



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년09월30일
(11) 등록번호 10-1657414
(24) 등록일자 2016년09월07일

(51) 국제특허분류(Int. Cl.)
G06F 9/50 (2006.01) G06N 7/00 (2006.01)
(52) CPC특허분류
G06F 9/5027 (2013.01)
G06F 1/3234 (2013.01)
(21) 출원번호 10-2015-0070331
(22) 출원일자 2015년05월20일
심사청구일자 2015년05월20일
(56) 선행기술조사문헌
JP05334102 A*
W02014023607 A1*
KR1020120126549 A*
KR1020150050689 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
이승룡
경기도 성남시 분당구 구미로144번길 25, 102동 102호(구미동, 삼환빌라)
부이던마오
경기도 용인시 기흥구 덕영대로 1732, 전자정보대학 351호(서천동, 경희대학교 국제캠퍼스)
(74) 대리인
김연권

전체 청구항 수 : 총 8 항

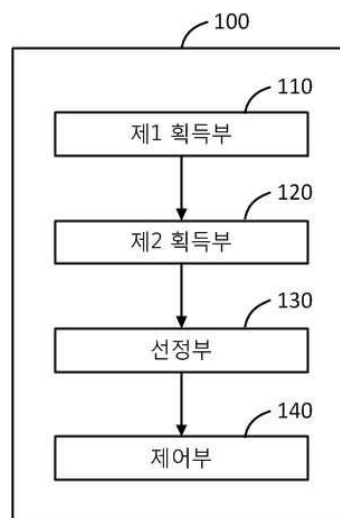
심사관 : 유진태

(54) 발명의 명칭 CPU 사용률 제어 장치 및 방법

(57) 요약

본 발명은 베이지안 학습(bayesian learning) 및 가우시안 프로세스 회귀(gaussian process regression)에 기초하여 각 코어들의 사용률을 획득하는 제1 획득부, 상기 획득된 사용률에 대응하는 상기 각 코어들의 평균 대기시간 및 차단율을 획득하는 제2 획득부, 상기 획득된 평균 대기시간 및 상기 획득된 차단율에 기초하여 소스 코어(source core)의 타겟 프로세싱(target processing) 및 목적지 코어(destination core)를 선정하는 선정부 및 상기 선정된 소스 코어의 타겟 프로세싱을 추출하여 상기 선정된 소스 코어를 비활성화되도록 제어하고, 상기 추출된 타겟 프로세싱을 상기 선정된 목적지 코어로 마이그레이팅(migrating)되도록 제어하는 제어부를 포함하는 CPU 사용률 제어 장치 및 방법을 개시한다.

대표도 - 도1



(52) CPC특허분류

G06F 1/3243 (2013.01)

G06F 11/2236 (2013.01)

G06N 7/005 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1415135342

부처명 산업통상자원부

연구관리전문기관 한국산업기술평가관리원

연구사업명 지식서비스산업핵심기술개발

연구과제명 퍼스널 빅데이터를 활용한 마이닝마인즈 핵심기술 개발

기 여 율 1/1

주관기관 경희대학교 산학협력단

연구기간 2014.06.01 ~ 2015.05.31

명세서

청구범위

청구항 1

베이지안 학습(bayesian learning) 및 가우시안 프로세스 회귀(gaussian process regression)에 기초하여 랜덤 변수의 유한 세트에 대응하는 각 코어들 별 프로세싱에 대한 입력 데이터의 평균 및 공분산을 획득하고, 상기 획득된 입력 데이터의 평균 및 공분산으로부터 가우시안 프로세스의 사후 분포(posterior distribution)를 획득하여 추정되는 출력 데이터로부터 상기 각 코어들의 사용률을 획득하는 제1 획득부;

PASTA(Poisson Arrival See Time Averages) 이론에 기초하여 상기 획득된 사용률에 대응하는 상기 각 코어들의 도착율에 대한 도착 평균시간을 획득하고, 폴라첵 힌친(Pollaczek-Khintchine) 이론에 기초하여 평균 대기시간을 획득하며, 상기 획득된 평균 대기시간을 통해 차단율을 획득하는 제2 획득부;

상기 획득된 평균 대기시간 및 상기 획득된 차단율에 기초하여 소스 코어(source core)의 타겟 프로세싱(target processing) 및 목적지 코어(destination core)를 선정하는 선정부; 및

상기 선정된 소스 코어의 타겟 프로세싱을 추출하여 상기 선정된 소스 코어를 비활성화되도록 제어하고, 상기 추출된 타겟 프로세싱을 상기 선정된 목적지 코어로 마이그레이팅(migrating)되도록 제어하는 제어부

를 포함하는 CPU 사용률 제어 장치.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 제2 획득부는

상기 각 코어들의 프로세싱에 대한 도착률과 서비스율에 대응하는 상기 평균 대기시간을 획득하고, 상기 획득된 평균 대기시간 및 상기 서비스율에 대응하는 상기 차단율을 획득하는

CPU 사용률 제어 장치.

