

一种基于图形数据库的快速电力网络拓扑分析方法^①

蒋厚明, 孙 昊, 孔 震

(国网电力科学研究院, 南京 210009)

摘 要: 提出了一种基于图形数据库的通用拓扑分析方法, 将多状态设备映射为多节点, 每个节点的父对象都为此设备, 将设备连接抽象为连接电缆节点和连接关系, 每种设备的电气属性都抽象为节点和关系的属性, 在将网络的拓扑信息入库后, 组合利用图形数据库中的多种图形分析算法, 并结合电力网络应用开发出相连节点连通性分析和停电范围检测算法, 基于此开发的电网停电范围监测系统和变电站智能开票操作票程序该方法具备高可用性和高性能, 取得了良好的效果。

关键词: 网络拓扑; 图形数据库; 节点; 属性; 海量节点

A Quick Electric Network Topology Technology Based on Graph Database

JIANG Hou-Ming, SUN Hao, KONG Zhen

(National Electricity Science Research Institute, Nanjing 210009, China)

Abstract: A general topological analysis method based on graph database, multi-state device is mapped to multiple nodes, each node has the parent property of this device, connection between the device is mapped to connection cable node and the relation between each connect device. Every device's attribute is mapped to properties of nodes and relationships in the network topology. After store these information into graphic database, combined use a variety of graphical analysis algorithm in graphic database, and author also developed electric network algorithm based on these inherited algorithms. Base on the method, monitoring systems and substation intelligent ticket operations have high availability and performance, and achieved good results.

Key words: network topology; graph database; node; property; mass node

1 引言

在电力信息系统中, 网络结构的描述和拓扑分析具有重要的地位, 甚至是很多业务应用的基石, 早期的网络拓扑分析都是基于特定的数据结构, 其普遍存在的缺点是效率低和扩展性差, 当网络中新增了某个设备或者某个电气元件具备新的拓扑特性时不能很好地自动适应, 可能需要进行大量的修改, 从系统的角度看稳定性和复用性都较弱, 不能很好地适应未来电力信息系统智能化发展的趋势, 因此如何从实际的电力系统中抽取通用的拓扑模型、如何利用此通用拓扑更好地进行快速的、智能的电力图形拓扑分析已成为业界热点。

图形数据库作为一个新兴的数据库技术, 其重点关注大数据量和数据内部之间复杂依赖的处理, 目前

在互联网社交应用中使用较多, 但作者发现其节点数量对拓扑分析速度影响很小的特性特别适合分析超多设备数量的电力网络, 据此本文提出了一种基于图形数据库的快速电力图形拓扑分析方法, 其构建了一套可用于电力拓扑的通用描述模型, 底层基于图形数据库进行构建, 前端封装出多种电力行业的图形拓扑算法, 应用程序无须关心后端具体实现, 很好地实现了效率和易用性的结合, 其在电力调度管理系统、电力生产管理操作票系统中都得到了很好的应用。

2 传统的电力网络拓扑分析方法

电力网络拓扑分析的方法很多, 但这些方法本质都是先用图的形式来表达电力系统模型^[1], 然后利用图的

^① 收稿时间:2012-08-19;收到修改稿时间:2012-09-17

拓扑算法来构建整个系统的拓扑关系,所有对于电力网络的拓扑分析都转向对于图形的拓扑分析,因此在进行图形拓扑分析之前需要将电力网络转换为图形,通常的方法是电力网络转化为一个有向图来描述: $G=(V,E(S))$. 其中 V 为节点集合, E 为有向边集合, S 为边的状态(断开与否). 有向图中的每个节点用来表示电力系统中的一个单元实体,例如其节点可表示公司、厂站、线路、母线、开关、刀闸、发电机、电容等等,每个节点都有自己的状态,其状态可决定节点的连通关系.

在将电力网络转换为图形后,即可利用图形拓扑算法来构建拓扑关系^[2],图形拓扑算法很多,也较为成熟,随着研究的深入,现已出现了结合设备物理特性的新型拓扑算法,这些都已经超出了了图形拓扑本身的范畴,下面对两种主流的拓扑算法做一个简要介绍.

