

# 基于案例推理的故障诊断算法<sup>①</sup>

孔 钦, 叶长青

(南京大学金陵学院 信息科学与工程学院, 南京 210089)

**摘 要:** 电力系统智能告警研究目前主要集中于规则推理领域, 未能充分利用系统内历史故障的案例进行有效的分析与故障处理. 本文初步研究了基于案例推理的故障诊断算法, 在此基础上提出了一种基于向量计算的案例匹配算法. 该算法通过对故障信息进行抽象, 并与系统案例库中的案例进行相似度计算的方式, 对故障信息与案例信息进行一一匹配. 案例库的建立基于经过人工分析的各类历史故障信息. 相似案例的故障性质、故障原因分析可对本次故障的分析、诊断和后续处理起到指导作用.

**关键词:** 智能告警; 事故分析; 案例推理; 模式匹配

## Fault Diagnosis Algorithm with Case-Based Reasoning

KONG Qin, YE Chang-Qing

(School of Information Science and Engineering, Jinling College, Nanjing University, Nanjing 210089, China)

**Abstract:** Power system intelligent alarm is mainly concentrated on the rule reasoning domain. This method does not effectively use the analysis of historical failure to analyze and handle the accident which is currently happening in the system. In this paper, after a preliminary study on fault diagnosis algorithm with case-based reasoning, a matching algorithm is proposed based on vector calculation. In this algorithm, the fault information in the system is abstracted and the similarity to the cases within system case library is then calculated to try to match the fault information to the case information one by one. All the history cases are obtained by manual analysis of historical failure information in the system case library. The fault properties and the cause of similar cases are helpful for current failure analysis, diagnosis and exception handling.

**Key words:** intelligent alarm; fault diagnosis; case-based reasoning; pattern matching

## 1 电力系统智能告警研究现状

### 1.1 电力系统事故特点

电力系统结构复杂, 地域分布广泛, 涉及的元件众多, 运行的过程中需要快速敏捷的动态响应. 任何意外的扰动若无正确应对措施, 都有可能导事故扩散并迅速波及全系统, 造成重大损失风险<sup>[1]</sup>. 如 2003 年美加大停电, 受影响人口达 5000 万. 同时由于电力设备投资巨大, 往往以数亿元乃至数十亿元计算, 如果发生事故, 也会影响系统本身安全. 因此电力系统发展了完善的继电保护理论与继电保护装置, 保护控制系统负责在事故发生后很短时间内(10-20ms)切除事

故设备, 并尽可能的采用重合闸或投入备用线路的方式, 将故障带来的损失控制在最小的范围内.

电力事故根据其产生缘由主要可分为以下几个类型:

①人为事故. 某些操作人员责任心不强或违反操作规程进行作业.

②天气事故. 天气对输电线路以及电力设备会产生巨大影响, 如雷击会造成线路暂时短路甚至更严重的事故. 再比如 2008 年南方雪灾造成某些省市的电力系统大面积瘫痪.

③设备缺陷. 如 XX 年某市发生大面积停电事故,

<sup>①</sup> 收稿时间:2015-05-04;收到修改稿时间:2015-07-06

经调查发现为 500kV 主变压器突然起火, 保护控制系统虽然即使切断了故障线路, 但因备用主变处于检修阶段, 无法投入运行, 导致多个区域大面积停电。

④用电环境或电网变化. 电力系统运行需要将系统频率和电压维持在一定范围内, 极限用电条件(度夏)或电网拓扑发生变化如果引起系统频率震荡, 严重时会引起全网解列。

由此我们可以得到以下结论:

①重大电力事故影响巨大, 会给人民财产造成重大损失。

②如能对事故进行快速分析, 辅助进行故障定位, 故障范围报告, 迅速的对故障进行诊断, 及时给出处理预案, 则能在事故早期有效. 国内外很多文献选取近年来国内外电网故障作为分析样本, 研究故障过程的发展模式、演变特点, 分析各自相继事件间的关联触发特征与系统状态发展态势. 并且采用人工智能<sup>[2]</sup>、模糊因果网络<sup>[3]</sup>、构建专家系统库<sup>[4]</sup>等各种手段和方法对故障进行诊断和处理。

