

동적 어텐션을 이용한 지식그래프 임베딩

Knowledge Graph Embedding with Dynamic Attention

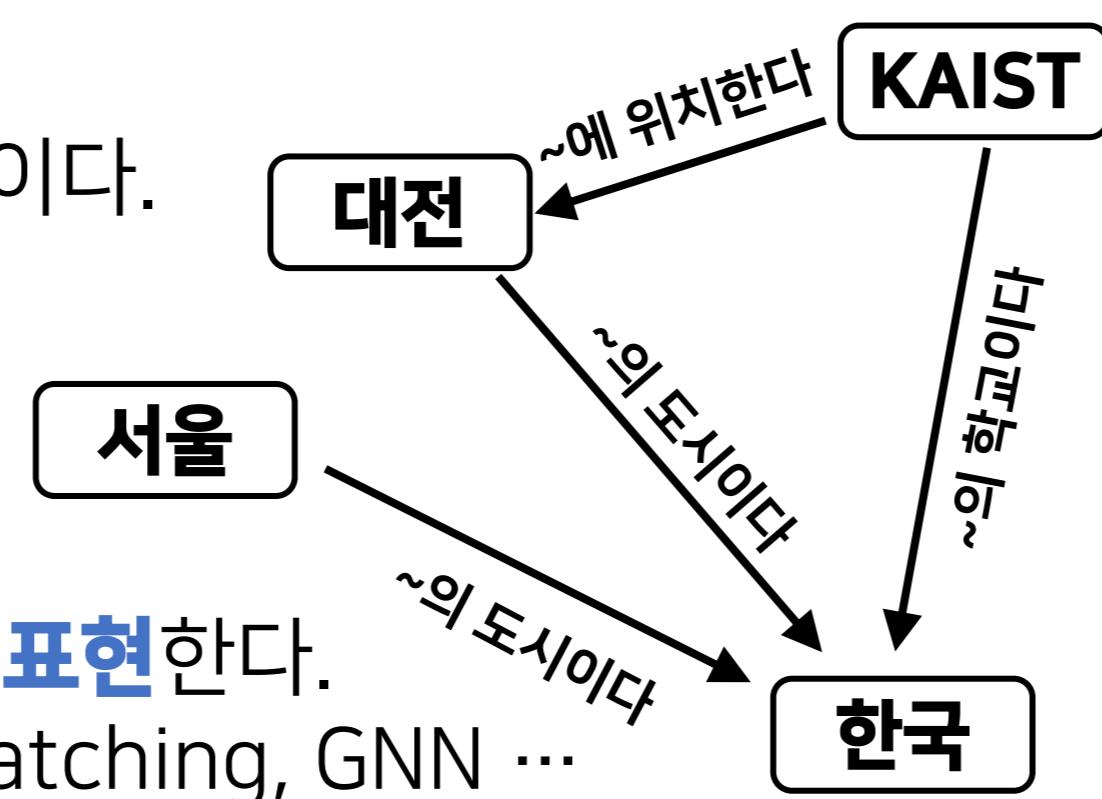
황민성 황지영
KAIST 전기및전자공학부 KAIST 전산학부
{hminsung, jjwhang}@kaist.ac.kr



서론

• 지식그래프

- 인간의 여러 **지식**을 **그래프**로 표현한 것이다.
- 각 지식들은 **삼중항**으로 표현한다.
- 삼중항: (시작개체, 관계, 도착개체)



• 지식그래프 임베딩

- 개체, 관계들을 연속적인 공간의 **벡터로 표현**한다.
- Translational Distance, Semantic Matching, GNN ...

