



(12) 发明专利申请

(10) 申请公布号 CN 104363106 A

(43) 申请公布日 2015.02.18

(21) 申请号 201410530709.2

(51) Int. Cl.

(22) 申请日 2014.10.09

H04L 12/24 (2006.01)

G06Q 50/06 (2012.01)

(71) 申请人 国网辽宁省电力有限公司信息通信
分公司

地址 110006 辽宁省沈阳市和平区宁波路
18号

(72) 发明人 陈硕 赵永彬 李巍 喻洪辉
刘树吉 于亮亮 邓春宇 卢斌
张靖欣 王鸥

(74) 专利代理机构 沈阳维特专利商标事务所
(普通合伙) 21229

代理人 甄玉荃

权利要求书1页 说明书3页

(54) 发明名称

一种基于大数据技术的电力信息通信故障预警分析方法

(57) 摘要

本发明属于数据分析领域,尤其涉及一种基于大数据技术的电力信息通信故障预警分析方法。该方法包括如下步骤:通过整理历史故障数据建立决策表,将历史故障数据的每个事件拆分为两个字段,所述字段包括故障特征和故障特征对应的多个故障原因,根据故障特征和对应的故障原因建立成行,每一行表示一个真实的专家诊断案例;根据整理的历史故障数据建立的决策表构建专家系统,输入专家系统的知识库内;在专家系统中建立推理机,当输入新的故障特征时,通过推理机分析出故障特征与专家系统内的决策表进行故障原因的推理并计算出故障原因的概率进行报警,推理机通过贝叶斯网络根据历史故障特征对每个故障原因建立朴素贝叶斯网络从而推断出故障原因。

1. 一种基于大数据技术的电力信息通信故障预警分析方法,其特征在于,该方法包括如下步骤:

1) 整理历史故障数据建立决策表,将历史故障数据的每个事件拆分为两个字段,所述字段包括故障特征和故障特征对应的多个故障原因,根据故障特征和对应的故障原因建立成行,每一行表示一个真实的专家诊断案例;

2) 根据步骤 1) 整理的历史故障数据建立的决策表构建专家系统,输入专家系统的知识库内;

3) 在专家系统中建立推理机,当输入新的故障特征时,通过推理机分析出故障特征与专家系统内的决策表进行故障原因的推理并计算出故障原因的概率进行报警,所述推理机通过贝叶斯网络根据历史故障特征对每个故障原因建立朴素贝叶斯网络,令历史故障特征的总和: $X = \{x_1, x_2, \dots, x_n\}$,历史故障原因的总和: $R = \{r_1, r_2, \dots, r_n\}$,建立每个历史故障原因 r_i 对应的历史故障特征组 $\{x_1, x_2, \dots, x_m\}$,历史故障特征组 $\{x_1, x_2, \dots, x_m\}$ 属于历史故障特征的总和: $X = \{x_1, x_2, \dots, x_n\}$ 的子集;

假设有故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$,那么故障特征属于某个故障原因 r_i 的概率为 $P(r_i|x_1, x_2, \dots, x_m) = \frac{P(x_1, x_2, \dots, x_m|r_i) * P(r_i)}{P(x_1, x_2, \dots, x_m)}$,其中, $P(r_i) = \frac{N_{r_i}}{N}$,

$P(x_1, x_2, \dots, x_m|r_i) = \frac{N_{r_i}^{x_1, x_2, \dots, x_m}}{N_{r_i}}$, $P(x_1, x_2, \dots, x_m) = \frac{N_{x_1, x_2, \dots, x_m}}{N}$, N 为所有历史故障数据的记录

数, N_{r_i} 为所有历史故障数据中出现故障原因 r_i 的个数, N_{x_1, x_2, \dots, x_m} 为所有历史故障数据中出现特征组 $\{x_1, x_2, \dots, x_m\}$ 的记录数, $N_{r_i}^{x_1, x_2, \dots, x_m}$ 为所有历史数据中出现故障原因 r_i 又出现故障特征 $x_s = \{x_1, x_2, \dots, x_m\}$ 的记录数。

2. 按照权利要求 1 所述的基于大数据技术的电力信息通信故障预警分析方法,,其特征在于,包括将故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 以及步骤 3) 中推理的故障原因形成历史故障事件更新步骤 1) 的决策表。