청구항 4

제1항에 있어서,

상기 선정부는

상기 획득된 평균 대기시간의 선정 범위에 만족하는 상기 소스 코어의 타겟 프로세싱 및 상기 획득된 차단율의 선정 범위에 만족하는 상기 목적지 코어를 선정하는

CPU 사용률 제어 장치.

청구항 5

제4항에 있어서,

상기 선정부는

상기 획득된 평균 대기시간의 선정 범위에 만족하지 않은 경우, 상기 제1 획득부에게 상기 각 코어들의 사용률에 대한 갱신을 요청하는

CPU 사용률 제어 장치.

청구항 6

제5항에 있어서,

상기 선정부는

상기 선정된 목적지 코어가 상기 소스 코어와 동일하거나, 상기 획득된 차단율의 선정 범위에 만족하지 않은 경우, 상기 목적지 코어에 대한 선정을 취소하고, 상기 제1 획득부에게 상기 각 코어들의 사용률에 대한 상기 갱신을 요청하는

CPU 사용률 제어 장치.

청구항 7

CPU 사용률 제어 장치의 동작 방법에 있어서,

베이지안 학습 및 가우시안 프로세스 회귀에 기초하여 랜덤 변수의 유한 세트에 대응하는 각 코어들 별 프로세싱에 대한 입력 데이터의 평균 및 공분산을 획득하고, 상기 획득된 입력 데이터의 평균 및 공분산으로부터 가우시안 프로세스의 사후 분포(posterior distribution)를 획득하여 추정되는 출력 데이터로부터 상기 각 코어들의 사용률을 획득하는 단계;

PASTA(Poisson Arrival See Time Averages) 이론에 기초하여 상기 획득된 사용률에 대응하는 상기 각 코어들의 도착율에 대한 도착 평균시간을 획득하고, 폴라첵 힌친(Pollaczek-Khintchine) 이론에 기초하여 평균 대기시간을 획득하며, 상기 획득된 평균 대기시간을 통해 차단율을 획득하는 단계;

상기 획득된 평균 대기시간 및 상기 획득된 차단율에 기초하여 소스 코어의 타겟 프로세싱 및 목적지 코어를 선정하는 단계; 및

상기 선정된 소스 코어의 타겟 프로세싱을 추출하여 상기 선정된 소스 코어를 비활성화되도록 제어하고, 상기 추출된 타겟 프로세싱을 상기 선정된 목적지 코어로 마이그레이팅되도록 제어하는 단계

를 포함하는 CPU 사용률 제어 방법.

청구항 8

삭제

청구항 9

제7항에 있어서,

상기 평균 대기시간 및 상기 차단율을 획득하는 단계는

상기 각 코어들의 프로세싱에 대한 도착률과 서비스율에 대응하는 상기 평균 대기시간을 획득하고, 상기 획득된 평균 대기시간 및 상기 서비스율에 대응하는 상기 차단율을 획득하는

CPU 사용률 제어 방법.

청구항 10

제7항 및 제9항 중 어느 한 항의 방법을 수행하기 위한 프로그램이 기록된 컴퓨터로 판독 가능한 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 CPU 사용률 제어 장치 및 방법에 관한 것으로, 보다 상세하게는 소스 코어의 타겟 프로세싱을 선정된 목적지 코어로 마이그레이팅하여 CPU 사용률을 제어하는 CPU 사용률 제어 장치 및 방법에 관한 것이다.

배경 기술

[0002] 가우시안 프로세스(gaussian process)는 확률적 프로세스를 추론하는 데이터 통신 분야, 네트워킹 분야 및 다양한 컴퓨터 과학 분야에서 널리 사용되고 있고, 타겟 시스템의 특성을 모델링 할 수 있는 비-파라메트릭(non-parametric) 및 확률적 접근으로 사용되고 있다.

