

【서지사항】

【서류명】	특허출원서
【참조번호】	SDP20234087
【출원구분】	특허출원
【출원인】	
【명칭】	연세대학교 산학협력단
【특허고객번호】	2-2005-009509-9
【대리인】	
【명칭】	특허법인시공
【대리인번호】	9-2023-100041-2
【지정된변리사】	권성현, 최인혜
【포괄위임등록번호】	2023-059479-9
【발명의 국문명칭】	메타튠 기반 데이터베이스 튜닝 방법 및 장치
【발명의 영문명칭】	DEVICE AND METHOD FOR TUNING DATABASE WITH METATUNE
【발명자】	
【성명】	박상현
【성명의 영문표기】	SANGHYUN PARK
【주민등록번호】	670101-1XXXXXX
【우편번호】	08004
【주소】	서울특별시 양천구 오목로 300, 204동 3701호
【발명자】	
【성명】	염찬호

【성명의 영문표기】 CHANHO YEOM
【주민등록번호】 970106-1XXXXXX
【우편번호】 03726
【주소】 서울특별시 서대문구 연희로10가길 51-12, 305호
【발명자】

【성명】 이지은
【성명의 영문표기】 JIEUN LEE
【주민등록번호】 960912-2XXXXXX
【우편번호】 03716
【주소】 서울특별시 서대문구 동교로 291, 101동 1004호
【발명자】

【성명】 서상민
【성명의 영문표기】 SANGMIN SEO
【주민등록번호】 930507-1XXXXXX
【우편번호】 63272
【주소】 제주특별자치도 제주시 고마로 44, 801호
【발명자】

【성명】 권세인
【성명의 영문표기】 SEIN KWON
【주민등록번호】 970908-2XXXXXX
【우편번호】 03708
【주소】 서울특별시 서대문구 연희로 93, 702호

【출원언어】 국어

【심사청구】 청구

【이 발명을 지원한 국가연구개발사업】

【과제고유번호】 1711193130

【과제번호】 2017-0-00477-007

【부처명】 과학기술정보통신부

【과제관리(전문)기관명】 정보통신기획평가원

【연구사업명】 SW컴퓨팅산업원천기술개발

【연구과제명】 (SW 스타랩) IoT 환경을 위한 고성능 플래시 메모리 스토리지 기반 인메모리 분산 DBMS 개발

【기여율】 1/2

【과제수행기관명】 연세대학교 산학협력단

【연구기간】 2023.01.01 ~ 2023.12.31

【이 발명을 지원한 국가연구개발사업】

【과제고유번호】 1711193986

【과제번호】 2020-0-01361-004

【부처명】 과학기술정보통신부

【과제관리(전문)기관명】 정보통신기획평가원

【연구사업명】 정보통신방송혁신인재양성

【연구과제명】 인공지능대학원지원(연세대학교)

【기여율】 1/2

【과제수행기관명】 연세대학교 산학협력단

【연구기간】 2023.01.01 ~ 2023.12.31

【취지】 위와 같이 특허청장에게 제출합니다.

대리인 특허법인시공 (서명 또는 인)

【수수료】

【출원료】 0 면 46,000 원

【가산출원료】 23 면 0 원

【우선권주장료】 0 건 0 원

【심사청구료】 9 항 625,000 원

【합계】 671,000원

【감면사유】 전담조직(50%감면)[1]

【감면후 수수료】 335,500 원

【발명의 설명】

【발명의 명칭】

메타튠 기반 데이터베이스 튜닝 방법 및 장치{DEVICE AND METHOD FOR TUNING DATABASE WITH METATUNE}

【기술분야】

【0001】 본 개시는 데이터베이스 튜닝 방법에 관한 것이다.