③电力系统重视事后分析, 对于事故原因与经验教训会进行仔细总结与认真分析, 对暴露出来的相关问题进行落实于整改。

④相关事故分析案例可作为其它事故分析的参考。

### 1.2 研究现状

目前电力系统智能告警的实现均是基于推理为主的预测判断. 在告警事件发生后, 可以根据每条告警信息作出推理, 给出告警信息的描述、发生原因、处理措施以及图解. 通过将自动化系统采集的每个告警信号与知识库中归纳的告警信号种类建立起关联关系, 即定义好每个信号所属的告警信号种类, 形成逐一事件推理判断关系. 推理判断可以人工灵活干预, 关联关系可以逐条定义, 也可以通过快速定义软件批量实现, 提高关联效率. 根据告警信号的类型、主次性、所属单元间隔与知识库的匹配情况等综合信息, 经过匹配和判断后, 给出一个推理结果. 近几年, 故障智能推理成为主流. 根据故障发生的关键条件, 结合接线方式、运行方式、开关变位及开关状态、遥测量、时序等综合判断, 给出当前故障的故障类型、故障结论及处理方式. 运行方式通过拓扑技术获得<sup>[5]</sup>。

本文讨论的是以规则推理为主的智能告警系统, 其实现方式主要有两种:

①基于纯规则推理. 手工输入规则至规则库。

②规则推理结合继电保护原理与系统拓扑. 在基于规则推导的基础上, 基于继电保护原理与系统实时拓扑搜索的保护动作链算法, 考虑不同保护类型的动作方案和动作原理, 同时根据故障前后的系统断面分析故障原因、影响范围和造成后果。

但两者存在以下不足:

①以往故障经验对于未来工作不能提供足够的指导和借鉴。

②系统缺乏相关的容错能力, 信号时序和信号缺失均会导致推理结果的不正确。

## 2 基于案例的推理

CBR: Case-based Reasoning, 就是基于案例的推理技术, 是一种相似或类比的推理方法, 它是通过访问知识库中过去同类问题的求解从而获得当前问题解决方案的一种推理模式, 即利用旧的事例或经验来解决新问题、评价新问题、解释异常情况或理解新情况<sup>[6]</sup>. 在 CBR 中, 一个问题的状态描述及其求解策略用一个案例(Case)表示, 案例库模拟人脑的记忆, 存储了一些过去的相关经历(案例), 案例本身则可以用语义网节点、规则、框架或对象实现<sup>[7]</sup>. 这些案例按一定的模式在知识库中组织, 以便在需要的时候能及时取出. CBR 技术直接利用以往解决问题的实例, 能有效地解决知识表达困难或无法表达的领域问题。

推理循环由 4 个基本过程组成, 即 4R 循环: Retrieve、Reuse、Revise、Retain, 分别对应着案例的提取、重用、改编和保存(学习). 其工作过程如图 1 所示。

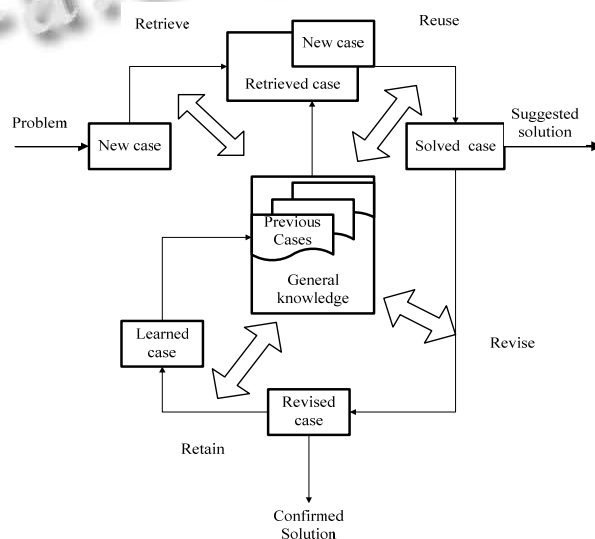


图 1 基于案例推理的过程

基于案例推理是不同于以往的基于规则的推理 RBR(Rule-based reasoning)模式,它克服了 RBR 的一些缺点.知识获取仅是简单的获取过去的案例非常方便.CBR 的知识库储存的是一个一个案例,修改其中的某一案例不会影响其他内容,这使得 CBR 的知识库维护变的十分简单.CBR 能够在解决新问题的同时,保存新案例到案例库中来实现自学习的功能.经过一段时间的使用,系统的准确性会得到提高.这是 CBR 系统最突出的优点.而 RBR 则没有这个功能,它必须借助于领域专家和知识工程师的合作<sup>[8]</sup>.

### 3 算法设计

#### 3.1 算法思路

对于电力系统发生的大量事故案例进行特征量提取,形成系统案例库.当电力事故发生后,对新事故进行特征量提取,与案例库事故案例进行比较,获取案例库中与当前事故最为相似的事故案例.

