \leftarrow 1$ 
6:   While  $\|\mathbf{m} - \mathbf{m}_{\text{old}}\| > \epsilon_2$ 
7:      $\mathbf{m}_{\text{old}} \leftarrow \mathbf{m}$ 
8:      $\mathbf{m} \leftarrow \sum_{k=1}^K \lambda_k \mathbf{G}_k \mathbf{G}_k^H \mathbf{m}^{(l)}$ 
9:     Find  $s_k, \forall k$  using Algorithm 1
10:     $\lambda_k \leftarrow \max(\lambda_k - \delta_j(\tilde{\varphi}_k(\mathbf{m}; \mathbf{m}^{(l)}) - s_k^2), 0), \forall k$ 
11:     $j \leftarrow j + 1$ 
12:   End
13:    $\mathbf{m}^{(l+1)} \leftarrow \mathbf{m}$ 
14:    $l \leftarrow l + 1$ 
15: End

```

[0076]

[0077]

도 3은 본 발명의 일실시예에 있어서, 기지국의 최적화 방법의 예를 도시한 흐름도이다. 본 실시예에 따른 최적화 방법은 서버로서의 기지국을 구현하는 컴퓨터 장치(200)에 의해 수행될 수 있다. 일례로, 컴퓨터 장치(200)의 프로세서(220)는 메모리(210)가 포함하는 운영체제의 코드나 적어도 하나의 컴퓨터 프로그램의 코드에 따른 제어 명령(instruction)을 실행하도록 구현될 수 있다. 여기서, 프로세서(220)는 컴퓨터 장치(200)에 저장된 코드가 제공하는 제어 명령에 따라 컴퓨터 장치(200)가 도 3의 방법이 포함하는 단계들(310 내지 350)을 수행하도록 컴퓨터 장치(200)를 제어할 수 있다.

[0078]

단계(310)에서 컴퓨터 장치(200)는 글로벌 모델 및 장치별 학습물을 전송할 수 있다. 일례로, 하나의 기지국이 60개의 로컬 장치와 통신하는 경우를 고려하면, 기지국은 60개의 로컬 장치 각각으로 글로벌 모델과 60개의 로컬 장치 각각에 대한 장치별 학습물을 전송할 수 있다.

[0079]

단계(320)에서 컴퓨터 장치(200)는 빔포밍 벡터 및 장치별 학습물을 최적화할 수 있다. 일례로, 컴퓨터 장치(200)는 복수의 로컬 장치로부터 전송된 신호가 집계된 정보 및 복수의 로컬 장치로부터 수신된 로컬 모델들이 집계된 정보간의 평균 제곱 오차(mean squared error, MSE)를 최소화하는 빔포밍 벡터 및 동적 학습물을 통해 빔포밍 벡터 및 장치별 학습물을 최적화할 수 있다. 여기서, 신호가 집계된 정보는 로컬 장치와 컴퓨터 장치간의 채널 정보, 간섭제거를 위한 스칼라 변수, 로컬 모델을 위한 송신 신호 및 가우시안 잡음에 의해 계산되는 수신신호와 빔포밍 벡터간의 곱에 기반하여 얻어질 수 있다. 수신신호에 대해서는 앞서 수학적 3을 통해 설명한 바 있다. 로컬 모델들이 집계된 정보는 복수의 로컬 장치 각각에 대한 동적 학습물의 역과 로컬 모델의 곱에 대한 평균에 기반하여 얻어질 수 있다. 이러한 로컬 모델들이 집계된 정보에 대해서는 앞서 수학적 2를 통해 설명한 바 있다. 또한, 컴퓨터 장치(200)는 빔포밍 벡터와 동적 학습물로 라그랑지안을 최소화한 후 쌍대 변수로 최대화하는 과정을 반복하여 빔포밍 벡터 및 장치별 학습물을 최적화할 수 있다. 앞서, 수학적 11, 표 1 및 수학적 12를 통해 설명한 바와 같이 라그랑지안을 최소화하는 빔포밍 벡터와 동적 학습물의 해를 찾은 후 라그랑주 듀얼 함수(Lagrange dual function)를 최대화하는 쌍대 변수의 해를 찾는 과정을 반복하여 빔포밍 벡터와 장치별 학습물의 최적화가 이루어질 수 있음에 대해 설명하였다.