- [0003] 예를 들어, 가우시안 프로세스는 모델링의 파라미터를 결정하는 것보다, 실제 기본 기능을 표현하는 파라미터를 채택되도록 도와주는 기능으로 사용될 수 있다.
- [0004] 또한, 가우시안 프로세스는 노이즈 데이터, 손상된 데이터 및 에러가 존재하는 데이터에 대한 적합한 선택을 제공할 수 있고, 화상 분류, 차원 감소, 자율 학습 문제 및 시계열적 예측 등 다양한 기계 학습 솔루션을 포함하는 다양한 어플리케이션에게 제공될 수 있다.
- [0005] 가우시안 프로세스 회기(gaussian process regression)는 각각의 예측값의 신뢰도를 나타내는 가우시안 분포(gaussian distribution)의 평균과 분산의 세트를 포함할 수 있고, 추론과 보간의 목적으로 사용될 수 있다.
- [0006] 또한, 가우시안 프로세스 회기는 선형 회기(linear regression), k-최근린(k-nearest neighbor), 다변량 회귀 스플라인(multi-layer perceptron), 다층 퍼셉트론(multi-layer perceptron) 및 지원 벡터 회귀(support vector regression) 방법에 비해 정확성 관점에서 보다 우수한 성능과 유연성을 제공할 수 있다.
- [0007] 이러한 가우시안 프로세스 회기는 데이터세트의 n개의 트레이닝 포인트를 계산할 때, 계산 복잡도가 증가하는 문제점이 존재하여 시스템의 전체 성능을 저하시키는 단점이 존재할 수 있고, 대규모 시스템에서 그 영향이 더욱 커질 수 있다.
- [0008] 종래에는 콜레스키 분해(cholesky decomposition) 및 감소된 랭크 공분산 매트릭스(reduced rank covariance matrix)를 이용하여 가우시안 프로세스 회기의 계산 복잡도를 줄이는 연구가 진행되고 있었으나, 대규모 시스템의 요구사항을 만족시키는데 어려움이 존재하였다.

선행기술문헌

비특허문헌

- [0009] (비특허문헌 0001) D. Petelin, B. Filip c, and J. Kocijan, "Optimization of gaussian process models with evolutionary algorithms," in Proceedings of the 10th International Conference on Adaptive and Natural Computing Algorithms - Volume Part I, ser. ICANNGA'11. Berlin, Heidelberg: Springer-Verlag, pp. 420-429, 2011.
- (비특허문헌 0002) R. Grande, G. Chowdhary, and J. How, "Nonparametric adaptive control using gaussian processes with online hyperparameter estimation" in Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, pp. 861-867, Dec 2013.

발명의 내용

해결하려는 과제

- [0010] 본 발명은 베이시안 학습 및 가우시안 프로세스 회귀에 기초하여 대규모 시스템의 각 코어들의 사용률을 제어하는 CPU 사용률 제어 장치 및 방법을 제공한다.
- [0011] 본 발명은 소스 코어의 타겟 프로세싱을 추출하여 소스 코어를 비활성화되도록 제어하고, 추출된 타겟 프로세싱을 목적지 코어로 마이그레이팅되도록 제어하여 대규모 시스템의 에너지 소비를 줄이는 CPU 사용률 제어 장치 및 방법을 제공한다.

과제의 해결 수단

- [0012] 본 발명의 실시예에 따른 CPU 사용률 제어 장치는 베이시안 학습(bayesian learning) 및 가우시안 프로세스 회귀(gaussian process regression)에 기초하여 각 코어들의 사용률을 획득하는 제1 획득부, 상기 획득된 사용률에 대응하는 상기 각 코어들의 평균 대기시간 및 차단율을 획득하는 제2 획득부, 상기 획득된 평균 대기시간 및 상기 획득된 차단율에 기초하여 소스 코어(source core)의 타겟 프로세싱(target processing) 및 목적지 코어(destination core)를 선정하는 선정부 및 상기 선정된 소스 코어의 타겟 프로세싱을 추출하여 상기 선정된 소스 코어를 비활성화되도록 제어하고, 상기 추출된 타겟 프로세싱을 상기 선정된 목적지 코어로 마이그레이팅(migrating)되도록 제어하는 제어부를 포함한다.
- [0013] 상기 제1 획득부는 상기 각 코어들 별 프로세싱에 대한 입력 데이터의 평균 및 공분산을 획득하고, 상기 획득된

입력 데이터의 평균 및 상기 획득된 공분산에 기초하여 사후 분포(posterior distribution)를 획득하며, 상기 획득된 사후 분포에 기초하여 상기 각 코어들 별 사용률을 획득할 수 있다.