因电力系统事故类型较多,情况复杂.对于案例匹配可划分为线路故障、主变故障、母线故障、馈线故障等几大类型.

为简化讨论,本文仅讨论线路故障,对于事故的隔离与处理均为就地保护装置完成,一旦事故期间系统运行状态发生变化,保护装置、测控装置将上送大量的信号.信号按分类可分成两大类:1)保护动作 SOE(Sequence of event),根据不同类型的保护原理可分为过流、差动、纵差等类型.2)开关变位 SOE,保护装置检测到故障后,会开放跳闸回路,对相关开关发出跳闸命令,开关的动作将以动作报告的形式上送至自动化系统.

保护装置的动作信号会引起开关的位置变化.一般来说,线路保护可分为主保护、后备保护、远后备保护三个层次.对于主保护,直接跳开相关开关,如为单相故障则会尝试重合开关一次;主保护未能成功隔离故障,则会启动后备保护,尝试跳开与本开关相连的开关;如后备保护失效则启动远后备保护.同时,由于保护动作信号与开关变位信号一般会有对应关系,即保护动作信号引起保护动作信号.这里可能有几种特殊情况:

①保护动作但开关并未动作,有可能是保护误动或者开关拒动.

②开关动作但保护并未动作,有可能是误报信号、人工操作、偷跳信号.

③对于某些瞬时性故障,例如雷电导致的接地,可能会迅速恢复.在此种情况下,保护动作会尝试在开关动作后重新合闸.

一般来说,电力系统保护可分为主保护、后备保护和远后备三个层次,其中主保护应考虑重合闸问题.同时相关开关的动作信号应与对于现有案例,所有信息均为已知,对于需要分析的案例,可根据电压等级、接线方式、故障动作序列、涉及设备进行匹配,将搜索到的案例的故障原因和处理方式供用户参考.

因此在计算案例相似度时,可考虑保护动作情况与开关动作情况两个向量,这两个向量有相应的对应关系,其中主保护应考虑保护动作和重合闸两种情况.设向量  $N\{N1,N2,N3,N4\}$ ,其中  $N1$  代表主保护动作情况、 $N2$  代表重合闸、 $N3$  为后备保护、 $N4$  为远后备保护.每个变量的值取值如果为 0 代表有相关保护动作未发生,其它取值表示不同的保护动作原理.再设置向量  $M\{M1,M2,M3,M4\}$ ,4 个变量均代表开关动作情况,0 代表未动作,1 代表动作.

#### 3.2 案例特征量与数据规整

一般来说,电力系统事故信号由继电保护动作信号与开关动作信号组成,继电保护装置通过检测系统电流、电压的变化判断是否发生故障,通过控制相关开关达到隔离故障、保护设备的目的.相关开关反应、信号序列均会反应出事故的不同性质.

对于事故信息,可从故障简述、变电站电压等级、接线方式、故障动作序列、故障涉及设备、故障设备、故障原因、处理方式等几个方面进行考查.

高压保护原理与中低压保护原理有显著差异,因此将变电站电压等级可规整为以下类型,如表 1 所示.

表 1 变电站电压等级

电压等级	规整值
110kV	1
220kV-500kV	2
750kV-1000kV	3

同时因变电站接线方式对继电保护配置方式有较大影响,同时变电站接线方式相对来说比较复杂,因此接线方式可用一个 4 维向量来描述,见表 2 所示.

表 2 接线方式

主变数量	高压侧母线	主变卷数	低压侧母线
1-10	双母/单母/带旁路/一个半接线/无母线	两卷/三卷	双母/单母/带旁路

一次故障可能涉及多个设备,可采用一个 4 维布尔型向量来描述,其中 0 代表未有此类型设备涉及本次故障,数量代表涉及的设备数量,见表 3 所示。

表 3 故障涉及设备数量

主变	开关	母线	进线	馈线
0-10	0-10	0-2	0-4	0-40

3.3 相似度计算

案例信息拟合可分静态信息拟合和故障信息拟合,静态信息指变电站电压等级、接线方式等,动态信息是指故障设备类型、故障信号等。

设案例 C, 则 C 可表示为以下元组 {e1, e2, e3, e4, e5}

例如对于某变电站发生单相接地故障,重合成功这一案例可表示为

CI={3,(2,4,3,2)^T,(0,0,0,1,0)^T,(1,1,0,0)^T,(1,1,0,0)^T}.