[0080]

단계(330)에서 컴퓨터 장치(200)는 목적함수가 수렴하는지 여부를 결정할 수 있다. 컴퓨터 장치(200)는 목적함수가 수렴하는 경우 단계(330)를 수행할 수 있으며, 목적함수가 수렴하지 않는 경우에는 단계(320)를 반복 수행할 수 있다. 이때, 목적함수가 수렴하는지 여부는 표 2에 나타난 바와 같이 빔포밍 벡터와 이전 반복에서의 빔

포밍 벡터의 차이가 ϵ_2 이하의 값으로 수렴하는지 여부를 의미할 수 있다.

- [0081] 단계(340)에서 컴퓨터 장치(200)는 로컬 모델들의 가중 평균으로 글로벌 모델을 갱신할 수 있다. 단계(310)에서 로컬 장치 각각으로 전송된 글로벌 모델은 이후 도 4에서 설명하는 바와 같이 로컬 장치 각각에서 장치별 학습률을 반영하여 업데이트될 수 있으며, 로컬 장치 각각에서 업데이트된 글로벌 모델로서 로컬 모델들이 생성될 수 있다. 단계(340)에서 컴퓨터 장치(200)는 로컬 장치 각각에서 로컬 장치 각각이 생성한 로컬 모델들을 수신할 수 있으며, 수신된 로컬 모델들의 가중 평균으로 글로벌 모델을 갱신할 수 있다. 일례로, 앞서 수학적 2를 통해 업데이트된 로컬 모델로부터 글로벌 모델을 얻을 수 있음을 설명한 바 있다.
- [0082] 단계(350)에서 컴퓨터 장치(200)는 학습 모델이 수렴하는지 여부를 결정할 수 있다. 학습 모델은 단계(340)에서 갱신된 글로벌 모델일 수 있으며, 이러한 학습 모델이 수렴하는지 테스트 데이터로 학습 모델을 평가했을 때 목표 성능을 달성하는지 또는 정해진 반복만큼 학습이 진행되었는지 여부를 의미할 수 있다. 컴퓨터 장치(200)는 학습 모델이 수렴하는 경우 도 3의 최적화 방법에 따른 과정을 종료할 수 있으며, 학습 모델이 수렴하지 않는 경우 단계(310)부터의 과정을 재수행할 수 있다.
- [0083] 도 4는 본 발명의 일실시예에 있어서, 로컬 장치의 최적화 방법의 예를 도시한 흐름도이다. 본 실시예에 따른 최적화 방법은 로컬 장치를 구현하는 컴퓨터 장치(200)에 의해 수행될 수 있다. 일례로, 컴퓨터 장치(200)의 프로세서(220)는 메모리(210)가 포함하는 운영체제의 코드나 적어도 하나의 컴퓨터 프로그램의 코드에 따른 제어 명령(instruction)을 실행하도록 구현될 수 있다. 여기서, 프로세서(220)는 컴퓨터 장치(200)에 저장된 코드가 제공하는 제어 명령에 따라 컴퓨터 장치(200)가 도 4의 방법이 포함하는 단계들(410 내지 440)을 수행하도록 컴퓨터 장치(200)를 제어할 수 있다.
- [0084] 단계(410)에서 컴퓨터 장치(200)는 글로벌 모델 및 장치별 학습률을 수신할 수 있다. 일례로, 컴퓨터 장치(200)는 도 3의 단계(310)에서와 같이, 서버로서의 기지국이 전송하는 글로벌 모델과 장치별 학습률을 수신할 수 있다.