• 연구의 목적

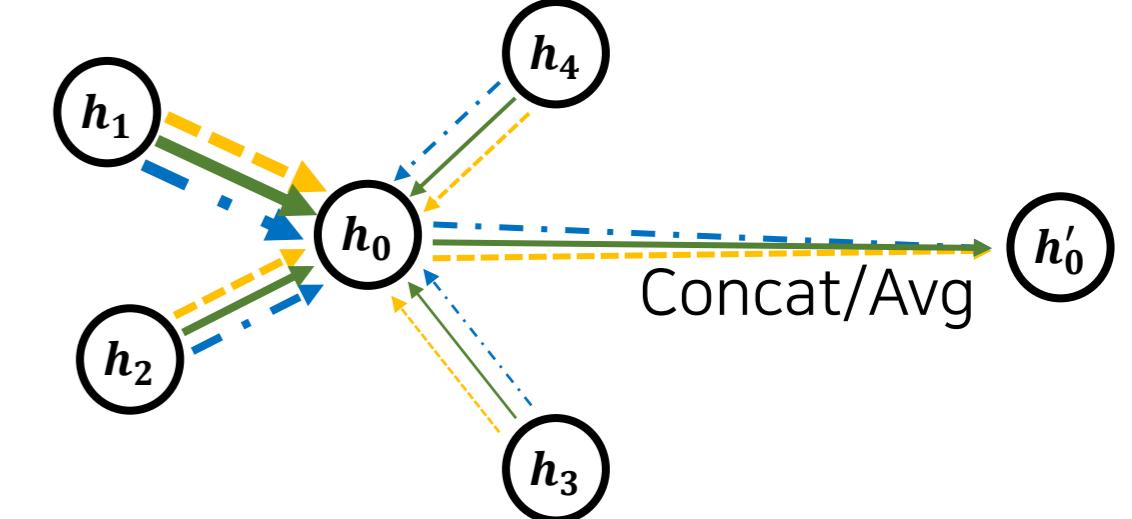
- 기존 어텐션 기반 지식그래프 임베딩 기법에서, **정적 어텐션 문제**가 발생하는지 확인한다.
- 동적 어텐션**을 적용하여 새로운 모델을 구현하고, 성능이 개선되는지 확인한다.

사전연구: 그래프 어텐션 네트워크(GAT)

• 그래프 어텐션 네트워크 (GAT)

- 어텐션(Attention)을 이용한 GNN 모델 중 하나이다.
- 이웃 노드의 임베딩을 취합할 때, **가중치**(=어텐션)를 부여한다.
- 여러 개의 어텐션 메커니즘을 동시에 사용하기도 한다.

$$e(\mathbf{h}_i, \mathbf{h}_j) = \sigma(\mathbf{a}^T \cdot \mathbf{W}[\mathbf{h}_i || \mathbf{h}_j])$$



• 정적 어텐션 문제(Static Attention Problem)

- 각 노드가 계산하는 이웃 노드 어텐션의 **순위가 일정**한 문제이다.
- 대부분의 노드가 특정 노드의 정보를 가장 많이 가져온다.
- 간단한 문제에서도 틀린 예측을 하는 경우가 있다.

