

异构 YANG 模型驱动的网络领域知识图谱构建

董永强^{1,3} 王 鑫¹ 刘永博¹ 杨 望^{2,3}

¹(东南大学计算机科学与工程学院 南京 211189)

²(东南大学网络空间安全学院 南京 211189)

³(计算机网络和信息集成教育部重点实验室(东南大学) 南京 211189)

(dongyq@seu.edu.cn)

Building Network Domain Knowledge Graph from Heterogeneous YANG Models

Dong Yongqiang^{1,3}, Wang Xin¹, Liu Yongbo¹, and Yang Wang^{2,3}

¹(School of Computer Science and Engineering, Southeast University, Nanjing 211189)

²(School of Cyber Science and Engineering, Southeast University, Nanjing 211189)

³(Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing 211189)

Abstract With the continuous expansion of network scale, network management and operation face great challenges of complexity and heterogeneity. The existing intelligent network operation approaches lack a unified data model at the knowledge level to guide the process of network big data. As a data modeling language, YANG has been used to model the configuration and state data transmitted by NETCONF protocol. This paper proposes an intelligent network operation scheme which builds network domain knowledge graph from heterogeneous YANG models. As per YANG language specification, the scheme proposes the basic principles of network domain ontology construction, forming an ontology structure containing 51 classes and more than 70 properties. Then, the heterogeneous YANG models from different standardization organizations and vendors are extracted and instantiated into network domain knowledge graph. Entity alignment methods are therein employed to explore the semantic co-reference relationships among uni-source YANG models. The acquired knowledge graph provides a unified semantic framework to organize massive network operation data, which thus eliminates the requirement to construct AIOps ontology manually. As such, the configuration management and operational maintenance of networks could be greatly simplified, enlightening new solutions for network performance optimization and anomaly detection problems.

Key words heterogeneous network; intelligent network operation; ontology; knowledge graph; YANG models

摘要 随着网络规模持续扩大,复杂且异构的网络环境给网络的自动化配置管理带来了严峻的挑战,现有的网络智能化运维方案缺少知识层面的统一数据模型,难以有效进行网络大数据处理.YANG 作为一种数据建模语言,用于对网络配置管理协议 NETCONF 传输的配置与状态数据进行建模.提出一种 YANG 模型驱动的网络领域知识图谱构建方案,该方案基于 YANG 语言规范,提出网络知识本体

收稿日期:2019-12-18;修回日期:2020-02-09

基金项目:国家自然科学基金项目(61971131);国家重点研发计划项目(2018YFB1800205)

This work was supported by the National Natural Science Foundation of China (61971131) and the National Key Research and Development Program of China (2018YFB1800205).

通信作者:王鑫(wxin@seu.edu.cn)

构建的基本原则,形成包含 51 个类、70 余种属性的本体结构;随后对来自不同标准化组织和厂商的异构 YANG 模型,进行数据抽取和实例化生成单源知识图谱,进而利用 YANG 模型之间存在的异构共指特性,采用实体对齐方法建立模型间的语义映射关系,形成网络领域知识图谱.该知识图谱可为网络运维大数据的生成与维护提供统一的语义框架,无须再进行手工的运维本体构建,从而极大地简化网络的配置管理与运行维护,为网络性能优化和异常检测等运维难题提供新的解决思路.

关键词 异构网络;智能网络运维;本体;知识图谱;YANG 模型

中图法分类号 TP393

随着网络技术的蓬勃发展,多元化终端类型及其接入方式的演变,网络中的被管对象和设备的种类、数量呈几何式增长,网络规模及其设备的异构性和复杂性也在不断增加,基于命令行接口(CLI)以及简单网络管理协议(SNMP)的网络运维管理方式变得非常困难.在此背景下,迫切需要高效、智能的方法来实现网络的自动化运维管理.

为了满足网络配置管理的要求,IETF 于 2006 年推出网络配置管理协议 NETCONF.该协议从概念上将消息传输划分为内容层、操作层、远程调用层和传输层,同时采用 XML 进行管理者和代理之间的数据传输和模块表达.其内容层采用 IETF 定制的标准 YANG 语言^[1]进行数据建模,这是一种建立网络数据模型的标准解决方法.基于 YANG 语言生成的 YANG 模型定义了一个数据层次结构,包括配置数据、状态数据、远程过程调用(RPCs)和通知,具有很强的可读性、理解和扩展性.NETCONF 和 YANG 的出现,极大地降低了人工运维的成本,网络设备所需的配置或网管人员所需的数据可以通过脚本化方式执行,网络运维更加趋于自动化.

然而,由于业务实现、市场需求等原因,不同厂商、标准化组织在基于 YANG 语言进行数据建模时,生成的 YANG 模型存在很大的异构性,通常在描述设备的同一种功能或概念时定义的结构(如节点名称、节点组织结构等)各不相同.异构的网络数据模型给网络运维互操作带来了挑战,网管人员必须了解异构的 YANG 模型细节才能进行信息交互,人力要求和学习成本较高,不利于智能化网络运维的推进.

Clark 等人早在 2003 年就提出网络“知识平面”的概念^[2],建议在网络数据平面之上建立一个协调统一的知识平面,通过知识平面可以自动从低级行为中提取高级目标,实现网络的自适应、自修复.文献[3]提出将网络设备、链路及其相关的网络状态信息作为知识存在知识库中,集中管理网络中产生

的运维大数据,实现网络智能化运维.但是该文献仅涉及知识存储,忽略了语义层面的信息,底层网络设备及其数据还是处于异构状态;且该文献提出的知识库是根据实验网络环境定制的,无法提供适应所有网络的统一全局视图.可见,在当今异构的网络环境下,亟需一个知识层面、与网络环境无关的统一数据模型来统领网络大数据处理,从根本上解决网络的异构描述问题.