【발명의 배경이 되는 기술】

【0003】 빅데이터 분석 및 서비스에 대한 수요가 증가함에 따라 비정형 데이터를

【0004】 효율적으로 처리하기 위한 데이터 저장 엔진이 주목받고 있다. 대량의 비정형 데이터를 관리하려면 이러한 매개변수를 조정하는 것이 중요하다. 그러나 데이터베이스 성능을 최적화하기 위한 적절한 configuration을 결정하는 것은 데이터베이스 관리자(DBA)에게 상당한 노력이 필요한데, 특히 최적의 configuration은 워크로드에 따라 달라지고 매번 추가 튜닝이 필요하기 때문이다. 또한, DBA는 최적의 configuration을 결정하기 위해 시행착오를 거치며 데이터베이스 관리 시스템(DBMS)을 벤치마킹하는 데 상당한 시간을 소비하기 때문에 업무의 효율성이 떨어진다.

【발명의 내용】

【해결하고자 하는 과제】

【0006】 본 개시에서는 상술한 문제를 해결하기 위하여 모델에 구애받지 않는 메타 학습 기법을 사용하는 자동 데이터베이스 튜닝 방법이 제공된다.

【과제의 해결 수단】

【0008】 본 개시의 일 실시예에 따르면, 메타튠 기반 데이터베이스 튜닝 방법은 데이터베이스의 제1 튜닝 세션에서 관찰된 복수의 워크로드 샘플을 획득하는 단계, 데이터베이스의 제2 튜닝 세션에서 타겟 워크로드와 복수의 워크로드 샘플 각각 사이의 거리를 산출하는 단계 - 제2 튜닝 세션은 제1 튜닝 세션 이후에 수행됨 -, 복수의 워크로드 샘플 중 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출하는 단계, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계 및 기계학습 모델을 이용하여 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계를 포함할 수 있다.

【0009】 일 실시예에 따르면, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계는, 제1 세트의 워크로드 샘플을 기초로, 기계학습 모델의 초기 파라미터를 제1 파라미터로 결정하는 단계를 포함할

수 있다.

【0010】 일 실시예에 따르면, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계는, 제2 세트의 워크로드 샘플을 기초로 기계학습 모델에 대한 파인 티닝을 수행하여 기계학습 모델의 파라미터를 제1 파라미터에서 제2 파라미터로 업데이트하는 단계를 더 포함할 수 있다.

【0011】 일 실시예에 따르면, 기계학습 모델은 ConvS2S(Convolution Sequence-to-sequence) 모델을 포함할 수 있다.

【0012】 일 실시예에 따르면, 거리는 Mahalanobis 거리를 나타낼 수 있다.

【0013】 일 실시예에 따르면, 기계학습 모델을 이용하여 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계는, 기계학습 모델을 기초로 데이터베이스를 위한 복수의 컨피규레이션(configuration) 중 성능이 가장 우수한 하나의 컨피규레이션을 추출하는 단계를 포함할 수 있다.

【0014】 일 실시예에 따르면, 복수의 컨피규레이션은 유전 알고리즘(GA, Genetic Algorithm)을 기초로 획득될 수 있다.

【0015】 본 개시의 다른 실시예에 따르면, 메타튠 기반 데이터베이스 튜닝 방법을 실행시키도록 컴퓨터로 판독 가능한 기록 매체에 기록된 컴퓨터 프로그램이 제공될 수 있다.

【0016】 본 개시의 또 다른 실시예에 따르면, 메타튠 기반 데이터베이스 튜닝 장치는 메모리 및 메모리와 연결되고, 메모리에 포함된 컴퓨터 판독 가능한 적어도 하나의 프로그램을 실행하도록 구성된 적어도 하나의 프로세서를 포함하고, 적어도 하나의 프로그램은, 데이터베이스의 제1 튜닝 세션에서 관찰된 복수의 워크로드 샘플을 획득하고, 데이터베이스의 제2 튜닝 세션에서 타겟 워크로드와 복수의 워크로드 샘플 각각 사이의 거리를 산출하고 - 제2 튜닝 세션은 제1 튜닝 세션 이후에 수행됨 -, 복수의 워크로드 샘플 중 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출하고, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키고, 기계학습 모델을 이용하여 데이터베이스를 위한 하나의 컨피규레이션을 획득하기 위한 명령어들을 포함할 수 있다.

