

王向前,朱佳,孟祥瑞,等.一种基于本体与关联规则的煤矿安全监控预警模型[J].矿业安全与环保,2019,46(3):27-31.
文章编号:1008-4495(2019)03-0027-05

开发设计

一种基于本体与关联规则的煤矿安全监控预警模型

王向前^{1,3},朱佳¹,孟祥瑞²,何叶荣⁴

(1. 安徽理工大学 经济与管理学院,安徽 淮南 232001; 2. 安徽理工大学 能源与安全学院,安徽 淮南 232001;
3. 复旦大学 管理科学与工程博士后流动站,上海 200433; 4. 安徽建筑大学 经济与管理学院,安徽 合肥 230601)

摘要:为了解决煤矿安全监控系统中的可靠性较差、信息管理混乱,以及产生的海量监测数据利用率较低等问题,提出了一种新的安全监控预警模型。利用本体对煤矿安全监控领域的信息进行系统化组织,使其便于共享与重用;利用关联规则技术对海量监测数据进行挖掘,得到大量隐藏的可用于预警的推理规则;将煤矿安全监控本体及推理规则与 Jena 推理机进行绑定形成具有推理机制的安全监控预警模型。利用该模型对实时数据进行推理实验,验证了该模型有效。结果表明,结合了本体与关联规则技术的煤矿安全监控预警模型可有效地整合监控领域信息,并在一定程度上提高了预警的效率和准确率。

关键词:煤矿安全监控预警模型;本体;关联规则;Jena 推理

中图分类号:TD76;TP182

文献标志码:A

A Model of Safety Monitoring and Early Warning for Coal Mine based on Ontology and Association Rules

WANG Xiangqian^{1,3}, ZHU Jia¹, MENG Xiangrui², HE Yerong⁴

(1. School of Economics and Management, Anhui University of Science and Technology, Huainan 232001, China;
2. School of Mining and Safety Engineering, Anhui University of Science and Technology, Huainan 232001, China;
3. Post-doctoral Station of Management Science and Engineering, Fudan University, Shanghai 200433, China;
4. School of Economics and Management, Anhui Jianzhu University, Hefei 230601, China)

Abstract: In coal mine safety monitoring system, there are many problems, such as poor reliability, information management confusion, low utilization rate of mass monitoring data. In order to solve these problems, a new monitoring and early warning model of coal mine safety is put forward. The model uses the ontology to systematize the information in the field of coal mine safety monitor, so it makes mine knowledge benefit sharing and reusing; then, the association rule is used to mine the mass monitoring data, and a large number of hidden reasoning rules which can be used for early-warning are obtained; finally, the coal mine safety monitoring ontology and reasoning rules are combined with the Jena reasoning machine, a safety monitoring and warning model with reasoning mechanism is formed. The validity of the model is verified by reasoning experiments with real-time data. The results show that combined with ontology and association rules, the coal mine safety monitoring and early warning model effectively integrate information in monitoring field and improve the accuracy of early-warning to some extent.

Keywords: coal mine safety monitoring and early warning system; ontology; association rule; Jena reasoning

收稿日期:2018-07-09;2019-05-17 修订

基金项目:国家自然科学基金项目(51474007, 51374117)

作者简介:王向前(1981—),男,安徽阜阳人,博士,副教授,主要从事管理信息系统及矿业管理工程方面的研究工作。E-mail:xqwaust@163.com。

随着科学技术的迅速发展,以及国家对煤矿安全生产的高度重视,越来越多的煤矿安全监控方法相继被研究出来并不断应用于煤矿安全监控系统中^[1]。杨维等^[2]通过分析矿井巷道的特点后,采用路由机制和拓扑机制开发出一种矿井巷道层次型无

线传感器网络应用于无线监控;焦尚彬等^[3]利用 ZigBee 无线传感器网络的无线监控方法,既可以采集环境参数信息又可以实时定位井下人员位置。然而这些改进过的煤矿安全监控系统依然存在一些问题,如信息管理较为混乱、数据利用率低、监控依据单一、可靠性较差等^[4]。

对于信息管理混乱问题,是因为煤矿安全监控系统中没有统一的标准去规范监控概念,因而难以集成和共享信息,造成积累的大量数据形成信息孤岛^[5]。而本体是明确的形式化的共享概念模型^[6],能够实现知识集成、共享与重用,且是机器可读的。另外,煤矿安全监控系统缺少对数据分析处理的能力,导致大量数据并没有得到有效的利用。而数据挖掘(DM)技术就可以对大量数据进行挖掘,得到监控参数间隐藏的信息用于进行联动预警,这样使得监控依据不再单一,对提高监控预警能力具有一定的意义。

