

OANet: 데이터베이스 성능 예측을 위한 Ortho Attention Net *

염찬호[○] 이지은 서주연 JIN HUIJUN 박상현[†]

연세대학교 컴퓨터과학과

{chanho0475, jieun199624, sjy9728s, jinhuijun, sanghyun}@yonsei.ac.kr

OANet: Orthogonal Attention Network for performance prediction of Database

Chanho Yeom[○] Jieun Lee Juyeon Seo JIN HUIJUN Sanghyun Park[†]

Dept. of Computer Science, Yonsei University

요 약

데이터베이스에는 데이터 처리량의 성능에 영향을 주는 많은 Knob들이 있고 데이터 처리량의 성능을 향상시키기 위해 Knob 튜닝을 할 때 그에 따른 성능을 신속하고 신뢰성 있게 예측하는 모델이 필요하다. 하지만 워크로드의 환경이 다를 경우 같은 Knob을 이용하여 벤치마크를 수행해도 결과가 다를 수 있다. 따라서 본 논문에서는 Knob과 워크로드 환경의 연관성도 고려할 수 있도록 주의 관심 메커니즘을 활용한 OANet을 제안한다. 실험을 통해 데이터 처리량 예측에 대한 R2 score를 다양한 기계학습 기법과 비교한 결과 가장 높은 점수를 보임으로써, 모델의 우수성을 보였다.

1. 서 론

스마트 시티화가 진행됨에 따라 도시 곳곳에서는 다양한 장치와 센서를 통해 여러 데이터를 수집한다. 수집된 데이터는 도시 운영 및 시민에게 제공되는 서비스의 효율성을 향상시키기 위해 사용된다[1, 2]. 이러한 빅데이터 환경에서는 방대한 양의 비정형 데이터가 생성되고 이를 효율적으로 처리할 수 있는 데이터베이스가 사용되고 있다[3, 4, 5].

데이터베이스는 수많은 종류의 Knob을 포함하고 있고, 각 Knob의 값의 변화에 따라 해당 데이터베이스의 성능에 영향을 준다[6]. 이에 많은 Knob을 튜닝하기 위한 여러 선행연구가 진행되어 오고 있다. Knob 튜닝의 주요한 프레임워크는 최적화 알고리즘을 통해 Knob을 선정하고 알고리즘 내에서는 성능을 예측하는 모델을 사용한다[6, 7, 8]. 모델로는 다양한 모델이 적용될 수 있는데 대표적으로 랜덤 포레스트 등의 기계 학습 모델을 적용한다. 모델의 정확도가 높을수록 전체적인 튜닝 결과에 영향을 주기 때문에, 본 논문에서는 높은 정확도를 도출하는 모델 개발을 목표로 한다.

Knob을 선정하더라도 이는 워크로드 환경에 따라 변화하게 되고, 이는 Knob과 워크로드는 상호의존적 관계라는 것

을 보여준다. 이에 본 논문은 상호의존적 특성을 반영하기 위해, 주의 관심 메커니즘[9]을 적용하고, 벡터 데이터인 Knob을 매트릭스 데이터로 표현할 수 있는 방법론을 제안한다. 결과적으로 워크로드 환경과 Knob과의 연관성을 학습하여 워크로드에 따른 성능 예측 모델을 제안한다.

본 논문은 모델의 우수성을 보여주기 위해, 실험을 통해 정량적, 정성적 평가를 진행한다. 실험을 위해 RocksDB 데이터베이스와 db_bench를 사용한 데이터셋을 사용한다. 또한, 다수의 기계학습 회귀 모델과의 R2 score 비교를 진행한 결과, 제안하는 모델이 가장 좋은 성능을 보였다. 그리고 주의 관심 메커니즘 가중치에 따른 그래프를 통해 Knob과 워크로드의 상호의존성을 정성적으로 보였다.

2. Ortho Attention Net

본 논문에서는 Knob들과 워크로드 환경 간의 연관성을 파악하여 결과 예측에 참고할 수 있도록 주의 관심

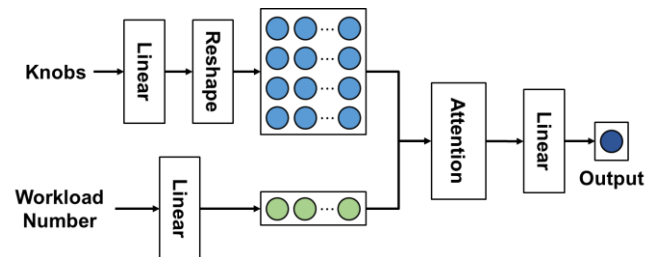


그림 1. 모델 흐름도

* 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(IITP-2017-0-00477, (SW 스타랩) IoT 환경을 위한 고성능 플래시 메모리 스토리지 기반 인메모리 분산 DBMS 연구개발)과 국토교통부의 스마트시티 혁신인재육성사업으로 지원을 받아 수행된 연구임.

[†] 교신 저자: sanghyun@yonsei.ac.kr

메커니즘[9]과 Soft Orthogonality Regularization[10]을 적용한 OANet(Ortho Attention Net)을 제안한다.