2.1 深度或广度搜索法

此种网络拓扑分析技术是利用堆栈技术进行搜索.一般是将拓扑结构表述为链表关系,用图论中的搜索技术,如深度优先搜索法和广度优先搜索法分析节点的连通性.这种方法一般需要建立反映拓扑结构的链表,通过处理链表实现拓扑分析,然后以搜索回溯的框架,利用堆栈记录划分.由于其基本算法采用“堆栈”原理-先进后出的搜索逻辑^[3],程序不可避免采用递归的实现形式,因此编程和维护较复杂,效率较低,应用于实时网络分析时,在运算时间上不能满足要求.

2.2 面向对象(OO)的启发式搜索算法

在电网的实际运行过程中,状态频繁发生变化的开关占少数,因此将追踪技术引入拓扑分析中,仅在开关状态发生改变时进行局部拓扑分析^[4],可以减少拓扑分析的计算量.在完成网络的初始拓扑分析并构筑了电网的结点树之后,当电网发生开关变位事件时,根据开关变位只造成局部电网拓扑发生变化的特点,采用启发式搜索算法进行电网结点树拓扑的跟踪.针对不同的变位事件,分开关“开”和“合”两种情况进行分析.实现拓扑跟踪 OO 模型的启发式拓扑分析方法,利用 OO 技术可扩展拓扑算法的适用范围.

小结:传统的电力网络拓扑分析方法无论是哪种,其算法复杂度都为 $O(n^2)$,随着设备数量的增加,其性能将会极大的下降,这对于现代电网拓扑分析来说,其最看重的实时性将会大打折扣^[5].

3 图形数据库

图形数据库作为 NoSQL 数据库的一种, NoSQL 数

据库处理的是规模化、速度和灵活性等问题,这些通常都是互联网应用所关注的,图形数据库在具备这些特性的同时还具备一个特长:找出不同数据之间存在的关系,而且寻找的速度与节点数量没有关系.其最典型的案例是图谱分析,即找出图形节点和图形节点之间的关系.图形数据库能够分析 TB 级的图形数据,与此同时还能够维护数据之间的关系,即便数据和关系不断变化也能如此.

图形数据库的基本概念如下:

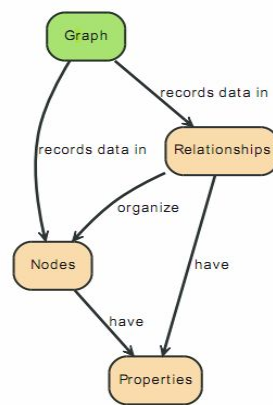


图 1 图形数据库组成

图形数据库主要由三部分组成:

- ① 节点(Node):从概念上来说,这类似于对象实例,拥有唯一的 ID.
- ② 关系(Relation):它连接了两个节点,此外还有方向和关系类型.
- ③ 属性(Property):它们是字符串类型的键值对 (Key/Value)值对,节点与关系都有属性.

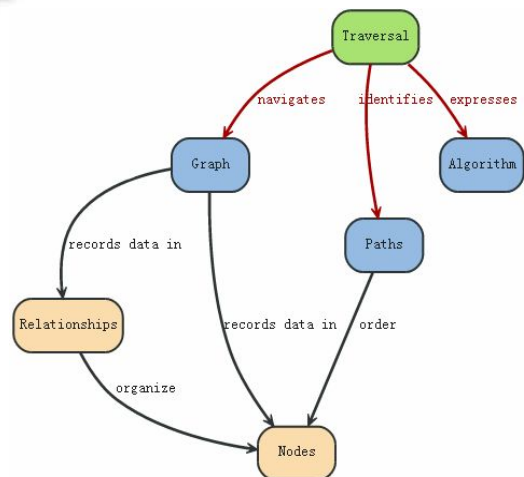


图 2 图形数据库查询过程

一次遍历(Traversal), 可以理解为通过一种算法从一些开始节点开始查询与其关联的节点, 比如想找到“相连的设备的下一个接地设备是什么?”, 又或者“如果断电了, 哪些设备会受到影响?”, 可通过 Traversal 查询获得结果.

