

【서지사항】

【서류명】	특허출원서
【참조번호】	SDP20244652
【출원구분】	특허출원
【출원인】	
【명칭】	연세대학교 산학협력단
【특허고객번호】	2-2005-009509-9
【대리인】	
【명칭】	특허법인시공
【대리인번호】	9-2023-100041-2
【지정된변리사】	조예찬, 손하윤
【포괄위임등록번호】	2023-059479-9
【발명의 국문명칭】	점진적으로 증가하는 저차원 합성 파라미터를 이용한 파라미터 튜닝 방법 및 장치
【발명의 영문명칭】	METHOD AND APPARATUS FOR TUNING PARAMETER USING PROGRESSIVELY INCREASING LOW-DIMENSIONAL SYNTHETIC PARAMETER
【발명자】	
【성명】	박상현
【성명의 영문표기】	SANGHYUN PARK
【국적】	KR
【주민등록번호】	670101-1XXXXXX
【우편번호】	08004

【주소】 서울특별시 양천구 오목로 300, 204동 3701호

【거주국】 KR

【발명자】

【성명】 이지은

【성명의 영문표기】 JIEUN LEE

【국적】 KR

【주민등록번호】 960912-2XXXXXX

【우편번호】 03716

【주소】 서울특별시 서대문구 동교로 291, 101동 1004호

【거주국】 KR

【발명자】

【성명】 서상민

【성명의 영문표기】 SANGMIN SEO

【국적】 KR

【주민등록번호】 930507-1XXXXXX

【우편번호】 63272

【주소】 제주특별자치도 제주시 고마로 44, 801호

【거주국】 KR

【발명자】

【성명】 염찬호

【성명의 영문표기】 CHANHO YEOM

【국적】 KR

【주민등록번호】 970106-1XXXXXX

【우편번호】 03726

【주소】 서울특별시 서대문구 연희로10가길 51-12, 305호

【거주국】 KR

【발명자】

【성명】 김휘군

【성명의 영문표기】 JINHUIJUN

【국적】 CN

【생년월일】 1995-10-31

【주소】 서울특별시 관악구 남부순환로159길 32 동방주택 204

【거주국】 KR

【발명자】

【성명】 권세인

【성명의 영문표기】 SEIN KWON

【국적】 KR

【주민등록번호】 970908-2XXXXXX

【우편번호】 03708

【주소】 서울특별시 서대문구 연희로 93, 702호

【거주국】 KR

【출원언어】 국어

【심사청구】 청구

【이 발명을 지원한 국가연구개발사업】

【과제고유번호】 2710008456
【과제번호】 II170477
【부처명】 과학기술정보통신부
【과제관리(전문)기관명】 정보통신기획평가원
【연구사업명】 SW컴퓨팅산업원천기술개발
【연구과제명】 (SW 스타랩) IoT 환경을 위한 고성능 플래시 메모리 스토리지 기반 인메모리 분산 DBMS 개발
【과제수행기관명】 연세대학교 산학협력단
【연구기간】 2024.01.01 ~ 2024.12.31
【취지】 위와 같이 특허청장에게 제출합니다.

대리인 특허법인시공

(서명 또는 인)

【수수료】

【출원료】	0 면	46,000 원
【가산출원료】	40 면	0 원
【우선권주장료】	0 건	0 원
【심사청구료】	11 항	727,000 원
【합계】		773,000원
【감면사유】	전담조직(50%감면)[1]	
【감면후 수수료】		386,500 원

【발명의 설명】

【발명의 명칭】

점진적으로 증가하는 저차원 합성 파라미터를 이용한 파라미터 튜닝 방법 및 장치{METHOD AND APPARATUS FOR TUNING PARAMETER USING PROGRESSIVELY INCREASING LOW-DIMENSIONAL SYNTHETIC PARAMETER}

【기술분야】

【0001】 본 개시는 점진적으로 증가하는 저차원 합성 파라미터를 이용한 파라미터 튜닝 방법 및 장치에 관한 것으로, 데이터 처리 시스템의 구성 파라미터를 자동으로 튜닝하는 방법 및 장치에 관한 것이다.

【발명의 배경이 되는 기술】

【0003】 빅데이터 등의 대규모 데이터를 처리하기 위한 데이터 처리 시스템이 사용 및 개발되고 있다. 이러한 데이터 처리 시스템은 유연한 구성(configuration) 세팅을 제공한다. 예를 들어, 구성은 수백 개의 파라미터를 포함할 수 있다. 일반적으로, 데이터베이스 관리자(database administrator)는 수백 개의 파라미터를 직접 튜닝하여 워크로드(workload)에 적합한 구성을 설정해야 한다. 그러나, 관리자가 직접 수많은 파라미터를 튜닝하는 것은 상당한 비용 및 시간이 소요되는 문제가 있다.

【발명의 내용】**【해결하고자 하는 과제】**

【0005】 본 개시는 상기와 같은 문제점을 해결하기 위한 파라미터 튜닝 방법, 컴퓨터 판독 가능 매체에 저장된 컴퓨터 프로그램, 컴퓨터 프로그램이 저장된 컴퓨터 판독 가능 매체 및 장치(시스템)를 제공한다.

【과제의 해결 수단】

【0007】 본 개시는 방법, 장치(시스템), 컴퓨터 판독 가능 매체에 저장된 컴퓨터 프로그램 또는 컴퓨터 프로그램이 저장된 컴퓨터 판독 가능 매체를 포함한 다양한 방식으로 구현될 수 있다.

【0008】 본 개시의 일 실시예에 따르면, 적어도 하나의 프로세서에 의해 수행되는 파라미터 튜닝 방법은, 고차원의 입력 공간의 구성 파라미터 및 구성 파라미터에 대응하는 성능 메트릭을 수집하는 단계, 구성 파라미터에 매핑되는 저차원의 타겟 공간 상의 합성 파라미터를 생성하는 단계, 합성 파라미터 및 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출하는 단계 및 후보 파라미터에 대한 성능 평가를 수행하는 단계를 포함한다.

【0009】 본 개시의 일 실시예에 따르면, 고차원의 입력 공간의 구성 파라미터 및 구성 파라미터에 대응하는 성능 메트릭을 수집하는 단계는, 데이터 처리 시스템의 마스터 노드에 구성 파라미터를 제공하는 단계 및 마스터 노드로부터 제공

된 구성 파라미터에 대응하는 성능 메트릭을 수신하는 단계를 포함한다.

【0010】 본 개시의 일 실시예에 따르면, 고차원의 입력 공간에 대응하는 저차원의 타겟 공간을 생성하는 단계를 더 포함한다.

【0011】 본 개시의 일 실시예에 따르면, 합성 파라미터 및 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출하는 단계는, 대체 모델을 기초로 샘플에 대응하는 목적 함수를 결정하는 단계 및 획득 함수를 기초로 목적 함수를 최적화하기 위한 후보 파라미터를 추출하는 단계를 포함한다.

【0012】 본 개시의 일 실시예에 따르면, 획득 함수를 기초로 목적 함수를 최적화하기 위한 후보 파라미터를 추출하는 단계는, EI 상에 노이즈에 관한 패널티를 결합한 AEI를 이용하는 획득 함수를 기초로 목적 함수를 최적화하기 위한 후보 파라미터를 추출하는 단계를 포함한다.

【0013】 본 개시의 일 실시예에 따르면, 후보 파라미터에 대한 성능 평가를 수행하는 단계는, 타겟 공간 상의 후보 파라미터를 입력 공간 상의 후보 구성 파라미터로 변환하는 단계, 변환된 후보 구성 파라미터를 데이터 처리 시스템의 마스터 노드에 제공하는 단계 및 마스터 노드로부터 제공된 후보 구성 파라미터에 대응하는 성능 메트릭을 수신하는 단계를 포함한다.

【0014】 본 개시의 일 실시예에 따르면, 튜닝 세션 동안 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계를 더 포함한다.

【0015】 본 개시의 일 실시예에 따르면, 튜닝 세션 동안 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계는, 평가 버킷 내에서 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계를 포함한다.

【0016】 본 개시의 일 실시예에 따르면, 튜닝 세션 동안 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계는, 타겟 차원수가 증가할수록 타겟 차원수에 할당되는 평가 버킷을 증가시키는 단계를 더 포함한다.

【0017】 본 개시의 일 실시예에 따른 상술된 방법을 컴퓨터에서 실행하기 위해 컴퓨터 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램이 제공된다.

