

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2018 年 5 月 3 日 (03.05.2018)



(10) 国际公布号  
**WO 2018/077152 A1**

- (51) 国际专利分类号:  
**G06F 11/07** (2006.01)
- (21) 国际申请号: PCT/CN2017/107407
- (22) 国际申请日: 2017 年 10 月 24 日 (24.10.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201610924584.0 2016 年 10 月 24 日 (24.10.2016) CN
- (71) 申请人: 中国银联股份有限公司 (CHINA UNIONPAY CO. LTD.) [CN/CN]; 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。
- (72) 发明人: 马平清(MA, Pingqing); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 王巍(WANG, Wei); 中国上海市浦东新

区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 韩智东(HAN, Zhidong); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 廉宜果(LIAN, Yiguo); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 戴月(DAI, Yue); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 朱雅蓉(ZHU, Yarong); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 李昂(LI, Ang); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。 朱伟(ZHU, Wei); 中国上海市浦东新区含笑路 36 号银联大厦 7 层, Shanghai 200135 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司 (TDIP & PARTNERS); 中国北京市海淀区知春路 7 号致真大厦 A1304-05 室, Beijing 100191 (CN)。

(54) Title: METHOD AND SYSTEM FOR DETECTING ANOMALY OF VARIABLES OF HOST

(54) 发明名称: 一种主机变量异常检测方法及其系统

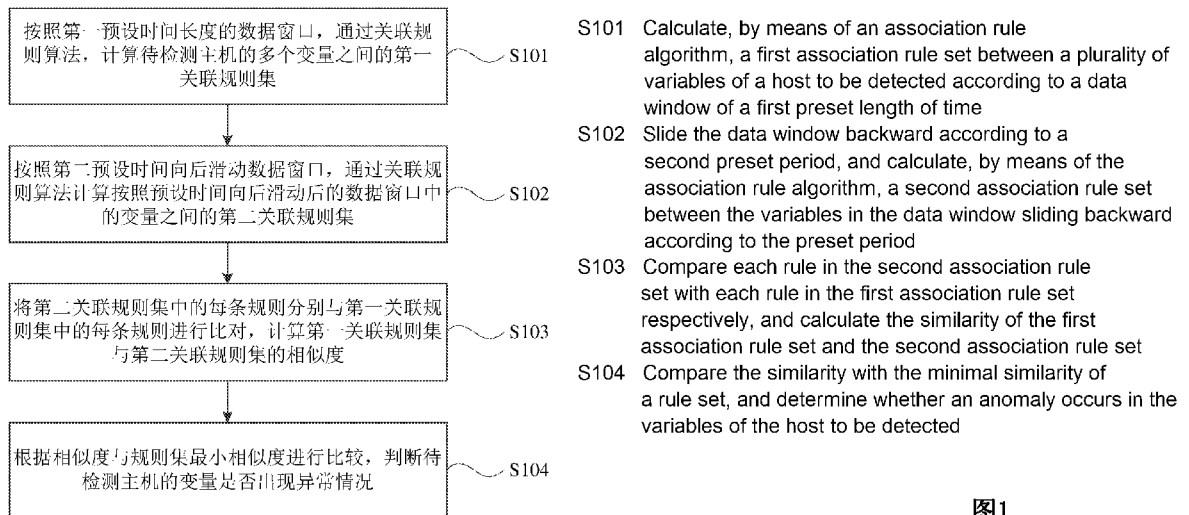


图1

(57) Abstract: The present invention provides a method and system for detecting the anomaly of variables of a host. The method comprises: calculating, by means of an association rule algorithm, a first association rule set between a plurality of variables of a host to be detected according to a data window of a first preset length of time; sliding the data window backward according to a second preset period, and calculating, by means of the association rule algorithm, a second association rule set between the variables in the data window sliding backward according to the preset period; comparing each rule in the second association rule set with each rule in the first association rule set respectively, and calculating the similarity of the first association rule set and the second association rule set; and comparing the similarity with the minimal similarity of a rule set, and determining whether an anomaly occurs in the variables of the host to be detected.

