软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, 2023,34(10):4518–4532 [doi: 10.13328/j.cnki.jos.006885]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

FS-Net: 面向时序知识图谱推理的频次统计网络*

刘康正[1,2,3,4], 赵 峰[1,2,3,4], 金 海[1,2,3,4]

[1](大数据技术与系统国家地方联合工程研究中心, 湖北 武汉 430074)
[2](服务计算技术与系统教育部重点实验室, 湖北 武汉 430074)
[3](集群与网格计算湖北省重点实验室, 湖北 武汉 430074)
[4](华中科技大学 计算机科学与技术学院, 湖北 武汉 430074)
通信作者: 赵峰, E-mail: zhaof@hust.edu.cn

摘 要: 时序知识图谱推理吸引了研究人员的极大关注. 现有的时序知识图谱推理技术通过建模历史信息取得了巨大的进步. 但是, 时变性问题和不可见实体(关系)问题仍然是阻碍时序知识图谱推理模型性能进一步提升的两大挑战; 而且由于需要对历史子图序列的结构信息和时间依赖信息进行建模, 传统的基于嵌入的方法往往在训练和预测过程中具有较高的时间消耗, 这极大地限制了推理模型在现实场景中的应用. 针对以上困境, 提出了一个用于时序知识图谱推理的频次统计网络, FS-Net. 一方面, FS-Net 不断基于最新的短期历史的事实频次统计, 动态地为变化的时间戳上的预测生成时变的得分; 另一方面, FS-Net 基于当前时间戳上的事实频次统计, 为预测补充历史不可见实体(关系); 特别地, FS-Net 不需要进行训练, 而且具有极高的时间效率. 在两个时序知识图谱基准数据集上的大量实验, 表明了 FS-Net 相较于基准模型的巨大提升.

关键词: 时序知识图谱; 事实频次统计; 时变性; 不可见信息
中图法分类号: TP18

中文引用格式: 刘康正, 赵峰, 金海. FS-Net: 面向时序知识图谱推理的频次统计网络. 软件学报, 2023, 34(10): 4518–4532. http://www.jos.org.cn/1000-9825/6885.htm
英文引用格式: Liu KZ, Zhao F, Jin H. FS-Net: Frequency Statistical Network for Temporal Knowledge Graph Reasoning. Ruan Jian Xue Bao/Journal of Software, 2023, 34(10): 4518–4532 (in Chinese). http://www.jos.org.cn/1000-9825/6885.htm

FS-Net: Frequency Statistical Network for Temporal Knowledge Graph Reasoning

LIU Kang-Zheng[1,2,3,4], ZHAO Feng[1,2,3,4], JIN Hai[1,2,3,4]

[1](National Engineering Research Center for Big Data Technology and System, Wuhan 430074, China)
[2](Key Laboratory of Services Computing Technology and System, Ministry of Education, Wuhan 430074, China)
[3](Cluster and Grid Computing Lab, Wuhan 430074, China)
[4](School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: Temporal knowledge graph (TKG) reasoning has attracted significant attention of researchers. Existing TKG reasoning methods have made great progress through modeling historical information. However, the time-variability problem and unseen entity (relation) problem are still two major challenges that hinder the further improvement of this field. Moreover, since the structural information and temporal dependencies of the historical subgraph sequence have to be modeled, the traditional embedding-based methods often have high time consumption in the training and predicting processes, which greatly limits the application of the reasoning model in real-world scenarios. To address these issues, this study proposes a frequency statistical network for TKG reasoning, namely FS-Net. On the one hand, FS-Net continuously generates time-varying scores for the predictions at the changing timestamps based on the latest short-term

* 基金项目: 国家自然科学基金(62072203)
本文由“知识赋能的信息系统”专题特约编辑高宏教授、陈华钧教授、赵翔教授、李瑞轩教授推荐
收稿时间: 2022-07-05; 修改时间: 2022-08-18, 2022-12-14; 采用时间: 2022-12-28; jos 在线出版时间: 2023-01-13

© 中国科学院软件研究所 http://www.jos.org.cn

historical fact frequency statistics. On the other hand, based on the fact frequency statistics at the current timestamp, FS-Net supplements the historical unseen entities (relations) for the predictions; specially, FS-Net does not need training, and has a very high time efficiency. The experiments on two TKG benchmark datasets demonstrate that FS-Net has a great improvement compared with the baseline models.

Key words: temporal knowledge graph; fact frequency statistics; time variability; unseen information

传统的知识图谱通过三元组(头实体,关系,尾实体)来表示事实并已经广泛应用于金融、医疗、智能开发管理[1,2]等领域, 但是现实世界事实的发生往往具有一定的时间约束. 如图 1 左图所示: 潘基文在 2004 年至 2006 年担任韩国外交部长官, 而在 2007 年至 2016 年担任联合国秘书长; 古特雷斯在 1995 年至 2002 年担任葡萄牙总理, 而在 2017 年至今接替潘基文担任联合国秘书长. 如果忽略时间信息, 将会造成知识图谱中时间相关性信息的混淆(如古特雷斯和潘基文). 时序知识图谱在三元组的基础上加入了时间维度, 通过四元组(头实体,关系,尾实体,时间戳)来对事实进行表示, 从而能够准确地表达更为丰富的时间语义信息.

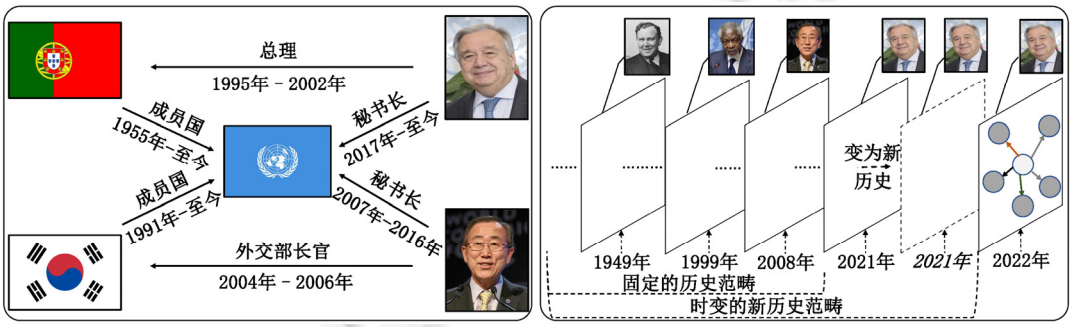


图 1 时序知识图谱、时变性问题和不可见实体问题图示

现实世界中的大多数时序知识图谱是不完整的, 时序知识图谱推理旨在一定的时间条件下对缺失的实体(头实体或尾实体)或关系进行预测. 时间约束的事实往往在一定的时间范围内长期有效, 如图 1 左图所示: 葡萄牙在 1955 年至今都是联合国的成员国, 而潘基文在 2007 年和 2016 年之间均担任联合国秘书长. 因此, 时序知识图谱上的推理预测往往在历史中具有重现的模式. 与之前聚焦于当前的预测时间戳进行建模的方法[3−5]不同, 最近的工作[6−12]通过对历史信息进行建模取得了较大的提升. 比如: RE-NET[7]将当前预测时间戳上的查询建模为历史子图序列上重复信息的条件概率; CyGNet[12]借用文本摘要的思想, 从历史中提取非重复的实体, 并为这些实体赋予较大的负得分; RE-GCN[9]设计了循环关系感知图卷积网络(recurrent relational GCN), 在建模一定长度的历史子图内部结构信息的同时, 通过时间门控单元(time gate unit)建模历史子图序列的时间依赖关系. 然而, 它们都忽视了时序知识图谱推理过程中存在的时变性问题.

如图 1 右图所示: 时序知识图谱实际上是一个顺序子图序列, 而这些子图在特定的时间戳上都是静态的. 以往的方法认为预测时间戳的历史范畴固定不变, 往往将历史信息约束为训练集和验证集的大小, 所以忽视了时序推理的时变性特征. 比如: 当要预测“2021 年的联合国秘书长是谁?”这个问题时, 在固定的历史范畴下, 2021 年的预测时间戳可以获得完整的历史信息; 然而随着时间的发展, 当要预测“2022 年的联合国秘书长是谁?”这一问题时, 2021 年的预测时间戳将会转变为新的历史时间戳(如图 1 右图虚线框所示). 因为时序知识图谱中历史事实重现的模式具有一定的时间范围, 比如古特雷斯的历史重复事实最早只能追溯到 2017 年, 所以, 由于靠近预测时间戳, 新的历史时间戳往往包含着极为重要的历史信息(我们在后文第 4.1 节和第 4.3 节对此进行了证明), 则时变的新历史范畴理应将新历史时间戳纳入其中而加以考虑. CEN[13]和 DA-Net[14]分别通过在线的学习策略和注意力机制来应对时变性挑战, 但却需要花费大量的时间用于训练和预测, 这极大地降低了模型的现实价值.