【0018】 본 개시의 일 실시예에 따른 컴퓨팅 장치는, 통신 모듈, 메모리 및 메모리와 연결되고, 메모리에 포함된 컴퓨터 판독 가능한 적어도 하나의 프로그램을 실행하도록 구성된 적어도 하나의 프로세서를 포함한다. 적어도 하나의 프로그램은, 고차원의 입력 공간의 구성 파라미터 및 구성 파라미터에 대응하는 성능 메트릭을 수집하고, 구성 파라미터에 매핑되는 저차원의 타겟 공간 상의 합성 파라미터를 생성하고, 합성 파라미터 및 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출하고, 후보 파라미터에 대한 성능 평가를 수행하기 위한 명령어들을 포함한다.

【발명의 효과】

【0020】본 개시의 다양한 실시예에서 컴퓨팅 장치는 저차원의 타겟 공간에서 파라미터 튜닝을 수행함으로써 컴퓨팅 자원을 효과적으로 감소시킬 수 있으며, 파라미터 튜닝 시 타겟 공간의 타겟 차원수를 점진적으로 증가시켜 활성 공간의 차원수를 만족하는 파라미터 튜닝을 수행할 수 있다.

【0021】본 개시의 다양한 실시예에서 AEI 기반의 획득 함수는 노이즈에 관계없이 최적의 성능을 갖는 후보 파라미터를 추출하여 제공할 수 있다.

【0022】본 발명의 효과는 이상에서 언급한 효과로 제한되지 않으며, 언급되지 않은 다른 효과들은 청구범위의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자("통상의 기술자"라 함)에게 명확하게 이해될 수 있을 것이다.

【도면의 간단한 설명】

【0024】본 개시의 실시예들은, 이하 설명하는 첨부 도면들을 참조하여 설명될 것이며, 여기서 유사한 참조 번호는 유사한 요소들을 나타내지만, 이에 한정되지는 않는다.

도 1은 본 개시의 일 실시예에 따른 컴퓨팅 장치의 내부 구성을 나타내는 기능적인 블록도이다.

도 2는 본 개시의 일 실시예에 따른 타겟 공간의 타겟 차원수가 점진적으로 증가하는 예시를 시각화한 이미지이다.

도 3은 본 개시의 일 실시예에 따른 AEI 기반의 획득 함수의 생성 과정을 나

타내는 예시적인 블록도이다.

도 4는 본 개시에 따른 파라미터 튜닝 프로세스를 단계 별로 구분한 예시를 나타내는 블록도이다.

도 5는 본 개시의 일 실시예에 따른 파라미터 튜닝 방법의 예시를 나타내는 흐름도이다.

도 6은 본 개시의 일 실시예에 따른 컴퓨팅 장치의 하드웨어 구성을 나타내는 블록도이다.

【발명을 실시하기 위한 구체적인 내용】

【0025】 이하, 본 개시의 실시를 위한 구체적인 내용을 첨부된 도면을 참조하여 상세히 설명한다. 다만, 이하의 설명에서는 본 개시의 요지를 불필요하게 흐릴 우려가 있는 경우, 널리 알려진 기능이나 구성에 관한 구체적 설명은 생략하기로 한다.

【0026】 첨부된 도면에서, 동일하거나 대응하는 구성요소에는 동일한 참조부호가 부여되어 있다. 또한, 이하의 실시예들의 설명에 있어서, 동일하거나 대응되는 구성요소를 중복하여 기술하는 것이 생략될 수 있다. 그러나, 구성요소에 관한 기술이 생략되어도, 그러한 구성요소가 어떤 실시예에 포함되지 않는 것으로 의도되지는 않는다.

【0027】 개시된 실시예의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러

나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명이 완전하도록 하고, 본 발명이 통상의 기술자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것일 뿐이다.

【0028】 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 개시된 실시예에 대해 구체적으로 설명하기로 한다. 본 명세서에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 관련 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서, 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 개시의 전반에 걸친 내용을 토대로 정의되어야 한다.

【0029】 본 명세서에서의 단수의 표현은 문맥상 명백하게 단수인 것으로 특정하지 않는 한, 복수의 표현을 포함한다. 또한, 복수의 표현은 문맥상 명백하게 복수인 것으로 특정하지 않는 한, 단수의 표현을 포함한다. 명세서 전체에서 어떤 부분이 어떤 구성요소를 포함한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다.

【0030】 본 개시에서, "포함하다", "포함하는" 등의 용어는 특징들, 단계들, 동작들, 요소들 및/또는 구성 요소들이 존재하는 것을 나타낼 수 있으나, 이러한

용어가 하나 이상의 다른 기능들, 단계들, 동작들, 요소들, 구성 요소들 및/또는 이들의 조합이 추가되는 것을 배제하지는 않는다.

【0031】 본 개시에서, 특정 구성 요소가 임의의 다른 구성 요소에 "결합", "조합", "연결" 되거나, "반응" 하는 것으로 언급된 경우, 특정 구성 요소는 다른 구성 요소에 직접 결합, 조합 및/또는 연결되거나, 반응할 수 있으나, 이에 한정되지 않는다. 예를 들어, 특정 구성 요소와 다른 구성 요소 사이에 하나 이상의 중간 구성 요소가 존재할 수 있다. 또한, 본 발명에서 "및/또는"은 열거된 하나 이상의 항목의 각각 또는 하나 이상의 항목의 적어도 일부의 조합을 포함할 수 있다.

【0032】 본 개시에서, "제1", "제2" 등의 용어는 특정 구성 요소를 다른 구성 요소와 구별하기 위해 사용되는 것으로, 이러한 용어에 의해 상술된 구성 요소가 제한되진 않는다. 예를 들어, "제1" 구성 요소는 "제2" 구성 요소와 동일하거나 유사한 형태의 요소일 수 있다.

【0033】 본 개시에서, "데이터 처리 시스템"은 대규모의 데이터를 처리하기 위한 데이터 분석 플랫폼을 지칭할 수 있으며, 예를 들어, 'Apache Spark'와 같은 병렬 및 분산 데이터베이스 시스템을 포함할 수 있다. 데이터 처리 시스템은 마스터 노드(master node) 및 워커 노드(worker node)를 포함할 수 있다. 또한, 데이터 처리 시스템의 구성 파라미터(configuration parameter)는 메모리 설정, 스케줄링, 자원 관리 등과 같은 시스템 실행을 제어하기 위한 것일 수 있다.

【0034】 본 개시에서, "파라미터 튜닝(parameter tuning)"은 데이터 처리 시스템의 성능을 향상시키기 위해, 해당 데이터 처리 시스템을 구성하는 파라미터들 중 적어도 일부를 최적화하는 작업을 지칭할 수 있다. 파라미터 튜닝은 선택 기반의 튜닝 방식과 생성 기반의 튜닝 방식을 포함할 수 있다. 여기서, 선택 기반의 튜닝 방식은 데이터 처리 시스템을 파라미터들 중 일부를 선택하여 튜닝하는 방식을 지칭하고, 생성 기반의 튜닝 방식은 고차원의 파라미터 공간을 저차원의 파라미터 공간으로 매핑하여 튜닝하는 방식을 지칭할 수 있다.

【0035】 도 1은 본 개시의 일 실시예에 따른 컴퓨팅 장치(100)의 내부 구성을 나타내는 기능적인 블록도이다. 일 실시예에 따르면, 컴퓨팅 장치(100)는 데이터 처리 시스템(120)의 구성 파라미터를 워크로드(workload)에 적합한 형태로 튜닝하기 위한 임의의 장치를 지칭할 수 있다. 예를 들어, 컴퓨팅 장치(100)는 제어기(controller)(102), 공간 생성기(space generator)(104), 데이터 저장소(data repository)(106), 획득 함수(acquisition function)(108), 대체 모델(surrogate model)(110) 등을 포함할 수 있다.

【0036】 일 실시예에 따르면, 컴퓨팅 장치(100)는 생성 기반의 파라미터 튜닝을 수행할 수 있다. 즉, 컴퓨팅 장치(100)는 데이터 처리 시스템(120)의 파라미터 공간을 저차원(low-dimensional)의 파라미터 공간으로 매핑한 후, 파라미터에 대한 효율적인 튜닝을 수행할 수 있다. 이 경우, 컴퓨팅 장치(100)는 노이즈(noise)에 강인한 베이저안 최적화(Bayesian optimization) 기반의 파라미터 튜닝을 수행할 수 있다.

【0037】 일 실시예에 따르면, 컴퓨팅 장치(100)는 데이터 처리 시스템(120)과 통신하며 파라미터 튜닝에 필요한 데이터 및/또는 정보를 주고받을 수 있다. 예를 들어, 컴퓨팅 장치(100)는 데이터 처리 시스템(120)의 마스터 노드에 성능 평가를 위한 구성 파라미터를 제공할 수 있다. 여기서, 구성 파라미터는 데이터 처리 시스템(120)의 고차원(high-dimensional)의 입력 공간(input space) 상의 파라미터일 수 있다. 이 경우, 컴퓨팅 장치(100)는 데이터 처리 시스템(120)의 마스터 노드로부터 제공된 구성 파라미터에 대응하는 성능 메트릭(performance metric)을 수신하여, 구성 파라미터 및 성능 메트릭 쌍을 수집할 수 있다.