4 基于图形数据库的网络拓扑分析方法

由于图形数据库能够存储海量节点, 并且节点的数量对于拓扑分析只有很轻微的影响, 此种特性决定了其非常适合电力网络拓扑分析, 理论上能够将电力网络分析的性能指标提高几个数量级, 下面将介绍如何利用图形数据库进行电力网络拓扑分析.

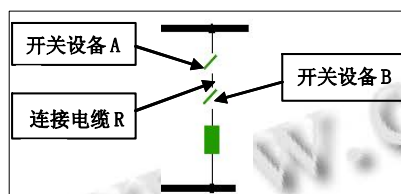
按照国际通用的 CIM 模型, 电网拓扑分析即是厂站内的开关组合成节点集合, 并根据网络中的支路(包括线路、母线、c 变压器等)的连接关系将电气网络拓扑图描述为一个点和关系组成的集合, 每个点和关系都有自己的属性, 具体描述规则如下:

① 点: 表示一个实体;例如开关等电气设备.

② 边: 表示关系, 点和点之间的连接关系, 通常是有方向性的. 例如某个电气设备 A 和电气设备 B 之间是单向连通的, 则表示为 A->B; 如果 A 和 B 之间是双向连通的, 则表示为 A<->B.

③ 属性: 表示点和边具备的属性. 例如设备的名称、电气特性、配置信息, 这些属性可动态设置, 不受限制.

下面以一个实例说明如何描述设备之间的关系:



上图为一个变电站中常用的母线之间的连通图, 其中开关 A 为开合式开关, 有两个状态(开、合), 为方便拓扑分析, 将 A 拆分为两个节点 A0、A1, 其属性[parent]都为 A, A0 和其中 A 一端的设备相连, A1 则和另一端的设备相连, 这些相邻关系都表示为关系存储在图形数据库中, 当 A 为闭合状态时, 只需将 A0 和 A1 连通, 如果处于断开状态, 则 A0 不和 A1 连通.

开关设备 B 拆分为 B0、B1, 如果为断开状态, 则 R0 关系连接属性, R0 其连接节点为 B1、A0, 当 A 处于连通状态时, R0 的连通属性也设为连通.

上图三个设备之间的拓扑关系可表述如下:

节点名称	所属设备	开合状态	连接电缆
A0	开关 A	开	开关 A 另一端电缆 (未列出)
A1	开关 A	开	R
B0	开关 B	开	R
B1	开关 B	开	开关 B 另一端连接电缆 (未列出)
R0	连接电缆 R	合	

以此类推, 每个设备有几种设备状态即可拆分为多个节点, 三相设备即有三个节点, 然后将每种设备与外界连接的设备表述为关系即可, 用此种描述规则可将网络中所有的设备和连接关系进行表示, 所有的设备描述信息都变为节点和关系的属性存储到图形数据库中.

当将网络设备转变为节点和关系存储到图形数据库后, 即可进行图形分析, 图形数据库自带很多图形分析算法, 其中比较有用的算法有:

两点之间最短距离算法

```
PathFinder<Path> finder =
```

```
GraphAlgoFactory.shortestPath(?Traversal.expanderForTypes( ExampleTypes.MY_TYPE,Direction.OUTGOING ), 15);
```

带权重计算的 Dijkstra 算法

```
PathFinder<WeightedPath> finder =
```

```
GraphAlgoFactory.dijkstra(Traversal.expanderForTypes( ExampleTypes.MY_TYPE, Direction.BOTH ), "cost");
```

```
WeightedPath path = finder.findSinglePath( nodeA, nodeB );
```

```
path.weight();
```

计算两点之间的最方便的 A*路径算法

```
PathFinder<WeightedPath> astar =
```

```
GraphAlgoFactory.aStar(?Traversal.expanderForAllTypes (), CommonEvaluators.doubleCostEvaluator( "length" ), estimateEvaluator );
```

```
WeightedPath path = astar.findSinglePath(nodeA, nodeB );
```

找出相连节点的连通性分析算法

```
PathFinder<WeightedPath> finder =
```

```
GraphAlgoFactory.connectedNodes(Traversal.expanderF
```

```
orAllTypes(),CommonEvaluators.doubleCostEvaluator(
"length"), estimateEvaluator);
```

```
WeightedPath path = finder.findSinglePath( nodeA,
* );
```

```
return path.Nodes;
```

通过组合这些算法可进行各种电力拓扑分析,例如找出与某个设备相连通的下一级设备,则只需指定遍历深度为 1,然后相连节点连通性分析算法即可.另外图形数据库通常都支持图形算法自定义扩展以适应特定应用场景.