鉴于此,本文提出一种由数据模型驱动的网络领域知识图谱构建方法,以消除异构 YANG 模型的互操作问题,为网络运维管理提供知识层面的统一描述.据我们所知,这是业界首次提出利用 YANG 语言本身生成网络知识本体,再基于异构 YANG 模型构建网络领域知识图谱.本文的主要贡献有 4 个方面:

- 1) 提出一种由 YANG 语言及其数据模型驱动的、基于领域知识图谱的异构网络智能运维方案;
- 2) 给出基于 YANG 语言规范生成网络知识本体的基本原则和构建方法;
- 3) 设计利用异构 YANG 模型进行知识抽取,并结合表示学习等手段构建网络领域知识图谱的方法;
- 4) 给出从网络领域知识图谱自动生成网络运维本体并开展智能化运维实践的方法.

1 相关工作

随着网络和人工智能技术的发展,近年来有不少工作将深度学习、强化学习等用于网络智能化运维,实现网络性能优化和故障异常检测等.文献[4]使用了深度强化学习对数据中心的长流和短流分别进行调度管理和队列阈值优化,实现对数据中心规模的流量调度的优化.文献[5]在网络效用最优化问题中使用了深度强化学习,采用了改进的策略和经验重放方法,避免了对通信网络模型的依赖.这些

工作强调机器学习手段在网络智能化运维中的应用,有意无意地忽略了网络的异构性。它们假设从网络中获取的运行和状态数据都可以互相理解,并将其直接存储,然后通过机器学习等方法进行处理和分析。这些方法较少涉及网络数据模型在网络运维管理中的运用,从而难以建立对网络中异构数据进行统一描述的全局视图。近年来,语义技术和知识图谱的发展,引发了人们对于建立统一全局视图的思考和研究。

文献[6]较早提出一种基于本体的 IP 网络设备智能配置管理模型,利用管理信息库(MIB),使用本体建模语言(OWL)集成配置管理信息,使用语义网规则语言(SWRL)定义行为信息,使用 OWL-S 定义服务;当达到触发条件时,管理器会根据 SWRL 定义的规则调用相关服务,然后基于 OWL-S 将服务执行到托管设备,实现智能管理配置。文献[7-8]提出从网络设备的配置手册中提取语法和语义信息同时利用专家知识,用 OWL 表示交换机/路由器配置域的关键概念和关系,基于语义简化和自动化设备的配置过程。

文献[3]提出一个知识图框架 KnowNet,该框架中的自定义知识来自实时的网络环境,并设计了专门为网络管理定制的原语,以简化网络运维工作。但是 KnowNet 只停留在知识存储和数据层面,没有考虑使用 OWL 等语义信息技术,也忽略了网络的异构性和网络底层运维数据模型的作用;该方案建立的知识库只适用于具体的网络,网络环境改变时需要重建知识库,可扩展性和可移植性较差。

随着 NETCONF 的普及应用,相关的智能化运维研究逐渐涌现。文献[9]提出基于本体、利用 NETCONF 解决物联网(IoT)环境下的互操作问题,该文作者将 IoT 设备的功能以本体的方式进行表述,并利用本体知识将其转化为 YANG 模型,通过 RPC 调用和响应实现对异构 IoT 设备的管理与控制。文献[10]同样基于 NETCONF 协议,利用本体解决电子医疗系统中的互操作问题。然而上述方法,只是借用了 YANG 的语法格式和模型特点,构建了特定环境下的物联网设备的本体模型,并不具有通用性。此外,文献[11]提出了一个融合不同厂商设备之间相同配置数据的本体;然而该文献只针对 Cisco 和 Juniper 两个关于 IPv6 配置的 YANG 模型,通过人工解析的方式构建网络本体,如此生成的本体针对性过强、可扩展性差,无法拓展到其他模型和网络环境。

综上所述,现有网络智能化运维技术忽视网络的异构性,导致从网络中获取的状态和配置数据彼此无法理解;本体均为手工生成且可扩展性太差、自动化程度过低;当今网络领域还没有一个能够在知识层面上对网络数据模型进行统一描述、可用于统领网络大数据处理的全局视图。为此,本文尝试利用 YANG 语言规范,构建首个数据模型驱动的网络领域知识图谱,作为知识层面的统一描述,解决网络大数据的异构性问题,促进网络运维管理的智能化。

2 基于网络领域知识图谱的异构网络智能运维方案

知识图谱是一种用图模型来描述知识的技术方法,由节点(实体)和边(属性)组成,包含描述抽象知识的本体层和描述具体事实的实例层。在语义技术框架中,知识图谱定义了对象之间的关系,简化了从异构数据中搜索信息的过程;可以促进知识共享和发现,能够在大型异构数据集上表达和发现潜在的关系^[12]。OWL 是 W3C 定义的用于描述本体的标准语言,采用 subject-predicate-object(SPO)三元组的形式表示对象之间的关系;SPARQL 查询语言则提供了对知识图谱中的各种关系进行搜索的手段。

本文提出一个能够在网络领域统领大数据处理的全局视图——YANG 模型驱动的网络领域知识图谱,从语义角度描述异构 YANG 模型的信息集成,实现网络运维管理中异构信息的统一描述。基于网络领域知识图谱的异构网络智能运维方案整体架构如图 1 所示,包含概念网络层面和物理网络层面。

1) 概念网络层——YANG 模型驱动的网络领域知识图谱构建

概念网络层对应通用网络的知识平面,负责构建基于 YANG 模型的网络领域知识图谱;该图谱不受具体网络环境的影响,也不受网络设备异构的影响。步骤有 3 个:

- ① 基于 YANG 语言的 RFC7950 规范构建通用网络知识本体;
- ② 对单源异构 YANG 模型(由不同厂商或标准化组织提供)进行知识抽取,实例化网络知识本体;
- ③ 利用 OWL 语义信息,根据已有的专家知识及节点映射关系进行训练,得出潜在的实例匹配关系,向实例化后的网络知识本体中添加相关映射规则,构建网络领域知识图谱。