【0038】 일 실시예에 따르면, 제어기(102)는 구성 파라미터 및 성능 메트릭 쌍을 수집하고, 이를 공간 생성기(104) 및 데이터 저장소(106)에 제공할 수 있다. 공간 생성기(104)는 고차원의 입력 공간에 대응하는 저차원의 타겟 공간(target space)을 생성할 수 있다. 또한, 공간 생성기(104)는 구성 파라미터에 매핑되는 저차원의 합성 파라미터(synthetic parameter)를 생성할 수 있다. 이와 같이 생성되는 저차원의 합성 파라미터를 이용하여 파라미터 튜닝을 수행하는 경우, 고차원의 구성 파라미터를 이용하는 경우보다 컴퓨팅 자원이 현저히 감소하여 튜닝 성능이 효과적으로 향상될 수 있다.

【0039】 일 실시예에 따르면, 공간 생성기(104)에 의해 생성되는 합성 파라미터는 데이터 저장소(106)에 제공될 수 있다. 즉, 데이터 저장소(106)는 데이터 처리 시스템(120)의 구성 파라미터, 성능 메트릭 및 합성 파라미터의 집합으로 구성된 샘플(sample)을 저장 및/또는 관리할 수 있다. 이와 같이 데이터 저장소(10

6)에 저장된 샘플들은 성능 평가를 위해 추출되어 사용될 수 있다.

【0040】 일 실시예에 따르면, 대체 모델(110) 및 획득 함수(108)를 기초로 베이지안 최적화가 수행될 수 있다. 여기서, 베이지안 최적화는 베이지안 확률에 기반한 최적화 방식으로, 입력 값에 대응하여 관측되는 값을 이용하여 최적화된 목적 함수(objective function)를 도출하는데 사용되는 최적화 기법을 지칭할 수 있다. 예를 들어, 대체 모델(110) 및 획득 함수(108)를 기초로 타겟 공간의 타겟 차원수(target dimensionality)를 점진적으로 증가시키는 방식으로 최적화가 수행될 수 있다.

【0041】 일 실시예에 따르면, 타겟 공간의 타겟 차원수는 실제로 튜닝되는 합성 파라미터의 수를 지칭할 수 있다. 또한, 활성 공간(active space)의 차원수는 튜닝 성능에 큰 영향을 미치는 최소한의 파라미터의 수를 나타내는 것으로 미지의 값을 가질 수 있다. 따라서, 튜닝 성능을 향상시키기 위해서는 타겟 차원수를 고정하지 않고 점진적으로 증가시키며 활성 공간의 차원수 이상으로 유지하는 것이 필요할 수 있다.

【0042】 일 실시예에 따르면, 대체 모델(110)은 합성 파라미터 및 성능 메트릭과 연관된 하나의 샘플을 기초로 피팅(fitting)될 수 있으며, 피팅 과정에서 관측되는 값을 기초로 목적 함수(objective function)가 추정될 수 있다. 이 경우, 획득 함수(108)는 관측된 값 및/또는 목적 함수의 노이즈 분산(noise variance)을 분석하여 최적의 성능을 달성할 가능성이 있는 후보 파라미터를 추출할 수 있다.

【0043】 일 실시예에 따르면, 추출된 후보 파라미터는 저차원의 타겟 공간

상에서 정의되는 것이어서, 성능 평가를 위해 고차원의 입력 공간 상의 파라미터로 변환될 수 있다. 예를 들어, 공간 생성기(104)는 후보 파라미터를 입력받아 입력 공간 상의 후보 구성 파라미터로 변환할 수 있다. 이 경우, 제어기(102)는 변환된 후보 구성 파라미터를 데이터 처리 시스템(120)의 마스터 노드에 제공하고, 마스터 노드로부터 제공된 후보 구성 파라미터에 대응하는 성능 메트릭을 수신하는 것에 의해 후보 파라미터에 대한 성능 평가를 수행할 수 있다.

【0044】 상술된 파라미터 튜닝 및 성능 평가 프로세스는 튜닝 세션(tuning session) 동안 반복적으로 수행될 수 있다. 예를 들어, 컴퓨팅 장치(100)는 평가 버짓(evaluation budget) 내에서 타겟 공간의 타겟 차원수를 점진적으로 증가시키며 파라미터 튜닝 및 성능 평가를 수행할 수 있다. 이 경우, 타겟 차원수가 증가할수록 타겟 차원수에 할당되는 평가 버짓은 증가할 수 있다.

【0045】 도 1에서는 컴퓨팅 장치(100)에 포함된 각각의 기능적인 구성이 구분되어 상술되었으나, 이는 발명의 이해를 돕기 위한 것일 뿐이며, 하나의 연산 장치에서 둘 이상의 기능을 수행할 수도 있다. 이와 같은 구성에 의해, 컴퓨팅 장치(100)는 저차원의 타겟 공간에서 파라미터 튜닝을 수행함으로써 컴퓨팅 자원을 효과적으로 감소시킬 수 있으며, 파라미터 튜닝 시 타겟 공간의 타겟 차원수를 점진적으로 증가시켜 활성 공간의 차원수를 만족하는 파라미터 튜닝을 수행할 수 있다.

【0046】 도 2는 본 개시의 일 실시예에 따른 타겟 공간의 타겟 차원수가 점진적으로 증가하는 예시를 시각화한 이미지(200)이다. 상술된 것과 같이, 타겟 공간의 타겟 차원수는 입력 공간의 차원수와 대응될 때까지 점진적으로 증가할 수 있

다. 이와 같이 타겟 차원수를 점진적으로 증가시키기 위해 랜덤 임베딩(random embedding) 및 분할(split)이 수행될 수 있다.

【0047】 일 실시예에 따르면, 입력 공간의 구성 파라미터는 하나의 합성 파라미터와 대응되고, 하나의 합성 파라미터는 하나 이상의 구성 파라미터와 대응될 수 있다. 이 경우, 구성 파라미터 및 합성 파라미터 사이의 임베딩(embedding)을 위해 임베딩 행렬이 사용될 수 있다. 예를 들어, 임베딩 행렬은 희소 투영 행렬(sparse projection matrix)을 포함할 수 있다. 여기서, 희소 투영 행렬은 각각의 열(column)에 0이 아닌 하나의 항목을 가지는 행렬을 지칭할 수 있다. 이러한 희소 투영 행렬을 기초로 임베딩을 수행하는 경우, 구성 파라미터는 하나의 합성 파라미터와 대응될 수 있다.

【0048】 구체적으로, $S: \mathcal{Z} \rightarrow \mathcal{X}$ 은 $z \in \mathcal{Z}$ 의 합성 파라미터 $z = \{\theta'_1, \dots, \theta'_d\}$ 를 $x \in \mathcal{X}$ 의 구성 파라미터 $x = \{\theta_1, \dots, \theta_D\}$ 로 매핑시킬 수 있다. 이 경우, 성능 메트릭은 $y = f(Sz) + \epsilon$ 로 표현될 수 있으며, $\epsilon \sim N(0, \sigma_\epsilon^2)$ 는 노이즈를 나타낼 수 있다. 또한, σ_ϵ^2 는 노이즈 분산(noise variance)을 나타낼 수 있다.

【0049】 일 실시예에 따르면, 합성 파라미터의 타겟 차원수는 분할에 의해 점진적으로 증가할 수 있다. 즉, 합성 파라미터는 분할될 수 있으며, 분할에 따라 임베딩 행렬은 업데이트되고, 이에 따라, 입력 공간은 타겟 공간과 재매칭될 수 있다. 이미지(300)에 도시된 예에서, 합성 파라미터 θ'_1 및 θ'_2 는 각각 2개의 합성 파라미터로 분할될 수 있다. 이 경우, 분할되어 생성된 합성 파라미터 θ'_{11} , θ'_{12} , θ'_{21}

및 θ'_{22} 의 값은 초기 합성 파라미터와 동일한 값을 가질 수 있다. 이와 같이 분할이 수행된 후 동일한 값이 할당되는 경우, 대체 모델의 관찰 결과가 새로운 타겟 차원수를 갖는 타겟 공간에서도 유지되어 관찰 데이터의 일관성이 유지될 수 있다.

【0050】 일 실시예에 따르면, 컴퓨팅 장치는 평가 버킷 내에서 타겟 공간의 타겟 차원수를 점진적으로 증가시킬 수 있으며, 이 때, 타겟 차원수가 증가할수록 타겟 차원수에 할당되는 평가 버킷을 증가시킬 수 있다. 예를 들어, 컴퓨팅 장치는 다음의 수학식 1과 같이 평가 버킷을 증가시킬 수 있다.