- [0085] 단계(420)에서 컴퓨터 장치(200)는 장치별 학습률을 반영하여 글로벌 모델을 업데이트하여 로컬 모델을 생성할 수 있다. 예를 들어, 컴퓨터 장치(200)는 앞서 수학적 1을 이용하여 로컬 모델을 생성할 수 있다.
- [0086] 단계(430)에서 컴퓨터 장치(200)는 주어진 에포크만큼 단계(420)를 반복할 수 있다. 일례로, 에포크가 10으로 주어진 경우, 컴퓨터 장치(200)는 단계(420)를 10회 반복하여 최종적으로 로컬 모델을 생성할 수 있다.
- [0087] 단계(440)에서 컴퓨터 장치(200)는 로컬 모델을 전송할 수 있다. 로컬 모델은 서버로서의 기지국으로 전송될 수 있다. 일례로, 도 3의 단계(340)에서 기지국이 개별 로컬 장치들로부터 로컬 모델들을 수신할 수 있으며, 수신된 로컬 모델들의 가중 평균으로 글로벌 모델을 갱신할 수 있다.
- [0088] 이하에서는 제안된 최적화 방법 및 장치의 성능을 검증하기 위해 기존 알고리즘 DLR-DC(Dynamic Learning Rate - Difference of Convex)와 집계 오류가 없다고 가정된 이상적인 알고리즘 Oracle-AirComp, 성능의 하한을 위해 랜덤하게 빔포밍 벡터를 선택하는 SLR(Signal-to-Leakage Ratio)-Random beamforming, 수학적 10을 CVX toolbox로 해결한 DLR-CVX, 표 2에 해당하는 DLR-Subgradient와 학습률이 미치는 영향을 알기 위해 s_k 를 1로 고정된 SLR-Subgradient의 성능을 비교한다.
- [0089] 도 5는 기지국의 안테나 개수가 16개, 각 로컬 장치의 안테나 개수가 2개 일 때 전체 장치 수에 따른 집계 오류를 보여주는 그래프이다. DLR-DC는 SLR-Subgradient보다 많은 집계 오류를 가진다. DLR-Subgradient는 저 복잡도에서 DLR-CVX의 성능과 일치하며 학습률을 최적화하지 않는 SLR-Subgradient보다 높은 성능을 달성한다. 또한 장치의 개수가 60개보다 많은 경우 제안하는 알고리즘들은 기존 알고리즘 DLR-DC보다 5dB이상의 이득을 달성한다.
- [0090] 도 6는 도 5와 같은 상황에서 전체 장치 수가 60개일 때 통신 라운드에 따른 인공지능 모델의 학습 성능을 보여준다. 이 실험에서는 MNIST(Modified National Institute of Standards and Technology) 데이터 셋을 사용하였으며 다층 퍼셉트론(multi-layer perceptron, MLP)으로 모델의 분류 성능을 평가하였다. 1번의 통신 라운드는 연합 학습이 1번 진행되는 모든 과정을 포함한다. SLR-Subgradient와 DLR-DC는 높은 집계 오류로 인해서 통신 환경이 좋지 않은 상황이 발생하면 인공지능 모델의 성능이 떨어진다. DLR-CVX과 DLR-Subgradient는 이상적인 알고리즘 Oracle-AirComp과 거의 동일한 성능을 달성한다.
- [0091] 도 7은 다양한 데이터와 학습모델을 가지고 실험하였을 때 100번째 통신 라운드에서의 학습 성능을 보여주는 표

이다. Fashion-MNIST 데이터 셋으로 합성곱 신경망(Convolutional Neural Networks, CNN)을 학습할 때 DLR-DC는 다른 알고리즘보다 약 10%p 낮은 예측 정확도를 가진다. 빔포밍 벡터와 학습물을 함께 최적화하는 제안 알고리즘은 모든 경우에 1%p 이내에서 Oracle-AirComp와 유사한 성능을 달성한다.