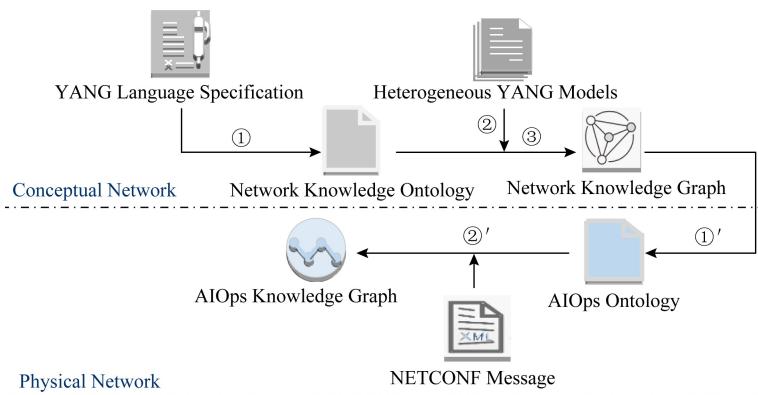


Fig. 1 Framework of heterogeneous network intelligent operation based on network domain knowledge graph

图 1 基于网络领域知识图谱的异构网络智能运维方案

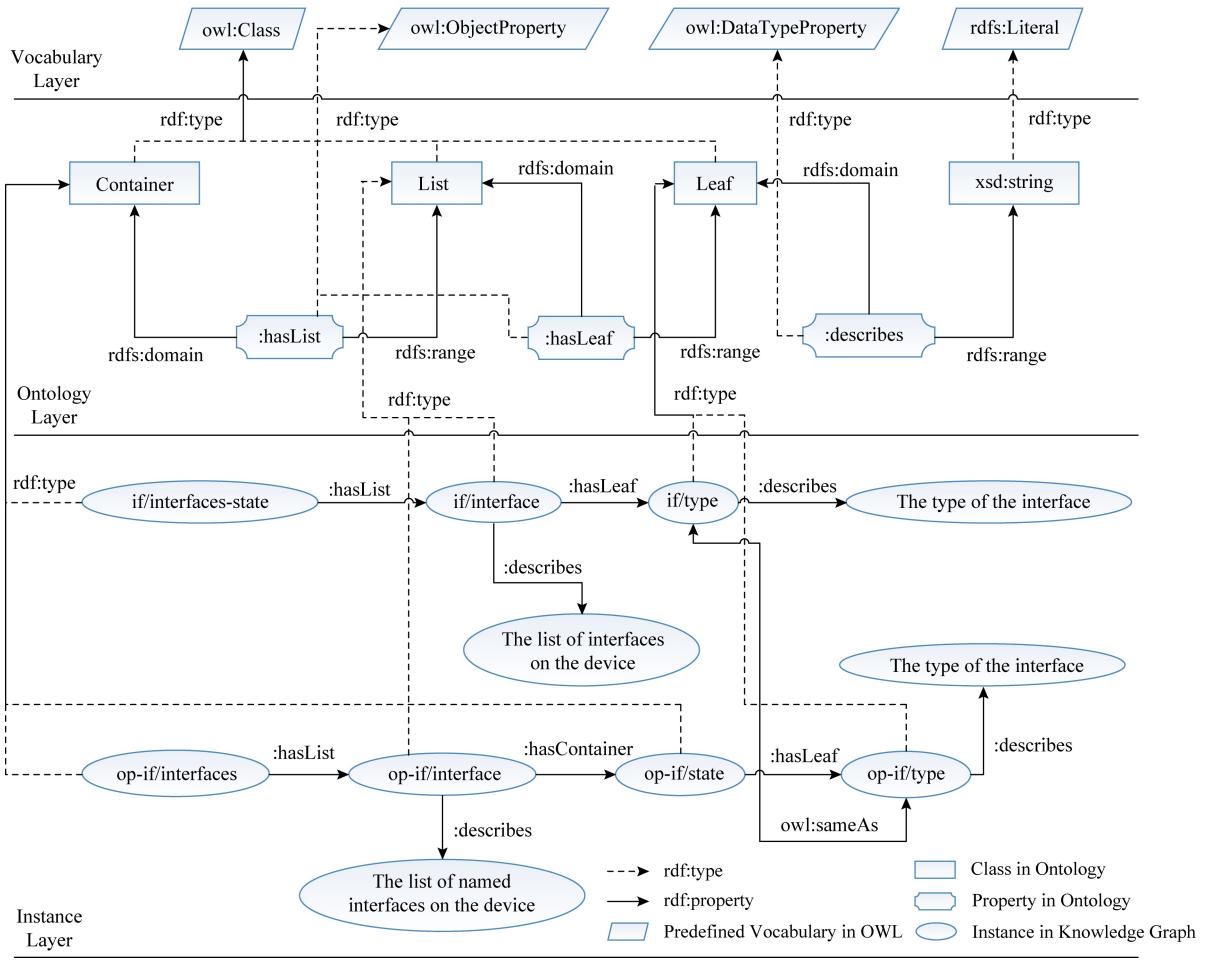


Fig. 2 Demonstration of network domain knowledge graph

图 2 网络领域知识图谱示意

构建的网络领域知识图谱如图 2 所示, 具体方法见第 3 节。

2) 物理网络层——基于领域知识图谱的网络运维本体生成与智能运维应用

物理网络层对应具体网络的控制与管理平面, 体现网络领域知识图谱在实际网络中的应用. 步骤有 2 个:

①' 基于网络领域知识图谱, 定义概念网络层

与物理网络层之间的映射规则,自动生成网络运维本体;

②' 利用 NETCONF 消息或其他动态数据,实例化网络运维本体,构建网络运维知识图谱,支撑性能优化、故障检测等运维需求.