【발명의 효과】

【0018】 본 개시의 일부 실시예에 따르면, 유전 알고리즘(GA)을 이용한 자동 튜닝 시스템을 제안하여 광범위한 검색 공간을 빠르고 정확하게 탐색할 수 있다.

【0019】 본 개시의 일부 실시예에 따르면, 성능 예측 모델을 활용하고 메타학습을 적용함으로써 데이터베이스를 반복적으로 실행하는 데 시간이 많이 걸리고 많은 양의 고품질 데이터가 필요하다는 문제를 크게 완화할 수 있다.

【도면의 간단한 설명】

【0021】 도 1은 본 개시의 일 실시예에 따른 메타튠 기반 데이터베이스 튜닝 방법을 나타내는 아키텍처이다.

도 2는 본 개시의 일 실시예에 따른 메타튠 기반 데이터베이스 튜닝 방법을 수행하는 정보 처리 시스템의 내부 구성을 나타내는 블록도이다.

도 3은 본 개시의 일 실시예에 따른 메타튠 기반 데이터베이스 튜닝 방법을 나타내는 흐름도이다.

【발명을 실시하기 위한 구체적인 내용】

【0022】 이하, 본 개시의 실시를 위한 구체적인 내용을 첨부된 도면을 참조하여 상세히 설명한다. 다만, 이하의 설명에서는 본 개시의 요지를 불필요하게 헐릴 우려가 있는 경우, 널리 알려진 기능이나 구성에 관한 구체적 설명은 생략하기로 한다.

【0023】 첨부된 도면에서, 동일하거나 대응하는 구성요소에는 동일한 참조번호가 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 의도 되지는 않는다.

【0024】 개시된 실시예의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러

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

【0025】 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 개시된 실시예에 대해 구체적으로 설명하기로 한다. 본 명세서에서 사용되는 용어는 본 개시에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 관련 분야에 종사하는 기술자의 의도 또는 판례, 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 개시에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 개시의 전반에 걸친 내용을 토대로 정의되어야 한다.

【0026】 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 '포함'한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다.

【0027】 본 개시의 일 실시예에 따르면, '메모리'는 전자 정보를 저장 가능한 임의의 전자 컴포넌트를 포함하도록 넓게 해석되어야 한다. '메모리'는 임의 액

세스 메모리(RAM), 판독-전용 메모리(ROM), 비-휘발성 임의 액세스 메모리(NVRAM), 프로그램가능 판독-전용 메모리(PROM), 소거-프로그램가능 판독 전용 메모리(EPROM), 전기적으로 소거가능 PROM(EEPROM), 플래쉬 메모리, 자기 또는 광학 데이터 저장장치, 레지스터들 등과 같은 프로세서-판독가능 매체의 다양한 유형들을 지칭할 수도 있다. 프로세서가 메모리로부터 정보를 판독하고/하거나 메모리에 정보를 기록할 수 있다면 메모리는 프로세서와 전자 통신 상태에 있다고 불린다.

【0028】 도 1은 본 개시의 일 실시예에 따른 메타튠 기반 데이터베이스 튜닝 방법을 나타내는 아키텍쳐이다. 본 개시에서 메타튠 기반 데이터베이스 튜닝 방법은 정보 처리 시스템(도 2에 후술됨)의 적어도 하나의 프로세서에 의해 수행될 수 있다. 한편, 도시된 바와 같이, 방법은 워크로드 선택 단계, 모델 트레이닝 단계 및 컨피규레이션 추천 단계로 구분될 수 있다.

【0029】 프로세서는 데이터베이스의 성능을 결정하는 컨피규레이션을 제공받기 위하여, 컨피규레이션 성능 예측 모델을 학습시킬 수 있다. 구체적으로, 프로세서는 데이터베이스와 연관된 워크로드(들)을 기초로 컨피규레이션의 성능을 예측하도록 구성되는 기계학습 모델을 학습시킬 수 있다. 이를 위해, 프로세서는 데이터베이스의 한 튜닝 세션에서 타겟 워크로드를 위한 워크로드 샘플(들)을 획득할 수 있다.