另一方面, 如果一个查询预测在某一个时间戳上第 1 次出现, 那么该预测将会面临不可见实体(关系)问题. 如图 1 右图所示: 当要预测“1949 年的联合国秘书长是谁?”这一问题时, 由于赖伊是第一任联合国秘书长,

© 中国科学院软件研究所 http://www.jos.org.cn

而在历史中也从来没有出现过“赖伊”“联合国”和“秘书长”的实体和关系，所以无法通过建模历史信息的方法对缺失的实体或者关系进行预测，这种情况被称为不可见实体(关系)问题. 最近的工作 xERTE[6]和 TITer[11]均试图从历史信息中归纳不可见实体的嵌入表示，一方面，它们忽视了对历史不可见关系的处理；另一方面，它们忽略了当前预测时间戳中存在的历史不可见信息.

在本文中，我们提出了一个频次统计网络 FS-Net 用于应对时变性和不可见实体(关系)两个挑战. 如图 2 所示：FS-Net 的时变性频次信息统计模块从一定长度的历史序列中提取重复出现的实体和关系，并基于历史中的频次统计信息对重复事实赋予一定的正得分，而对非重复事实赋予一定的负得分，通过不断时变地更新进行频次统计的历史时间戳范围，FS-Net 能够将新出现的历史信息纳入其中加以考虑；另一方面，基于当前预测时间戳的频次统计信息，FS-Net 的不可见频次信息统计模块从中提取历史不可见的实体和关系，并同样根据它们出现的频次赋予预测的得分. 如图 2 中右侧所示：当前预测时间戳中的历史不可见实体 t_3 和关系 r_5 被反馈给相应的预测，而相应的预测得分由灰色的负值根据更新的频次统计转变为一定的正值. 可以看到：FS-Net 完全基于频次统计对未知的实体或者关系进行预测，无需训练，具有极低的时间损耗(我们在后文第 4.4 节证明了这一点).

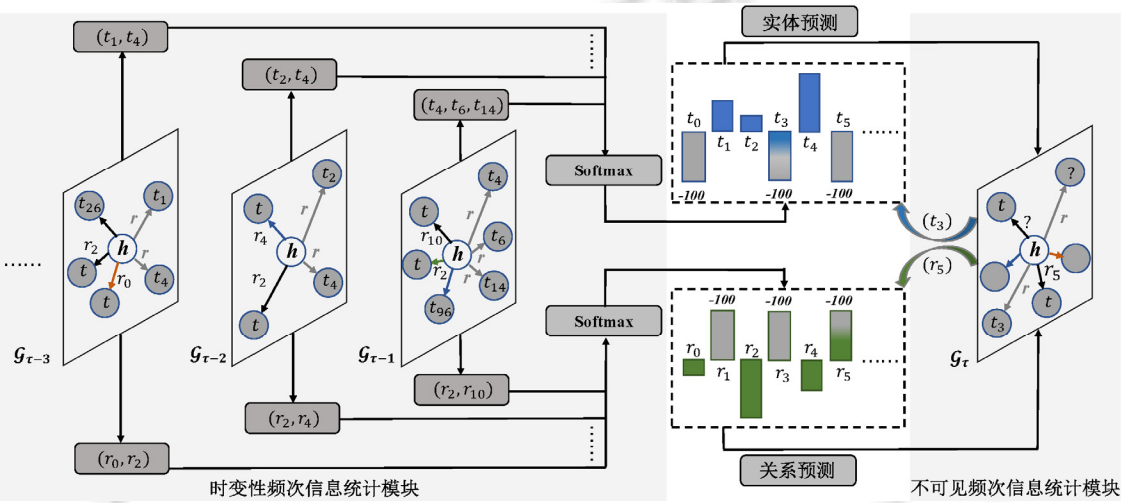


图 2 面向时序知识图谱推理的频次统计网络 FS-Net 模型框架图

我们工作的贡献可以归纳为以下几点.

- 证明了基于短期历史时间戳和当前预测时间戳的事实频次统计能够有效地解决困扰时序知识图谱推理预测的两大挑战：时变性挑战和不可见实体(关系)挑战. 基于这个研究，我们提出了频次统计网络(frequency statistical network, FS-Net)用于时序知识图谱推理. 特别地，FS-Net 不仅能够进行实体预测，还能够对未知的关系进行预测；
- 与以往的基于训练的模型不同，FS-Net 是一个完全基于统计的模型，而不需要进行任何训练；另一方面，与以往的基于嵌入的方法不同，FS-Net 仅仅通过一个与历史长度相关的循环累加就能够完成对未知实体和关系的预测，具有较低的时间复杂度，我们进一步对比证明了 FS-Net 相较于基准模型，在训练和预测阶段具有极高的效率；
- 在时序知识图谱推理领域两个公开的数据集 YAGO 和 WIKI 上对 FS-Net 模型的性能进行了验证. 实验结果表明：与丰富的基准模型相比，我们提出的 FS-Net 模型在具有较高的时间效率的同时，实现了时序推理任务表现上的较大提升.

本文第 1 节介绍时序知识图谱推理的相关方法和研究现状. 第 2 节介绍本文构建的面向时序知识图谱推理的频次统计网络模型. 第 3 节介绍主要实验的细节和与基准模型的结果对比. 第 4 节进一步设计实验并进

© 中国科学院软件研究所 http://www.jos.org.cn

行分析和讨论. 最后, 第 5 节总结全文.

1 时序知识图谱推理相关工作

时序知识图谱推理方法根据是否建模时间信息可以划分为静态的和动态的两类. 另一方面, 时间序列挖掘的方法也在知识追踪(knowledge tracing)领域广为应用.

1.1 静态的时序知识图谱推理方法

静态的推理方法忽视了时序知识图谱中的时间信息, 而只对去除了时间维度的事实三元组进行建模. 静态的推理方法在时序概念被提出之前即得到广泛研究. 基于翻译的方法包括 TransE[15]、TransH[16]、TransR[17]、TransD[18]等, 它们将实体和关系投影到低维向量空间, 并使头实体和关系的嵌入向量加和尽可能地接近尾实体的嵌入. RotatE[19]将实体和关系投影到复数空间, 并将关系定义为复数平面从一个实体到另一个实体的旋转. 此外, 基于矩阵分解的方法包括 ComplEx[20]、DistMult[21]等. ComplEx[20]对复数空间中的关系矩阵进行特征值分解, 而 DistMult[21]将关系的嵌入定义为对角矩阵. 基于卷积网络的方法包括 ConvE[22]、Conv-TransE[23]等, 它们将实体和关系表示为两个矩阵, 并引入卷积核来进行卷积操作. 基于小样本学习的方法包括 GMatching[24]、MetaR[25]、KnowCo-Tuning[26]、MaKEr[27]等, 它们为了解决知识图谱推理过程中存在的长尾问题, 旨在通过少量的训练样本完成对关系特定缺失三元组的预测. 基于图卷积网络的方法包括 R-GCN[28]、Comp-GCN[29]等, 它们通过聚合邻域的结构信息来学习知识图谱中的实体表示.

1.2 动态的时序知识图谱推理方法

动态的推理方法由于考虑了时序知识图谱中的时间信息而取得了更好的性能. TTransE[5]通过将时间附加到关系的嵌入中进行推理. HyTE[3]将时间信息建模为超平面, 通过将实体和关系投影到时间戳特定的超平面进行推理. TA-DistMult[4]也将事实的时间信息集成到关系的嵌入中. 然而, 仅仅关注当前预测时间戳无法捕获预测事实在历史中的重现模式, 所以最近的一些工作试图通过建模历史信息来进行推理, 并在性能上取得了较大的提升. RE-NET[7]将事实的发生建模为一定长度的历史子图序列的条件概率. xERTE[6]通过生成一定跳数范围内的推理图来进行推理. CyGNet[12]受启发于文本摘要, 通过复制机制抑制非重复事实在预测中所发挥的作用, 该方法对固定范畴的历史中频次为 0 的事实赋予较低的负得分, 首次将系统的频次统计方法引入时序知识图谱推理领域, 并证明了其有效性. 但是, 一方面, 它忽略了非负频次信息所带来的价值; 另一方面, 它没有考虑历史发展的时变性问题. CluSTeR[8]通过强化学习提取预测查询的相关子图, 然后通过图卷积网络对相关子图序列进行建模. RE-GCN[9]通过递归的关系图卷积神经网络对一定长度的历史子图序列的演化模式进行建模. TITer[11]同样通过强化学习来对给定查询预测路径中的演化模式进行抽取. TLogic[10]基于时间逻辑规则来对查询预测的路径进行约束.