3 YANG 模型驱动的网络领域知识图谱构建

3.1 基于 YANG 语言规范构建网络知识本体

YANG 语言是用来描述 YANG 模型的数据建模语言,由 IETF 定制,按照 NETCONF 协议的层次结构来描述相关内容,包括配置数据、状态数据、RPC 和通知;YANG 语言将网络配置与状态数据转化为树形结构,可以表示具有复杂层次关系的数据.RFC7950 规范了 YANG 语言用于构建数据模型的 68 种声明,部分声明存在自己的子声明,声明与声明之间有互相嵌套的关系,RFC7950 还对每个声明进行了基数限制,如表 1、表 2 所示(部分声明).

Table 1 Substatements of module Statement

表 1 module 声明的部分子声明

substatement	cardinality
import	0..n
container	0..n
:	:
prefix	1

Table 2 Substatements of container Statement

表 2 container 声明的部分子声明

substatement	cardinality
description	0..1
container	0..n
:	:
leaf	0..n

为了更好地展现 YANG RFC7950 中复杂的关系,同时丰富 YANG 的语义信息,我们以 OWL 为本体建模语言,利用 protégé 本体构建工具进行网络知识本体建模.构建过程遵循 6 个原则,其中,小写字母开头的单词表示 RFC7950 中定义的声明名,如 module 等;大写字母开头的单词表示网络知识本体中的类名,如 Module 等;大驼峰式命名的单词表示网络知识本体中的类名,如 BelongsTo 等;小驼峰式命名的单词表示网络知识本体中的关系名,如 hasModule 等.

1) 将 RFC7950 中具有子声明的所有声明定义为本体中的类,类之间的对象属性关系命名为关系本身的含义,如表 1 展示的有关 module 和 container 声明的三元组关系为

Module \rightarrow hasContainer \rightarrow Container

2) 将 RFC7950 中没有子声明的所有声明转为数据类型属性关系的命名,如表 1 中定义的 module 和 prefix 声明之间的关系:

Module \rightarrow hasPrefix \rightarrow xsd:string

3) RFC7950 中定义了一些特殊的用于引用其他声明的声明,如 uses, 该声明用于引用某个已经定义了的 grouping 信息.对于这类声明,对象属性关系除了既定声明之间的关系,还需要加上被引用的关系.如 module, grouping, uses 声明之间的关系定义为:

Module \rightarrow hasGrouping \rightarrow Grouping

Module \rightarrow hasUses \rightarrow Uses

Uses \rightarrow toUse \rightarrow Grouping

其中 uses, grouping 是 module 的子声明,而 grouping 不是 uses 的子声明.也就是说,除了考虑子声明的关系,还需要考虑隐性的引用声明关系.类似的处理方式还包括 6 个声明:import, augment, include, refine, belongs-to, type.

4) OWL 提供了 2 种描述枚举的方法:枚举类(通过直接枚举类的成员来描述类)和枚举数据类型(定义属性的值域)^[13].可以用枚举类来表示 RFC7950 中定义的具有枚举性质的声明,其他任何个体都不能被声明属于这个类.此外,可以用枚举数据类型定义 RFC7950 中一些具有确定值的声明,比如 status 声明的值是“current”、“deprecated”、“obsolete”之一,可以用 oneOf 结构表示这样的声明值.owl:oneOf 的三元组中的 subject 是一个 owl:DataRange 的空白节点,object 是值域构成的字面量列表;其他任何数据类型值都不属于这个枚举数据类型.

5) RFC7950 在定义声明的同时,会给声明设置基数限制,在本体建模时,需要利用 OWL 的基数限制原则,利用 min, exactly, max 等词语进行语义上的约束.如表 1 中的 prefix 声明只能有一个,可以这样描述该关系:

Module hasPrefix exactly 1 xsd:string

6) 需要遵循总的本体构建原则:本体模型中位置最“高”的类描述个体共有的知识,即子类继承父类所有属性,因此属性应该尽可能定义在通用类中.

根据这一原则,可以为 RFC7950 中定义的具有公共子声明的声明建立通用父类,在父类上添加具体属性关系及基数限制,让子类自动继承父类,从而简化本体结构.通用类构建示例如图 3 所示,其中连接线上的数字代表 2 个关系之间的基数限制.

基于上述 6 个原则,根据 YANG RFC7950 构建的网络知识本体一共包括 51 个类、44 个对象属性和 28 个数据类型属性,如图 2 和图 4 所示.图 2 本体层部分为构建的网络知识本体(局部,且省略命名空间);图 4 为利用 protégé 工具构建本体后的可视化关系图,图 4 中方框代表本体层的所有类,所有

实线代表“has subclass”(箭头指向子类),其他的虚线代表属性.