【0052】 【수학식 1】

$$m_i = \left\lceil \frac{b(m_D - n_{init})d_i}{d_0((b+1)^{n_s+1} - 1)} \right\rceil$$

【0054】 여기서, m_i 는 i 번째 타겟 공간에서의 평가 버킷을 나타내고, d_i 는 타겟 차원수를 나타낼 수 있다. 또한, d_0 는 초기 타겟 차원수를 나타내고, n_{init} 은 랜덤하게 생성된 초기 샘플 수를 나타낼 수 있다. 또한, n_s 는 입력 공간의 차원수에 도달하기 위해 필요한 분할 횟수를 나타낼 수 있다.

【0055】 일 실시예에 따르면, 컴퓨팅 장치는 $\lceil \log_{b+1} D/d_0 \rceil$ 를 이용하여 n_s 를 산출할 수 있다. 또한, 컴퓨팅 장치는 분할 인자(split factor),

$b = \text{round}(\log_{n_s} \frac{D}{d_0} - 1)$ 를 초기화하고, $d_0(b+1)^{n_s}$ 를 기초로 d_i 를 산출할 수 있다.

이와 같은 프로세스에 의해 타겟 차원수와 평가 버짓은 함께 증가할 수 있으며, 저 차원의 공간에서는 더 적은 탐색이 수행되고, 차원이 증가할수록 더 많은 탐색이 수행될 수 있다.

【0056】 도 3은 본 개시의 일 실시예에 따른 AEI 기반의 획득 함수(330)의 생성 과정을 나타내는 예시적인 블록도이다. 일 실시예에 따르면, 획득 함수는 대체 모델과 연관된 목적 함수를 기초로 최적해(예: 최적의 후보 파라미터)를 도출하기 위한 합성 파라미터를 추출 및 제안하기 위한 것일 수 있다. 또한, AEI 기반의 획득 함수(330)는 노이즈에 강인한 성능을 갖는 획득 함수를 나타낼 수 있다. 예를 들어, 노이즈는 네트워크 I/O, 디스크 I/O, 클라우드의 공유 하드웨어 환경 등에 의해 발생할 수 있다.

【0057】 일 실시예에 따르면, 노이즈를 고려하지 않는 노이즈 프리 획득 함수(310)는 다음의 수학적 식 2와 같이 표현될 수 있다.

【0059】 【수학적 식 2】

$$\alpha_{EI}(x_t|f^*) = \mathbb{E}[\max(0, f^* - y_t) | y_t \sim \mathcal{GP}(x_t | \mathcal{D}_{t-1})]$$

【0061】 여기서, $f^* = \min_{k \in \mathcal{D}_{t-1}} f(x_t)$ 는 관측치(예: 후보 파라미터 및 성능 메트릭의 쌍)들이 저장된 \mathcal{D}_{t-1} 에서 최소값을 갖는 현재까지의 최적해를 나타낼 수 있다. 그러나, 이러한 노이즈 프리 획득 함수(310)는 노이즈를 전혀 고려하지 않아 최적의 성능을 나타내는 최적해를 적절히 선택하지 못하는 단점이 있다.

【0062】 일 실시예에 따르면, EI 기반의 획득 함수(320)는 EI(expected improvement)를 이용하는 획득 함수를 지칭할 수 있다. 여기서, EI는 착취(exploitation) 전략 및 탐색(exploration) 전략을 모두 반영하여 현재까지 조사된 최대 관측치보다 더 큰 관측치를 도출할 확률 및 최대 관측치와의 차이값(magnitude), 즉 EI 값을 산출하기 위한 것일 수 있다. 예를 들어, EI 기반의 획득 함수(320)는 다음의 수학적 식 3과 같이 표현될 수 있다.

【0064】 【수학적 식 3】

$$\alpha_{EI}(z_t|T) = \mathbb{E}[\max(0, T - y_t) | y_t \sim \mathcal{GP}(z_t | \mathcal{D}_{t-1})]$$

【0066】 여기서, 최적해 T는 다음의 수학적 식 4와 같이 표현될 수 있다.

【0068】 【수학식 4】

$$T = \min_{k \in \mathcal{D}_{t-1}} \mu_{t-1}(x_k)$$

【0070】 여기서, μ_{t-1} 는 대체 모델의 사후 평균(posterior mean)을 나타낼 수 있다. 즉, EI 기반의 획득 함수(320)는 대체 모델의 사후 평균을 이용하여 최솟값을 구하는 것에 의해 노이즈를 고려한 최적해를 산출할 수 있다. 그러나, 이러한 EI 기반의 획득 함수(320)는 현재까지의 노이즈를 반영할 수 있을 뿐 미래 관찰에 대한 노이즈를 고려하지 않는 단점이 있다.

【0071】 이를 해결하기 위해, AEI 기반의 획득 함수(330)가 이용될 수 있다. 예를 들어, AEI 기반의 획득 함수(330)는 EI 값에 패널티를 곱하는 것에 의해 산출되는 관측치를 이용하는 것일 수 있다. AEI 기반의 획득 함수(330)는 다음의 수학식 5와 같이 표현될 수 있다.

【0073】 【수학식 5】

$$\alpha_{AEI}(z_t|T^*) = \alpha_{EI}(z_t|T^*) \times \left(1 - \frac{\sigma_\epsilon}{\sqrt{\sigma_{t-1}^2(z_t) + \sigma_\epsilon^2}}\right)$$

【0075】 여기서, σ_ϵ 는 노이즈의 분산(variance)을 나타낼 수 있다. 이 경우, AEI 기반의 획득 함수(330)는 T 를 대신하여 다음의 수학식 6과 같은 최적해 T^* 를 이용할 수 있다.

【0077】 【수학식 6】

$$T^* = \min_{k \in \mathcal{D}_{t-1}} (\mu_{t-1}(z_k) + \lambda \sigma_{t-1}(z_k))$$

【0079】 최적해 T^* 는 모델의 불확실성을 나타내는 표준편차 항 σ_{t-1} 을 통합하여 노이즈에 민감하지 않을 수 있다. 여기서, λ 는 하이퍼 파라미터(hyper parameter)를 나타낼 수 있다.

【0080】 이와 같이 AEI 기반의 획득 함수(330)를 이용하는 경우, 패널티 항에 의해 노이즈가 강할 경우 기존의 EI 값은 감소하고, 노이즈가 적을수록 AEI 값(AEI 기반의 획득 함수(330)에 의해 산출되는 확률값 및 차이값)은 기존의 EI 값과 유사해질 수 있다. 이와 같은 과정에 의해, AEI 기반의 획득 함수(330)는 노이즈에 관계없이 최적의 성능을 갖는 후보 파라미터를 추출하여 제공할 수 있다.

【0081】 도 4는 본 개시에 따른 파라미터 튜닝 프로세스를 단계 별로 구분한 예시를 나타내는 블록도이다. 도시된 것과 같이, 파라미터 튜닝은 초기화 단계

(410), 최적화 단계(420), 차원 증가 단계(430) 및 리샘플링(resampling) 단계(440)로 구분될 수 있다.

【0082】 일 실시예에 따르면, 초기화 단계(410)는 다양한 하이퍼 파라미터를 초기화하는 것을 지칭할 수 있다. 예를 들어, 컴퓨팅 장치(도 1의 100)는 초기 타겟 차원수 d_0 , 최대 평가 버짓 m_D , 분할 요인(split factor) b 의 수, 초기 샘플 크기 n_{init} 등을 초기화할 수 있다. 그리고 나서, 컴퓨팅 장치는 합성 파라미터 및 성능 메트릭의 쌍으로 구성된 초기 샘플 \mathcal{D}_0 을 추출할 수 있다.

【0083】 일 실시예에 따르면, 최적화 단계(420)는 최적의 성능을 갖는 구성 파라미터를 추출하기 위한 최적해를 구하는 것을 지칭할 수 있다. 일 실시예에 따르면, 대체 모델은 평가 버짓 내에서 루프(loop)를 반복하여 샘플 \mathcal{D}_{j-1} 를 기초로 피팅될 수 있다. 이 경우, AEI 기반의 획득 함수는 다음 탐색을 위한 후보 파라미터 z_j 를 추출할 수 있다. 후보 파라미터 z_j 는 임베딩 행렬 S 에 의해 후보 구성 파라미터로 변환되고 평가되는 것에 의해 관측될 수 있다. 이와 같이 관측이 수행되는 것에 의해 관측치를 포함하는 데이터셋이 업데이트될 수 있다.