[0014] 상기 제2 획득부는 상기 각 코어들의 프로세싱에 대한 도착률과 서비스율에 대응하는 상기 평균 대기시간을 획득하고, 상기 획득된 평균 대기시간 및 상기 서비스율에 대응하는 상기 차단율을 획득할 수 있다.

[0015] 상기 선정부는 상기 획득된 평균 대기시간의 선정 범위에 만족하는 상기 소스 코어의 타겟 프로세싱 및 상기 획득된 차단율의 선정 범위에 만족하는 상기 목적지 코어를 선정할 수 있다.

[0016] 본 발명의 일측에 따르면, 상기 선정부는 상기 획득된 평균 대기시간의 선정 범위에 만족하지 않은 경우, 상기 제1 획득부에게 상기 각 코어들의 사용률에 대한 갱신을 요청할 수 있다.

[0017] 본 발명의 다른 일측에 따르면, 상기 선정부는 상기 선정된 목적지 코어가 상기 소스 코어와 동일하거나, 상기 획득된 차단율의 선정 범위에 만족하지 않은 경우, 상기 목적지 코어에 대한 선정을 취소하고, 상기 제1 획득부에게 상기 각 코어들의 사용률에 대한 상기 갱신을 요청할 수 있다.

[0018] 본 발명의 실시예에 따른 CPU 사용률 제어 방법은 베이지안 학습 및 가우시안 프로세스 회귀에 기초하여 각 코어들의 사용률을 획득하는 단계, 상기 획득된 사용률에 대응하는 상기 각 코어들의 평균 대기시간 및 차단율을 획득하는 단계, 상기 획득된 평균 대기시간 및 상기 획득된 차단율에 기초하여 소스 코어의 타겟 프로세싱 및 목적지 코어를 선정하는 단계 및 상기 선정된 소스 코어의 타겟 프로세싱을 추출하여 상기 선정된 소스 코어를 비활성화되도록 제어하고, 상기 추출된 타겟 프로세싱을 상기 선정된 목적지 코어로 마이그레이팅되도록 제어하는 단계를 포함한다.

[0019] 상기 사용률을 획득하는 단계는 상기 각 코어들 별 프로세싱에 대한 입력 데이터의 평균 및 공분산을 획득하고, 상기 획득된 입력 데이터의 평균 및 상기 획득된 공분산에 기초하여 사후 분포를 획득하며, 상기 획득된 사후 분포에 기초하여 상기 각 코어들 별 사용률을 획득할 수 있다.

[0020] 상기 평균 대기시간 및 상기 차단율을 획득하는 단계는 상기 각 코어들의 프로세싱에 대한 도착률과 서비스율에 대응하는 상기 평균 대기시간을 획득하고, 상기 획득된 평균 대기시간 및 상기 서비스율에 대응하는 상기 차단율을 획득할 수 있다.

발명의 효과

[0021] 본 발명은 베이지안 학습 및 가우시안 프로세스 회귀에 기초하여 대규모 시스템의 각 코어들의 사용률을 제어할 수 있다.

[0022] 본 발명은 소스 코어의 타겟 프로세싱을 추출하여 소스 코어를 비활성화되도록 제어하고, 추출된 타겟 프로세싱을 목적지 코어로 마이그레이팅되도록 제어하여 대규모 시스템의 에너지 소비를 줄일 수 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 실시예에 따른 CPU 사용률 제어 장치를 도시한 블록도이다.

도 2는 도 1에서 복수 개의 코어에 대한 평균 대기시간과 서비스율의 흐름을 도시한 예이다.

도 3은 본 발명의 실시예에 따른 CPU 사용률 제어 방법을 도시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0024] 이하 첨부 도면들 및 첨부 도면들에 기재된 내용들을 참조하여 본 발명의 실시예를 상세하게 설명하지만, 본 발명이 실시예에 의해 제한되거나 한정되는 것은 아니다.

[0025] 한편, 본 발명을 설명함에 있어서, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는, 그 상세한 설명을 생략할 것이다. 그리고, 본 명세서에서 사용되는 용어(terminology)들은 본 발명의 실시예를 적절히 표현하기 위해 사용된 용어들로서, 이는 사용자, 운용자의 의도 또는 본 발명이 속하는 분야의 관례 등에 따라 달라질 수 있다. 따라서, 본 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0026] 도 1은 본 발명의 실시예에 따른 CPU 사용률 제어 장치를 도시한 블록도이다.