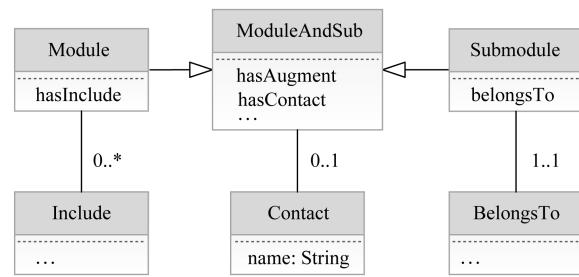


Fig. 3 Generic class example

图 3 通用类示例

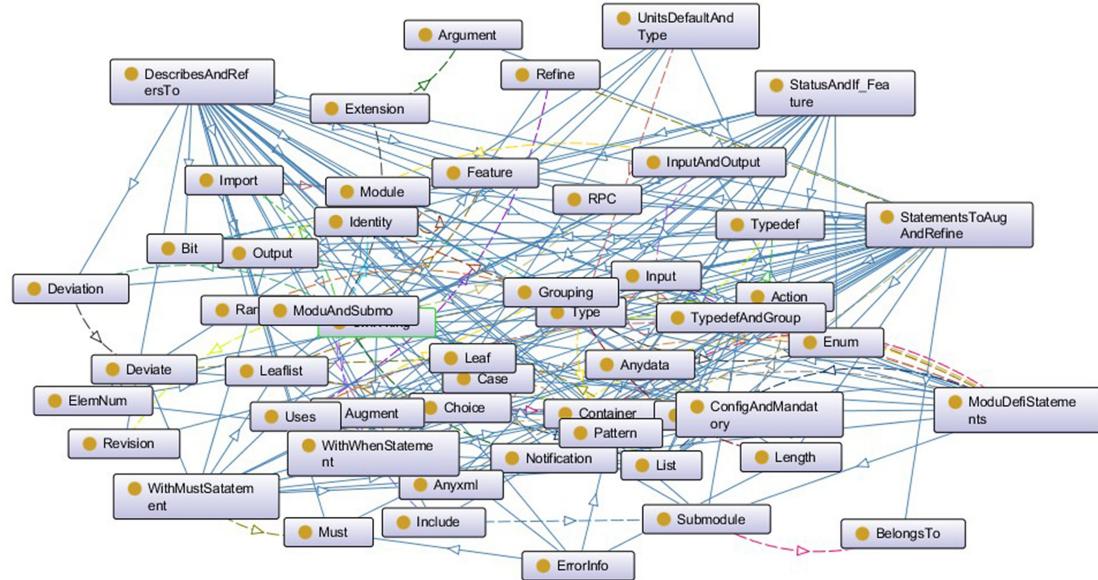


Fig. 4 Visualization of network knowledge ontology

图 4 网络知识本体可视化

3.2 基于 YANG 模型构建单源知识图谱

根据实际业务需求,即使采用统一的 YANG 语言规范,不同厂商/标准化组织定义的 YANG 模型也各不相同,导致了 YANG 模型的异构性.因此,在构建网络领域知识图谱之前,首先要根据各厂商/标准化组织提供的 YANG 模型,对网络知识本体进行实例化操作得到单源知识图谱.

单源知识图谱构建过程如图 5 和算法 1 所示.

YANG 模型通过树形结构来描述配置和状态数据的层次结构以及各个变量的数据类型,直接解析 YANG 模型文件存在一定困难;而业界针对 XML 的解析和处理工具比较成熟,因此可以首先借助 Pyang^[14] 等工具将 YANG 无损地转化为 Yin 格式文件(XML 格式),然后制定解析规则,对其进行相应 XML 解析;最后把解析结果用于网络知识本体实例化,就能得到单源知识图谱.

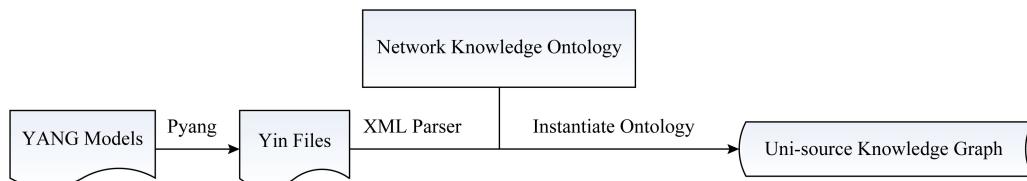


Fig. 5 Construction of knowledge graph based on uni-source YANG models

图 5 单源 YANG 模型知识图谱构建

经过算法 1 生成的单源知识图谱如图 2 和图 6 所示,其中图 2 实例层部分具有同一命名前缀的实体构成单源知识图谱的实例层(局部,且省略命名空间及实例在 XML 文件中的路径前缀);图 6 是单源知识图谱的可视化结果,其中有圆形标记的方框代

表本体层的类,有菱形标记的方框代表实例,连线代表“has subclass”(类→类)或“has instance”(类→实例),左下侧方框中显示的是一个实例在知识图谱中的详细信息,包括实例的 URI、注释、对象属性断言等。

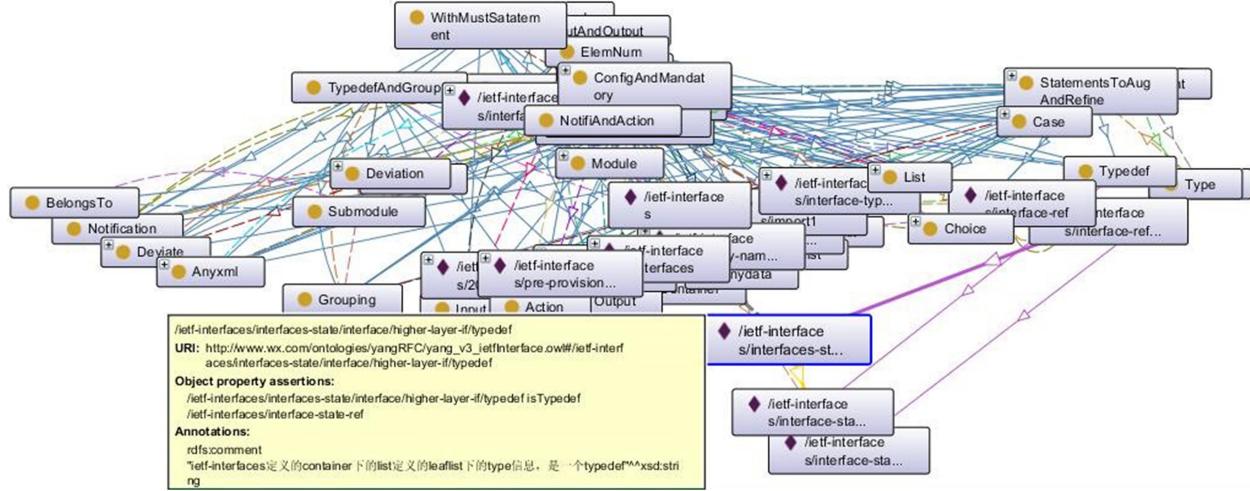


Fig. 6 Visualization of knowledge graph based on uni-source YANG models

图 6 单源 YANG 模型知识图谱可视化

算法 1. 基于 YANG 模型的单源知识图谱构建.