但是, 上述这些工作都忽略了时序知识图谱推理过程中存在的时变性问题. CEN[13]尝试通过在线的学习策略来应对该挑战, 但是这样的选择需要不断地对训练集进行循环训练, 并且也对验证集和测试集进行训练和预测以获得实体和关系的时变性嵌入表示, 而这极大地降低了模型的时间效率. DA-Net[14]首先提取历史重复事实, 然后采用注意力机制与频次统计信息相结合的方式来应对时变性挑战, 历史重复事实的抽取过程与基于嵌入的注意力计算过程都需要耗费大量的时间, 不具备时效性. 与此相反, FS-Net 是一个完全基于统计而非嵌入的模型, 在关注时变性问题的同时, 也同样使推理的时间损耗降到了较低的水准. 另一方面, xERTE[6]和 TITer[11]均试图从历史信息中归纳不可见实体和关系的嵌入表示, 而忽略了当前预测时间戳中存在的历史不可见信息. 而 FS-Net 基于当前时间戳的频次统计信息对历史不可见信息进行抽取, 并将其纳入预测的考虑之中.

1.3 知识追踪的时间序列挖掘方法

知识追踪是构建自适应教育系统的核心和关键, 旨在基于学生行为序列进行建模, 预测学生对知识的掌握程度. 知识追踪问题可以描述为: 给定某个学生的观测序列$\{x_0,x_1,\ldots,x_{t-1},x_t\}$, 预测其下次表现 x_{t+1}. 其中,

© 中国科学院软件研究所 http://www.jos.org.cn

$x_t=\{q_t,a_t\}$, 而 q_t 代表回答的问题知识点, $a_t=\{0,1\}$ 代表对应的回答是否正确. 因此, 知识追踪模型本质上可以被视为一个基于学生学习历史的时间序列预测问题. 另一方面, 最近的时序知识图谱动态推理方法[6−12]通过对历史子图序列进行建模, 取得了巨大的提升. 与我们聚焦的研究对象相似, 知识追踪的研究(建模)对象为给定学生的历史观测序列, 所以, 知识追踪领域的时间序列挖掘方法对于我们进行时序子图序列的建模有一定的借鉴意义.

一直以来, 有大量的工作投入到了知识追踪领域. 贝叶斯知识追踪(BKT)[30]将学生的知识状态设为一组二元变量, 分别代表其对知识的掌握与否, 并采用隐马尔可夫模型(hidden Markov model)[31]来建模每个学生的知识状态时间序列. 在 BKT 的基础上, KT-IDEM[32]进一步考虑了学生学习知识过程中的练习难度, 而一些个性化的贝叶斯知识追踪模型[33]则进一步考虑了学生之间的个体差异性. 然而, 早期的 BKT 模型都假设每个问题的回答只需要一类学习到的知识, 因此, 它们不能处理需要多类知识的问题. 为了应对这一挑战, 动态贝叶斯网络(dynamic Bayesian network)[34]被应用于对多种知识和不同类别知识之间的依赖性进行联合建模. 另外, 一系列基于深度学习的方法又使知识追踪模型的性能获得了进一步的提升. DKT-DSC[35]通过 K-means 将学生回答问题的表现聚类成组, 并在表现随时间动态变化时更新聚类信息. SKVMN[36]采用一种改进的 LSTM (hop-LSTM)对学生知识序列进行顺序建模, 同时保持与 DKVMN[37]相同的键值存储结构. SAKT[38]首次在知识追踪模型中引入注意力机制[39]. 基于图算法的知识追踪模型进一步与时序知识图谱领域中的历史子图序列挖掘方法相聚拢. GKT[40]将学生回答问题所需知识结合成图, 图中的边代表知识之间的依赖关系. 他们将知识追踪问题重新组织为时间序列节点分类问题, 并使用传统的图神经网络(graph neural network)来解决该问题. GIKT[41]进一步将问题表示为图中的节点, 通过学习节点嵌入的方式进行答案预测. SKT[42]旨在捕获知识之间的相似、先决条件等关系, 进而增强图中节点嵌入的语义表示.

2 频次统计网络 FS-Net

本节我们介绍所提出的时序知识图谱推理模型 FS-Net. 首先对符号和任务进行定义; 然后, 依次介绍 FS-Net 的总体设计框架以及构成 FS-Net 的两个模块: 时变性频次信息统计模块和不可见频次信息统计模块; 最后, 对 FS-Net 在训练和预测过程中的计算复杂度进行分析.

2.1 符号和任务定义

在一个时序知识图谱$\mathcal{G}$中, 我们定义实体集为$\mathcal{E}$, 大小为 E; 定义关系集为$\mathcal{R}$, 大小为 R; 定义时间戳集合为$\mathcal{T}$, 大小为 T. 定义时序知识图谱中的一个事实四元组为(h,r,t,τ), 其中, h 表示头实体, r 表示关系, t 表示尾实体, τ表示时间戳, 并且有$\{h,t\}\in\mathcal{E}$, $r\in\mathcal{R}$, $\tau\in\mathcal{T}$, 则时序知识图谱$\mathcal{G}$可以表示为$\mathcal{G}=\{\mathcal{G}_0,\mathcal{G}_1,\ldots,\mathcal{G}_{T-1}\}$, 其中, 每一个时序子图$\{\mathcal{G}_\tau|0\leqslant\tau\leqslant T-1\}$由发生在时间戳$\tau$的所有事实三元组$(h,r,t)$组成.

时序知识图谱推理可以进一步细分为以下两个子任务.

(1) 实体预测. 在某一个特定的时间戳τ的条件下, 对缺失的头实体$(?,r,t,\tau)$或者尾实体$(h,r,?,\tau)$进行预测;

(2) 关系预测. 在某一个特定的时间戳τ的条件下, 对缺失的关系$(h,?,t,\tau)$进行预测.

对于某一个时间戳τ上的实体预测或者关系预测任务, 以尾实体预测$(h,r,?,\tau)$为例, 设其缺失的实体为 o, 则如前文关于图 1 例子的阐述, 需要补全的缺失事实(h,r,o,τ)往往在距离当前时间戳τ 一定的历史时间范围$\{t|\tau-k\leqslant t<\tau\}$内有效, 其中, k 表示历史时间戳的长度, 即缺失事实(h,r,o,τ)在一定长度 k 的历史子图$\{\mathcal{G}_t|\tau-k\leqslant t<\tau\}$中存在重现的模式$\{(h,r,o,t)|\tau-k\leqslant k<\tau\}$. 更进一步, 基于后文第 4.1 节和第 4.3 节中我们对时变性问题的研究, 随着时间的演进, 对于新的预测时间戳$\tau+1$ 上的预测任务$(h,r,?,\tau+1)$(同样设缺失尾实体为 o), 原本的预测时间戳τ将变为历史时间戳, 而时变的新历史时间戳τ上包含着重要的历史信息. 这是因为缺失事实$(h,r,o,\tau+1)$的有效时间范围$\{t|\tau+1-k\leqslant t\leqslant\tau+1\}$是连续的, 它存在于长度为$k$的历史子图和当前时间戳子图中. 特别地, 当 k 足够小时, 对应于缺失事实$(h,r,o,\tau+1)$的有效期足够短, 而在前一个时变的新历史时间戳τ中仍然存在着有用的重复事实(h,r,o,τ), 所以往往最新的时变历史时间戳上包含着重要的重复信息.

© 中国科学院软件研究所 http://www.jos.org.cn

2.2 总体设计框架

图 2 展示了我们提出的 FS-Net 模型的整体框架, 它由时变性频次信息统计模块和不可见频次信息统计模块两个部分组成. 对于时间戳τ(时序子图$\mathcal{G}_\tau$)上的实体预测任务$(h,r,?,\tau)$或者关系预测任务$(h,?,t,\tau)$, 为了能够捕获新产生的历史时间戳上重要的重复信息, 与之前的工作将历史范围设定为固定的训练集大小并忽略时变性问题不同, 我们时变地为预测时间戳提取最新的一定长度的历史时间戳上的频次统计信息, 并随着时间的演进, 根据时变性的频次信息统计为不同的时间戳上的预测任务不断更新得分; 为了能够为当前时间戳τ上的预测任务补充时变性的历史信息中仍然不存在的缺失事实, 如图 2 中所示的(h,r,t_3,τ)和(h,r_5,t,τ), 我们又根据当前时间戳τ上的频次统计, 将需要预测的缺失实体 t_3 和缺失关系 r_5 反馈给时变性频次信息统计模块, 并根据它们在当前时间戳上的频次统计赋予预测的得分.

2.3 时变性频次信息统计模块

由于时序知识图谱中事实的时效性, 所以对于一个时间戳τ上的实体预测$(h,r,?,\tau)$, 或者关系预测$(h,?,t,\tau)$, 同样的事实往往在一定的历史长度 k 上重复出现. 时变性, 即随着预测时间戳的改变, 历史信息的范畴也不断改变, 并不断有新的历史信息产生. 所以, 为了捕获时序知识图谱的时变性特征, 我们动态地随着预测时间戳的改变而更新频次统计的历史范畴, 确保长度为 k 的历史子图始终包含最新的历史信息.