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 本发明提供一种主机变量异常检测方法及系统, 该方法包括: 按照第一预设时间长度的数据窗口, 通过关联规则算法, 计算待检测主机的多个变量之间的第一关联规则集; 按照第二预设时间向后滑动数据窗口, 通过关联规则算法计算按照预设时间向后滑动后的数据窗口中变量之间的第二关联规则集; 将第二关联规则集中的每条规则分别与第一关联规则集中的每条规则进行比对, 计算第一关联规则集与第二关联规则集的相似度; 根据相似度与规则集最小相似度进行比较, 判断待检测主机的变量是否出现异常情况。

## 一种主机变量异常检测方法及系统

### 相关申请的交叉引用

本申请要求 2016 年 10 月 24 日递交，申请号为 CN201610924584.0，标题为“一种主机变量异常检测方法及系统”的中国专利的优先权，上述专利申请的全部内容被引用  
5 作为参考。

### 技术领域

本发明是关于数据异常检测技术，具体地，是关于一种主机变量异常检测方法及系统。

### 10 背景技术

关联规则挖掘是指找出支持度大于等于  $\text{min-sup}$  并且置信度大于等于  $\text{min-conf}$  的所有规则， $\text{min-sup}$  和  $\text{min-conf}$  是对应的支持度和置信度阈值。

由于关联规则挖掘能够发现海量数据中不同属性之间的有趣关系，因此其应用范围较为广泛。有文献提出了一种模糊加权关联规则挖掘方法，结合模糊集合、Apriori 数据挖掘算法和时间序列分析，挖掘报警关联规则。并将其应用于大规模工业生产中，有效的抑制了工业报警的泛滥。有文献运用关联规则和马尔可夫模型的思想对 IP 定位数据集进行挖掘，以此提高 IP 定位的准确率和覆盖率。有文献运用 Apriori 算法，挖掘上市公司之间的信用风险传染，将其应用于上市公司风险评估中，发现产生强关联规则的上市公司之间信用风险传染较为明显，为上市公司风险控制提供决策支持。有文献应用关联规则挖掘算法，对民航快递市场销售数据进行挖掘，为民航快递市场营销提供决策支持。  
15  
20

然而，基于上述算法所进行的异常检测所具备的缺点主要有以下几点：

1. 准确率偏低，其漏检率偏高。
2. 仅仅支持单一变量的检测，无法对多变量进行检测。
3. 通过设定阈值来完成的，而阈值是经验值，且都是静态的，其科学性较差。

25

### 发明内容

本发明实施例的主要目的在于提供一种主机变量异常检测方法及系统，以应对多种不同的异常检测需求，提高异常检测准确率。

为了实现上述目的，本发明实施例提供一种主机变量异常检测方法，所述的主机变量异常检测方法包括：按照第一预设时间长度的数据窗口，通过关联规则算法，计算待检测主机的多个变量之间的第一关联规则集；按照第二预设时间向后滑动所述的数据窗口，通过关联规则算法计算按照预设时间向后滑动后的数据窗口中所述变量之间的第二关联规则集；将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度；根据所述相似度与规则集最小相似度进行比较，判断所述待检测主机的变量是否出现异常情况。

本发明实施例还提供一种主机变量异常检测系统，所述的主机变量异常检测系统包括：第一关联规则集计算单元，用于按照第一预设时间长度的数据窗口，通过关联规则算法，计算待检测主机的多个变量之间的第一关联规则集；第二关联规则集计算单元，用于按照第二预设时间向后滑动所述的数据窗口，通过关联规则算法计算按照预设时间向后滑动后的数据窗口中所述变量之间的第二关联规则集；相似度计算单元，用于将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度；异常情况判断单元，用于根据所述相似度与规则集最小相似度进行比较，判断所述待检测主机的变量是否出现异常情况。

本发明实施例还提供一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述方法。

本发明实施例还提供一种计算机可读存储介质，所述计算机可读存储介质存储有执行上述方法的计算机程序。

本发明实施例的有益效果在于，能够实现将滑动窗口技术与关联规则算法相结合，确定待检测主机的多个变量之间的规则集，并通过滑动窗口滑动前后数据的相似度来判断待检测主机的变量是否出现异常情况，相较于现有的异常检测技术，能够提高异常检测的覆盖率以及准确度，且此方法具有普适性，能够适用于多种数据异常检测场合。