1 关联规则

关联规则技术是数据挖掘领域中应用广泛的技术之一^[7],通过算法对数据进行分析,挖掘出数据间相互关联的信息。其定义为^[8],首先设 $I = (i_1, i_2, i_3, \dots, i_n)$ 为由 n 个不同数据项组成的数据项集合,简称项集;然后,给定一个事务数据库 $D = (T_1, T_2, T_3, \dots, T_n)$,其中 $T \subseteq I$,是部分数据项的集合;最后定义的关联规则就是形如 $X \Rightarrow Y(s(\%), c(\%))$ 的规则,其中 $X \subseteq I, Y \subseteq I, X \cap Y = \Phi$ 并且 X, Y 是 T 中的数据项或项集。 s 是规则的支持度(support),即同时包含 X 和 Y 的事务占事务数据库 D 的比率为 s ,公式为 $\text{support}(X \Rightarrow Y) = P(X \cup Y)$; c 是规则的置信度(confidence),即事务数据库 D 中包含 X 的事务也包含 Y 的比率为 c ,公式为 $\text{confidence}(X \Rightarrow Y) = P(Y/X)$ 。根据用户定义的最小支持度和最小置信度可以确定频繁项集,以及强关联规则:若支持度 $\text{support}(X \Rightarrow Y) \geq \text{minsupport}$,则称 X 为频繁项集;若 $\text{support}(X \Rightarrow Y) \geq \text{minsupport}$, $\text{confidence}(X \Rightarrow Y) \geq \text{minconfidence}$,则称 $X \Rightarrow Y$ 为强关联规则。

处理关联规则的算法丰富,笔者选用 Apriori 算法。Apriori 算法^[9]是常用且具有代表性的挖掘算法。该算法的思想简单、易于理解,且操作容易实现,这也是该算法至今还在广泛使用的原因。

2 煤矿安全监控预警模型设计

利用 Jena 推理机^[10]对关联规则进行推理预警,

因为规则推理是基于事件与规则匹配得到推理结果,而每个规则是相互独立的,具有很强的模块性^[11],无法描述知识的结构性,使得一些信息无法被表达出来,影响推理效率;而本体可以较好地表达出领域中概念的结构,也解决了因为同义词或近义词造成的推理难题。因此,构建了基于本体与关联规则挖掘的煤矿安全监控预警模型,如图1所示。

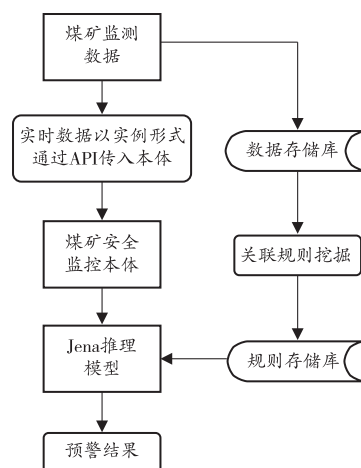


图1 基于本体与关联规则挖掘的煤矿安全监控预警模型

关键步骤描述如下^[12-14]:

1)由传感器监测的实时数据将以实例形式通过API传入煤矿安全监控本体中进行预处理,这样会使数据更适用于本体推理,再将处理好的数据传入Jena推理模型;同时累积实时监测数据,备份到数据仓库中以便关联规则挖掘使用。

2)定时将存储的大量历史监测数据利用 Apriori 算法进行关联规则挖掘,得到可用于推理的强关联规则加入规则存储库中。

3)将推理机与需要进行推理的煤矿安全监控本体、推理规则绑定在一起,得到进行推理或检索所需的包含推理机制的模型。当实时数据传入推理模型中,可以通过Jena推理模型处理得到预警结果,并对超过限值的结果进行报警。

3 模型实现的关键技术

3.1 煤矿安全监控本体构建

为了保障本体的质量,查阅了包括《煤矿安全规程》等资料,得到煤矿安全监控系统中涉及到的一些概念。为了能够保证本体的完整性及实现推理一致性,本体概念需要运用英文表示^[15]。

为了呈现概念间非层次结构,还需要定义本体的属性,主要属性包括对象属性(object property)和数值属性(datatype property)。在煤矿安全监控本体