사전연구: 동적 어텐션

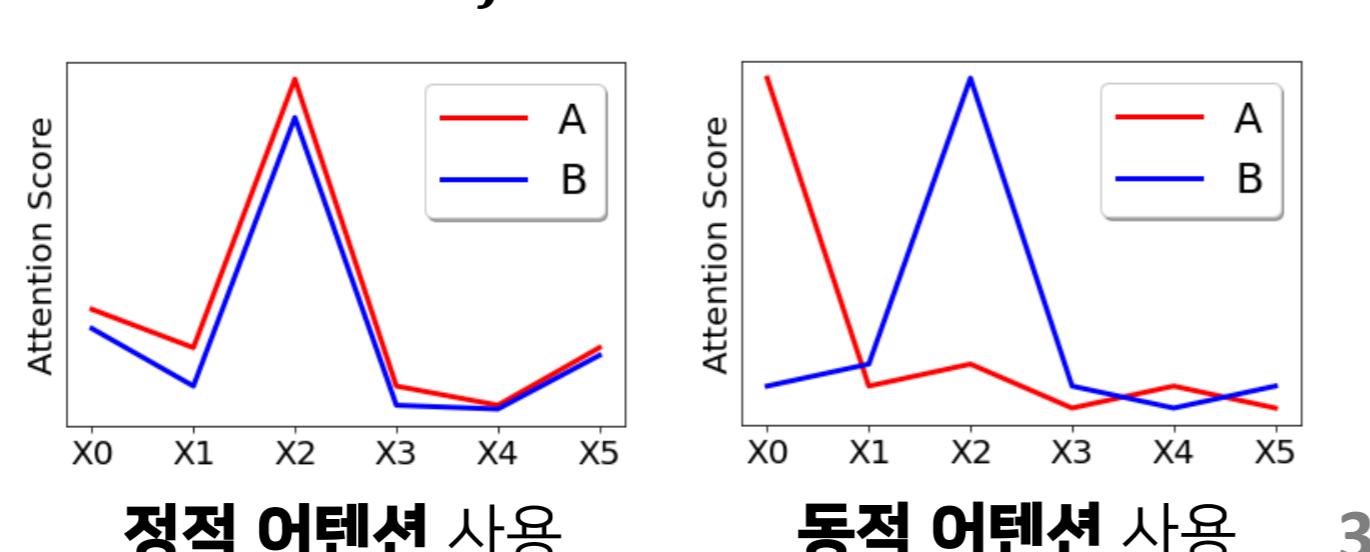
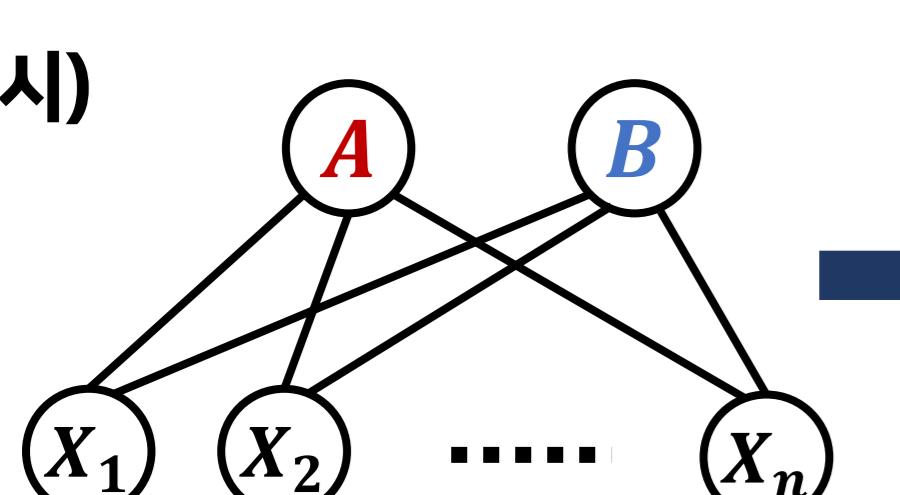
• 동적 어텐션(Dynamic Attention)

- 정적 어텐션 문제를 개선하기 위해 새롭게 제시된 어텐션이다.
- 각 노드가 계산하는 이웃 노드 어텐션의 **순위가 모두 다르게** 된다.
- 정적 어텐션은 두 개의 선형 레이어(\mathbf{a}^T, \mathbf{W})가 **연속**적으로 곱해지기 때문에 발생한다.
- 비선형 활성화 함수(σ)를 두 선형 레이어 사이로 이동하여 계산한다.

정적 어텐션(GAT) : $e(\mathbf{h}_i, \mathbf{h}_j) = \sigma(\mathbf{a}^T \cdot \mathbf{W}[\mathbf{h}_i || \mathbf{h}_j])$

동적 어텐션(GATv2) : $e(\mathbf{h}_i, \mathbf{h}_j) = \mathbf{a}^T \cdot \sigma(\mathbf{W}[\mathbf{h}_i || \mathbf{h}_j])$

예시)



1

사전연구: 어텐션 기반 지식그래프 임베딩 (KBAT)

• GNN 기반 지식그래프 임베딩

- GNN과 마찬가지로 연결된 개체 및 관계의 임베딩을 **취합**한다.
- 개체 및 관계의 임베딩을 취합하는 방식에 따라 여러 모델이 존재한다.