[0027] 도 1을 참조하면, CPU 사용률 제어 장치(100)는 제1 획득부(110), 제2 획득부(120), 선정부(130) 및 제어부(140)를 포함한다.

[0028] 제1 획득부(110)는 베이지안 학습(bayesian learning) 및 가우시안 프로세스 회귀(gaussian process regression)에 기초하여 각 코어들의 사용률을 획득한다.

[0029] 보다 상세하게는, 제1 획득부(110)는 각 코어들 별 프로세싱에 대한 입력 데이터의 평균 및 공분산을 획득하고, 획득된 입력 데이터의 평균 및 획득된 공분산에 기초하여 사후 분포(posterior distribution)를 획득하며, 획득된 사후 분포에 기초하여 각 코어들 별 사용률을 획득할 수 있다.

[0030] 실시예에 따르면, 제1 획득부(110)는 하기 수식 (1) 내지 (3)을 통하여 랜덤 변수의 유한 세트 ($Y=[Y_1, Y_2, Y_3, \dots, Y_n]$)에 대응하는 각 코어들 별 프로세싱에 대한 입력 데이터 ($x=[x_1, x_2, x_3, \dots, x_n]$)의 평균 및 공분산을 획득할 수 있다.

[0031] [수식 1]

$$f(y|x) \sim \mathcal{GP}(\mu(x), k(x, x'))$$

[0032]

[0033] [수식 2]

$$\mu(x) = \mathbb{E}(f(x))$$

[0034]

[0035] [수식 3]

$$k(x, x') = \mathbb{E}((f(x) - m(x))(f(x') - m(x')))$$

[0036]

[0037] 여기서, $\mathcal{GP}$ 는 가우시안 프로세스이고, $\mu(x)$ 는 입력 데이터인 시간 위치 x 에서 계산되는 평균함수이며, $k(x, x')$ 는 입력 데이터에 대한 공분산 또는 커널 함수일 수 있다. x' 는 x 의 다음 시간 위치를 의미한다.

[0038] 실시예에 따르면, 제1 획득부(110)는 수식 (4)를 통하여 가우시안 프로세스의 사후 분포를 획득할 수 있다.

[0039] [수식 4]

$$p\left(\begin{bmatrix} y \\ y_* \end{bmatrix}\right) = \mathcal{GP}\left(\begin{bmatrix} \mu(x) \\ \mu(x_*) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(x, x') & \mathbf{K}(x, x_*) \\ \mathbf{K}(x_*, x) & \mathbf{K}(x_*, x_*) \end{bmatrix}\right)$$

[0040]

[0041] 여기서, $p\left(\begin{bmatrix} y \\ y_* \end{bmatrix}\right)$ 는 가우시안 프로세스의 사후 분포이고, y 는 출력 데이터이며, y_* 는 추정된 출력 데이터이다.

[0042] 또한, $\mathbf{K}(x_*, x_*)$ 는 $k(x_*, x_*)$ 이고, $\mathbf{K}(x, x_*)$ 는 $k(x_1, x_*), k(x_2, x_*), \dots, k(x_n, x_*)$ 이며, $\mathbf{K}(x_*, x)$ 는 $\mathbf{K}(x, x_*)^T$ 이다.

[0043] 실시예에 따르면, 제1 획득부(110)는 수식 (5) 내지 수식 (9)를 통하여 추정된 출력 데이터(y_*)를 획득할 수 있다.

[0044] [수식 5]

$$m_{\star} = \mu(x_{\star}) + \mathbf{K}(x_{\star}, x) \mathbf{K}(x, x')^{-1} (y - \mu(x))$$

[0045]

[0046] [수식 6]

$$C_{\star} = \mathbf{K}(x_{\star}, x_{\star}) - \mathbf{K}(x_{\star}, x) \mathbf{K}(x, x')^{-1} \mathbf{K}(x, x_{\star})$$

[0047]

[0048] [수식 7]

$$p(y_{\star}) \sim \mathcal{GP}(m_{\star}, C_{\star})$$

[0049]

[0050] [수식 8]

$$\bar{y}_{\star} = \mathbf{K}(x_{\star}, x) \mathbf{K}(x, x')^{-1} y$$

[0051]

[0052] [수식 9]

$$\text{var}(y_{\star}) = \mathbf{K}(x_{\star}, x_{\star}) - \mathbf{K}(x_{\star}, x) \mathbf{K}(x, x')^{-1} \mathbf{K}(x, x_{\star})$$

[0053]

[0054] 여기서, $m_{\star}$ 는 사후 분포에 대한 평균이고, $C_{\star}$ 는 사후 분포에 대한 공분산이며, $\bar{y}_{\star}$ 는 $y_{\star}$ 의 평균이고, $\text{var}(y_{\star})$ 는 $y_{\star}$ 의 분산이다.