中定义的对象属性有 has time、locate in、has result、is part of 等,其定义域和值域为本体模型中概念。数据属性主要描述概念具有的特性,包括 has CO value、has CH₄ value、alarm threshold、analog valuetype 等。与对象属性不同的是,数据属性的值域不是概念,而是 int、float、boolean、integer、string、decimal 等数据类型,这些需要根据填入的实例类型进行适当选择^[16]。

然后,根据《煤矿安全规程》等相关规定对本体概念进行实例化,并定义实例的属性及属性值。最后利用 Protégé4.3 及 Java 工具实现煤矿安全监控本体的构建^[15]。煤矿安全监控本体如图2所示。

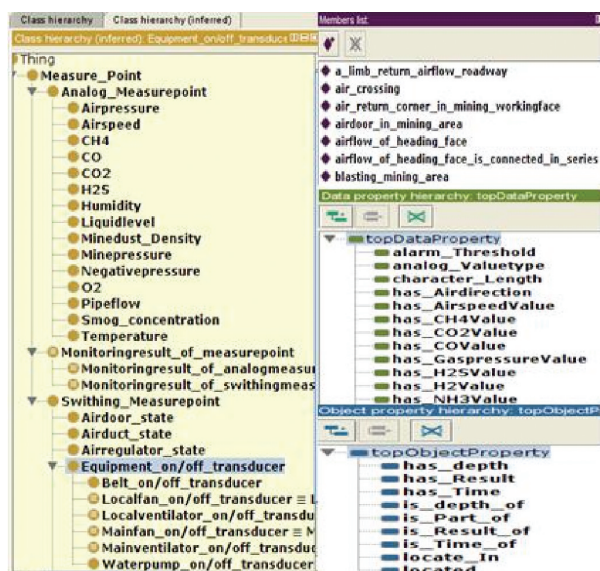


图2 煤矿安全监控本体

3.2 Jena 推理规则

在 Jena 内部定义了一些基于本体特性的通用规则,这些规则可用于检验概念的可满足性、不同概念之间的联系等^[17]。但是,煤矿安全监控领域的概念间或实例间还存在一些特定的关系,而这些是利用通用规则表达不了的^[17]。因此,本推理预警模型需要自定义规则并创建特定的推理机。自定义规则由 body 和 head 构成,具体的规则形式举例如下:

Rule1: (? x rdf:type O:CH₄) (? x O:located In ? y) (? y rdf:type O:return airflow roadway of mining area) (? y O:has CH₄ value ? z) greater Than (? z 1.0%) → (? x O:has State 'alarm')。

若甲烷传感器 x 设在 y 地,y 地属于采区回风巷,采区回风巷的瓦斯浓度(CH₄ 体积分数)限值是 1.0%,若 x 的瓦斯浓度值超过限值,则 x 测点的状态属于危险状态,应该报警。

4 实验验证

4.1 关联规则挖掘

本实验从中煤新集二矿 2017 年 11 月 21 日至 30 日的煤矿安全监控系统连续监测的数据中选取了 960 个数据进行关联规则操作,部分监测数据见表 1。

表1 新集二矿安全监测数据(部分)

序列	风速/ (m · s ⁻¹)	CO 浓度/ (mg · m ⁻³)	温度/ ℃	瓦斯浓度/ %
1	1.80	1.25	25.3	0.31
2	1.85	1.25	26.5	0.28
3	1.90	1.25	25.7	0.45
4	2.01	160.00	22.3	0.35
5	1.95	1.25	21.7	0.37
6	1.90	1.25	23.7	0.25
7	1.85	1.25	25.4	0.41
8	1.90	1.25	26.0	0.44
9	1.90	1.25	24.6	0.39
10	1.86	1.25	26.2	0.16
⋮	⋮	⋮	⋮	⋮

在关联规则操作实验中,由于 Apriori 属于布尔型算法,处理不了连续的量化属性,需要先对数据进行处理使其离散化^[18]。由于要搜寻这种安全监测数据之间隐藏的关联规则,所以选取了矿井内同一位置的监测数据并忽略了时间属性,设定属性风速、CO 浓度、温度及瓦斯浓度的符号分别为 A、B、C、D。按照风速值的范围可以将其分为[0,1.85),[1.85,2.00),[2.00,+∞)3 组,对应的符号分别为 A1,A2,A3;按照 CO 浓度值的范围可以将其分为[0,24),[24,+∞)2 组,对应的符号分别为 B1,B2;按照温度值的范围可以将其分为(0,22),[22,26),[26,+∞)3 组,对应的符号分别为 C1,C2,C3;按照瓦斯浓度值的范围可以将其分为[0,0.3%),[0.3%,0.4%),[0.4%,+∞)3 组,对应的符号分别为 D1,D2,D3。