## 附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为根据本发明实施例的主机变量异常检测方法的流程图；

图 2A 至图 2B 为根据本发明实施例的主机变量异常检测系统的结构示意图；

图 3 为根据本发明实施例的相似度计算单元的结构示意图。

## 具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

本发明实施例提供一种主机变量异常检测方法及系统。以下结合附图对本发明进行详细说明。

本发明实施例提供一种主机变量异常检测方法，如图 1 所示，该主机变量异常检测方法主要包括以下步骤：

步骤 S101：按照第一预设时间长度的数据窗口，通过关联规则算法，计算待检测主机的多个变量之间的第一关联规则集；

步骤 S102：按照第二预设时间向后滑动数据窗口，通过关联规则算法计算按照预设时间向后滑动后的数据窗口中的变量之间的第二关联规则集；

步骤 S103：将第二关联规则集中的每条规则分别与第一关联规则集中的每条规则进行比对，计算第一关联规则集与第二关联规则集的相似度；

步骤 S104：根据相似度与规则集最小相似度进行比较，判断待检测主机的变量是否出现异常情况。

通过上述的步骤 S101 至步骤 S104，本发明实施例的主机变量异常检测方法，实现将滑动窗口技术与关联规则算法相结合，确定待检测主机的多个变量之间的规则集，并通过滑动窗口滑动前后数据的相似度来判断待检测主机的变量是否出现异常情况，相较于现有的异常检测技术，能够提高异常检测的覆盖率以及准确度，且此方法具有普适性，能够适用于多种数据异常检测场合。

以下结合具体实施例，对本发明实施例的主机变量异常检测方法做进一步说明。

上述的步骤 S101，按照第一预设时间长度的数据窗口，通过关联规则算法，计算待检测主机的多个变量之间的第一关联规则集。

在实际应用中，待检测主机中可能包含有多种变量，如下表 1 所示。

表 1

| 序号 | 主机名   | 监控变量                                                                               |
|----|-------|------------------------------------------------------------------------------------|
| 1  | HOST1 | CPUCpuUtil, CPUSysTime, CPUUserTime, CPUWio, DSKPercentBusy, SWPTotSwapUsedPercent |
| 2  | HOST2 | CPUCpuUtil, CPUSysTime, CPUUserTime, CPUWio, DSKPercentBusy, SWPTotSwapUsedPercent |
| 3  | HOST3 | CPUCpuUtil, CPUSysTime, CPUUserTime, CPUWio, DSKPercentBusy, SWPTotSwapUsedPercent |
| 4  | HOST4 | CPUCpuUtil, CPUSysTime, CPUUserTime, CPUWio, DSKPercentBusy, SWPTotSwapUsedPercent |
| 5  | HOST5 | CPUCpuUtil, CPUSysTime, CPUUserTime, CPUWio, DSKPercentBusy, SWPTotSwapUsedPercent |

在本发明实施例中，是选取上述表格中一主机的两个变量，按照第一预设时间长度（例如是一天）的数据窗口，通过关联规则算法，计算待检测主机的两个变量之间的第一关联规则集 R1，如下表 2 所示。其中规则前项、规则后项、规则前项的分箱数、规则后项的分箱数、规则最小支持度、规则最小置信度均以参数形式给出，便于后期数据验证时的调整。

需要说明的是，在此实施例中，给出的示例是针对主机中的两个变量计算关联规则的过程。但实际应用中，也可是针对三个或三个以上的变量，本发明并不以此为限。并且，是选用 Apriori 算法作为上述的关联规则算法，但本发明并不以此为限，可根据实际情况选用不同的关联规则算法。