【0030】 프로세서는 상술한 기계학습 모델의 초기 파라미터를 결정하기 위해 사용되는 워크로드 샘플(들)을 획득할 수 있다. 구체적으로, 프로세서는 데이터베이스의 이전 튜닝 세션에서 관찰된 복수의 워크로드 샘플을 획득할 수 있다. 그리

고 나서, 프로세서는 복수의 워크로드 샘플 중 현재 튜닝 세션에서의 타겟 워크로드와 유사한 일부 워크로드 샘플(들)을 추출할 수 있다. 예를 들어, 프로세서는 복수의 워크로드 샘플 각각과 타겟 워크로드 사이의 거리를 산출할 수 있다. 그리고 나서, 프로세서는 복수의 워크로드 샘플 중 타겟 워크로드와의 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출할 수 있다.

【0031】 프로세서는 추출된 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 데이터베이스의 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시킬 수 있다. 구체적으로, 프로세서는 제1 세트의 워크로드 샘플을 기초로, 기계학습 모델의 초기 파라미터를 제1 파라미터로 결정할 수 있다. 그리고 나서, 프로세서는 제2 세트의 워크로드 샘플을 기초로, 기계학습 모델의 파라미터를 제1 파라미터에서 제2 파라미터로 업데이트할 수 있다. 일반적으로, 이전 튜닝 세션에서 관찰된 제1 세트의 워크로드 샘플은 데이터의 양이 충분하고, 산출된 거리를 기초로 타겟 워크로드와 유사한 일부 워크로드만이 선별된 것이므로, 이를 기초로 기계학습 모델을 학습시키는 것은 성능 측면에서 유리하다.

【0032】 추가적으로, 프로세서는 제2 세트의 워크로드 샘플을 기초로 기계학습 모델에 대한 파인 튜닝을 수행하여 기계학습 모델의 파라미터를 업데이트할 수 있다. 이 경우, 제2 세트의 워크로드 샘플은 제1 세트의 워크로드 샘플보다 양적인 측면에서 충분하지 않지만, 타겟 워크로드를 위한 샘플들이라는 점에서 기계학습 모델의 파인 튜닝에 적합하다.

【0033】 상술한 방법에 따라 기계학습 모델의 학습이 완료된 후, 프로세서는 데이터베이스의 튜닝을 위한 복수의 컨피규레이션을 획득할 수 있다. 그리고 나서, 프로세서는 복수의 컨피규레이션 중 기계학습 모델을 통해 데이터베이스의 성능을 가장 우수하게 만드는 하나의 컨피규레이션을 추천받을 수 있다. 이러한 구성에 의해, 프로세서는 복수의 컨피규레이션 각각의 성능을 파악하기 위하여 매번 데이터베이스를 탐색할 필요가 없어지므로, 데이터베이스의 튜닝에 요구되는 연산 부담은 감소하게 된다.

【0034】 도 2는 본 개시의 일 실시예에 따른 메타튠 기반 데이터베이스 튜닝 방법을 수행하는 정보 처리 시스템(200)의 내부 구성을 나타내는 블록도이다. 정보 처리 시스템(200)은 메모리(210), 프로세서(220), 통신 모듈(230) 및 입출력 인터페이스(240)를 포함할 수 있다. 도 2에 도시된 바와 같이, 정보 처리 시스템(200)은 각각의 통신 모듈(230)을 이용하여 네트워크를 통해 정보 및/또는 데이터를 통신할 수 있도록 구성될 수 있다.

【0035】 메모리(210)는 비-일시적인 임의의 컴퓨터 판독 가능한 기록매체를 포함할 수 있다. 일 실시예에 따르면, 메모리(210)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 정보 처리 시스템(200)에 포함될 수 있다. 또한, 메모리(210)에는 운

영체제와 적어도 하나의 프로그램 코드(예를 들어, 정보 처리 시스템(200)에 설치되어 구동되는 정보 처리 시스템(200)의 제어를 위한 코드)가 저장될 수 있다.