[0055] 제2 획득부(120)는 획득된 사용률에 대응하는 상기 각 코어들의 평균 대기시간 및 차단율을 획득한다.

[0056] 보다 상세하게는 제2 획득부(120)는 각 코어들의 프로세싱에 대한 도착률과 서비스율에 대응하는 평균 대기시간을 획득하고, 획득된 평균 대기시간 및 서비스율에 대응하는 차단율을 획득할 수 있다.

[0057] 또한, 제2 획득부(120)는 PASTA(Poisson Arrival See Time Averages) 이론에 기초하여 각 코어들의 도착율에 대한 도착 평균시간 또는 평균 대기시간을 획득할 수 있고, 폴라첵-힌친(Pollaczek-Khintchine) 이론에 기초하여 평균 대기시간을 획득할 수 있다.

[0058] 여기서, 도착률은 전술한 수식 (1) 내지 (10)에 의해 연산되는 $y_{\star}$ 일 수 있다.

[0059] 실시예에 따르면, 제2 획득부(120)는 수식 (11)을 통하여 각 코어들의 평균 대기시간을 획득할 수 있고, 수식 (12)를 통하여 차단율을 획득할 수 있다.

[0060] [수식 11]

$$\mathbf{E}_j(W) = \frac{\lambda_{\tau}^{1/\mu^2}}{2(1 - \lambda_{\tau}^{1/\mu})}$$

[0061]

[0062] [수식 12]

$$\mathbf{BR}_j = \frac{\mathbf{E}_j(W)}{\mu}$$

[0063]

[0064] 여기서, $\mathbf{E}_j(W)$ 는 j번째 코어에 대한 평균 대기시간이고, λ_{τ} 는 도착 프로세스에서의 시간(τ)에 대한 도착률(arrival rate)이며, μ 는 코어에 대한 서비스율(service rate)이며, $\mathbf{BR}_j$ 는 j번째 코어에 대한 차단율(blocking rate)이다.