对于实体预测$(h,r,?,\tau)$(这里以尾实体预测为例, 头实体预测类似), 历史重复事实可以表示为

$$\{(h,r,t_i)|(h,r,t_i)\in\mathcal{G}_j,j<\tau\},$$

其中, $\{t_i\}$为历史中重复出现的实体.

而非重复事实可以表示为$\{(h,r,t_i)|(h,r,t_i)\notin\mathcal{G}_j,j<\tau\}$. 我们通过一个多热的向量 $\{f_i^{(h,r)}\in\mathbb{R}^E\mid\tau-k\leqslant i<\tau\}$ 来记录预测查询在长度为 k 的每一个历史时间戳上的重现模式, 其中, 如果事实(h,r,t_q)在第 i 个时间戳重现过, 那么 $f_i^{(h,r)}$ 的第 q 维为 1; 否则为 0. 因此, 长度为 k 的历史子图的实体频次统计信息可以表示为

$$F_\tau^{(h,r)}=f_{\tau-k}^{(h,r)}+f_{\tau-k+1}^{(h,r)}+\ldots+f_{\tau-1}^{(h,r)} \tag{1}$$

其中, $F_\tau^{(h,r)}\in\mathbb{R}^E$ 的第 q 个维度表示为实体 t_q 在长度为 k 的历史中出现的频次. 该模块进一步基于历史频次信息的统计对非重复的实体赋予一定的负得分, 可以表示为 $Ne_\tau^{(h,r)}\in\mathbb{R}^E$; 而对于重复的实体, 根据其出现的频次赋予一定的正得分:

$$Po_\tau^{(h,r)}=Softmax(F_\tau^{(h,r)}) \tag{2}$$

同样, 对于关系预测$(h,?,t,\tau)$, 历史重复事实可以表示为$\{(h,r_i,t)|(h,r_i,t)\in\mathcal{G}_j,j<\tau\}$, 其中, $\{r_i\}$是历史重复关系. 类似地, 多热向量 $\{f_i^{(h,t)}\in\mathbb{R}^R\mid\tau-k\leqslant i<\tau\}$ 可以用来记录查询在长度为 k 的每一个历史时间戳上的重现模式, 其中, 如果事实(h,r_p,t)在第 i 个时间戳重现过, 则有 $f_i^{(h,t)}$ 的第 p 维为 1; 否则为 0. 那么长度为 k 的历史子图的关系频次统计信息可以表示为

$$F_\tau^{(h,t)}=f_{\tau-k}^{(h,t)}+f_{\tau-k+1}^{(h,t)}+\ldots+f_{\tau-1}^{(h,t)} \tag{3}$$

其中, $F_\tau^{(h,t)}\in\mathbb{R}^R$ 的第 p 个维度表示为关系 r_p 在长度为 k 的历史中出现的频次. 同样地, 该模块对非重复的关系赋予一定的负得分, 表示为 $Ne_\tau^{(h,t)}\in\mathbb{R}^R$; 而对于重复关系, 根据其频次赋予一定的正得分:

$$Po_\tau^{(h,t)}=Softmax(F_\tau^{(h,t)}) \tag{4}$$

其中, $Po_\tau^{(h,t)}\in\mathbb{R}^R$. 该模块最终的得分可以表示为正得分和负得分的加和:

$$\mathcal{S}c_{history}=\begin{cases}Ne_\tau^{(h,r)}+Po_\tau^{(h,r)}, & \text{实体预测}\\ Ne_\tau^{(h,t)}+Po_\tau^{(h,t)}, & \text{关系预测}\end{cases} \tag{5}$$

2.4 不可见频次信息统计模块

该模块基于当前预测时间戳的频次统计信息, 将历史不可见实体和关系反馈给时变性频次信息统计模块所获得的得分.

© 中国科学院软件研究所 http://www.jos.org.cn

我们同样通过一个多热向量来记录当前时间戳的频次统计信息. 对于实体预测$(h,r,?,\tau)$, 假设其缺失的尾实体为t, 则该多热向量可以表示为$C_\tau^{(h,r)}$. 如果事实(h,r,t_q)在当前预测时间戳出现过, 那么$C_\tau^{(h,r)}$中第q个相应的维度为 1. 而与$F_\tau^{(h,t)}$中的记录不同, 对于历史中没有出现过的实体, $C_\tau^{(h,r)}$中的相应维度为 1, 而不是 0, 这是因为当前时间戳τ下存在推理的完整事实(h,r,t,τ), 缺失的实体 t 虽然为历史不可见信息, 却一定存在于当前时间戳τ, 进而可以将历史不可见实体纳入预测的考虑. 同样地, 不可见频次信息统计模块为频次统计为 0 的实体赋予一定的负得分:

$$\mathcal{S}c_{current} = zero(C_\tau^{(h,r)}) + (-\delta) \tag{6}$$

其中, δ是一个任务特定的常量, 我们设定其为 50. 对于关系预测$(h,?,t,\tau)$, 其处理过程与上述实体预测类似, 针对关系预测的该模块得分可以表示为

$$\mathcal{S}c_{current} = zero(C_\tau^{(h,t)}) + (-\delta) \tag{7}$$

所以, 对于 FS-Net 模型, 最终基于长度为 k 的历史时间戳和当前预测时间戳的频次统计信息所获得的进行实体和关系预测的得分可以表示为

$$\mathcal{S}c=\mathcal{S}c_{history}+\mathcal{S}c_{current} \tag{8}$$

其中, 对于实体预测任务, $\mathcal{S}c\in\mathbb{R}^E$; 而对于关系预测任务, $\mathcal{S}c\in\mathbb{R}^R$. 预测的过程可以看作一个进行多标签分类的过程, 而 Sc 中的每一个维度均表示为预测相应的实体或关系作为缺失事实的概率.

2.5 计算复杂度分析

FS-Net 极高的时间效率主要有以下两个方面的原因: (1) 基于统计而无需训练, 在训练过程的时间消耗为 0; (2) 预测过程中的计算复杂度较低. 下面我们对 FS-Net 模型在预测阶段的时间损耗进行了分析. 像之前工作[12]的做法一样, 我们采用大小为 $E*R\times E$ 的稀疏矩阵(对于关系预测为 $E*E\times R$ 的稀疏矩阵)对长度为 k 的历史子图的重复模式(频次信息)进行记录. 比如: 事实(h,r,t,τ)在第τ个时间戳上存在, 那么对应于时序子图$\mathcal{G}_\tau$的记录稀疏矩阵中坐标$(h*r,t)$对应的元素为 1, 否则为 0, 则历史频次信息的抽取所消耗的时间复杂度为 O(k). 时变性频次信息统计模块的得分仅仅通过一个与历史长度 k 相关的循环累加, 将相应记录频次信息的稀疏矩阵相加即可获得, 所以该模块的时间复杂度为 O(k). 类似地, 针对当前预测时间戳的不可见频次信息统计模块的时间复杂度为 O(1). 进而在预测过程中, 以尾实体预测$(h,r,?,\tau)$为例, 基于长度为 k 的历史子图$\{\mathcal{G}_{\tau-k}, \mathcal{G}_{\tau-k+1},\ldots,\mathcal{G}_{\tau-1}\}$对应的稀疏矩阵的加和, 最终所得的加和稀疏矩阵中保存着预测时间戳中每一个事实在历史的 k 个子图中出现情况与频次统计. 比如: 对于尾实体预测$(h,r,?,\tau)$, 我们通过简单地读取最终的加和稀疏矩阵中的第 $h*r$ 行, 就可以完成查找匹配并获得维度为 E 的多热向量, 该向量即为针对该尾实体预测的多分类实体得分. 所以, FS-Net 的训练和预测的计算复杂度为 O(k). 我们进一步在后文第 4.3 节证明了 k 取值为较小的常数即可以满足 FS-Net 正常的推理和预测.

3 主实验

本节在两个公开的时序知识图谱数据集上对我们提出的 FS-Net 模型的性能进行评估.

3.1 数据集介绍

我们用来进行模型性能评估的时序知识图谱数据集包括 YAGO 和 WIKI: YAGO 数据集是传统的大型知识图谱 YAGO3[43]的时序子图, 而 WIKI 也来自于维基百科中的时序信息. 在以往的工作[3−14]中, YAGO 和 WIKI 两个数据集已被广泛应用于时序知识图谱推理模型性能的评估. 根据这些以往的工作[3,6−14], 我们按照 80%/10%/10%的比例将两个数据集划分为训练集、验证集和测试集, 其中, 根据传统的定义, 训练集和验证集用于进行训练, 而测试集用于进行预测. 数据集的详细信息见表 1.