【0036】 이러한 소프트웨어 구성요소들은 메모리(210)와는 별도의 컴퓨터에서 판독 가능한 기록매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록매체는 이러한 정보 처리 시스템(200)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(230)을 통해 메모리(210)에 로딩될 수도 있다. 예를 들어, 적어도 하나의 프로그램은 개발자들 또는 어플리케이션의 설치 파일을 배포하는 파일 배포 시스템이 네트워크를 통해 제공하는 파일들에 의해 설치되는 컴퓨터 프로그램에 기반하여 메모리(210)에 로딩될 수 있다.

【0037】 프로세서(220)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(210) 또는 통신 모듈(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.

【0038】 통신 모듈(230)은 네트워크를 통해 정보 처리 시스템(200)이 다른 시스템(일례로 별도의 클라우드 시스템 등)과 통신하기 위한 구성 또는 기능을 제공할 수 있다. 일례로, 정보 처리 시스템(200)의 프로세서(220)의 제어에 따라 제

공되는 제어 신호나 명령이 통신 모듈(230)과 네트워크를 거쳐 다른 시스템에 수신될 수 있다.

【0039】 입출력 인터페이스(240)는 정보 처리 시스템(200)과 연결되거나 정보 처리 시스템(200)이 포함할 수 있는 입력 또는 출력을 위한 장치(미도시)와의 인터페이스를 위한 수단일 수 있다. 도 2에서는 입출력 인터페이스(240)가 프로세서(220)와 별도로 구성된 요소로서 도시되었으나, 이에 한정되지 않으며, 입출력 인터페이스(240)가 프로세서(220)에 포함되도록 구성될 수 있다.

【0040】 정보 처리 시스템(200)은 도 2의 구성요소들보다 더 많은 구성요소들을 포함할 수 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다.

【0041】 일 실시예에 따르면, 프로세서(220)는 신경망의 학습 및 추론을 위한 프로그램을 동작하도록 구성될 수 있다. 이 때, 해당 프로그램과 연관된 코드가 메모리(210)에 로딩될 수 있다. 프로그램이 동작되는 동안에, 프로세서(220)는 입출력 장치(미도시)로부터 제공된 정보 및/또는 데이터를 입출력 인터페이스(240)를 통해 수신하거나 통신 모듈(230)을 통해 다른 시스템으로부터 정보 및/또는 데이터를 수신할 수 있으며, 수신된 정보 및/또는 데이터를 처리하여 메모리(210)에 저장할 수 있다. 또한, 이러한 정보 및/또는 데이터는 통신 모듈(230)을 통해 다른 시스템에 제공될 수 있다.

【0042】 프로세서(220)는 복수의 다른 시스템으로부터 수신된 정보 및/또는 데이터를 관리, 처리 및/또는 저장하도록 구성될 수 있다. 일 실시예에 따르면, 프로세서(220)는 다른 시스템으로부터 수신한 이미지, 이미지로부터 추출된 특징 등을 저장, 처리 및 전송할 수 있다. 추가적으로 또는 대안적으로, 프로세서(220)는 네트워크와 연결된 별도의 클라우드 시스템, 데이터베이스 등으로부터 인공 신경망의 학습 및 추론을 위해 이용되는 알고리즘을 실행하기 위한 프로그램 등을 저장 및/또는 업데이트하도록 구성될 수 있다.

【0043】 도 3은 본 개시의 일 실시예에 따른 메타튠 기반 데이터베이스 튜닝 방법(300)을 나타내는 흐름도이다. 방법(300)은 정보 처리 시스템의 적어도 하나의 프로세서(예: 프로세서(220))에 의해 수행될 수 있다. 한편, 도시된 바와 같이 방법(300)은 메타튠 기반 데이터베이스 튜닝 방법은 데이터베이스의 제1 튜닝 세션에서 관찰된 복수의 워크로드 샘플을 획득하는 단계(S310)로 개시될 수 있다.