应用 python 及其开发环境工具 pycharm 实现 Apriori 算法进行关联规则挖掘^[19]。本实验参考领域专家意见并通过大量的调试,设置了基准参数:最小支持度为 0.5,最小置信度为 0.8,得到的运行结果如表 2 所示。

表2 运行结果

关联规则	支持度 (support)	置信度 (confidence)
A2→B1	0.863 070 539	0.995 215 311

表 2(续)

关联规则	支持度 (support)	置信度 (confidence)
A2→C2→B1	0.726 141 079	0.994 318 182
D2→B1	0.547 717 842	0.992 481 203
C2→B1	0.829 875 519	0.990 099 010
B1→C2→A2	0.726 141 079	0.875 000 000
C2→A2	0.730 290 456	0.871 287 129
B1→A2	0.863 070 539	0.870 292 887
A2→C2	0.730 290 456	0.842 105 263
A2→B1→C2	0.726 141 079	0.841 346 154
B1→C2	0.829 875 519	0.836 820 084

根据煤矿安全预警内容得到的关联规则进行举例描述:

Rule2: A2→B1, 说明当风速达到 1.85 ~ 2.00 m/s, CO 浓度值为 0 ~ 24 mg/m³ 的可能性为 0.863, 而这种情况发生的可能性为 0.995。

Rule3: A2→C2→B1, 说明当风速达到 1.85 ~ 2.00 m/s, 温度值达到 22 ~ 26 ℃, CO 浓度值为 0 ~ 24 mg/m³ 的可能性为 0.726, 而这种情况发生的可能性为 0.994。

4.2 推理实验

以 Rule1 为例进行 Jena 推理, 其过程如下: 首先, 将本体持久化存入数据库 MySQL 中, 以方便使用查询语句。接下来, 在本体中传入一个实时监测数据, 其具体描述为: CH₄1 传感器设置在 (located In) return airflow roadway of mining area1 和 CH₄1 传感器上显示甲烷的浓度值 (has CH₄value) 是 0.02, 其中 CH₄1 是 CH₄ 传感器的实例, return airflow roadway of mining area1 是 return airflow roadway of mining area (采区回风巷) 的实例, 则该实时监测数据的 OWL 描述为:

```

<ClassAssertion>
  <Class IRI="#CH4" />
  <NamedIndividual IRI="#CH41" />
</ClassAssertion>
<ClassAssertion>
  <Class IRI="#ReturnAircourse" />
  <NamedIndividual IRI="#return airflow roadway of mining area1" />
</ClassAssertion>
<ObjectPropertyAssertion>
  <ObjectProperty IRI="#locate In" />
  <NamedIndividual IRI="#CH41" />
  <NamedIndividual IRI="#return airflow roadway of mining area1" />

```

```
</ObjectPropertyAssertion>
```

```
<DataPropertyAssertion>
```

```
  <DataProperty IRI="#has CH4value" />
```

```
  <NamedIndividual IRI="#CH41" />
```

```
  <Literal datatypeIRI="#xsd:float">0.02
```

```
</Literal>
```

```
</DataPropertyAssertion>
```

然后, 结合 Rule1 进行基于煤矿安全监控本体的 Jena 推理: 先将本体从 MySQL 中取出, 再添加推理规则, 创建推理机进行推理, 最后可以通过查询语句得到结果。上述推理过程的程序代码如下:

```

/* 设置规则 */
String uri = "http://www.semanticweb.org/user/ontologies...";
String rule = "[ rule... ";
/* 查询语句 */
String queryString = "PREFIX... ";
/* 创建推理机 */
Reasoner reasoner = new GenericRuleReasoner (Rule.parseRules (rule));
InfModel inf = ModelFactory.createInfModel (reasoner, model);
Query query = QueryFactory.create (queryString);
QueryExecution qe = QueryExecutionFactory.create (query, inf);
ResultSet results = qe.execSelect ();
/* 打印结果 */
ResultSetFormatter.out (System.out, results, query);
qe.close ();

```

最后, 由于实例传感器 CH₄1 在 return airflow roadway of mining area1 中的监测结果为 0.02, 超过了限值, 所以得到结果是: CH₄1 现在处于报警状态 (alarm)。