其中向量 e3, e4, e5 有很强的相关性, e1, e2 信息相对固定,在实际应用中可考虑对 e2 进行相似度计算,并结合对案例进行初步计算,在筛选过的结果里进行计算。

常用的相似度计算有以下方法:

① 欧拉距离两个 n 维向量 a(x11, x12, ..., x1n) 与 b(x21, x22, ..., x2n) 间的欧氏距离

d12 = sqrt(sum\_{k=1}^n (x1k - x2k)^2)

② 马氏距离

有 M 个样本向量 X1 ~ XM, 协方差矩阵记为 S, 均值记为向量 mu, 则其中样本向量 X 到 mu 的马氏距离表示为:

D(X) = sqrt((X - mu)^T S^-1 (X - mu))

③ 夹角余弦

类似的,对于两个 n 维样本点 a(x11, x12, ..., x1n) 和 b(x21, x22, ..., x2n), 可以使用类似于夹角余的概念来衡量它们间的相似程度。

cos(theta) = (a · b) / (|a| |b|)

即:

cos(theta) = (sum\_{k=1}^n x1k x2k) / (sqrt(sum\_{k=1}^n x1k^2) sqrt(sum\_{k=1}^n x2k^2))

④ 杰卡德距离

杰卡德距离可用如下公式表示:

J\_s(A, B) = 1 - J(A, B) = (|A union B| - |A intersection B|) / |A union B|

杰卡德距离用两个集合中不同元素占所有元素的

比例来衡量两个集合的区分度。

3.4 知识提取

无论是知识的获取或故障的抽象化,均需要实现从现实故障信息到数学模型的抽象化. 知识的提取可分为两个部分:

- ① 基本信息提取, 包括变电站的接线方式、电压等级、设备运行状态。
- ② 故障信息提取, 包括故障涉及设备、保护动作情况、开关动作情况等。

4 算法实现与验证

4.1 算法流程

算法可分为以下流程:

- ① 根据元数据模型对故障信息进行数据抽象, 获取故障相关的事件列表, 并对事件进行预处理, 根据系统元数据模型进行事件分类, 生成表达式元组。
- ② 根据电压等级、涉及设备类型信息对故障信息进行快速分类. 本步骤是做初步匹配在案例库中搜索出与故障系统运行环境类似的案例。

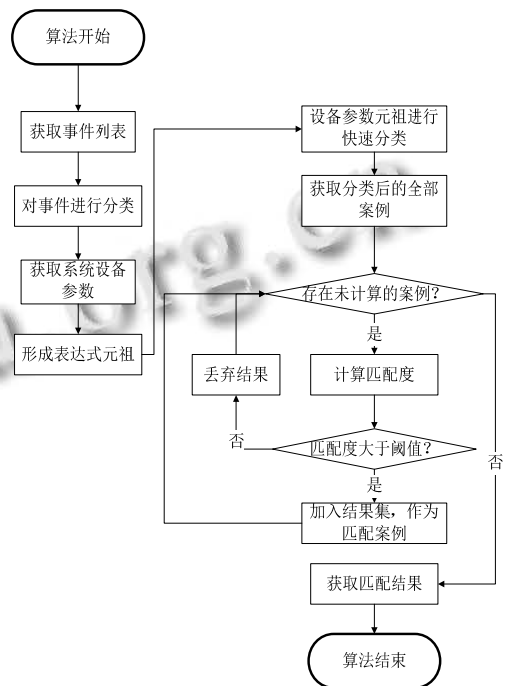


图 2 算法流程图

③ 对分类过的故障进行相似度计算, 使用欧拉距离 d12 = sqrt(sum\_{k=1}^n (x1k - x2k)^2)、夹角余弦 cos(theta) = (a · b) / (|a| |b|)、杰卡

德距离公式  $J_s(A, B) = 1 - J(A, B) = \frac{|A \cup B| - |A \cap B|}{\|A \cup B\|}$  进行计

算.

④对全部相似度计算结果进行排序, 获取相似度大于阈值的案例匹配信息, 选取匹配度最大的案例作为分析结果.

详细算法如图 2 所示.

#### 4.2 案例验证

案例矩阵: 实际计算中发现并不能保证矩阵可逆, 因此不满足马氏距离的要求, 因此, 采用欧拉距离、夹角余弦、杰卡德距离进行计算. 采用 matlab 进行矩阵计算.

```
[1 1 0 0 1 1 0 0;
1 1 1 0 1 1 1 0;
1 0 1 0 1 0 1 0;
1 0 1 1 1 0 1 1;
1 1 0 0 1 0 0 0;
1 1 1 1 1 0 1 1
1 1 0 0 1 1 0 0;
1 1 1 0 1 0 0 0;
1 1 1 1 1 0 0 1;
1 0 0 0 1 0 0 0;
1 1 0 0 1 1 0 0;
1 1 1 0 1 0 1 0;
1 1 1 0 1 0 1 0]
```

其中第一行是匹配目标, 其它行均为案例库信息, 对其分别进行距离计算, 如表 4 所示.