输入:单源 YANG 模型、网络知识本体、声明-属性映射表;

输出:单源知识图谱 $YANG_KG_U$.

- ① 借助 Pyang, 将 YANG 模型转为 Yin 文件;
- ② 拓扑排序, 得到按 Yin 文件的依赖嵌套关系排序的集合 $classSet$;
- ③ while $classSet \neq \emptyset$ do
- ④ for $i=0$ to $classSet.size()$ do
- ⑤ 找到集合中 XML 标签与声明-属性映射表中对应的映射元素;
- ⑥ 使用 Jena^[15] 定义的方法构建映射元素的实例与关系;
- ⑦ end for
- ⑧ end while
- ⑨ 保存 $YANG_KG_U$ 并返回.

3.3 基于单源知识图谱构建网络领域知识图谱

网络知识本体基于统一的 YANG 语言规范构建,因此 4.2 节构建的单源知识图谱具有统一的上层本体.然而由于实际业务需求不同,单源 YANG 模型存在异构共指现象:不同名实例可能指代同一个实

体,相同名实例可能指代不同实体.以 ietf-interfaces 模型^① 和 openconfig-interfaces 模型^② 为例,两者在定义“type”时,虽然概念结构、节点层次等不一样,但共同指向“端口的类型”这一含义.异构共指现象的存在,为基于单源知识图谱构建网络领域知识图谱提供了实现基础:如果能够发现异构 YANG 模型之前的节点匹配关系,就可以利用 OWL 语义等价关系(如 owl:sameAs 等)向单源知识图谱之间添加映射规则,从而构成网络领域知识图谱.

本实验室前期工作^[16] 中给出了一个寻找 YANG 模型间匹配节点的初步方案:首先对 YANG 模型进行解析和扁平化处理得到 XML 格式的配置树,根据配置树中节点的字符串信息构建缩略词表并进行语义消歧,然后计算不同模型的结构相似度和节点相似度,最终得到节点间的匹配关系.但是由于缺乏全局统一的视角,该方案只能得到少部分节点的匹配关系.本文在这些节点匹配关系基础上,利用单源知识图谱中丰富的语义信息,采用基于表示学习的实体对齐方法^[17],自动寻找更多潜在的匹配节点对,丰富单源知识图谱间的语义映射关系,构建网络领域知识图谱.

① <https://github.com/YangModels/yang/blob/master/standard/ietf/RFC/ietf-interfaces.yang>

② <http://www.netconfcentral.org/modules/openconfig-interfaces>

如图 7 所示,首先将待匹配的单源知识图谱映射到低维空间,得到它在低维空间中的向量表示。根据翻译模型^[18],知识图谱中每个三元组($s \rightarrow p \rightarrow o$)的关系 p 都可以看作从 s 到 o 的翻译,通过不断地调整其表示向量 S, P 和 O ,使得 $S + P$ 尽可能等于 O ,即 $S + P \approx O$ 。为了减少计算量并增加随机性,模型训练需要使用负采样技术,使得对负样本三元组 $S + P \neq O$ 。然后,利用文献[16]提供的已有匹配对训

练转换矩阵,将异构知识图谱的向量表示映射到统一向量空间。最后,在该空间中利用欧氏距离或其他方式度量实体向量间的对齐程度。模型训练完成后,将其用于实例对齐预测,对齐程度越大,则表示节点实例之间的相似程度越高。按照以上过程不断迭代学习预测,找到单源知识图谱中所有潜在的等价实例关系,添加对应的映射规则,构建网络领域知识图谱。

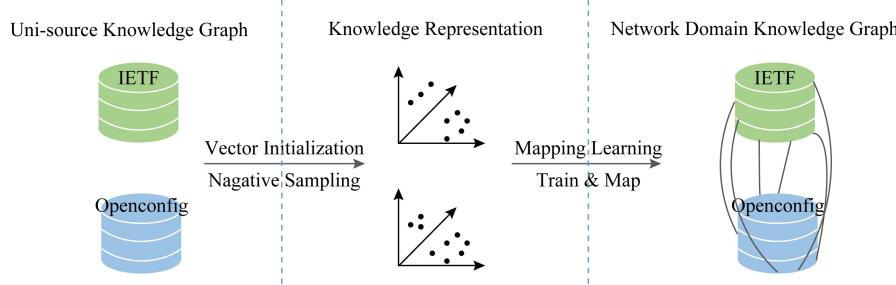


Fig. 7 Framework of network domain knowledge graph construction

图 7 网络领域知识图谱构建框架

相比文献[16],基于知识图谱表示学习的实体对齐方法将待匹配的单源知识图谱映射到同一空间,算法可以在该过程中自动获取并推理所需特征,不需要显式的规则约束;同时,向量表示蕴含了原来知识图谱内的所有节点、属性(关系)特征和结构信息,不依赖字符串的文本信息,从而获取 YANG 模型的深层次特征,并根据所有模型构建全局统一视图,能够挖掘单源知识图谱之间更多的语义匹配关系。

以 ieft-interfaces 模型中待匹配的节点:/if: interfaces-state/if: interface/ip: ipv6/ip: mtu 为例,原方案将其错误映射到 openconfig-interfaces 模型中的节点:/oc-if: interfaces/oc-if: interface/oc-if: state /oc-if: mtu。通过图 7 所示方法,则发现其正确映射节点应为:/oc-if: interfaces/oc-if: interface/oc-tun: tunnel/oc-ip: ipv6/oc-ip: state/oc-ip: mtu。

采用以上方法,挖掘并标记等价实例之后生成的网络领域知识图谱(局部),如图 2 的实例层所示,其中,标记“if/”的椭圆形代表 IETF 的 YANG 模型,标记“op-if/”的椭圆形代表 OpenConfig 的 YANG 模型。可以看到,2 个组织的异构 YANG 模型因为 owl:sameAs 等价属性建立了映射关系,从而形成联合的网络领域知识图谱。