- [0065] 선정부(130)는 획득된 평균 대기시간 및 획득된 차단율에 기초하여 소스 코어(source core)의 타겟 프로세싱(target processing) 및 목적지 코어(destination core)를 선정한다.
- [0066] 보다 상세하게는, 선정부(130)는 획득된 평균 대기시간의 선정 범위에 만족하는 소스 코어의 타겟 프로세싱 및 획득된 차단율의 선정 범위에 만족하는 목적지 코어를 선정할 수 있다.
- [0067] 예를 들어, 선정부(130)는 평균 대기시간이 가장 낮은 코어가 소스 코어로 선정될 수 있고, 차단율을 만족하는 코어(예를 들어, 프로세스 허용이 가능한 코어 또는 프로세스 차단이 되지 않은 코어 등)가 목적지 코어로 선정될 수 있다.
- [0068] 본 발명의 일측에 따르면, 선정부(130)는 획득된 평균 대기시간의 선정 범위에 만족하지 않은 경우, 제1 획득부(110)에게 각 코어들의 사용률에 대한 갱신을 요청할 수 있다.
- [0069] 본 발명의 다른 일측에 따르면, 선정부(130)는 선정된 목적지 코어가 소스 코어와 동일하거나, 획득된 차단율의 선정 범위에 만족하지 않은 경우, 목적지 코어에 대한 선정을 취소하고, 제1 획득부(110)에게 각 코어들의 사용률에 대한 갱신을 요청할 수 있다.
- [0070] 제어부(140)는 선정된 소스 코어의 타겟 프로세싱을 추출하여 선정된 소스 코어를 비활성화되도록 제어하고, 추출된 타겟 프로세싱을 선정된 목적지 코어로 마이그레이팅(migrating)되도록 제어한다.
- [0071] 예를 들어, 제어부(140)는 선정된 소스 코어를 스탠드-바이 모드(stand-by mode) 또는 아이들 모드(idle mode) 상태로 비활성화되도록 제어할 수 있다.
- [0072] 여기서, 제어부(140)는 대규모 시스템에 포함되는 운영체제 또는 큐잉 이론(queueing theory) 등 다양한 시스템 환경 조건을 고려하여 선정된 소스 코어의 타겟 프로세싱을 추출하여 선정된 소스 코어를 비활성화되도록 제어하고, 추출된 타겟 프로세싱을 선정된 목적지 코어로 마이그레이팅되도록 제어할 수 있다.
- [0073] 도 2는 도 1에서 복수 개의 코어에 대한 평균 대기시간과 서비스율의 흐름을 도시한 예이다.
- [0074] 도 2를 참조하면, 제어부(140)는 복수 개의 코어 중에서 소스 코어의 타겟 프로세싱을 추출하여 선정된 소스 코어를 비활성화되도록 제어하고, 추출된 타겟 프로세싱을 선정된 목적지 코어로 마이그레이팅(migrating)되도록 제어할 수 있다.
- [0075] 따라서, CPU 사용률 제어 장치는 대규모 시스템 상에 존재하는 코어의 에너지 소비를 줄일 수 있고, 효율적으로 코어의 사용률을 제어할 수 있다.
- [0076] 도 3은 본 발명의 실시예에 따른 CPU 사용률 제어 방법을 도시한 흐름도이다.
- [0077] 도 3을 참조하면, CPU 사용률 제어 장치는 단계 310에서, 베이지안 학습 및 가우시안 프로세스 회귀에 기초하여 각 코어들의 사용률을 획득한다.
- [0078] 보다 상세하게는, CPU 사용률 제어 장치는 단계 310에서, 각 코어들 별 프로세싱에 대한 입력 데이터의 평균 및 공분산을 획득하고, 획득된 입력 데이터의 평균 및 획득된 공분산에 기초하여 사후 분포를 획득하며, 획득된 사후 분포에 기초하여 각 코어들 별 사용률을 획득할 수 있다.
- [0079] CPU 사용률 제어 장치는 단계 320에서, 획득된 사용률에 대응하는 상기 각 코어들의 평균 대기시간 및 차단율을 획득한다.
- [0080] 보다 상세하게는 CPU 사용률 제어 장치는 단계 320에서, 각 코어들의 프로세싱에 대한 도착률과 서비스율에 대응하는 평균 대기시간을 획득하고, 획득된 평균 대기시간 및 서비스율에 대응하는 차단율을 획득할 수 있다.
- [0081] CPU 사용률 제어 장치는 단계 330에서, 획득된 평균 대기시간 및 획득된 차단율에 기초하여 소스 코어의 타겟 프로세싱 및 목적지 코어를 선정한다.
- [0082] 보다 상세하게는, CPU 사용률 제어 장치는 단계 330에서, 획득된 평균 대기시간의 선정 범위에 만족하는 소스 코어의 타겟 프로세싱 및 획득된 차단율의 선정 범위에 만족하는 목적지 코어를 선정할 수 있다.
- [0083] 예를 들어, CPU 사용률 제어 장치는 단계 330에서, 평균 대기시간이 가장 낮은 코어가 소스 코어로 선정될 수

있고, 차단율을 만족하는 코어(예를 들어, 프로세스 허용이 가능한 코어 또는 프로세스 차단이 되지 않은 코어 등)가 목적지 코어로 선정될 수 있다.

[0084] 본 발명의 일측에 따르면, CPU 사용률 제어 장치는 단계 330에서, 획득된 평균 대기시간의 선정 범위에 만족하지 않은 경우, 각 코어들의 사용률에 대한 갱신을 수행할 수 있다.

[0085] 본 발명의 다른 일측에 따르면, CPU 사용률 제어 장치는 단계 330에서, 선정된 목적지 코어가 소스 코어와 동일하거나, 획득된 차단율의 선정 범위에 만족하지 않은 경우, 목적지 코어에 대한 선정을 취소하고, 각 코어들의 사용률에 대한 갱신을 수행할 수 있다.

[0086] CPU 사용률 제어 장치는 단계 340에서, 선정된 소스 코어의 타겟 프로세싱을 추출하여 선정된 소스 코어를 비활성화되도록 제어하고, 추출된 타겟 프로세싱을 선정된 목적지 코어로 마이그레이팅되도록 제어한다.

[0087] 따라서, CPU 사용률 제어 장치는 대규모 시스템 상에 존재하는 코어의 에너지 소비를 줄일 수 있고, 효율적으로 코어의 사용률을 제어할 수 있다.

[0088] 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. 상기 매체에 기록되는 프로그램 명령은 실시예를 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. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0089]

[0090] 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.