2.1 모델 구조

OANet의 구조는 그림 1과 같다. Knob과 workload 정보를 통해서 데이터 처리량(Rate)을 예측하는 것을 목표로 하고, 다음과 같이 변수를 정의한다. x 는 Knob, \hat{y} 는 모델을 통해 예측한 데이터 처리량 그리고 wk 는 워크로드 환경 정보를 가리킨다.

$$X = \{x_1, x_2, \dots, x_n\}$$

$$\hat{y} = OANet(X, wk_x)$$

본 논문에서 제안하는 모델은 벡터 데이터인 Knob에 대해 주의 관심 모델을 적용하기 위해서 매트릭스 데이터로 변환하는 과정을 포함한다.

$$h_x = \sigma(XW_L + b_L) \quad (1)$$

$$\tilde{h}_x = Reshape(h_x) \quad (2)$$

$$W_L \in \mathbb{R}^{n \times g \times k}, h_x \in \mathbb{R}^{1 \times g \times k}, \tilde{h}_x \in \mathbb{R}^{g \times k} \quad (3)$$

수식 (1)을 통해 선형 변환으로 차원을 늘리고, 수식 (2)와 같이 h_x 를 2차원의 \tilde{h}_x 로 Reshape한다. 이를 통해, Knob의 정보를 k 차원을 갖는 g 개의 벡터로 이루어진 매트릭스로 표현한다. 각 g 개의 벡터는 서로 독립적인 Knob의 특성을 포함해야 한다. σ 는 sigmoid 활성화 함수이며, 변수의 차원은 수식 (3)과 같다.

워크로드 환경의 정보는 one-hot 벡터를 통해 표현한다. 예를 들어 첫 번째 워크로드는 $wk_1 = \{1, 0, 0, 0\}$ 로 표현하고 one-hot 벡터의 차원수는 사용한 워크로드 환경 개수이다. wk 또한, 선형 변환을 통해 \tilde{h}_x 와 차원을 맞춰준다, $h_{wk} \in \mathbb{R}^{1 \times k}$.

앞서 구한 knob의 정보와 workload의 정보의 연관성을 학습하기 위해 주의 관심 메커니즘을 적용한다.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

$$\alpha = \sigma(Attention(h_{wk}, \tilde{h}_x, \tilde{h}_x)) \quad (5)$$

$$\hat{y} = \alpha W_y + b_y \quad (6)$$

수식 (5)를 통해 워크로드 환경 정보에 대한 knob의 특성을 학습하고 수식 (6)을 통해 데이터 처리량을 예측한다.

2.2 목적 함수

제안한 모델에서 \tilde{h}_x 의 각 벡터들의 독립성을 보장하기 위해 목적 함수에 SOR 정규항(Soft Orthogonality Regularization)을 추가한다. SOR 정규항은 벡터간의 직교성을 이용하고 수식 (8)와 같이 계산된다.

$$SOR(M) = \lambda \|M \cdot M^T - I\|^2 \quad (7)$$

$$Loss = MSE(y, \hat{y}) + SOR(\tilde{h}_x) \quad (8)$$

3. 실험 및 결과

3.1 실험 환경

본 논문에서는 (RW 9:1), (RW 1:1), (RW 1:9), (UPDATE) 4개의 workload에서 RocksDB의 Knob들 중 22개의 Knob에 대해 db_bench로 성능측정을 한 8만여개의 데이터 셋으로 실험을 진행하였다.

표 1. 모델 간의 실험 성능 비교

Method	Model	RATE
Linear Regression	Lasso	0.3929
	Ridge	0.5399
	Elastic-Net	0.5822
	SGD	0.7022
Non-linear Regression	Random Forest	0.7550
	Adaboost	0.7442
	SVR	0.9467
Neural Net	Single Neural Net	0.9516
	OANet	0.9623

표 2. OANet의 하이퍼파라미터 성능 비교

k	g	RATE
16	4	0.9618
	8	0.9617
	16	0.9609
64	4	0.9568
	8	0.9623
	16	0.9555
128	4	0.9532
	8	0.8991
	16	0.9229

3.2 실험 결과

본 논문에서는 기존의 기계학습에서 대표적으로 사용되는 Lasso, Ridge, Elastic-Net, SGD, Random Forest, Adaboost, SVR과 제안한 모델의 성능을 비교하였다. 표 1은 OANet과 기존 방법론의 R2 score 결과를 보여준다. OANet이 실제 값과 가장 비슷하게 예측하였다. 또한, 비선형 기반의 기계 학습 모델이 선형 기반의 모델들보다 성능이 크게 향상된 것을 알 수 있고 OANet을 포함한 인공 신경망 기반의 모델이 비선형 기반의 모델보다 좋은 성능을 보인다. SVR 모델의 경우 인공 신경망 모델과 가장 근접한 점수를 보이지만, 예측하기 위해 소요되는 시간이 OANet보다 15배 느리게 나타나므로 향후 Knob 튜닝 연구에 적용하는데 한계가 있다.