【0044】 프로세서는 데이터베이스의 제2 튜닝 세션에서 타겟 워크로드와 복수의 워크로드 샘플 각각 사이의 거리를 산출할 수 있다(S320). 이 경우, 제2 튜닝 세션은 제1 튜닝 세션 이후에 수행된 세션으로 구성될 수 있다. 그리고 나서, 프로세서는 복수의 워크로드 샘플 중 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출할 수 있다(S330).

【0045】 프로세서는 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시킬 수 있다(S340). 그리고 나서, 프로세서는 기계

학습 모델을 이용하여 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계를 포함할 수 있다(S350).

【0046】 일 실시예에 따르면, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계(S340)는, 제1 세트의 워크로드 샘플을 기초로, 기계학습 모델의 초기 파라미터를 제1 파라미터로 결정하는 단계를 포함할 수 있다.

【0047】 일 실시예에 따르면, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계(S340)는, 제2 세트의 워크로드 샘플을 기초로 기계학습 모델에 대한 파인 투닝을 수행하여 기계학습 모델의 파라미터를 제1 파라미터에서 제2 파라미터로 업데이트하는 단계를 더 포함할 수 있다.

【0048】 일 실시예에 따르면, 기계학습 모델은 ConvS2S(Convolution Sequence-to-sequence) 모델을 포함할 수 있다.

【0049】 일 실시예에 따르면, 거리는 Mahalanobis 거리를 나타낼 수 있다.

【0050】 일 실시예에 따르면, 기계학습 모델을 이용하여 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계는, 기계학습 모델을 기초로 데이터베이스를 위한 복수의 컨피규레이션(configuration) 중 성능이 가장 우수한 하나의 컨피규레이션을 추출하는 단계를 포함할 수 있다.

【0051】 일 실시예에 따르면, 복수의 컨파규레이션은 유전 알고리즘(GA, Genetic Algorithm)을 기초로 획득될 수 있다.

【0052】 본 개시의 앞선 설명은 통상의 기술자들이 본 개시를 행하거나 이용하는 것을 가능하게 하기 위해 제공된다. 본 개시의 다양한 수정예들이 통상의 기술자들에게 쉽게 자명할 것이고, 본원에 정의된 일반적인 원리들은 본 개시의 취지 또는 범위를 벗어나지 않으면서 다양한 변형예들에 적용될 수도 있다. 따라서, 본 개시는 본원에 설명된 예들에 제한되도록 의도된 것이 아니고, 본원에 개시된 원리들 및 신규한 특징들과 일관되는 최광의의 범위가 부여되도록 의도된다.

【0053】 비록 예시적인 구현예들이 하나 이상의 독립형 컴퓨터 시스템의 맥락에서 현재 개시된 주제의 양태들을 활용하는 것을 언급할 수도 있으나, 본 주제는 그렇게 제한되지 않고, 오히려 네트워크나 분산 컴퓨팅 환경과 같은 임의의 컴퓨팅 환경과 연계하여 구현될 수도 있다. 또 나아가, 현재 개시된 주제의 양상들은 복수의 프로세싱 칩들이나 디바이스들에서 또는 그들에 걸쳐 구현될 수도 있고, 스토리지는 복수의 디바이스들에 걸쳐 유사하게 영향을 받게 될 수도 있다. 이러한 디바이스들은 PC들, 네트워크 서버들, 및 핸드헬드 디바이스들을 포함할 수도 있다.

【0054】 본 명세서에서는 본 개시가 일부 실시예들과 관련하여 설명되었지만, 본 발명이 속하는 기술분야의 통상의 기술자가 이해할 수 있는 본 개시의 범위를 벗어나지 않는 범위에서 다양한 변형 및 변경이 이루어질 수 있다는 점을 알아야 할 것이다. 또한, 그러한 변형 및 변경은 본 명세서에서 첨부된 특허

2023-12-29

청구의 범위 내에 속하는 것으로 생각되어야 한다.