© 中国科学院软件研究所 http://www.jos.org.cn

表 1 时序知识图谱实验数据集

数据集	实体数目	关系	训练集数目	验证集数目	测试集数目	时间戳粒度(年)
YAGO	10 623	10	161 540	19 523	20 026	1
WIKI	12 554	24	539 286	67 538	63 110	1

3.2 评估指标及基准模型

我们将 FS-Net 与丰富的基准模型进行了对比，这些模型包括静态的推理方法，如 DistMult[21]、ConvE[22]、ComplEx[20]、Conv-TransE[23]、RotatE[19]和 R-GCN[28]. 对比的动态的推理方法包括 TTransE[5]、HyTE[3]和 TA-DistMult[4]. 一些动态的推理方法通过建模历史信息获得了很大的提升，这些方法包括 RE-NET[7]、CyGNet[12]、xERTE[6]、RE-GCN[9]、TITer[11]和 CEN[13]. 这些对比的基准模型详见第 1 节相关工作中的介绍.

在主实验中，我们使用经典的指标来评估 FS-Net 与基准模型的性能，包括平均倒数排名(MRR)、排名前 3 和排名前 10 的命中率(Hits@3/10)，它们都表示着缺失的实体或者关系在预测结果中的排名. 对于实体预测任务，我们报告头实体预测和尾实体预测的平均结果. 以往的工作[6,8−11,13]已经证明：对于评估指标而言，传统的静态过滤设置(filtered setting)[15−29]由于在过滤过程中忽视了时序的特点，所以是不合理的.

与以往工作[9]一样，不失一般性，我们采用原始配置(raw setting)的评估指标，而不进行任何过滤操作，以报告模型的最真实表现. 主实验的部分结果同样采用以往工作[9]中报告的数值.

3.3 实验细节

我们通过 PyTorch 框架实现了 FS-Net 模型，并在 Tesla V100 的 GPU 设备上进行实体和关系的预测任务. 除去第 2 节介绍的模型参数以外，对于 YAGO 和 WIKI 两个数据集，我们设定实体预测任务的历史长度 k 为 1；对于关系预测任务，我们设定 k 为 3. 设定预测阶段批处理的大小(batch size)为每一个预测时间戳的大小. 对于静态的推理方法，因为它们在建模时都忽视了时序动态性，所以与以往动态工作[3−14]的处理方法一样，我们将时序知识图谱数据集中事实的时间维度去除，继而符合这些静态模型的输入格式，然后在一个具有混淆语义的静态知识图谱上进行静态方法的推理预测评估.

对于 xERTE[6]、RE-GCN[9]、TITer[11]和 CEN[13]这些重要的基准模型，我们在原始的评估配置下，采用它们的开源代码和默认参数在与 FS-Net 相同的实验环境里对 YAGO 和 WIKI 两个数据集上的结果进行了复现. 对于应对时变性挑战的 CEN[13]模型，我们分别报告了它在离线设置(offline setting)和在线设置(online setting)下的表现，其中，其在线设置是为了解决时变性问题而提出的.

3.4 主要实验结果

本节我们分别在实体预测和关系预测任务上对 FS-Net 和基准模型的性能进行比较.

- 实体预测

如表 2 所示，我们将最好的结果加粗，次好的结果加下划线.

可以看到，FS-Net 的表现要远好于静态的推理模型，这是因为静态的推理方法完全忽略了时序知识图谱中的时间信息. 然而，一些动态推理方法的表现甚至还要差于静态的方法，比如 TTransE, HyTE 和 TA-DistMult. 这是因为这些方法仅仅关心时间信息的嵌入表示方法，更多地聚焦于当前预测时间戳的建模，而忽略了包含在历史中的大量有价值的信息. 与最近的建模历史信息的动态推理方法 RE-NET, CyGNet, xERTE, RE-GCN, TITer 和 CEN(offline)相比，FS-Net 的表现仍然要优于这些模型，因为它们都忽视了时变性的问题. 与尝试解决时变性问题的 CEN(online)方法相比，FS-Net 的表现仍然较优. 这是因为基于嵌入的 CEN (online)方法没有考虑不可见实体问题，而无法获得历史不可见实体的合理的嵌入表示，进而影响了模型的表现. 我们也观察到，TITer 的表现甚至要好于 CEN (online). 一方面，这是因为 TITer 尝试建模了历史不可见信息；另一方面，与其他模型在整个实体集范围内进行排名评估不同，TITer 通过强化学习的方法直接在候选集中进行排名评估，而候选集的大小要远小于实体集，所以往往能收获更好的性能表现.

© 中国科学院软件研究所 http://www.jos.org.cn

表 2 FS-Net 与时序知识图谱推理方法在原始配置下的实体预测性能比较(以百分比的形式展现)

Model	YAGO			WIKI		
	MRR	Hits@3	Hits@10	MRR	Hits@3	Hits@10
DistMult	44.05	49.70	59.94	27.96	32.45	39.51
ConvE	41.22	47.03	59.90	26.03	30.51	39.18
ComplEx	44.09	49.57	59.64	27.69	31.99	38.61
Conv-TransE	46.67	52.22	62.52	30.89	34.30	41.45
RotatE	42.08	46.77	59.39	26.08	31.63	38.51
R-GCN	20.25	24.01	37.30	13.96	15.75	22.05
TTransE	26.10	36.28	47.73	20.66	23.88	33.04
HyTE	14.42	39.73	46.98	25.40	29.16	37.54
TA-DistMult	44.98	50.64	61.11	26.44	31.36	38.97
RE-NET	46.81	52.71	61.93	30.87	33.55	41.27
CyGNet	46.72	52.48	61.52	30.77	33.83	41.19
xERTE	64.29	74.50	87.38	52.85	60.96	71.89
RE-GCN	62.72	70.63	82.05	51.01	57.35	68.64
TITer	64.97	74.80	87.44	57.36	63.80	72.52
CEN(offline)	62.96	70.98	82.42	51.28	57.65	69.41
CEN(online)	63.46	71.85	83.46	52.09	59.05	70.74
FS-Net	**70.30**	**80.96**	**94.39**	**62.81**	**72.70**	**84.53**

- 关系预测

如表 3 所示, 同样地, 我们将最好的结果加粗, 次好的结果加下划线. 由于两个数据集中的关系数目较少(YAGO 数据集包含 10 个关系, 而 WIKI 数据集包含 12 个关系), 所以我们仅报告最具代表性的 MRR 指标的结果. 以往大量的工作研究了时序知识图谱推理中的实体预测任务, 而在关系预测任务中的研究不多. 具体而言, 像以往工作[9]那样, 对于静态的推理模型, 我们对比了 ConvE[22]和 Conv-TransE[23]; 而对于动态的推理模型, 我们对比了 R-GCRN[44]和 RE-GCN[9], 其中, GCRN[44]是同构图模型, RE-NET[7]通过用 R-GCN[28]代替模型中的图卷积网络(GCN)而将其推广到 R-GCRN. 相较于在实体预测任务上的提升, FS-Net 相较于基准模型的提升相对较小. 这是因为关系的数目要远小于实体数目, 所以关系预测相对容易. 基准模型中, 动态方法的表现要相对优于静态的方法, 这也是因为它们建模了时间信息. RE-GCN 由于捕获了历史子图序列的演化模式而显著优于其他基准模型, 但是同样地, 它忽略了时序知识图谱推理过程中的时变性问题和不可见关系问题. 因此, FS-Net 在关系预测任务上的表现也要显著优于包括 RE-GCN 在内的所有基准模型.

表 3 FS-Net 与时序知识图谱推理方法在原始配置下的关系预测性能比较(以百分比的形式展现)

Model	YAGO	WIKI
ConvE	91.33	78.23
Conv-TransE	90.98	86.64
R-GCRN	90.18	88.88
RE-GCN	94.20	98.17
FS-Net	**94.59**	**99.39**

4 分析和讨论

为了进一步评估和分析面向时序知识图谱推理的频次统计网络 FS-Net 的有效性, 我们在难度更大的实体预测任务上继续研究了以下 5 个问题.

- RQ1: FS-Net 的时变性频次信息统计模块建模时变性特征的表现如何?
- RQ2: FS-Net 的不可见频次信息统计模块捕获不可见实体的表现如何?
- RQ3: FS-Net 的关键参数——历史长度 k 的取值对模型的表现有何影响?
- RQ4: FS-Net 的时间效率如何?
- RQ5: FS-Net 擅长什么场景条件下的时序推理任务?