【0084】 일 실시예에 따르면, 차원 증가 단계(430)는 현재 타겟 차원수 d_i 가 입력 차원의 차원수 D 보다 적은 경우, 타겟 차원수를 증가시키는 것을 지칭할 수 있다. 이와 같이 타겟 차원수가 증가하는 경우, 임베딩 행렬 S 는 임베딩 차원을 증가시키도록 업데이트될 수 있으며, 타겟 차원수 d_i 는 임베딩 행렬 S 의 로우(row)의

수로 설정될 수 있다. 또한, 평가 버짓 m_D 이 다시 산출될 수 있다.

【0085】 일 실시예에 따르면, 리샘플링 단계(440)는 타겟 차원수가 입력 차원의 차원수에 도달했을 때, 새로운 초기 샘플이 리샘플링될 수 있다. 또한, 이를 기초로 평가가 재수행되어 데이터셋이 재설정될 수 있다. 이와 같은 프로세스에 의해, 컴퓨팅 장치는 타겟 차원수의 점진적인 증가 및 노이즈에 강인한 획득 함수를 이용하여 최적의 파라미터를 추출할 수 있다.

【0086】 도 5는 본 개시의 일 실시예에 따른 파라미터 튜닝 방법(500)의 예를 나타내는 흐름도이다. 파라미터 튜닝 방법(500)은 프로세서(예를 들어, 컴퓨팅 장치의 적어도 하나의 프로세서)에 의해 수행될 수 있다. 파라미터 튜닝 방법(500)은 프로세서가 고차원의 입력 공간의 구성 파라미터 및 구성 파라미터에 대응하는 성능 메트릭을 수집함으로써 개시될 수 있다(S510). 예를 들어, 프로세서는 데이터 처리 시스템의 마스터 노드에 구성 파라미터를 제공하고, 마스터 노드로부터 제공된 구성 파라미터에 대응하는 성능 메트릭을 수신할 수 있다.

【0087】 일 실시예에 따르면, 프로세서는 고차원의 입력 공간에 대응하는 저차원의 타겟 공간을 생성하고, 구성 파라미터에 매핑되는 저차원의 타겟 공간 상의 합성 파라미터를 생성할 수 있다(S520). 또한, 프로세서는 합성 파라미터 및 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출할 수 있다(S530).

【0088】 일 실시예에 따르면, 프로세서는 대체 모델을 기초로 샘플에 대응하는 목적 함수를 결정하고, 획득 함수를 기초로 목적 함수를 최적화하기 위한 후보 파라미터를 추출할 수 있다. 예를 들어, 프로세서는 EI 상에 노이즈에 관한 패널티를 결합한 AEI를 이용하는 획득 함수를 기초로 목적 함수를 최적화하기 위한 후보 파라미터를 추출할 수 있다.

【0089】 일 실시예에 따르면, 프로세서는 튜닝 세션 동안 타겟 공간의 타겟 차원수를 점진적으로 증가시키며 후보 파라미터를 추출할 수 있다. 예를 들어, 프로세서는 평가 버킷 내에서 타겟 공간의 타겟 차원수를 점진적으로 증가시킬 수 있다. 이 경우, 프로세서는 타겟 차원수가 증가할수록 타겟 차원수에 할당되는 평가 버킷을 증가시킬 수 있다.

【0090】 일 실시예에 따르면, 프로세서는 후보 파라미터에 대한 성능 평가를 수행할 수 있다(S540). 프로세서는 타겟 공간 상의 후보 파라미터를 입력 공간 상의 후보 구성 파라미터로 변환하고, 변환된 후보 구성 파라미터를 데이터 처리 시스템의 마스터 노드에 제공할 수 있다. 이 경우, 프로세서는 마스터 노드로부터 제공된 후보 구성 파라미터에 대응하는 성능 메트릭을 수신할 수 있다.

【0091】 도 6은 본 개시의 일 실시예에 따른 컴퓨팅 장치(100)의 하드웨어 구성을 나타내는 블록도이다. 컴퓨팅 장치(100)는 메모리(610), 프로세서(620), 통신 모듈(630) 및 입출력 인터페이스(640)를 포함할 수 있으며, 도 6에 도시된 바와 같이, 컴퓨팅 장치(100)는 통신 모듈(630)을 이용하여 네트워크를 통해 정보 및/또는 데이터를 통신할 수 있도록 구성될 수 있다.

【0092】 메모리(610)는 비-일시적인 임의의 컴퓨터 판독 가능한 기록매체를 포함할 수 있다. 일 실시예에 따르면, 메모리(610)는 RAM(random access memory), ROM(read only memory), 디스크 드라이브, SSD(solid state drive), 플래시 메모리(flash memory) 등과 같은 비소멸성 대용량 저장 장치(permanent mass storage device)를 포함할 수 있다. 다른 예로서, ROM, SSD, 플래시 메모리, 디스크 드라이브 등과 같은 비소멸성 대용량 저장 장치는 메모리와는 구분되는 별도의 영구 저장 장치로서 컴퓨팅 장치(100)에 포함될 수 있다. 또한, 메모리(610)에는 운영체제와 적어도 하나의 프로그램 코드가 저장될 수 있다.

【0093】 이러한 소프트웨어 구성요소들은 메모리(610)와는 별도의 컴퓨터에서 판독 가능한 기록매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록매체는 이러한 컴퓨팅 장치(100)에 직접 연결가능한 기록 매체를 포함할 수 있는데, 예를 들어, 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록매체를 포함할 수 있다. 다른 예로서, 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록매체가 아닌 통신 모듈(630)을 통해 메모리(610)에 로딩될 수도 있다. 예를 들어, 적어도 하나의 프로그램은 개발자들 또는 애플리케이션의 설치 파일을 배포하는 파일 배포 시스템이 통신 모듈(630)을 통해 제공하는 파일들에 의해 설치되는 컴퓨터 프로그램에 기반하여 메모리(610)에 로딩될 수 있다.

【0094】 프로세서(620)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(610) 또

는 통신 모듈(630)에 의해 다른 사용자 단말(미도시) 또는 다른 외부 시스템으로 제공될 수 있다.

【0095】 통신 모듈(630)은 네트워크를 통해 사용자 단말(미도시)과 컴퓨팅 장치(100)가 서로 통신하기 위한 구성 또는 기능을 제공할 수 있으며, 컴퓨팅 장치(100)가 외부 시스템(일례로 별도의 클라우드 시스템 등)과 통신하기 위한 구성 또는 기능을 제공할 수 있다. 일례로, 컴퓨팅 장치(100)의 프로세서(620)의 제어에 따라 제공되는 제어 신호, 명령, 데이터 등이 통신 모듈(630)과 네트워크를 거쳐 사용자 단말 및/또는 외부 시스템의 통신 모듈을 통해 사용자 단말 및/또는 외부 시스템으로 전송될 수 있다.

【0096】 또한, 컴퓨팅 장치(100)의 입출력 인터페이스(640)는 컴퓨팅 장치(100)와 연결되거나 컴퓨팅 장치(100)가 포함할 수 있는 입력 또는 출력을 위한 장치(미도시)와의 인터페이스를 위한 수단일 수 있다. 도 6에서는 입출력 인터페이스(640)가 프로세서(620)와 별도로 구성된 요소로서 도시되었으나, 이에 한정되지 않으며, 입출력 인터페이스(640)가 프로세서(620)에 포함되도록 구성될 수 있다. 컴퓨팅 장치(100)는 도 6의 구성요소들보다 더 많은 구성요소들을 포함할 수 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다.

【0097】 컴퓨팅 장치(100)의 프로세서(620)는 복수의 사용자 단말 및/또는 복수의 외부 시스템으로부터 수신된 정보 및/또는 데이터를 관리, 처리 및/또는 저장하도록 구성될 수 있다.

【0098】 상술된 방법 및/또는 다양한 실시예들은, 디지털 전자 회로, 컴퓨터

하드웨어, 펌웨어, 소프트웨어 및/또는 이들의 조합으로 실현될 수 있다. 본 개시의 다양한 실시예들은 데이터 처리 장치, 예를 들어, 프로그래밍 가능한 하나 이상의 프로세서 및/또는 하나 이상의 컴퓨팅 장치에 의해 실행되거나, 컴퓨터 판독 가능한 기록 매체 및/또는 컴퓨터 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 상술된 컴퓨터 프로그램은 컴파일된 언어 또는 해석된 언어를 포함하여 임의의 형태의 프로그래밍 언어로 작성될 수 있으며, 독립 실행형 프로그램, 모듈, 서브 루틴 등의 임의의 형태로 배포될 수 있다. 컴퓨터 프로그램은 하나의 컴퓨팅 장치, 동일한 네트워크를 통해 연결된 복수의 컴퓨팅 장치 및/또는 복수의 상이한 네트워크를 통해 연결되도록 분산된 복수의 컴퓨팅 장치를 통해 배포될 수 있다.