【청구범위】

【청구항 1】

데이터베이스의 제1 투닝 세션에서 관찰된 복수의 워크로드 샘플을 획득하는 단계;

상기 데이터베이스의 제2 투닝 세션에서 타겟 워크로드와 상기 복수의 워크로드 샘플 각각 사이의 거리를 산출하는 단계 - 상기 제2 투닝 세션은 상기 제1 투닝 세션 이후에 수행됨 -;

상기 복수의 워크로드 샘플 중 상기 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출하는 단계;

상기 제1 세트의 워크로드 샘플 및 상기 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계; 및

상기 기계학습 모델을 이용하여 상기 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계

를 포함하는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 투닝 방법.

【청구항 2】

제1항에 있어서,

상기 제1 세트의 워크로드 샘플 및 상기 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계는,

상기 제1 세트의 워크로드 샘플을 기초로 상기 기계학습 모델의 초기 파라미터를 제1 파라미터로 결정하는 단계
를 포함하는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 투닝 방법.

【청구항 3】

제2항에 있어서,
상기 제1 세트의 워크로드 샘플 및 상기 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계는,

상기 제2 세트의 워크로드 샘플을 기초로 상기 기계학습 모델에 대한 파인 투닝을 수행하여 상기 기계학습 모델의 파라미터를 상기 제1 파라미터에서 제2 파라미터로 업데이트하는 단계

를 더 포함하는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 투닝 방법.

【청구항 4】

제1항에 있어서,

상기 기계학습 모델은 ConvS2S(Convolution Sequence-to-sequence) 모델을 포함하는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 튜닝 방법.

【청구항 5】

제1항에 있어서,

상기 거리는 Mahalanobis 거리를 나타내는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 튜닝 방법.

【청구항 6】

제1항에 있어서,

상기 기계학습 모델을 이용하여 상기 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계는,

상기 기계학습 모델을 기초로 상기 데이터베이스를 위한 복수의 컨피규레이션(configuration) 중 성능이 가장 우수한 하나의 컨피규레이션을 추출하는 단계를 포함하는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 튜닝 방법.

【청구항 7】

제6항에 있어서,

상기 복수의 컨피규레이션은 유전 알고리즘(GA; Genetic Algorithm)을 기초로 획득되는, 적어도 하나의 프로세서에 의해 수행되는 메타튠 기반 데이터베이스 튜닝 방법.

【청구항 8】

제1항 내지 7항 중 어느 한 항에 기재된 메타튠 기반 데이터베이스 튜닝 방법을 실행시키도록 컴퓨터로 판독 가능한 기록 매체에 기록된 컴퓨터 프로그램.

【청구항 9】

메모리; 및

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

상기 적어도 하나의 프로그램은,

데이터베이스의 제1 튜닝 세션에서 관찰된 복수의 워크로드 샘플을 획득하고,

상기 데이터베이스의 제2 튜닝 세션에서 타겟 워크로드와 상기 복수의 워크로드 샘플 각각 사이의 거리를 산출하고 - 상기 제2 튜닝 세션은 상기 제1 튜닝 세션 이후에 수행됨 -,

상기 복수의 워크로드 샘플 중 상기 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출하고,

상기 제1 세트의 워크로드 샘플 및 상기 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키고,

상기 기계학습 모델을 이용하여 상기 데이터베이스를 위한 하나의 컨피규레이션을 획득하기 위한 명령어들을 포함하는, 메타튠 기반 데이터베이스 튜닝 장치.