表 2

| host_naiparameter1 | parameter2 | frontItem      | behindItem | Support | Confidence    |
|--------------------|------------|----------------|------------|---------|---------------|
| HOST1              | CPUCpuUtil | DSKPercentBusy | A1         | E2      | 0.0007 1      |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B1      | 0.0105 0.0893 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B10     | 0.0028 0.0238 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B11     | 0.0014 0.0119 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B12     | 0.0014 0.0119 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B13     | 0.0028 0.0238 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B2      | 0.056 0.4762  |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B3      | 0.028 0.2381  |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B4      | 0.0063 0.0536 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B5      | 0.0035 0.0298 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A10        | B6      | 0.0014 0.0119 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A11        | B1      | 0.0147 0.0778 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A11        | B10     | 0.0021 0.0111 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A11        | B11     | 0.0056 0.0296 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A11        | B12     | 0.0063 0.0333 |
| HOST1              | CPUCpuUtil | DSKPercentBusy | A11        | B13     | 0.0028 0.0148 |

上表中，frontItem 表示规则前项，behindItem 表示规则后项，Support 表示支持度，Confidence 表示置信度。

上述的 Apriori 算法是一种最有影响的布尔关联规则算法，因此，基于上述任务的 Apriori 算法基本思想是：首先找出所有的频繁项集，选出的项集需要满足预先设定的最

小支持度。然后使用找到的频繁项集产生强关联规则，其中所产生的规则必须满足最小置信度。

上述步骤 S102，按照第二预设时间向后滑动数据窗口，通过关联规则算法计算按照预设时间向后滑动后的数据窗口中的变量之间的第二关联规则集。

5 在本发明实施例中，是将滑动窗口技术与关联规则算法相结合。例如，对于表 1 中，Host1 主机的两个变量 CPUUserTime 及 DSKPercentBus，上述的第一关联规则集即是获取上述两个变量在一天内的数据（当天 00:00 至次日 00:00）的规则集。而基于滑动窗口，按预设时间（例如是 5 分钟）向后滑动数据窗口，则是获取上述两个变量在当天 00:05 至次日 00:05 的数据，并通过 Apriori 算法计算生成第二关联规则集 R2。

10 上述步骤 S103，将第二关联规则集中的每条规则分别与第一关联规则集中的每条规则进行比对，计算第一关联规则集与第二关联规则集的相似度。

具体地，是利用规则集比对算法，对第二关联规则集 R2 中的每条规则与第一关联规则集 R1 中的每条规则进行比对。其中比对内容为：规则前项、规则后项、规则最小支持度、规则最小置信度，比对方法为：规则支持度和规则置信度比较时，可以按照偏差进行设置，即第二关联规则集 R2 中的一条规则和第一关联规则集 R1 中一条规则的前后项相同，如果这两条规则的支持度偏差 S(Dif)及置信度偏差 C(Dif)分别在给定支持度偏差阈值范围及置信度偏差阈值范围内（S(Dif)和 C(Dif)是可参数化的，根据验证结果进行调整），则这两条规则是相同的。然后，统计确定第一关联规则集 R1 与第二关联规则集 R1 中相同规则的数量 S(Sum)。

20 然后，根据第一关联规则集中规则的总数及相同规则的数量计算第一关联规则集与第二关联规则集的相似度。

具体地，在一实施例中，是结合第一关联规则集 R1 中规则的总数 R1(Sum)来计算第一关联规则集 R1 及第二关联规则集 R2 的相似度 S。可通过以下公式计算第一关联规则集 R1 及第二关联规则集 R2 的相似度 S： $S = S(\text{Sum}) / R1(\text{Sum}) \times 100\%$ ，计算结果如下表 3 所示。

表 3

| host_name | parameter1 | parameter2     | window          | percent     |
|-----------|------------|----------------|-----------------|-------------|
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:05 | 0.986577181 |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:10 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:15 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:20 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:25 | 0.993243243 |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:30 | 0.993288591 |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:35 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:40 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:45 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:50 | 0.993333333 |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 0:55 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:00 | 1           |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:05 | 0.825806452 |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:10 | 0.947712418 |

通过上表可以看出，在此实施例中，是以每 5 分钟作为预设时间，向后移动滑动窗口，从而获取第二关联规则集 R2。

然后，即可通过步骤 S104，将计算得到的第一关联规则集 R1 及第二关联规则集 R2 的相似度 S 与规则集最小相似度进行比较，判断待检测主机的变量是否出现异常情况。