【0099】 상술된 방법 및/또는 다양한 실시예들은, 입력 데이터를 기초로 동작하거나 출력 데이터를 생성함으로써, 임의의 기능, 함수 등을 처리, 저장 및/또는 관리하는 하나 이상의 컴퓨터 프로그램을 실행하도록 구성된 하나 이상의 프로세서에 의해 수행될 수 있다. 예를 들어, 본 개시의 방법 및/또는 다양한 실시예는 FPGA(Field Programmable Gate Array) 또는 ASIC(Application Specific Integrated Circuit)과 같은 특수 목적 논리 회로에 의해 수행될 수 있으며, 본 개시의 방법 및/또는 실시예들을 수행하기 위한 장치 및/또는 시스템은 FPGA 또는 ASIC와 같은 특수 목적 논리 회로로서 구현될 수 있다.

【0100】 컴퓨터 프로그램을 실행하는 하나 이상의 프로세서는, 범용 목적 또는 특수 목적의 마이크로 프로세서 및/또는 임의의 종류의 디지털 컴퓨팅 장치의

하나 이상의 프로세서를 포함할 수 있다. 프로세서는 읽기 전용 메모리, 랜덤 액세스 메모리의 각각으로부터 명령 및/또는 데이터를 수신하거나, 읽기 전용 메모리와 랜덤 액세스 메모리로부터 명령 및/또는 데이터를 수신할 수 있다. 본 발명에서, 방법 및/또는 실시예들을 수행하는 컴퓨팅 장치의 구성 요소들은 명령어들을 실행하기 위한 하나 이상의 프로세서, 명령어들 및/또는 데이터를 저장하기 위한 하나 이상의 메모리 디바이스를 포함할 수 있다.

【0101】 일 실시예에 따르면, 컴퓨팅 장치는 데이터를 저장하기 위한 하나 이상의 대용량 저장 장치와 데이터를 주고받을 수 있다. 예를 들어, 컴퓨팅 장치는 자기 디스크(magnetic disc) 또는 광 디스크(optical disc)로부터 데이터를 수신하거나/수신하고, 자기 디스크 또는 광 디스크로 데이터를 전송할 수 있다. 컴퓨터 프로그램과 연관된 명령어들 및/또는 데이터를 저장하기에 적합한 컴퓨터 판독 가능한 저장 매체는, EPROM(Erasable Programmable Read-Only Memory), EEPROM(Electrically Erasable PROM), 플래시 메모리 장치 등의 반도체 메모리 장치를 포함하는 임의의 형태의 비 휘발성 메모리를 포함할 수 있으나, 이에 한정되지 않는다. 예를 들어, 컴퓨터 판독 가능한 저장 매체는 내부 하드 디스크 또는 이동식 디스크와 같은 자기 디스크, 광 자기 디스크, CD-ROM 및 DVD-ROM 디스크를 포함할 수 있다.

【0102】 사용자와의 상호 작용을 제공하기 위해, 컴퓨팅 장치는 정보를 사용자에게 제공하거나 디스플레이하기 위한 디스플레이 장치(예를 들어, CRT (Cathode Ray Tube), LCD(Liquid Crystal Display) 등) 및 사용자가 컴퓨팅 장치 상에 입력

및/또는 명령 등을 제공할 수 있는 포인팅 장치(예를 들어, 키보드, 마우스, 트랙볼 등)를 포함할 수 있으나, 이에 한정되지 않는다. 즉, 컴퓨팅 장치는 사용자와의 상호 작용을 제공하기 위한 임의의 다른 종류의 장치들을 더 포함할 수 있다. 예를 들어, 컴퓨팅 장치는 사용자와의 상호 작용을 위해, 시각적 피드백, 청각 피드백 및/또는 촉각 피드백 등을 포함하는 임의의 형태의 감각 피드백을 사용자에게 제공할 수 있다. 이에 대해, 사용자는 시각, 음성, 동작 등의 다양한 제스처를 통해 컴퓨팅 장치로 입력을 제공할 수 있다.

【0103】 본 발명에서, 다양한 실시예들은 백엔드 구성 요소(예: 데이터 서버), 미들웨어 구성 요소(예: 애플리케이션 서버) 및/또는 프론트 엔드 구성 요소를 포함하는 컴퓨팅 시스템에서 구현될 수 있다. 이 경우, 구성 요소들은 통신 네트워크와 같은 디지털 데이터 통신의 임의의 형태 또는 매체에 의해 상호 연결될 수 있다. 예를 들어, 통신 네트워크는 LAN(Local Area Network), WAN(Wide Area Network) 등을 포함할 수 있다.

【0104】 본 명세서에서 기술된 예시적인 실시예들에 기반한 컴퓨팅 장치는, 사용자 디바이스, 사용자 인터페이스(UI) 디바이스, 사용자 단말 또는 클라이언트 디바이스를 포함하여 사용자와 상호 작용하도록 구성된 하드웨어 및/또는 소프트웨어를 사용하여 구현될 수 있다. 예를 들어, 컴퓨팅 장치는 랩톱(laptop) 컴퓨터와 같은 휴대용 컴퓨팅 장치를 포함할 수 있다. 추가적으로 또는 대안적으로, 컴퓨팅 장치는, PDA(Personal Digital Assistants), 태블릿 PC, 게임 콘솔(game console), 웨어러블 디바이스(wearable device), IoT(internet of things) 디바이스,

VR(virtual reality) 디바이스, AR(augmented reality) 디바이스 등을 포함할 수 있으나, 이에 한정되지 않는다. 컴퓨팅 장치는 사용자와 상호 작용하도록 구성된 다른 유형의 장치를 더 포함할 수 있다. 또한, 컴퓨팅 장치는 이동 통신 네트워크 등의 네트워크를 통한 무선 통신에 적합한 휴대용 통신 디바이스(예를 들어, 이동 전화, 스마트 전화, 무선 셀룰러 전화 등) 등을 포함할 수 있다. 컴퓨팅 장치는, 무선 주파수(RF; Radio Frequency), 마이크로파 주파수(MWF; Microwave Frequency) 및/또는 적외선 주파수(IRF; Infrared Ray Frequency)와 같은 무선 통신 기술들 및/또는 프로토콜들을 사용하여 네트워크 서버와 무선으로 통신하도록 구성될 수 있다.

【0105】 본 발명에서 특정 구조적 및 기능적 세부 사항을 포함하는 다양한 실시예들은 예시적인 것이다. 따라서, 본 개시의 실시예들은 상술된 것으로 한정되지 않으며, 여러 가지 다른 형태로 구현될 수 있다. 또한, 본 발명에서 사용된 용어는 일부 실시예를 설명하기 위한 것이며 실시예를 제한하는 것으로 해석되지 않는다. 예를 들어, 단수형 단어 및 상기는 문맥상 달리 명확하게 나타내지 않는 한 복수형도 포함하는 것으로 해석될 수 있다.

【0106】 본 발명에서, 달리 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함하여 본 명세서에서 사용되는 모든 용어는 이러한 개념이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 또한, 사전에 정의된 용어와 같이 일반적으로 사용되는 용어들은 관련 기술의 맥락에서의 의미와 일치하는 의미를 갖는 것으로 해석되어야 한다.

【0107】본 명세서에서는 본 발명이 일부 실시예들과 관련하여 설명되었지만, 본 개시의 발명이 속하는 기술분야의 통상의 기술자가 이해할 수 있는 본 개시의 범위를 벗어나지 않는 범위에서 다양한 변형 및 변경이 이루어질 수 있다. 또한, 그러한 변형 및 변경은 본 명세서에 첨부된 특허청구의 범위 내에 속하는 것으로 생각되어야 한다.

【부호의 설명】

【0109】 100: 컴퓨팅 장치

102: 제어기

104: 공간 생성기

106: 데이터 저장소

108: 획득 함수

110: 대체 모델

120: 데이터 처리 시스템

【청구범위】

【청구항 1】

적어도 하나의 프로세서에 의해 수행되는 파라미터 튜닝 방법으로서,

고차원의 입력 공간(input space)의 구성 파라미터(configuration parameter) 및 상기 구성 파라미터에 대응하는 성능 메트릭(performance metric)을 수집하는 단계;

상기 구성 파라미터에 매핑되는 저차원의 타겟 공간(target space) 상의 합성 파라미터(synthetic parameter)를 생성하는 단계;

상기 합성 파라미터 및 상기 성능 메트릭과 연관된 샘플(sample)을 기초로 베이저안 최적화(Bayesian optimization)를 수행하여 후보 파라미터(candidate parameter)를 추출하는 단계; 및

상기 후보 파라미터에 대한 성능 평가를 수행하는 단계;

를 포함하는 파라미터 튜닝 방법.