4 基于领域知识图谱开展网络智能化运维

实际网络环境中,不断运行的网络会实时动态地产生数量庞大的数据,包含设备、链路的配置与状

态信息等。网络领域知识图谱只是概念层面网络的抽象表示,即表征了有关网络的构成要素、实体功能、协议描述等信息。本节阐述如何根据网络领域知识图谱生成适用于实际网络场景、能统领网络大数据的全局视图——网络运维本体和运维知识图谱,并分析将其用于网络智能化运维的可行途径。

4.1 基于领域知识图谱构建运维本体

分析 NETCONF 消息可知,不论在 YANG 模型中嵌套多少层,所有有关网络设备配置和状态的数据都存放在 leaf 节点或 leaf-list 节点标签中。因此,只需要将网络领域知识图谱实例中的所有 leaf 节点和 leaf-list 节点抽出,同时保留节点的数据类型信息,即可自动化生成网络运维本体。具体过程如算法 2 所示。

算法 2. 网络运维本体的自动化生成。

```

输入: 网络领域知识图谱 YANG_KG;
输出: 网络运维本体 AIOps_Ontology.

① 获取 YANG_KG 中的所有 hasLeaflist 与
    hasLeaf 关系,存入关系列表 relist;
② for node in relist do
    ③ 获取 leaf 节点列表 leafSet 及对应类型信息;
    ④ 获取 leaf 节点间的 owl:sameAs 关系;
    ⑤ end for
    ⑥ for node in leafSet do
        ⑦ 根据列表信息自动创建类与属性关系;
        ⑧ 根据 owl:sameAs 关系自动生成 owl:
            equivalentClass 关系;

```

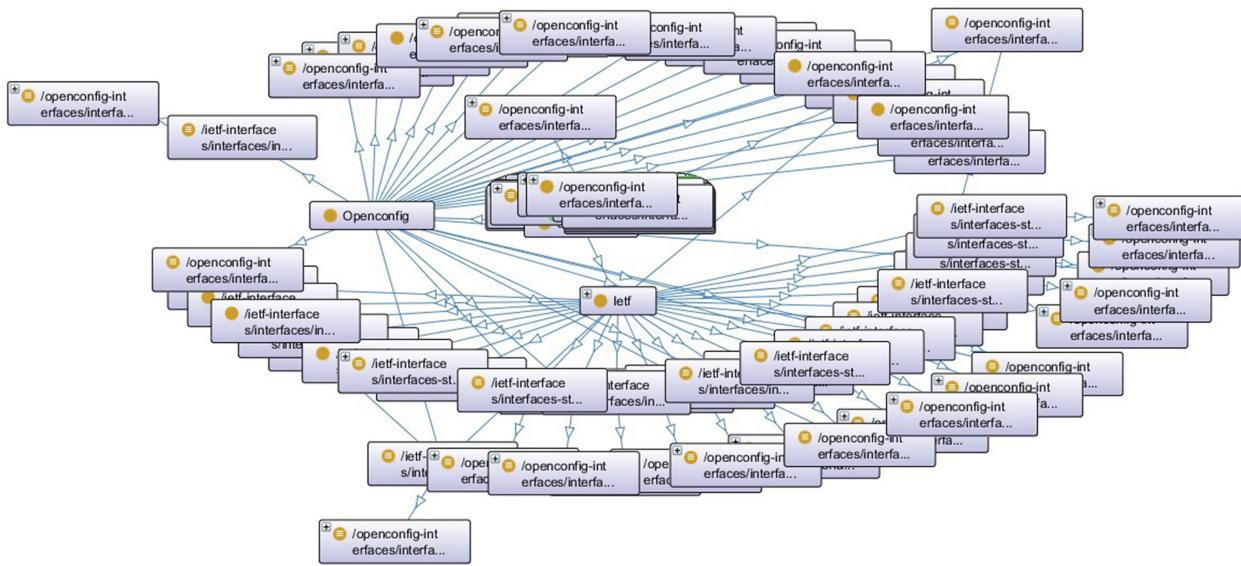


Fig. 8 AIOps Ontology

图 8 网络运维本体

⑨ end for

⑩ 保存 *AIOps_Ontology* 并返回.

基于算法 2 生成的网络运维本体如图 8 所示. 从图 8 中我们可以看出网络运维本体以 OpenConfig 和 IETF 为中心向四周散开, 2 个组织的本体类通过 owl:equivalentClass 进行关联, 为网络底层的数据模型提供了全局统一且稳定的视角. 通过该本体, 无论实际网络环境如何变化, 网络运维本体都会保持统一, 具有良好的扩展性和可迁移性.

4.2 运维知识图谱的构建及应用

NETCONF 协议通过传递 XML 消息对网络设备进行管理, 它包含了网络及其设备的所有配置与状态信息; 根据实际网络场景中的 XML 消息对网络运维本体实例化操作, 即可得到基于实际网络场景的运维知识图谱; 其构建过程与算法 1 类似, 不再赘述.

网络领域知识图谱与运维知识图谱之间存在语义映射关系, 因此网管人员不需要再关心网络的异构环境, 运维知识图谱会根据物理网络层和概念网络层之间的映射关系, 自主地将网管人员下发的网络配置管理命令与操作中的异构性信息消除, 自动将其转换成设备可理解的命令与操作, 为网络运维管理的智能化提供支撑.

在此基础上, 网络运维本体和运维知识图谱给解决网络性能优化以及异常检测等运维难题带来了新思路. 运维知识图谱不仅能存储实际网络中当前与历史的配置与状态数据, 同时能提供知识图谱独有的语义信息, 把知识与规则引入路由规划与流量

优化算法设计中, 结合机器学习框架, 根据实时的网络状态, 利用知识图谱中的断言与规则, 可以智能地完成动作空间的低维映射和调度规则的生成, 实现网络性能优化; 同时, 还可以从海量的运维数据中不断学习, 从历史信息中总结网络流量规律, 对网络行为和状态做出预判, 指导完成未来的故障检测与定位, 智能地完成异常检测.