표 2는 제안한 모델에 대해 2.1에서 서술한 변수 k 와 g 의 값을 조정 및 성능을 확인한 결과이다. 해당 결과를 통해 가장 좋은 성능을 보인 k 는 64, g 는 8로 선정하였다.

그림 2는 OANet이 주의 관심 메커니즘 적용 시에 워크로드 환경과 g 개의 Knob 그룹 정보에 대한 가중치 그래프이다. 워크로드마다 관련된 영향을 미치는 Knob의 특성이 모두 다른 것을 볼 수 있다. 예를 들어, 그림 2의 (a)는 데이터 처리량을 예측하기 위해 3번째 특성에

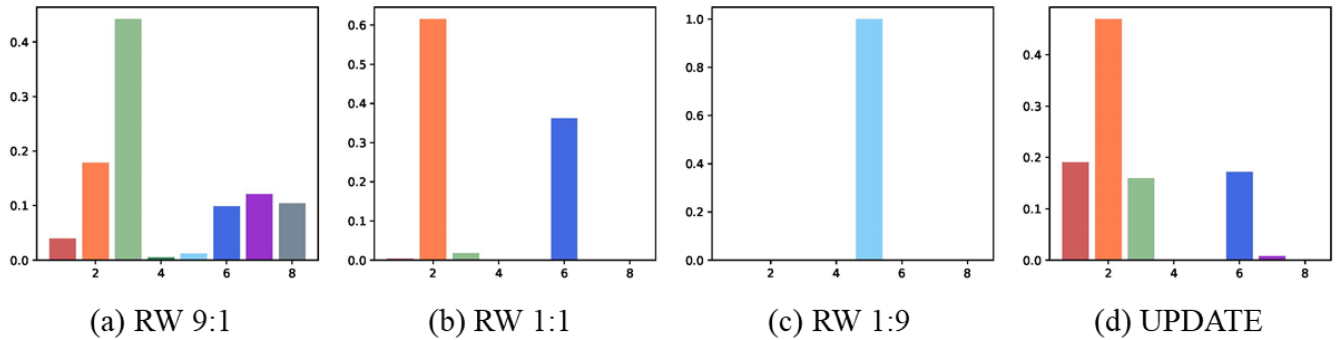


그림 2. 주의 관심 가중치 그래프

가중치를 많이 두는 반면에 (b)와 (d)는 2번째, (c)는 오롯이 5번째 특성에 가중치를 둔다. 이를 통해, 학습된 Knob 특성끼리 독립적인 특성을 지녔고, 워크로드 환경과 Knob 특성이 상호의존성 관계임을 보여준다.

4. 결 론

데이터베이스의 Knob 튜닝을 할 때 같은 Knob 수치에 대해서도 어떤 워크로드에서 수행되는지에 따라 결과가 상이하기 때문에 예측에 어려움이 있다는 문제점이 있다. 본 논문에서는 Knob과 워크로드 환경 간의 상호의존성을 파악할 수 있도록 주의 관심 기반의 인공 신경망을 제안하였고 성능을 확인하기 위해 RocksDB에서 생성한 데이터셋을 통해 실험하였다. 기존의 기계학습 예측 모델들과 비교하여 학습 시간과 성능이 향상된 것을 검증하였다.

향후에는 본 논문에서 제안한 예측 모델을 통해 Knob 튜닝을 해서 데이터베이스의 성능을 향상시키는 연구를 진행해볼 예정이다.

참 고 문 헌

- [1] Mama Nsangou Mouchili, et al., "Smart City Data Analysis", Proceeding of the First International Conference on Data Science, E-Learning and Information Systems, Article No.33, Pages 1-6, 2018
- [2] M. Dalla Cia et al., "Using Smart City Data in 5G Self-Organizing Networks", in IEEE Internet of Things Journal, vol. 5, no. 2, pp. 645-654, 2018
- [3] RocksDB, website. <http://rocksdb.org/>.
- [4] Redis. <https://redis.io>
- [5] MySQL. <https://github.com/mysql>
- [6] Dana Van Aken et al., "Automatic Database Management System Tuning Through Large-scale Machine Learning", SIGMOD '17 : Proceedings of the 2017 ACM International Conference on Management of Data, Pages 1009-1024, 2017
- [7] 김휘균, 최원기, 최종환, 성한승, 박상현, "데이터베이

스 성능 향상을 위한 기계학습 기반의 RocksDB 파라미터 분석 연구", 2020 온라인 추계학술박표대회 논문집, 제27권, 제2호

- [8] 서주연, 이지은, 김경훈, JIN HUIJUN, 박상현, "비선형 기계학습 기반의 Redis 파라미터 튜닝 연구", 한국정보과학회 2021컴퓨터종합학술대회 논문집, p.69-71
- [9] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. "Attention is all you need", In Advances in neural information processing systems, pp. 5998-6008, 2017
- [10] Bansal, Nitin, Xiaohan Chen, and Zhangyang Wang., "Can we gain more from orthogonality regularizations in training deep CNNs?", arXiv preprint arXiv:1810.09102, 2018