[0092] 이와 같이 본 발명의 실시예들에 따르면, 연합 학습을 위한 무선 계산에서 빔포밍 벡터와 학습물을 동시에 최적화하는 방법 및 장치를 제공할 수 있다.

[0093] 이상에서 설명된 시스템 또는 장치는 하드웨어 구성요소, 또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 어플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0094] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록매체에 저장될 수 있다.

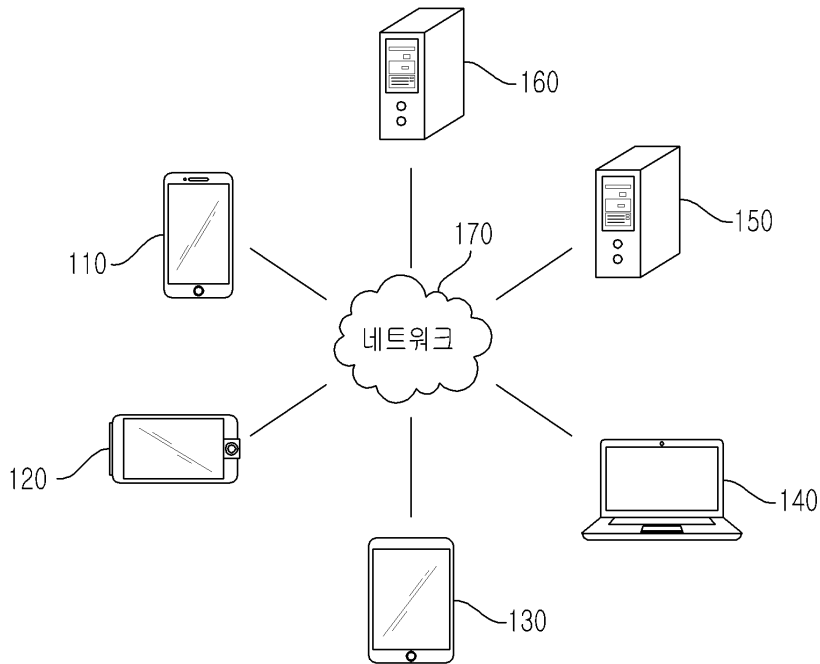