3. 按照权利要求 1 所述的基于大数据技术的电力信息通信故障预警分析方法,,其特征在于,当步骤 3) 中分析的属于某个故障原因 r_i 的概率低明显不属于故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 的故障原因 r_i 时,更新决策表时,将故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 中对应的故障原因 r_i 删除。

一种基于大数据技术的电力信息通信故障预警分析方法

技术领域

[0001] 本发明属于数据分析领域,尤其涉及一种基于大数据技术的电力信息通信故障预警分析方法。

背景技术

[0002] 随着信息通信设备的监控种类逐渐增加、数据类型日益丰富和获取途径的逐步完备,信息通信设备所监控的数据量快速增长。目前 IMS 系统的发展,各个业务系统监控指标多达数十种;根据调研测算,仅仅对于目前辽宁省电力公司就有 300 多设备,对于结构化数据分析而言,每台设备大概具有 100 个监测点,每个监测点涵盖 10 个指标,则采集指标点的数目为 30 万。采集指标涵盖性能、故障、配置数据等,每采集点 20 个字节。每五分钟采集一次,则每天采集的数据流约为 2G。对于半结构化数据分析而言,日志数据的采集点约为 600 个,每个采集点每天产生的日志信息约为 10M,则总采集点的日志数据为 6G。这还不涵盖设备每天产生的海量的非结构化数据包括 GIS、视频等信息。形成为面向新型可用性的大数据分析带来极大的挑战;另一方面,为缩短故障响应时间,未来各级各类系统的采集频率的提高和监控范围的提升必将进一步提升数据分析和处理的难度。这些都为一体化描述涵盖业务运行系统、中间件、数据库和主机等监控对象带来新的挑战。同时,面向这类关联对象的故障定位也必将面临信息难以汇集、关联关系难以完备描述的困难,从而为面向故障的定位和预警分析带来新的挑战。因而,面向监控数据的海量化、监控范围的综合化和告警及预警特征的多样化,需要引入高效的基于大数据的分布式处理架构,提供高效的数据抽取和分析手段,需要为分析和监控度量提供一致和实时的数据输出。

发明内容

[0003] 本发明所要解决的技术问题在于提供一种基于大数据技术的电力信息通信故障预警分析方法,旨在解决电力系统数据庞大中,正确分析出异常故障信息。

[0004] 本发明是这样实现的,一种大数据故障报警分析方法,该方法包括如下步骤:

[0005] 1) 整理历史故障数据建立决策表,将历史故障数据的每个事件拆分为两个字段,所述字段包括故障特征和故障特征对应的多个故障原因,根据故障特征和对应的故障原因建立成行,每一行表示一个真实的专家诊断案例;

[0006] 2) 根据步骤 1) 整理的历史故障数据建立的决策表构建专家系统,输入专家系统的知识库内;

[0007] 3) 在专家系统中建立推理机,当输入新的故障特征时,通过推理机分析出故障特征与专家系统内的决策表进行故障原因的推理并计算出故障原因的概率进行报警,所述推理机通过贝叶斯网络根据历史故障特征对每个故障原因建立朴素贝叶斯网络,令历史故障特征的总和 $X = \{x_1, x_2, \dots, x_n\}$,历史故障原因的总和 $R = \{r_1, r_2, \dots, r_n\}$,建立每个历史故障原因 r_i 对应的历史故障特征组 $\{x_1, x_2, \dots, x_m\}$,历史故障特征组 $\{x_1, x_2, \dots, x_m\}$ 属于历史故障特征的总和 $X = \{x_1, x_2, \dots, x_n\}$ 的子集;

[0008] 假设有故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 那么故障特征属于某个故障原因 r_i 的概率为 $P(r_i|x_1, x_2, \dots, x_m) = \frac{P(x_1, x_2, \dots, x_m|r_i) * P(r_i)}{P(x_1, x_2, \dots, x_m)}$, 其中, $P(r_i) = \frac{N_{r_i}}{N}$,

$$P(x_1, x_2, \dots, x_m|r_i) = \frac{N_{r_i}^{x_1, x_2, \dots, x_m}}{N_{r_i}}, P(x_1, x_2, \dots, x_m) = \frac{N_{x_1, x_2, \dots, x_m}}{N}$$