4.1 时变性问题分析

为了回答 RQ1, 我们将 YAGO 数据集的测试事实按照时间戳划分为 5 个单独的集合{183,184,185,186,187}, 其中, 每一个数字代表一个年份. 随着时间的演变, 我们依次报告模型在每一个时间戳上的表现. 以往

© 中国科学院软件研究所 http://www.jos.org.cn

的方法忽视了时变性问题, 将训练集和验证集作为固定的历史空间, 而未考虑历史范畴的时变. 为了直观地反映 FS-Net 的时变性频次信息统计模块建模时变性特征的能力, 我们在研究该问题的同时未考虑不可见实体问题(即不通过不可见频次信息统计模块补充预测时间戳的不可见信息).

如图 3 所示: 由于受空间限制, 我们同样采用最具代表性的 MRR 指标进行比较. 其中, FS-Net 代表我们提出的 FS-Net 模型的时变性频次信息统计模块的表现; FS-Net (w.FH)代表将历史范畴固定为训练集和验证集大小情况下 FS-Net 模型的表现; CEN (online)代表以往的尝试建模时变性信息的在线设置下的 CEN 模型. 可以看到: 在第 1 个预测时间戳(183), 由于与验证集相邻, 所以当前情况下 3 个对比的模型具有相同的历史空间. 因此, FS-Net 和 FS-Net (w.FH)在 183 时间戳的表现相同, 而 CEN (online)表现也不相上下. 但是随着时间的发展演变, 未考虑时变性信息的 FS-Net (w.FH)的表现显著下降, 而考虑了时变性信息的 FS-Net 和 CEN (online)持续具有较好的表现, 而且两个模型在线型趋势上的一致性证明了 FS-Net 模型解决时变性问题的有效性. 另一方面, CEN (online)着重建模时序子图序列的演化模式, 执着于获得历史实体的准确嵌入表示, 而忽视了时序知识图谱更为直接的重现模式. 因此, FS-Net 的时变性频次信息统计模块的表现随时间持续优于 CEN (online), 这也证明了重现模式在时变性背景下, 特别是时间戳粒度较大(1 年)的情况下, 作用更加凸显.

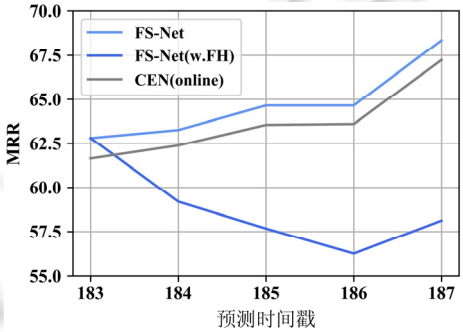


图 3 在 YAGO 数据集上的时变性问题分析

4.2 不可见实体问题分析

为了回答 RQ2, 我们分别在去除不可见实体信息和补充不可见实体信息的情况下, 测试 FS-Net 的不可见频次信息统计模块所发挥的作用. 由于 FS-Net 由两个模块组成, 去除不可见频次信息统计模块的 FS-Net 模型实际上是时变性频次信息统计模块. 因此, 我们同时与解决时变性问题的 CEN (online)模型和解决不可见实体问题的 TITer 模型进行比较. 由于空间原因, 我们同样仅报告最具代表性的 MRR 指标.

如图 4 所示, FS-Net 代表我们所提出的完整模型, 而 FS-Net (wo.UE)代表不包含不可见频次信息统计模块的 FS-Net 模型.

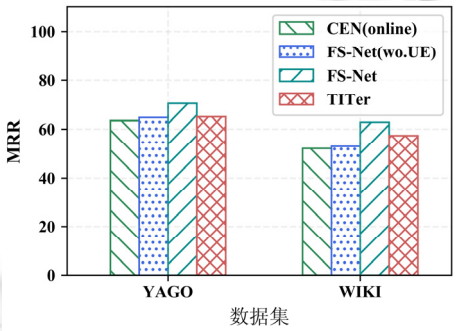


图 4 在 YAGO 和 WIKI 数据集上的不可见实体问题分析

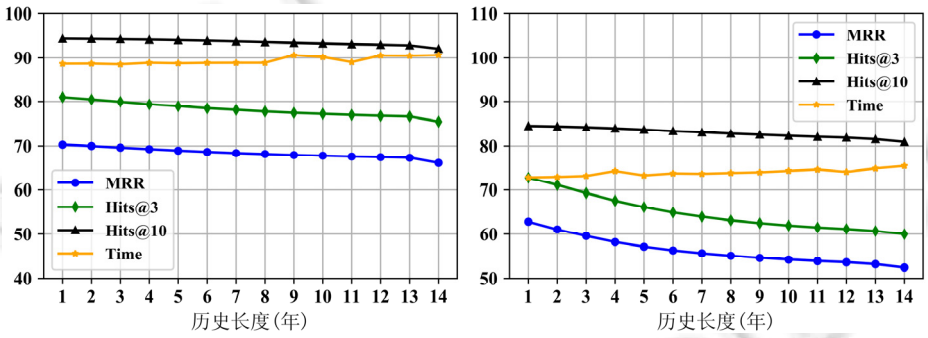
© 中国科学院软件研究所 http://www.jos.org.cn

可以看到, FS-Net (wo.UE)的性能还要略优于 CEN (online). 如第 4.1 节所述, 这证明了 FS-Net 的时变性频次信息统计模块的有效性. 在不加入不可见频次信息统计模块之前, FS-Net (wo.UE)的性能差于 TITer, 这是因为 TITer 考虑并建模了历史不可见信息. 而且如第 3.4 节所述, 基于强化学习方法的 TITer 往往能够收获到更好的结果数据. 当加入不可见频次信息统计模块之后, FS-Net 模型的性能实现了一定的提升, 并要优于 TITer 模型的表现. 这是因为 TITer 执着于从历史子图中对不可见信息进行归纳, 而忽视了预测时间戳中存在的更为直接的历史不可见信息. 这也证明了 FS-Net 的不可见频次信息统计模块在解决不可见实体问题时的优越性. 具体而言, 由于当前时间戳中针对不可见实体的预测查询往往没有在历史中出现过, 所以这些历史不可见实体即便仅仅在当前预测时间戳的频次统计中出现过一次, 也能对历史非重复的预测提供可靠的帮助.

4.3 灵敏度分析

为了回答 RQ3, 我们分别在 YAGO 和 WIKI 数据集上对历史长度 k 的取值进行了研究. 我们设定参数 k 的取值区间为{1,2,3,4,5,6,7,8,9,10,11,12,13,14}. 在报告模型的 MRR、Hits@3、Hits@10 性能指标表现的同时, 我们也统计了模型在 k 取值发生变化的情况下, 预测过程所消耗的时间. 需要注意的是: 为了能够同时展现模型的预测时间与性能表现, 对于 YAGO 数据集, 所消耗时间的统计在原始数值的基础上加了 70; 而对于 WIKI 数据集, 时间统计在原始数值的基础上加了 10. 所以, 图 5 仅刻画了预测时间消耗随历史长度 k 的变化趋势.

如图 5 所示: 在两个数据集上, 随着 k 的增加, 模型的性能都呈下降趋势. 这是因为时序知识图谱中的事实都具有一定的时效性, 有价值的历史信息往往聚集在距离预测时间戳较近的历史时间戳内. 显而易见: 随着历史长度 k 的增加, 进行频次统计的时间消耗将会增大, 即图中的“Time”曲线都呈上升趋势. 因此, 时序知识图谱的时效特性又极大地降低了 FS-Net 进行频次统计所消耗的时间, 历史长度 k 往往取较小的数值即可满足 FS-Net 进行时序推理的有效性. 所以, FS-Net 对参数 k 敏感也证明了模型在预测阶段具有极高的时间效率.



(a) YAGO 数据集 (b) WIKI 数据集

图 5 FS-Net 的性能和时间效率随历史长度 k 的变化研究(灵敏度分析)

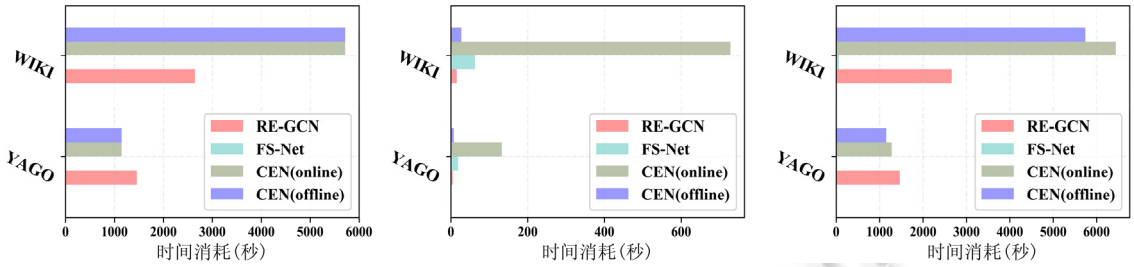
4.4 时间效率对比分析

为了回答 RQ4, 我们分别对比了 FS-Net、RE-GCN、CEN (online)和 CEN (offline)这 4 个模型在训练过程、预测过程和训练加预测过程的时间消耗.