• 어텐션 기반 지식그래프 임베딩

- GNN 기반 지식그래프 임베딩 기법 중 하나이다.
- 개체 및 관계로 이루어진 삼중항에 서로 다른 **가중치**를 부여한다.

• KBAT

- 대표적인 어텐션 기반 지식그래프 임베딩 모델이다.
- 삼중항의 두 개체 및 관계의 임베딩을 **연결**하여 어텐션을 계산한다.
- 직접적으로 연결되지 않은 개체의 임베딩도 함께 계산한다.

$$c_{ijk} = \mathbf{W}[\mathbf{e}_h || \mathbf{e}_t || \mathbf{r}] \rightarrow b_{ijk} = \text{LeakyReLU}(\mathbf{a}^T c_{ijk})$$

2

4

동적 어텐션 기법의 지식그래프 임베딩 적용

• 어텐션 기반 지식그래프 임베딩 개선

- KBAT 모델의 어텐션은 **정적 어텐션**이다.
- 세 개의 임베딩을 연결 한 뒤, 두 개의 선형 레이어가 **연속**적으로 곱해지기 때문이다.
- 동적 어텐션을 위해 비선형 활성화 함수의 위치를 이동하여 계산한다.
- 동적 어텐션을 사용하는 새로운 **KBATv2** 모델을 구현하였다.

KBAT

- $c_{ijk} = \mathbf{W}[\mathbf{e}_h || \mathbf{e}_t || \mathbf{r}]$
- $b_{ijk} = \text{LeakyReLU}(\mathbf{a}^T c_{ijk})$

기존 모델의 formulation

KBATv2

- $c_{ijk} = \mathbf{W}[\mathbf{e}_h || \mathbf{e}_t || \mathbf{r}]$
- $b_{ijk} = \mathbf{a}^T \text{LeakyReLU}(c_{ijk})$

새롭게 제안한 모델의 formulation

5

실험 결과

• 실험 결과

	FB15K-237		
측정값	Hits@10 (↑)	MR (↓)	MRR (↑)
KBAT	0.6215	206.7	0.4640
KBATv2	0.6296	177.9	0.4585
Gain (↑)	1.29%	13.93%	-1.19%
총 Gain (↑)		14.03%	

• 결과 분석

- Hits@10, MR의 경우 KBATv2의 성능이 더 높게 나타났다.
- MRR의 경우 약간 감소하는 것으로 나타났다.
- 총 Gain이 양수로 측정되어, 모델 성능이 **증가**하였다.

실험 방법

• 실험 방법

- Link Prediction을 진행 후 결과를 비교하였다.
- 측정값: Hits@10, MR, MRR
- 기존 모델과 비교한 Gain을 계산하여 모델의 성능을 측정한다.
- $Gain(metric) = sign(metric) \frac{Score(KBATv2) - Score(KBAT)}{Score(KBAT)} \times 100\%$
- $sign(MR) = -1, sign(Hits@10) = sign(MRR) = +1$
- 모든 측정값은 **filtered** 방식으로 계산하였다.

• 데이터셋 통계

데이터셋	개체 (Entity)	관계 (Relation)	삼중항 (Triplet)		
			Train	Validation	Test
FB15K-237	14,541	237	272,115	17,535	20,466

6

결론 및 향후 연구

• 결론

- 기존 지식그래프 임베딩 모델에 **동적 어텐션 기법**을 적용하여 성능이 증가하는지 확인하였다.
- 대표적인 지식그래프 벤치마크 데이터셋 FB15K-237에 대해서 **동적 어텐션 기법**이 **효과적**임을 확인하였다.
- 실제로 **정적 어텐션 문제**가 발생하는지 확인하지는 못하였다.

• 향후 연구

- KBAT 모델에서 계산하는 어텐션의 순위를 직접 **시각화**하여, **정적 어텐션 문제**가 발생하는지 확인한다.
- 다른 지식그래프 벤치마크 데이터셋에서도 **동적 어텐션 기법**이 효과를 보일지 확인한다.

7

8