【청구항 2】

제1항에 있어서,

상기 고차원의 입력 공간의 구성 파라미터 및 상기 구성 파라미터에 대응하는 성능 메트릭을 수집하는 단계는,

데이터 처리 시스템의 마스터 노드(master node)에 상기 구성 파라미터를 제

공하는 단계; 및

상기 마스터 노드로부터 상기 제공된 구성 파라미터에 대응하는 상기 성능 메트릭을 수신하는 단계;

를 포함하는 파라미터 튜닝 방법.

【청구항 3】

제1항에 있어서,

상기 고차원의 입력 공간에 대응하는 저차원의 타겟 공간을 생성하는 단계;

를 더 포함하는 파라미터 튜닝 방법.

【청구항 4】

제1항에 있어서,

상기 합성 파라미터 및 상기 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출하는 단계는,

대체 모델(surrogate model)을 기초로 상기 샘플에 대응하는 목적 함수(objective function)를 결정하는 단계; 및

획득 함수(acquisition function)를 기초로 상기 목적 함수를 최적화하기 위한 상기 후보 파라미터를 추출하는 단계;

를 포함하는 파라미터 튜닝 방법.

【청구항 5】

제4항에 있어서,

상기 획득 함수를 기초로 상기 목적 함수를 최적화하기 위한 후보 파라미터를 추출하는 단계는,

EI(expected improvement) 상에 노이즈(noise)에 관한 패널티(penalty)를 결합한 AEI(augmented expected improvement)를 이용하는 상기 획득 함수를 기초로 상기 목적 함수를 최적화하기 위한 후보 파라미터를 추출하는 단계;

를 포함하는 파라미터 튜닝 방법.

【청구항 6】

제1항에 있어서,

상기 후보 파라미터에 대한 성능 평가를 수행하는 단계는,

상기 타겟 공간 상의 후보 파라미터를 상기 입력 공간 상의 후보 구성 파라미터로 변환하는 단계;

상기 변환된 후보 구성 파라미터를 데이터 처리 시스템의 마스터 노드에 제공하는 단계; 및

상기 마스터 노드로부터 상기 제공된 후보 구성 파라미터에 대응하는 성능

메트릭을 수신하는 단계;

를 포함하는 파라미터 튜닝 방법.

【청구항 7】

제1항에 있어서,

튜닝 세션(tuning session) 동안 상기 타겟 공간의 타겟 차원수(target dimensionality)를 점진적으로 증가시키는 단계;

를 더 포함하는 파라미터 튜닝 방법.

【청구항 8】

제7항에 있어서,

상기 튜닝 세션 동안 상기 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계는,

평가 버짓(evaluation budget) 내에서 상기 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계;

를 포함하는 파라미터 튜닝 방법.

【청구항 9】

제8항에 있어서,

상기 튜닝 세션 동안 상기 타겟 공간의 타겟 차원수를 점진적으로 증가시키는 단계는,

상기 타겟 차원수가 증가할수록 상기 타겟 차원수에 할당되는 상기 평가 버킷을 증가시키는 단계;

를 더 포함하는 파라미터 튜닝 방법.

【청구항 10】

제1항 내지 제9항 중 어느 한 항에 따른 방법을 컴퓨터에서 실행하기 위해 컴퓨터 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램.

【청구항 11】

컴퓨팅 장치로서,

통신 모듈;

메모리; 및

상기 메모리와 연결되고, 상기 메모리에 포함된 컴퓨터 판독 가능한 적어도 하나의 프로그램을 실행하도록 구성된 적어도 하나의 프로세서

를 포함하고,

상기 적어도 하나의 프로그램은,

고차원의 입력 공간의 구성 파라미터 및 상기 구성 파라미터에 대응하는 성

능 메트릭을 수집하고,

상기 구성 파라미터에 매핑되는 저차원의 타겟 공간 상의 합성 파라미터를 생성하고,

상기 합성 파라미터 및 상기 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출하고,

상기 후보 파라미터에 대한 성능 평가를 수행하기 위한 명령어들을 포함하는 컴퓨팅 장치.

【요약서】**【요약】**

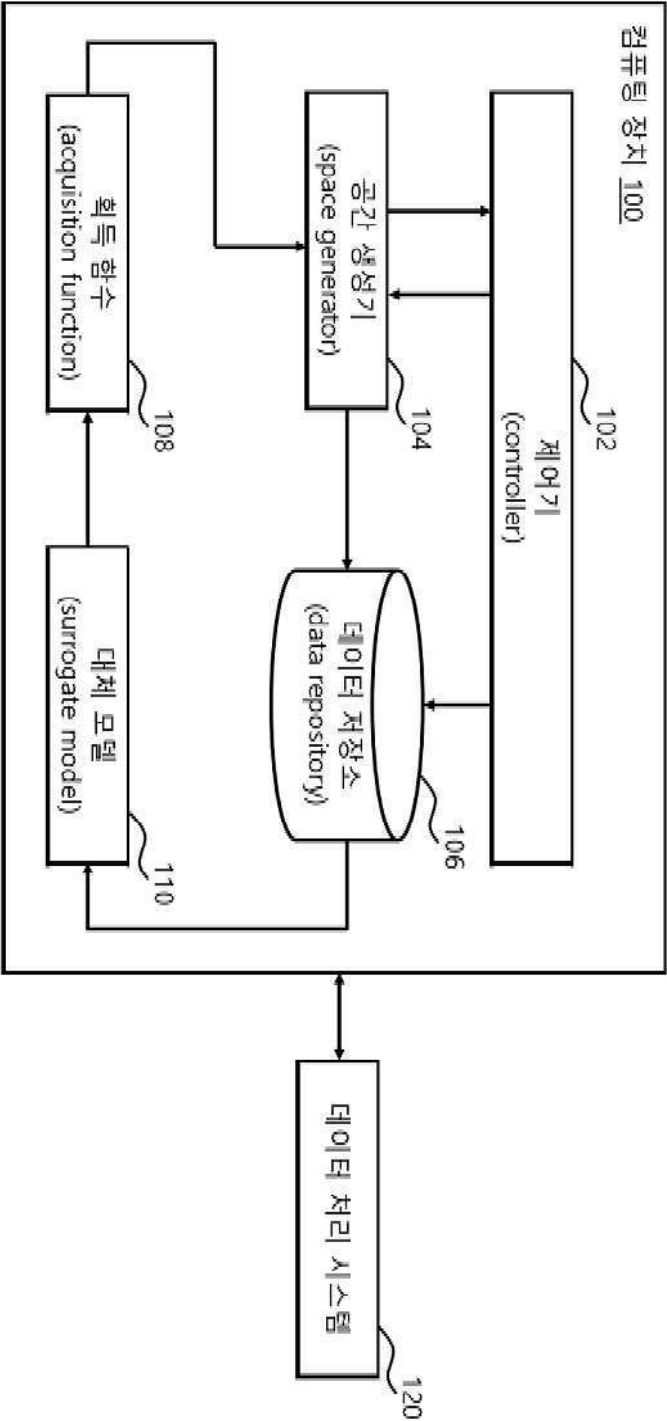
본 개시는 파라미터 튜닝 방법에 관한 것이다. 파라미터 튜닝 방법은, 고차원의 입력 공간의 구성 파라미터 및 구성 파라미터에 대응하는 성능 메트릭을 수집하는 단계, 구성 파라미터에 매핑되는 저차원의 타겟 공간 상의 합성 파라미터를 생성하는 단계, 합성 파라미터 및 성능 메트릭과 연관된 샘플을 기초로 베이지안 최적화를 수행하여 후보 파라미터를 추출하는 단계 및 후보 파라미터에 대한 성능 평가를 수행하는 단계를 포함한다.