[0091]

[0092] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

부호의 설명

[0093] 100: CPU 사용률 제어 장치

110: 제1 획득부

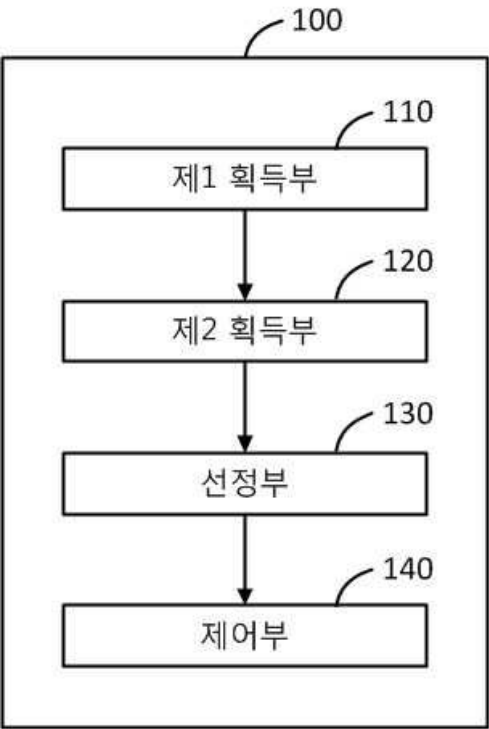
120: 제2 획득부

130: 선정부

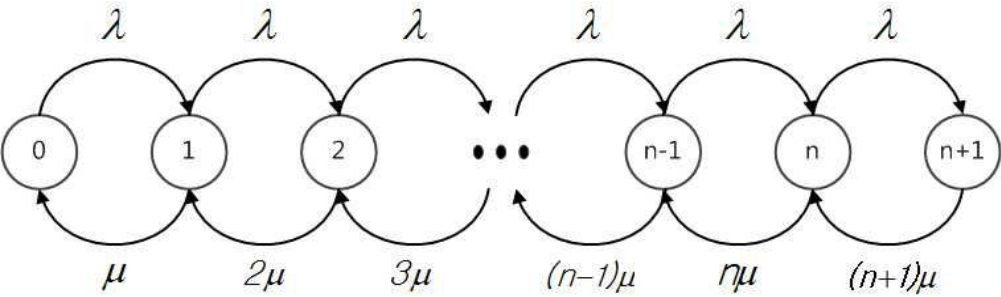
140: 제어부

도면

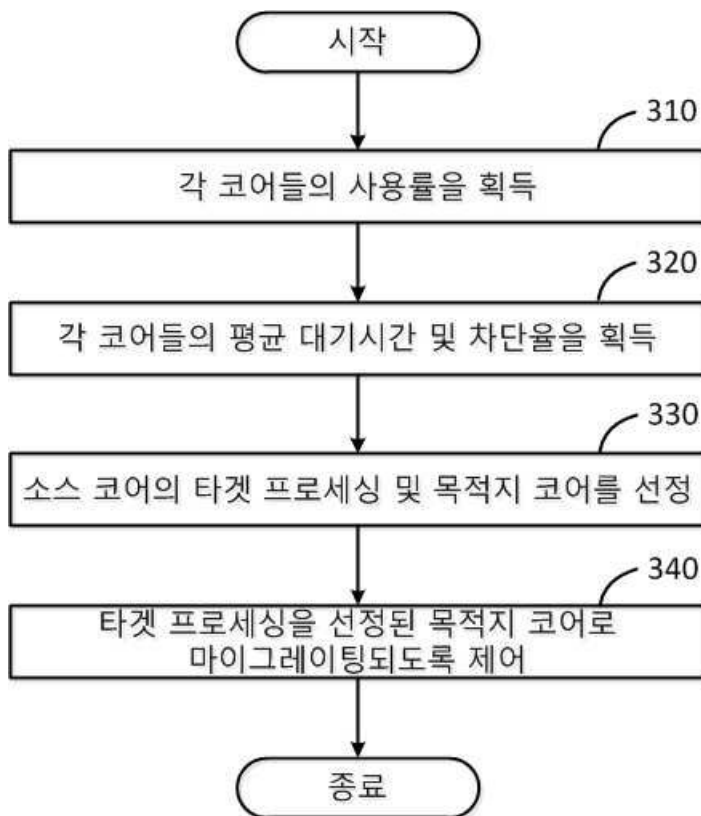
도면1



도면2



도면3



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1, 7

【변경전】

PASTA(Possion Arrival See Time Averages)

【변경후】

PASTA(Poisson Arrival See Time Averages)