[0095] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수개 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 애플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0096] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

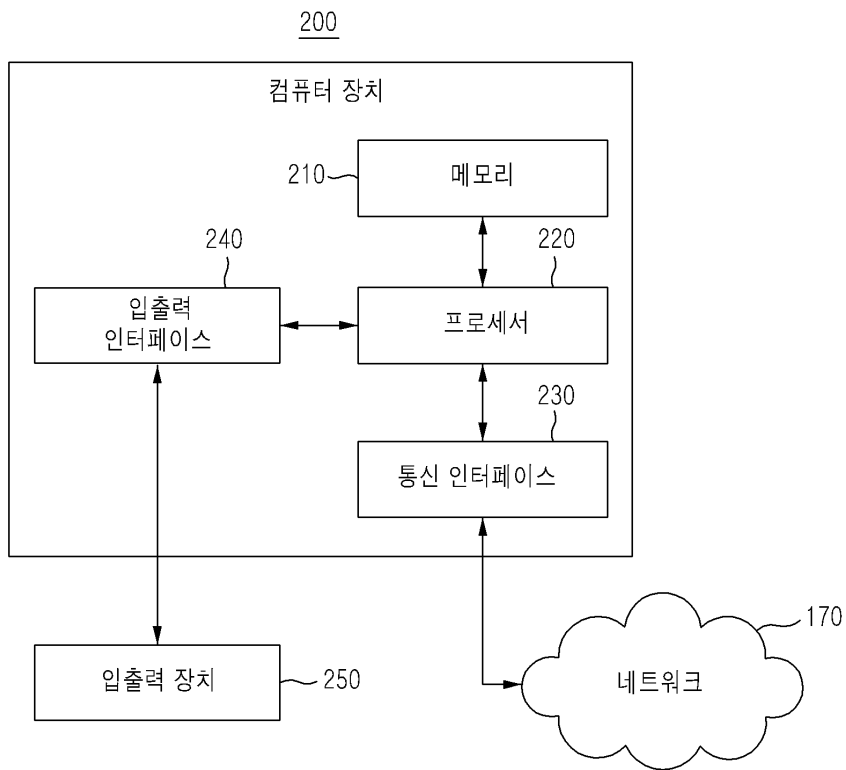
[0097] 그러므로, 다른 구현들, 다른 실시예들 및 청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

도면

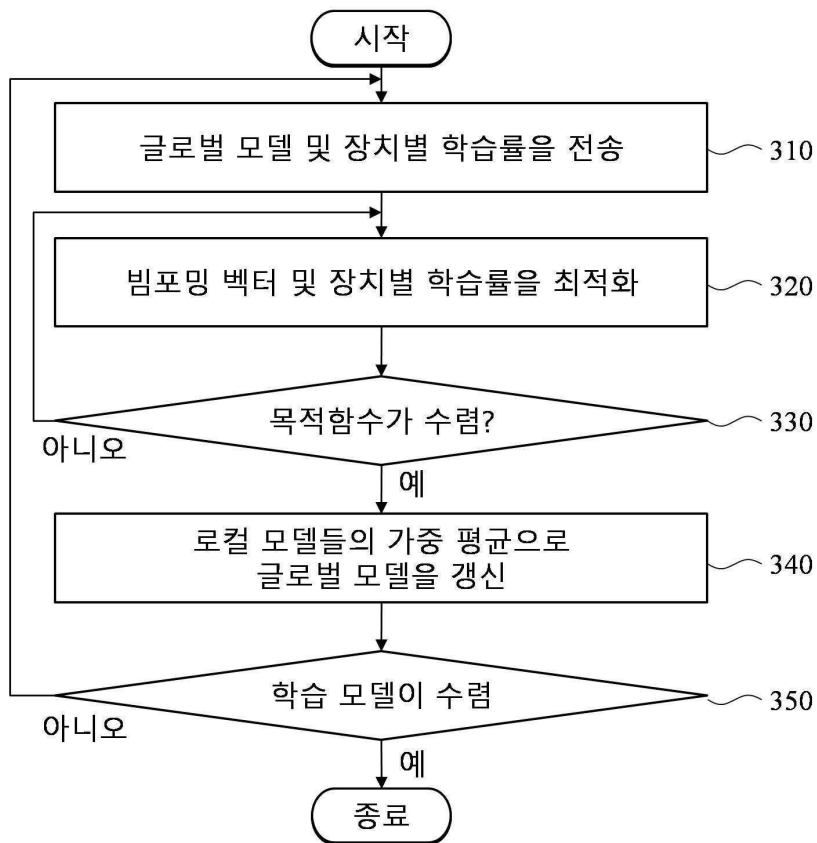
도면1