5 实际验证

基于图形数据库的新型电气网络拓扑分析方法在性能方面得到了很大的提升,主要体现在对于大型网络的分析,其性能几乎不受网络节点数量的影响,下面通过实际案例进行具体性能比对.

采用Java作为开发工具,图形数据库采用Neo4J,传统的网络拓扑分析采用广度优先算法(其与深度优先的算法复杂度一样,其对于有 n 个顶点 e 条边的拓扑图的时间复杂度都是 $O(n^2)$),编制基于 CIM 的电网拓扑分析软件,以某电网为例,网络主要设备有: 102 座变电站, 453 个母线, 1109 个开关, 810 个连接节点, 205 个线路, 198 台高压变压器, 176 个负荷,网络中的其它设备类型和数量较多,不一一描述,作为单一状态设备处理.

运行机器为双核 1.7G,内存为 4G,每种算法随机选取开始设备执行 10 次,取平均运行时间和占用内存,执行操作选择最常用的筛选相连接设备和停电范围检测.

首先选择只分析到变电站,全网范围共有约 200 个变电站,所有参与分析设备大概有 2000 个,其运行结果如下:

网络拓扑算法	筛选相连接设备	停电范围监测	占用内存
基于广度优先算法	10 秒	20 秒	35M
基于图形数据库 拓扑算法	2 秒	3 秒	50M

其次分析全网设备,全网共有重要设备 1 万个,辅助设备 5 万个,其中三相开关 853 个、四相开关 12 个、二相开关大约 5000 个,其运行结果如下:

网络拓扑算法	筛选相连接设备	停电范围监测	占用内存
基于广度优先算法	620 秒	1800 秒	450M
基于图形数据库 拓扑算法	2.5 秒	3.5 秒	55M

在此,已经可以看出在小数据量时传统的广度优先遍历和图形数据拓扑算法相差不大,但在大数据量情况下,广度优先耗费时间和内存呈几何级增长,而图形数据库几乎无明显增长.

为了验证基于图形数据库的拓扑分析算法的极限,作者使用计算机虚拟了一千万个设备,其中随机产生三相设备、四项设备、二相设备各 200 万个,并虚拟了 20 万个母线以及其它常用电器设备,其运行结果如下:

网络拓扑算法	筛选相连接设备	停电范围监测	占用内存
基于广度优先算法	不可用	不可用	不可用
基于图形数据库	4 秒	5 秒	58M

从此实验可看出,图形数据库分析所花费的时间、内存占用和设备数据并非线性关系,设备数量的增长只会稍微增加查询时间,其无论是小数据量还是超大数据量都能保持很好的效率和内存占用.

6 结论

本文在分析图形数据特性的基础上提出了基于图形数据库的快速电力网络拓扑分析方法,针对电力网络具有多状态设备的特性提出了一种实用的电力设备描述规则,并且在图形数据库底层图形分析算法扩展出多种具备电力行业特色的拓扑算法,可广泛应用于停电分析、智能操作票管理系统、线损分析等高级应用软件中.

参考文献

- 1 朱文东. 电力网络局部拓扑的快速算法. 电网技术, 1996,20(3):30-33.
- 2 陈树柏,左恺,张良震. 网络图论及其应用. 北京:科学出版社,1982.
- 3 Kambale P, Machauer JJ. The dispatcher training simulator for metropolitan Edison company, IEEE Trans on Power Systems,1996,11(2),
- 4 Vehsakul PD, Dabbaghch I. A topology based algorithm for tracking network connectivity. IEEE Trans on Power Systems, February 1995(2):339-345.
- 5 董张卓,秦红霞,孙启宏,彭俊松. 采用面向对象技术和方法的电力系统网络拓扑的快速跟踪(一、二). 中国电机工程学报,1998,18(3):178-181.