表 4 基于案例推理算法得出的距离值

序号	欧式距离	余弦夹角	杰卡德距离
1	0	0	0
2	1.4142	0.1835	0.3333
3	2	0.5	0.6667
4	2.4495	0.5918	0.75
5	1	0.134	0.25
6	2.2361	0.4331	0.625
7	0	0	0
8	1.4142	0.25	0.4
9	2	0.3876	0.5714
10	1.4142	0.2929	0.5
11	0	0	0
12	1.7321	0.3292	0.5

对欧式距离进行最大-最小值标准化, 得表 5.

表 5 最大-最小标准化的距离值

欧式距离	余弦夹角	杰拉德距离
0	0	0
0.5773	0.1835	0.3333
0.8165	0.5	0.6667
1	0.5918	0.75
0.4082	0.134	0.25
0.9129	0.4331	0.625
0	0	0
0.5773	0.25	0.4
0.8165	0.3876	0.5714
0.5773	0.2929	0.5
0	0	0
0.7071	0.3292	0.5
0.7071	0.3292	0.5

对数据进行多线段的线段回归, 可得图 3 如下.

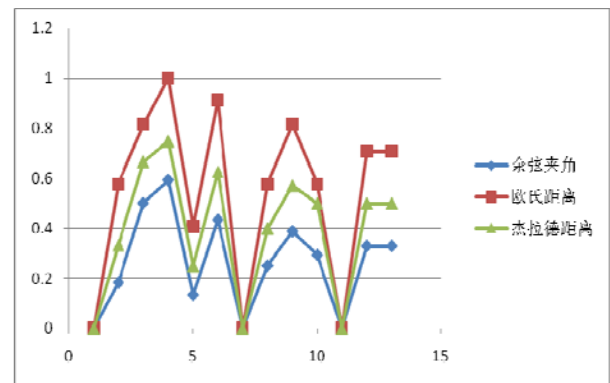


图 3 相似度计算比较示意图

从图中可以发现: 三种方法具有比较好的一致性, 但相比而言欧氏距离有比较好的分辨率, 但也发现算法的一些不足, 例如三种方式对案例库第一个案例, 后备保护动作并未做出比较好的区分度.

经分析这是由于将两组向量合并进行求值导致的两组向量取值互相干扰所致, 为解决该问题, 可考虑采用以下两种措施之一进行改进:

①对于每个向量组给予一个加权系数, 代表不同的向量具有不同的权重.

②对于向量分开计算相似度, 再根据加权系数进行计算总体相似度.

#### 5 结语

本文初步研究了基于案例推理的智能告警算法, 主

要思想是尝试将电力事故进行数学抽象,通过计算两组向量之间的距离来比较电力事故与案例之间的相似性,本文研究了特征量的提取、相似度的计算方法,并通过多种算法的计算验证,论证了本方法有较好的识别度与一致性,但方法在某些情况识别度尚有待提高。

#### 参考文献

- 1 刘友波,胥威汀,丁理杰,刘俊勇,胡斌,宋兆欧,许立维.电力系统连锁故障分析理论与应用(二)—关键特征与研究启示.电力系统保护与控制,2013,41(10):146-155.
- 2 张建华,基于人工智能方法判断环境信息的变电站视频及环境监控系统升级方案设计.广东电力,2011,24(12):97-99, 125.
- 3 徐国钧,刘永胜,华靓,胡晓琴,杨占胜,包拯民,王予疆,周念成.基于模糊因果网络的智能变电站故障诊断应用.电力系统保护与控制,2013,41(8):93-98.
- 4 王强,王超,刘远龙.青岛电网智能告警专家系统应用分析.能源技术经济,2011,23(5):14-17.
- 5 姬书君,朱学科,李淮海.变电站智能告警专家系统研究.华东电力,2011,39(5):773-775.
- 6 Kolodner JL. Case-Based Reasoning. San Mateo CA: Morgan Kaufmann Publishers, 1993.
- 7 Amodt A, Plaza E. Case-based reasoning: foundational issues, methodological variations, and system approaches. The Artificial Intelligence Community, 1994, 7(1): 39-59.
- 8 Leake DB. Case-base Reasoning: Experiences, Lessons, and Future Direction. Menlo Park: AAAI Press, 1996.