如图 6(a)所示: 我们提出的基于频次统计的 FS-Net 模型不需要进行训练, 所以在训练过程中的时间消耗为 0; 对于 CEN (online)和 CEN (offline)模型的训练时间统计, 我们采用它们提供的开源代码中的默认训练轮次(epoch); 对于 RE-GCN 模型, 由于它的开源代码中没有明确规定训练轮次, 所以我们统计它进行 30 轮训练(与 CEN 相同)的时间消耗. CEN (online)和 CEN (offline)在训练过程中时间消耗较大的原因是, 它们经历了“Pre-train”和“Curriculum train”两个训练步骤. 而如图 6(b)所示: 在预测过程中, CEN (offline)和 RE-GCN 模型由于没有考虑解决时变性问题而具有较高的时间效率. 然而, FS-Net 的时间效率要远高于同样考虑时变性问题的 CEN (online)模型. 因为 FS-Net 完全基于统计, 而且如第 4.3 节所述, 进行统计操作的历史长度参数 k 取

© 中国科学院软件研究所 http://www.jos.org.cn

值很小, 这使得 FS-Net 在预测阶段同样具有极高的时间效率. 总的来说, 如图 6(c)所示, 我们提出的面向时序知识图谱推理的 FS-Net 模型具有远低于卓越的基准模型的训练和预测时间损耗.



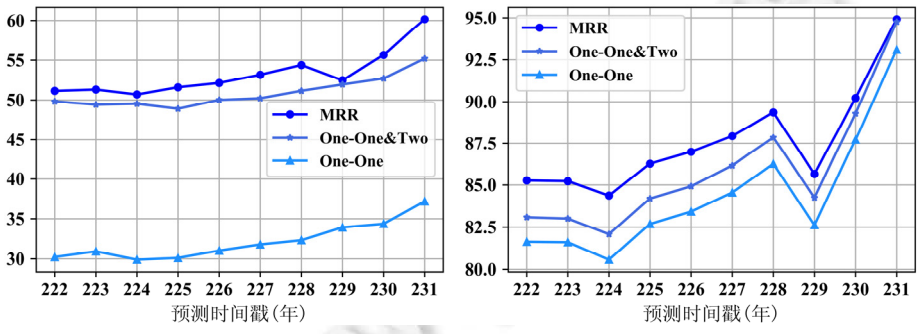
(a) 训练过程 (b) 预测过程 (c) 训练过程和预测过程

图 6 在 YAGO 和 WIKI 数据集上的时间效率对比分析

4.5 FS-Net的推理场景(局限性)分析

为了回答 RQ5, 我们在 WIKI 数据集上进行了实验. 根据第 2 节、第 3 节中对 FS-Net 模型和实验细节的介绍, 以时间戳τ上的尾实体预测$(h,r,?,\tau)$为例, FS-Net 从时变的前一个时间戳τ–1 上提取重复出现的历史事实$\{(h,r,t_i,\tau-1)\}$, 并从当前时间戳τ上补充历史不可见信息$\{(h,r,t_j,\tau)\}$, 最终的预测尾实体从候选集$\{t_i\cup t_j\}$中选取. 显然, 如果预测事实$(h,r,?,\tau)$为一对一关系事实, 那么$|\{t_i\cup t_j\}|=1$, 缺失尾实体 $t\in\{t_i\cup t_j\}$将为预测的唯一候补; 反之, 如果预测事实$(h,r,?,\tau)$为一对多关系事实, 那么$|\{t_i\cup t_j\}|>2$, 过多的候选实体将会使 FS-Net 产生困惑.

我们首先将 WIKI 数据集的测试集部分按照预测时间戳进行划分; 然后, 分别报告每一个预测时间戳中一对一关系事实和一对二关系事实所占的比例, 以及该预测时间戳的 MRR 指标性能表现. 如图 7 所示, 横坐标代表预测时间戳的年份, “One-One&Two”代表一对一和一对二关系事实所占当前预测时间戳事实总数的比例, “One-One”代表一对一关系事实所占当前预测时间戳事实总数的比例. 可以观察到: 在 WIKI 数据集的每一个预测时间戳上实体预测和关系预测的表现中, MRR 指标的统计与一对一关系事实和一对二关系事实所占的比例呈强相关, 这与我们前面的分析相对应. 同样, 对于尾实体预测$(h,r,?,\tau)$, 如果其为一对多关系事实, 则历史候选实体$\{t_i\}$和历史不可见候选实体$\{t_j\}$中将包含多个冲突候选, 而 FS-Net 对这种场景是不友好的. 因此, FS-Net 高效的建模擅长对一对一关系、一对二关系等事实场景下的稀疏图进行推理预测; 而一对多、多对多关系事实的稠密图场景, 也是我们未来的一个重要研究内容.



(a) 实体预测 (b) 关系预测

图 7 WIKI 数据集上 FS-Net 的推理场景(局限性)分析

5 总 结

本文提出了一种面向时序知识图谱推理的频次统计网络(FS-Net), 该方法时变地对一定长度的历史事实频次进行统计, 并基于此为不同时间戳的预测查询生成时变的得分; 而且, 基于预测时间戳上的事实频次统

© 中国科学院软件研究所 http://www.jos.org.cn

计, 该方法为查询预测补充相应的历史不可见信息. 实验结果表明, FS-Net 模型能够很好地应对时序知识图谱推理领域的时变性和不可见实体(关系)两大挑战. 另一方面, FS-Net 在性能表现超过卓越基准模型的同时, 也具有远低于它们的时间消耗. 因此, 基于统计的 FS-Net 模型能够高效、准确地完成时序知识图谱上的推理任务, 进而相较于基准模型具有更好的应用价值.

References:

[1] Wang F, Liu JP, Liu B, Qian TY, Xiao YH, Peng ZY. Survey on construction of code knowledge graph and intelligent software development. Ruan Jian Xue Bao/Journal of Software, 2020, 31(1): 47−66 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5893.htm [doi: 10.13328/j.cnki.jos.005893]

[2] Liu BZ, Wang X, Liu PK, Li SZ, Zhang XW, Yang YJ. KGDB: Knowledge graph database system with unified model and query language. Ruan Jian Xue Bao/Journal of Software, 2021, 32(3): 781−804 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6181.htm [doi: 10.13328/j.cnki.jos.006181]

[3] Dasgupta SS, Ray SN, Talukdar P. Hyte: Hyperplane-based temporally aware knowledge graph embedding. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. 2018. 2001−2011.

[4] Garcia-Duran A, Dumančić S, Niepert M. Learning sequence encoders for temporal knowledge graph completion. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. 2018. 4816−4821.

[5] Jiang TS, Liu TY, Ge T, Sha L, Chang BB, Li SJ, Sui ZF. Towards time-aware knowledge graph completion. In: Proc. of the 26th Int'l Conf. on Computational Linguistics: Technical Papers (COLING 2016). 2016. 1715−1724.

[6] Han Z, Chen P, Ma YP, Tresp V. Explainable subgraph reasoning for forecasting on temporal knowledge graphs. In: Proc. of the Int'l Conf. on Learning Representations. 2020.

[7] Jin WJ, Qu M, Jin XS, Ren X. Recurrent event network: Autoregressive structure inferenceover temporal knowledge graphs. In: Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing. 2020. 6669−6683.

[8] Li ZX, Jin XL, Guan SP, Li W, Guo JF, Wang YZ, Cheng XQ. Search from history and reason for future: Two-stage reasoning on temporal knowledge graphs. In: Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int'l Joint Conf. on Natural Language Processing (Vol.1: Long Papers). 2021. 4732−4743.

[9] Li ZX, Jin XL, Li W, Guan SP, Guo JF, Shen HW, Wang YZ, Cheng XQ. Temporal knowledge graph reasoning based on evolutional representation learning. In: Proc. of the 44th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. 2021. 408−417.

[10] Liu YS, Ma YP, Hildebrandt M, Joblin M, Tresp V. Tlogic: Temporal logical rules for explainable link forecasting on temporal knowledge graphs. In: Proc. of the AAAI Conf. on Artificial Intelligence. 2022, 36(4): 4120−4127.

[11] Sun HH, Zhong JL, Ma YP, Han Z, He K. Timetraveler: Reinforcement learning for temporal knowledge graph forecasting. In: Proc. of the 2021 Conf. on Empirical Methods in Natural Language Processing. 2021. 8306−8319.

[12] Zhu CC, Chen MH, Fan CJ, Cheng GQ, Zhang Y. Learning from history: Modeling temporal knowledge graphs with sequential copy-generation networks. In: Proc. of the AAAI Conf. on Artificial Intelligence, Vol.35. 2021. 4732−4740.