5 总 结

本文提出一种基于 YANG 语言及其数据模型构建网络领域知识图谱的方案, 详细阐述了其构建方法与应用场景, 得到用于描述网络知识和刻画网络状态的全局视图, 从而能够统领网络大数据处理, 解决网络异构带来的运维难题, 为数据驱动的网络管理与运行优化提供一种新的视角和方法. 本文构建的网络领域知识图谱的特点主要有: 自动化程度高、能屏蔽异构因素、数据查询简洁且高效、可扩展性强. 未来将在网络领域知识图谱的完善与开放、知识图谱驱动的配置级与控制级运维等方面开展进一步研究.

参 考 文 献

- [1] Bjorklund M. RFC 7950: The YANG 1.1 Data Modeling Language [OL]. IETF, 2016 [2019-12-01]. <https://tools.ietf.org/html/rfc7950>

- [2] Clark D, Partridge C, Ramming J C, et al. A knowledge plane for the Internet [C] // Proc of the 2003 ACM SIGCOMM Conf on Applications, Technologies, Architectures, and Protocols for Computer Communication. New York: ACM, 2003: 3-10
- [3] Quinn R, Kunz J, Syed A, et al. KnowNet: Towards a knowledge plane for enterprise network management [C] // Proc of the 2016 IEEE/IFIP Network Operations & Management Symp. Piscataway, NJ: IEEE, 2016: 249-256
- [4] Chen Li, Lingys J, Chen Kai, et al. Auto: Scaling deep reinforcement learning for datacenter-scale automatic traffic optimization [C] // Proc of the 2018 Conf of the ACM Special Interest Group on Data Communication. New York: ACM, 2018: 191-205
- [5] Xu Zhiyuan, Tang Jian, Meng Jingsong, et al. Experience-driven networking: A deep reinforcement learning based approach [C] // Proc of IEEE Conf on Computer Communications (IEEE INFOCOM 2018). Piscataway, NJ: IEEE, 2018: 1871-1879
- [6] Xiao Debao, Xu Hui. An integration of ontology-based and policy-based network management for automation [C] // Proc of the 2006 Int Conf on Computational Intelligence for Modelling Control and Automation. Piscataway, NJ: IEEE, 2006: 21-27
- [7] Anny M, Marcelo Y, Serral-Gracià R, et al. Ontology-based information extraction from the configuration command line of network routers [M] // Mining Intelligence and Knowledge Exploration. Berlin: Springer, 2014
- [8] Martinez A, Yannuzzi M, Serral-Gracià R, et al. An ontology-based information extraction system for bridging the configuration gap in hybrid SDN environments [C] // Proc of 2015 IFIP/IEEE Int Symp on Integrated Network Management (IM). Piscataway, NJ: IEEE, 2015: 441-449
- [9] Sahlmann K, Scheffler T, Schnor B, Managing IoT device capabilities based on oneM2M ontology descriptions [C] // Proc of the 16 GI/ITG KuVS Fachgespräch Sensornetze. Hamburg, Germany: HAW Hamburg, 2017: 23-26
- [10] Jin Wenquan, Hyeon K D. Design and implementation of e-health system based on semantic sensor network using IETF YANG [J]. Sensors, 2018, 18(2): 629-648
- [11] Ismail H, Hamza H, Shaimaa M, et al. Semantic enhancement for network configuration management [C] // Proc of IEEE Global Conf on Internet of Things. Piscataway, NJ: IEEE, 2018: 1-5
- [12] Liu Qiao, Li Yang, Duan Hong, et al. Knowledge graph construction techniques [J]. Journal of Computer Research and Development, 2016, 53(3): 582-600 (in Chinese)
(刘娇, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600)
- [13] Peterson D. W3C XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes [OL]. 2012 [2020-02-07]. <https://www.w3.org/TR/xmlschema11-2/>
- [14] Bjorklund M. Pyang [OL]. [2020-02-07]. <https://github.com/mbj4668/pyang>
- [15] Apache Software Foundation. Apache Jena [OL]. [2020-02-07]. <http://jena.apache.org/>
- [16] Sun Tao. Automatic adaptation scheme of YANG models based on configuration three similarity [D]. Nanjing: Southeast University, 2018 (in Chinese)
(孙涛. 基于配置树相似度的 YANG 模型自动适配方案[D]. 南京: 东南大学, 2018)
- [17] Zhu Jizhao, Qiao Jianzhong, Lin Shukuan. Entity alignment algorithm for knowledge graph of representation learning [J]. Journal of Northeastern University: Natural Science, 2018, 39(11): 1535-1539 (in Chinese)
(朱继召, 乔建忠, 林树宽. 表示学习知识图谱的实体对齐算法 [J]. 东北大学学报: 自然科学版, 2018, 39(11): 1535-1539)
- [18] Liu Zhiyuan, Sun Maosong, Lin Yankai, et al. Knowledge representation learning: A review [J]. Journal of Computer Research and Development, 2016, 53(2): 247-261 (in Chinese)
(刘知远, 孙茂松, 林衍凯, 等. 知识表示学习研究进展[J]. 计算机研究与发展, 2016, 53(2): 247-261)



Dong Yongqiang, born in 1973. PhD, associate professor. Member of CCF. His main research interests include network architecture, security and management, wireless networks and mobile computing.



Wang Xin, born in 1994, Master candidate. Her main research interests include knowledge graph and intelligent network operations.



Liu Yongbo, born in 1994. PhD candidate. His main research interests include AIops and deterministic networking.



Yang Wang, born in 1979. PhD, assistant professor. Member of CCF. His main research interests include network security monitoring, botnet detection and security event response.