【요약서】

【요약】

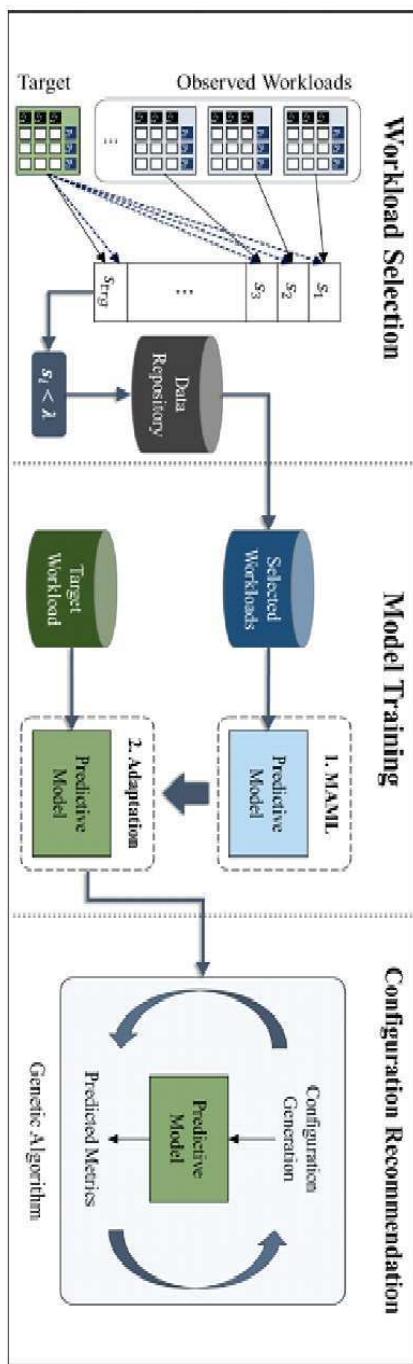
본 개시의 일 실시예에 따르면, 메타튠 기반 데이터베이스 튜닝 방법은 데이터베이스의 제1 튜닝 세션에서 관찰된 복수의 워크로드 샘플을 획득하는 단계, 데이터베이스의 제2 튜닝 세션에서 타겟 워크로드와 복수의 워크로드 샘플 각각 사이의 거리를 산출하는 단계 - 제2 튜닝 세션은 제1 튜닝 세션 이후에 수행됨 -, 복수의 워크로드 샘플 중 거리가 사전 결정된 임계치 미만인 제1 세트의 워크로드 샘플을 추출하는 단계, 제1 세트의 워크로드 샘플 및 타겟 워크로드를 위하여 사전 획득된 제2 세트의 워크로드 샘플을 기초로, 컨피규레이션 성능 예측을 위해 마련된 기계학습 모델을 학습시키는 단계 및 기계학습 모델을 이용하여 데이터베이스를 위한 하나의 컨피규레이션을 획득하는 단계를 포함할 수 있다.

【대표도】

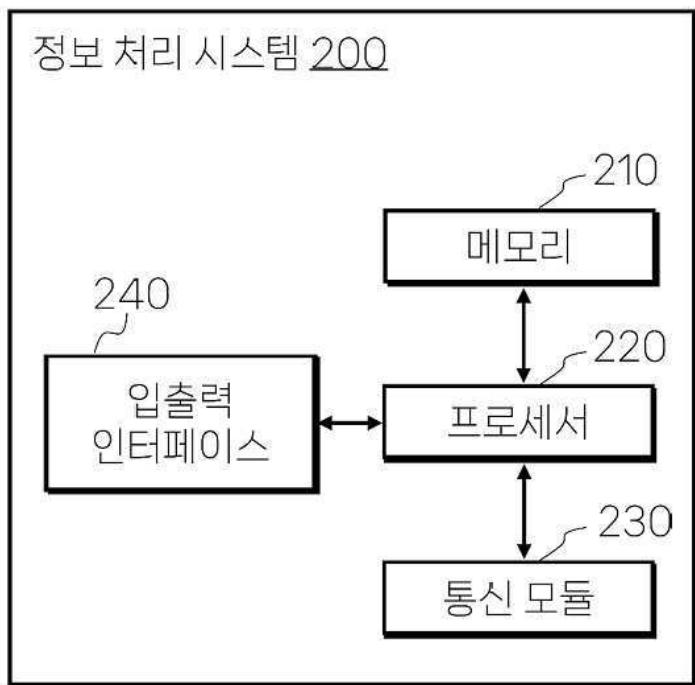
도 1

【도면】

【도 1】



【도 2】



【도 3】