[13] Li ZX, Guan SP, Jin XL, Peng WH, Lyu YJ, Zhu Y, Bai L, Li W, Guo JF, Cheng XQ. Complex evolutional pattern learning for temporal knowledge graph reasoning. In: Proc. of the 60th Annual Meeting of the Association for Computational Linguistics (Vol.2: Short Papers). 2022. 290−296.

[14] Liu KZ, Zhao F, Chen HX, Li YC, Xu GD, Jin H. DA-Net: Distributed attention network for temporal knowledge graph reasoning. In: Proc. of the 31st ACM Int'l Conf. on Information & Knowledge Management. 2022. 1289−1298.

[15] Bordes A, Usunier N, Garcia-Duran A, Weston J, Yakhnenko O. Translating embeddings for modeling multi-relational data. In: Advances in Neural Information Processing Systems. 2013. 26.

[16] Wang Z, Zhang JW, Feng JL, Chen Z. Knowledge graph embedding by translating on hyperplanes. In: Proc. of the AAAI Conf. on Artificial Intelligence, Vol.28. 2014.

[17] Lin Y, Liu Z, Sun M, Liu Y, Zhu X. Learning entity and relation embeddings for knowledge graph completion. In: Proc. of the 29th AAAI Conf. on Artificial Intelligence. 2015.

© 中国科学院软件研究所 http://www.jos.org.cn

[18] Ji G, He S, Xu L, Liu K, Zhao J. Knowledge graph embedding via dynamic mapping matrix. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing (Vol.1: Long Papers). 2015. 687−696.

[19] Sun ZQ, Deng ZH, Nie JY, Tang J. Rotate: Knowledge graph embedding by relational rotation in complex space. In: Proc. of the Int'l Conf. on Learning Representations. 2018.

[20] Trouillon T, Welbl J, Riedel S, Gaussier É, Bouchard G. Complex embeddings for simple link prediction. In: Proc. of the Int'l Conf. on Machine Learning. 2016. 2071−2080.

[21] Yang BS, Wen-tau Yih S, He XD, Gao JF, Deng L. Embedding entities and relations for learning and inference in knowledge bases. In: Proc. of the Int'l Conf. on Learning Representations 2015. 2015.

[22] Dettmers T, Minervini P, Stenetorp P, Riedel S. Convolutional 2D knowledge graph embeddings. In: Proc. of the AAAI Conf. on Artificial Intelligence, Vol.32. 2018.

[23] Shang C, Tang Y, Huang J, Bi JB, He XD, Zhou BW. End-to-end structure-aware convolutional networks for knowledge base completion. In: Proc. of the AAAI Conf. on Artificial Intelligence, Vol.33. 2019. 3060−3067.

[24] Xiong W, Yu M, Chang S, Guo X, Wang WY. One-shot relational learning for knowledge graphs. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. 2018. 1980−1990.

[25] Chen M, Zhang W, Zhang W, Chen Q, Chen H. Meta relational learning for few-shot link prediction in knowledge graphs. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. 2019. 4217−4226.

[26] Zhang NY, Xie X, Chen X, Deng SM, Ye HB, Chen HJ. Knowledge collaborative fine-tuning for low-resource knowledge graph completion. Ruan Jian Xue Bao/Journal of Software, 2022, 33(10): 3531−3545 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6628.htm [doi: 10.13328/j.cnki.jos.006628]

[27] Chen M, Zhang W, Yao Z, Chen X., Ding M, Huang F, Chen H. Meta-learning based knowledge extrapolation for knowledge graphs in the federated setting. arXiv:2205.04692, 2022.

[28] Schlichtkrull M, Kipf TN, Bloem P, Van Den Berg R, Titov I, Welling M. Modeling relational data with graph convolutional networks. In: Proc. of the European Semantic Web Conf. 2018. 593−607.

[29] Vashishth S, Sanyal S, Nitin V, Talukdar P. Composition-based multi-relational graph convolutional networks. In: Proc. of the Int'l Conf. on Learning Representations. 2019.

[30] Corbett AT, Anderson JR. Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 1994, 4(4): 253−278.

[31] Rabiner L, Juang B. An introduction to hidden Markov models. IEEE ASSP Magazine, 1986, 3(1): 4−16.

[32] Pardos ZA, Heffernan NT. KT-IDEM: Introducing item difficulty to the knowledge tracing model. In: Proc. of the Int'l Conf. on User Modeling, Adaptation, and Personalization. Berlin, Heidelberg: Springer, 2011. 243−254.

[33] Yudelson MV, Koedinger KR, Gordon GJ. Individualized Bayesian knowledge tracing models. In: Proc. of the Int'l Conf. on Artificial Intelligence in Education. Berlin, Heidelberg: Springer, 2013. 171−180.

[34] Käser T, Klingler S, Schwing AG, Gross M. Dynamic Bayesian networks for student modeling. IEEE Trans. on Learning Technologies, 2017, 10(4): 450−462.

[35] Minn S, Yu Y, Desmarais MC, Zhu F, Vie JJ. Deep knowledge tracing and dynamic student classification for knowledge tracing. In: Proc. of the 2018 IEEE Int'l Conf. on Data Mining (ICDM). IEEE, 2018. 1182−1187.

[36] Abdelrahman G, Wang Q. Knowledge tracing with sequential key-value memory networks. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. 2019. 175−184.

[37] Zhang J, Shi X, King I, Yeung DY. Dynamic key-value memory networks for knowledge tracing. In: Proc. of the 26th Int'l Conf. on World Wide Web. 2017. 765−774.

[38] Pandey S, Karypis G. A self-attentive model for knowledge tracing. In: Proc. of the 12th Int'l Conf. on Educational Data Mining, EDM 2019. Int'l Educational Data Mining Society, 2019. 384−389.

[39] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L. Attention is all you need. In: Advances in Neural Information Processing Systems. 2017. 30.

© 中国科学院软件研究所 http://www.jos.org.cn

[40] Nakagawa H, Iwasawa Y, Matsuo Y. Graph-based knowledge tracing: modeling student proficiency using graph neural network. In: Proc. of the 2019 IEEE/WIC/ACM Int'l Conf. on Web Intelligence (WI). IEEE, 2019. 156−163.

[41] Yang Y, Shen J, Qu Y, Liu Y, Wang K, Zhu Y, Zhang W, Yu Y. GIKT: A graph-based interaction model for knowledge tracing. In: Proc. of the Joint European Conf. on Machine Learning and Knowledge Discovery in Databases. Cham: Springer, 2020. 299−315.

[42] Tong S, Liu Q, Huang W, Huang Z, Chen E, Liu C, Ma H, Wang S. Structure-based knowledge tracing: An influence propagation view. In: Proc. of the 2020 IEEE Int'l Conf. on Data Mining (ICDM). IEEE, 2020. 541−550.

[43] Mahdisoltani F, Biega J, Suchanek F. YAGO3: A knowledge base from multilingual wikipedias. In: Proc. of the 7th Biennial Conf. on Innovative Data Systems Research. 2014.

[44] Seo YJ, Defferrard M, Vandergheynst P, Bresson X. Structured sequence modeling with graph convolutional recurrent networks. In: Proc. of the Int'l Conf. on Neural Information Processing. 2018. 362−373.

附中文参考文献:

[1] 王飞, 刘井平, 刘斌, 钱铁云, 肖仰华, 彭智勇. 代码知识图谱构建及智能化软件开发方法研究. 软件学报, 2020, 31(1): 47−66. http://www.jos.org.cn/1000-9825/5893.htm [doi: 10.13328/j.cnki.jos.005893]

[2] 刘宝珠, 王鑫, 柳鹏凯, 李思卓, 张小旺, 杨雅君. KGDB: 统一模型和语言的知识图谱数据库管理系统. 软件学报, 2021, 32(3): 781−804. http://www.jos.org.cn/1000-9825/6181.htm [doi: 10.13328/j.cnki.jos.006181]

[26] 张宁豫, 谢辛, 陈想, 邓淑敏, 叶宏彬, 陈华钧. 基于知识协同微调的低资源知识图谱补全方法. 软件学报, 2022, 33(10): 3531−3545. http://www.jos.org.cn/1000-9825/6628.htm [doi: 10.13328/j.cnki.jos.006628]

刘康正(1999－), 男, 博士生, 主要研究领域为时序知识图谱推理, 事件预测.

赵峰(1976－), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为知识图谱, 信息检索, 数据挖掘, 自然语言处理.

金海(1966－), 男, 博士, 教授, 博士生导师, IEEE Fellow, CCF 会士, 主要研究领域为计算机体系结构, 计算系统虚拟化, 集群计算和云计算, 网络安全, 大数据.

© 中国科学院软件研究所 http://www.jos.org.cn