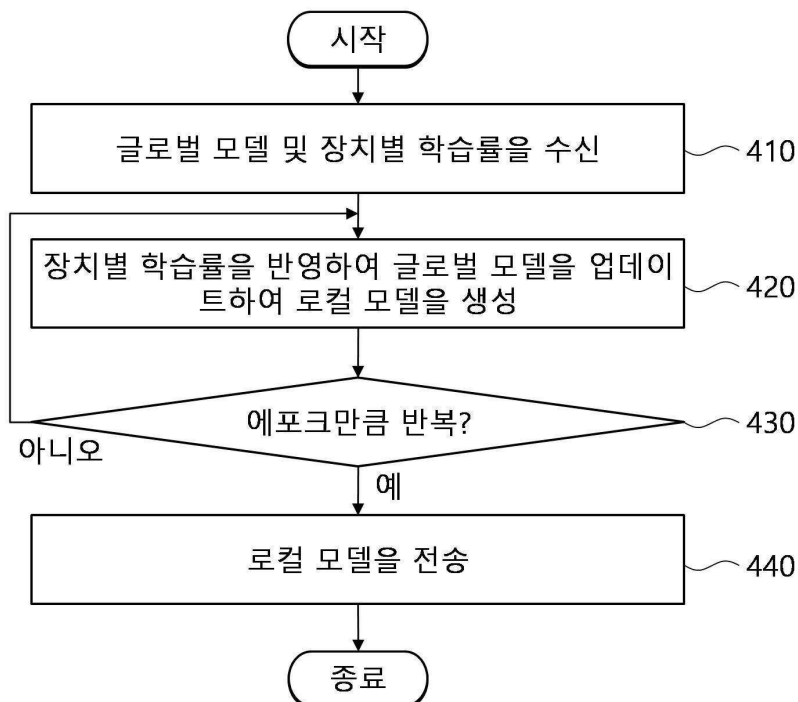
도면2



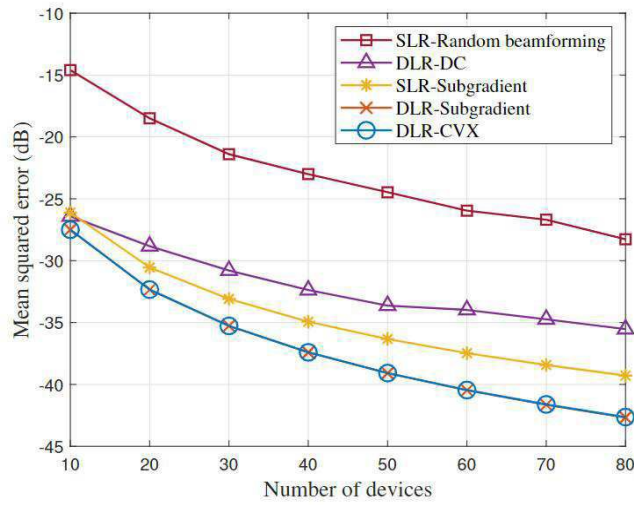
도면3



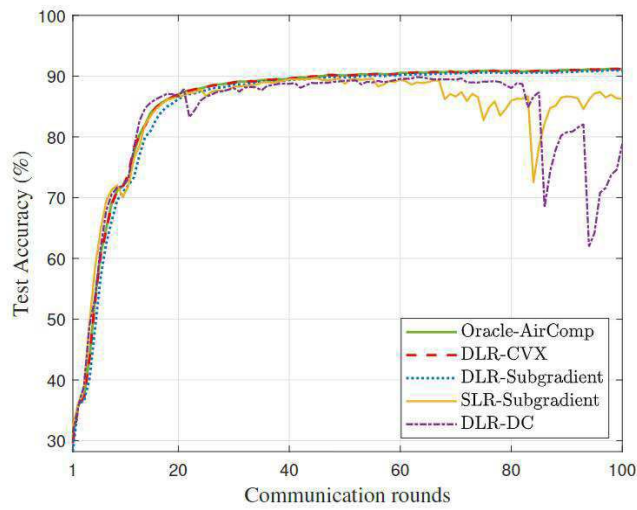
도면4



도면5



도면6



도면7

Dataset	MNIST		Fashion-MNIST	
	MLP	CNN	MLP	CNN
DLR-DC	80.03%	96.02%	68.30%	70.48%
SLR-Subgradient	85.75%	96.18%	73.67%	80.78%
DLR-Subgradient	91.01%	98.00%	75.21%	82.81%
DLR-CVX	91.04%	97.99%	75.33%	82.80%
Oracle-AirComp	91.05%	98.13%	76.36%	83.80%