【대표도】

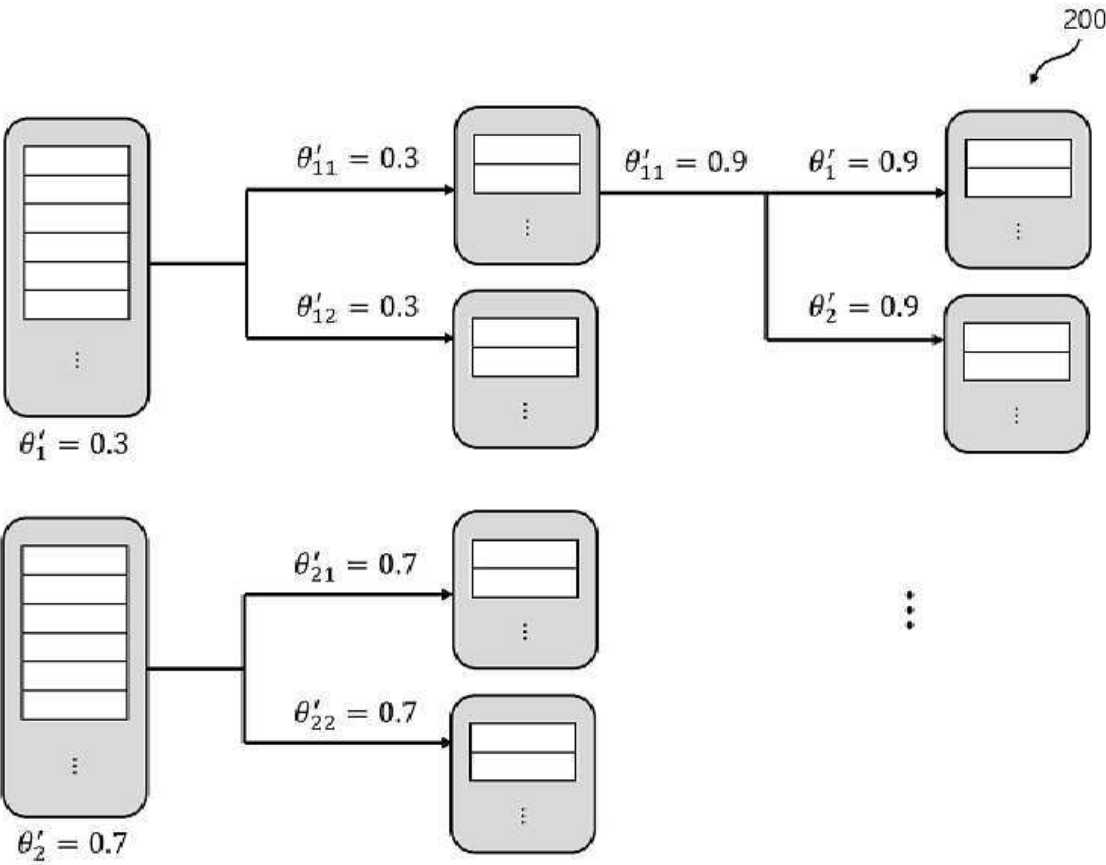
도 1

【도면】

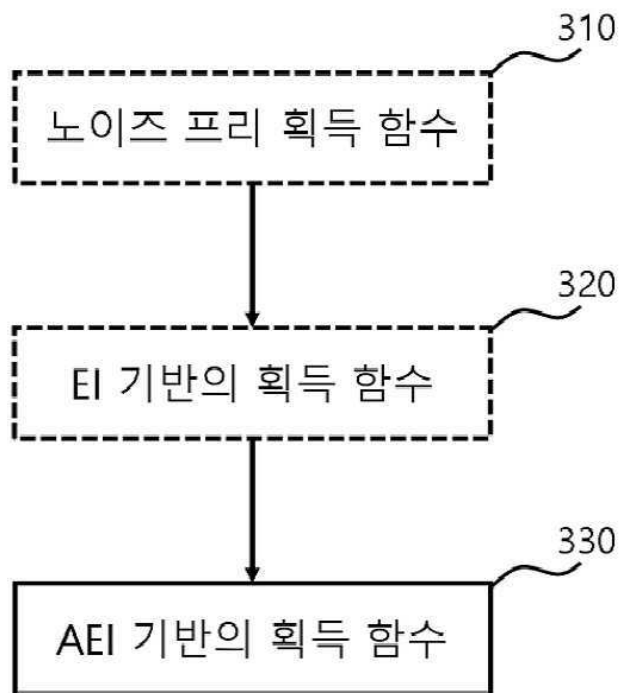
【도 1】



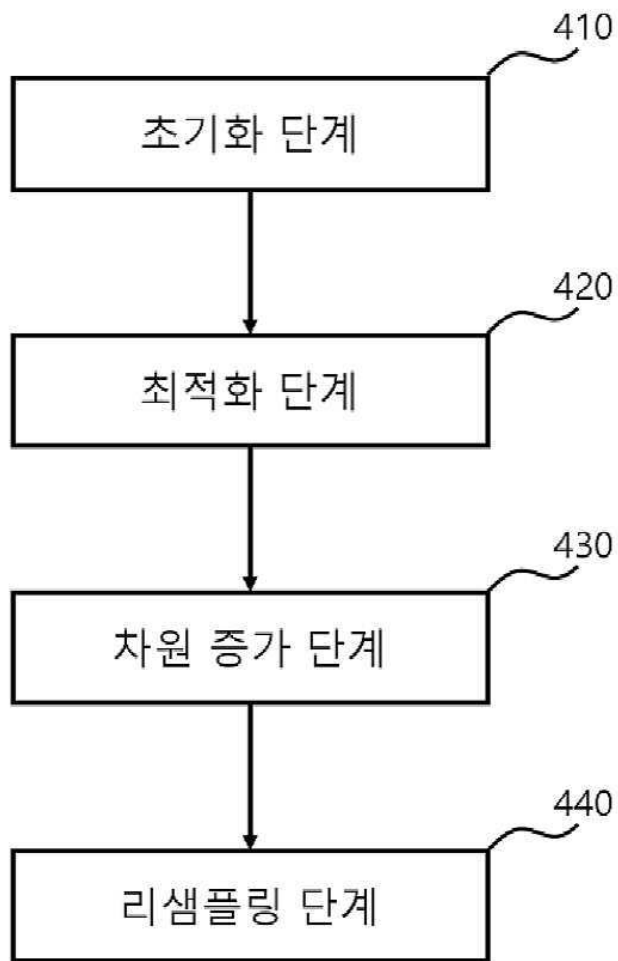
【도 2】



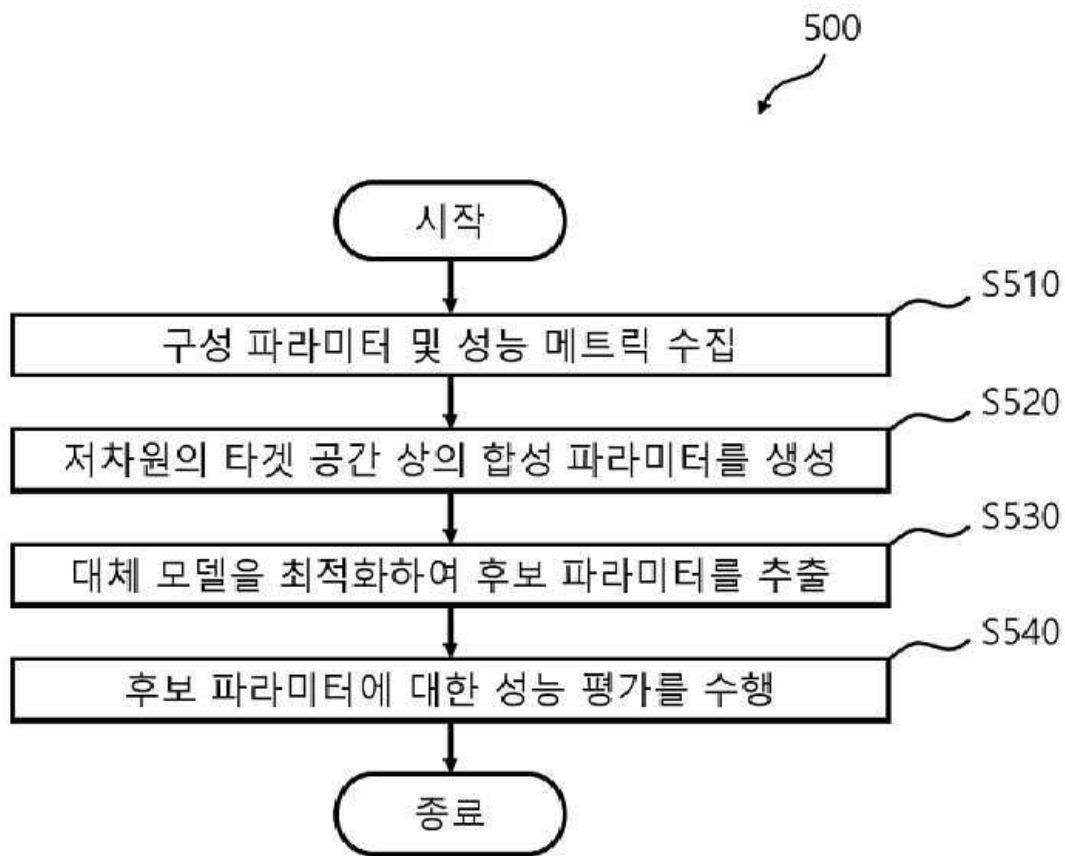
【도 3】



【도 4】



【도 5】



【도 6】