N 为所有历史故障数据的记录数, N_{r_i} 为所有历史故障数据中出现故障原因 r_i 的个数, N_{x_1, x_2, \dots, x_m} 为所有历史故障数据中出现特征组 $\{x_1, x_2, \dots, x_m\}$ 的记录数, $N_{r_i}^{x_1, x_2, \dots, x_m}$ 为所有历史数据中出现故障原因 r_i 又出现故障特征 $x_s = \{x_1, x_2, \dots, x_m\}$ 的记录数。

[0009] 进一步地, 包括将故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 以及步骤 3) 中推理的故障原因形成历史故障事件更新步骤 1) 的决策表。

[0010] 进一步地, 当步骤 3) 中分析的属于某个故障原因 r_i 的概率低明显不属于故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 的故障原因 r_i 时, 更新决策表时, 将故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 中对应的故障原因 r_i 删除。

[0011] 本发明与现有技术相比, 有益效果在于: 本发明解决了在电力系统中采集的庞大的数据中分析出故障原因的问题, 采用贝叶斯网络推理并计算出故障原因的概率进行报警, 具有精度高, 处理的数据步骤少等优点

具体实施方式

[0012] 为了使本发明的目的、技术方案及优点更加清楚明白, 以下结合实施例, 对本发明进行进一步详细说明。应当理解, 此处所描述的具体实施例仅仅用以解释本发明, 并不用于限定本发明。

[0013] 一种大数据故障报警分析方法, 该方法包括如下步骤:

[0014] 1) 整理历史故障数据建立决策表, 将历史故障数据的每个事件拆分为两个字段, 所述字段包括故障特征和故障特征对应的多个故障原因, 根据故障特征和对应的故障原因建立成行, 每一行表示一个真实的专家诊断案例;

[0015] 决策表如下表 1 所示的例子:

[0016] 表 1:

[0017]

故障原因	出现次数	故障特征			
		X_1	X_2	X_n
R_1	N_1	0	1	1
R_2	N_2	1	1	0
.....
R_n	N_n	1	0	1

[0018] 2) 根据步骤 1) 整理的历史故障数据建立的决策表构建专家系统, 输入专 家系统

的知识库内；

[0019] 3) 在专家系统中建立推理机,当输入新的故障特征时,通过推理机分析出故障特征与专家系统内的决策表进行故障原因的推理并计算出故障原因的概率进行报警,所述推理机通过贝叶斯网络根据历史故障特征对每个故障原因建立朴素贝叶斯网络,令历史故障特征的总和 : $X = \{x_1, x_2, \dots, x_n\}$,历史故障原因的总和 : $R = \{r_1, r_2, \dots, r_n\}$,建立每个历史故障原因 r_i 对应的历史故障特征组 $\{x_1, x_2, \dots, x_m\}$,历史故障特征组 $\{x_1, x_2, \dots, x_m\}$ 属于历史故障特征的总和 : $X = \{x_1, x_2, \dots, x_n\}$ 的子集；

[0020] 假设有故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$,那么故障特征属于某个故障原因 r_i 的概率为 $P(r_i|x_1, x_2, \dots, x_m) = \frac{P(x_1, x_2, \dots, x_m|r_i) * P(r_i)}{P(x_1, x_2, \dots, x_m)}$, 其中, $P(r_i) = \frac{N_{r_i}}{N}$,

$$P(x_1, x_2, \dots, x_m|r_i) = \frac{N_{r_i}^{x_1, x_2, \dots, x_m}}{N_{r_i}}, \quad P(x_1, x_2, \dots, x_m) = \frac{N_{x_1, x_2, \dots, x_m}}{N},$$

N 为所有历史故障数据的记录数, N_{r_i} 为所有历史故障数据中出现故障原因 r_i 的个数, N_{x_1, x_2, \dots, x_m} 为所有历史故障数据中出现特征组 $\{x_1, x_2, \dots, x_m\}$ 的记录数, $N_{r_i}^{x_1, x_2, \dots, x_m}$ 为所有历史数据中出现故障原因 r_i 又出现故障特征 $x_s = \{x_1, x_2, \dots, x_m\}$ 的记录数。

[0021] 包括将故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 以及步骤 3) 中推理的故障原因形成历史故障事件更新步骤 1) 的决策表。

[0022] 当步骤 3) 中分析的属于某个故障原因 r_i 的概率低明显不属于故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 的故障原因 r_i 时,更新决策表时,将故障特征的样本 $x_s = \{x_1, x_2, \dots, x_m\}$ 中对应的故障原因 r_i 删除。

[0023] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,均应包含在本发明的保护范围之内。