5 结语

1) 通过对煤矿安全监控系统中的有关内容进行归纳和抽象, 构建了煤矿安全监控本体, 并利用 Jena 自定义规则语法对《煤矿安全规程》等相关资料中搜集的规范规定进行构造, 得到可进行预警的推理规则, 再利用 Jena 推理机结合本体和推理规则构建了煤矿安全监控预警模型。

2) 通过实验结果可得, 与其他的本体推理预警技术相比, 该模型的推理规则不是一成不变的, 而是

通过关联规则技术不断学习的,是与时俱进的。并且,原有的煤矿安全监控模型仅是将实时数据与静态限值进行比较得到预警信息,如果传感器不灵敏,就会出现失误,更严重的会出现煤矿事故;而该模型在某个传感器出现误差时可以通过关联规则利用其他参数联合参与报警,从而在一定程度上减少了因为传感器信息量不足而造成的失误,提高了预警的正确率。

3)本体构建是一个长期过程,本文构建的本体并不完善,还需根据专家意见对概念、属性、结构关系及实例等进行进一步完善。为保证推理精确度,还需收集大量的数据,挖掘出更多的推理规则。这些问题都需要进一步地研究和处理。

参考文献:

- [1] 王涛. 我国煤矿矿井监控系统的现状与发展[J]. 煤炭科学技术,2000,28(9):43-45.
- [2] 杨维,王彬. 矿井巷道层次型无线监测无线传感器网络的实现[J]. 煤炭学报,2008,33(1):94-98.
- [3] 焦尚彬,宋丹,张青,等. 基于 ZigBee 无线传感器网络的煤矿监测系统[J]. 电子测量与仪器学报,2013,27(5):436-442.
- [4] 牛立东,赵希平,徐焱,等. 煤矿瓦斯安全监控系统的改进[J]. 太原理工大学学报,2010(2):197-200.
- [5] 郭华. 煤矿瓦斯监控系统的本体模型研究[D]. 太原:太原科技大学,2014.
- [6] STUDER R, BENJAMINS V R, FENSEL D. Knowledge engineering, principles and methods [J]. Data and Knowledge Engineering,1998,25(1-2):161-197.
- [7] 张治斌,姜亚南,郭政慧. 关联规则挖掘技术在煤矿安全预警系统中的应用研究[J]. 工矿自动化,2009(9):24-26.
- [8] 韩家炜. 数据挖掘概念与技术[M]. 北京:机械工业出版社,2013.
- [9] AGRAWAL R, SRIKANT R. Fast algorithms for mining association rules in large databases[C]//Proceedings of the 20th International Conference on Very Large Data Bases. IEEE, 1994:487-499.
- [10] 丁晟春,顾德访. Jena 在实现基于 Ontology 的语义检索中的应用研究[J]. 现代图书情报技术,2005(10):5-9.
- [11] 郑昌兴,刘喜文. 基于规则推理和案例推理的应用模型构建研究——以地震类突发事件为例[J]. 情报理论与实践,2016,39(2):108-112.
- [12] 任维武,胡亮,赵阔. 基于数据挖掘和本体的入侵警报关联模型[J]. 吉林大学学报(工学版),2015,45(3):899-906.
- [13] 谭春辉,王晓. 基于本体与 Web 挖掘的企业网上信任危机预防模型研究[J]. 情报科学,2011,29(10):1559-1564.
- [14] 郭渊,周敬勇. 集成本体和 CBR 的数据挖掘建模技术及其在工艺规划中应用研究[J]. 机械科学与技术,2017,36(4):579-585.
- [15] 刘婷,潘理虎,陈立潮,等. 基于 Jena 推理机制的采煤工作面本体模型推理[J]. 煤矿安全,2017,48(8):102-105.
- [16] 赵佳宝,谈晓洁,焦小澄,等. 基于本体的工业综合预报警系统设计与实现[J]. 系统工程与电子技术,2010,32(10):2145-2150.
- [17] HOSER B, HOTH O A, JASCHKE R, et al. Semantic network analysis of ontologies[C]//Proceedings of ESWC 2006. Berlin: Springer,2006: 514-529.
- [18] 孙国营. 煤矿隐患数据挖掘模型及适用算法研究[D]. 徐州:中国矿业大学,2015.
- [19] 张良均,王路. Python 数据分析与挖掘实战[M]. 北京:机械工业出版社,2015.

(责任编辑:李 琴)