其中，利用历史数据验证，调整相关参数，得到不同规则集的最小相似度 S(Min)。例如是，观察历史数据，在实际生产环境中没有发生故障和事件时，其相应的 S(Dif)为 0.006，C(Dif)为 0.03，S(Min)为 85%。

将相似度 S 与规则集最小相似度 S(Min)进行比较，如果  $S \geq S(\text{Min})$ ，则判定待检测主机的变量正常；如果  $S < S(\text{Min})$ ，则判定待检测主机的变量出现异常情况。

结合上表 3 中的数据可以看出，对于规则集最小相似度 S(Min)为 85%时，HOST1 的变量 CPUUserTime 及 DSKPercentBus 的异常情况出现在 2015/10/29 1:05，其余时间均为正常。

需要说明的是，上述实施例中所列举的具体时间、数值及百分比等数据均为举例说明，而并非用以限制本发明。在具体实施时，上述的具体时间、数值及百分比等数据均可根据实际情况和实际需要进行调整。

并且，上文中所提及的变量也仅是实例说明，在实际应用中，主机中可包含多种变量，针对不同的变量之间进行异常检测，均可通过本发明实施例的主机变量异常检测方法的上述过程实现，在此不再赘述。

以上即是本发明实施例的主机变量异常检测方法在进行异常检测的主要过程，通过上述过程，融合关联分析和滑动窗口技术，在实际应用中，根据验证结果，调整相应的

参数，得出不同滑动窗口规则集之间的相似度，然后与最小相似度进行比较，判断变量的异常情况。能够有效提高异常检测准确率及检测覆盖率，并减小检测漏报率。

在一实施例中，本发明实施例的主机变量异常检测方法还可包括如下步骤：当确定待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照  
5 预设时间向后滑动后的数据窗口中所述变量的数据。

如果出现异常情况，用上一个没有异常的滑动窗口期数据替换本窗口期的数据，然后重新计算规则集。用最新的窗口期规则集和上一个窗口期的规则集对比，形成当前窗口期的规则集相似度。

例如，对于上表 3 中的数据而言，当 2015/10/29 1:05 的数据出现异常时，用  
10 2015/10/29 1:00 的数据替换 2015/10/29 1:05 的数据，并用替换后的 2015/10/29 1:05 的数据与 2015/10/29 1:10 的数据再进行相似度计算。

基于本发明实施例的主机变量异常检测方法，在实际应用过程中进行了实际测试，并与传统数据中心的异常检测结果进行比对。具体测试结果如下表 4 及表 5 所示。

表 4

| host_name | parameter      | date_time       | value_str |
|-----------|----------------|-----------------|-----------|
| HOST1     | DSKPercentBusy | 2015/10/30 1:43 | 89.5157   |
| HOST1     | DSKPercentBusy | 2015/10/30 2:19 | 90.8924   |
| HOST1     | DSKPercentBusy | 2015/10/30 2:37 | 90.2954   |
| HOST2     | DSKPercentBusy | 2015/10/30 1:43 | 85.2514   |
| HOST3     | DSKPercentBusy | 2015/10/30 1:43 | 90.6892   |
| HOST4     | DSKPercentBusy | 2015/10/30 1:43 | 89.5157   |
| HOST5     | DSKPercentBusy | 2015/10/30 1:43 | 88.8795   |

表 4 中示出的是 2015 年 10 月份传统监控异常事件表。

上表中，host\_name 代表主机名，parameter 代表主机的变量，date\_time 代表时间点，value\_str 代表该时间点下该主机变量的具体值。

表 5

| host_name | parameter1 | parameter2     | windows         | percent  | isalarm | S(min) |
|-----------|------------|----------------|-----------------|----------|---------|--------|
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:05 | 0.825806 | 1       | 0.85   |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 2:35 | 0.641509 | 1       | 0.85   |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:05 | 0.143617 | 1       | 0.85   |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:45 | 0.636943 | 1       | 0.85   |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:50 | 0.367089 | 1       | 0.85   |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 2:20 | 0.374194 | 1       | 0.85   |
| HOST1     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 2:40 | 0.381579 | 1       | 0.85   |
| HOST2     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:05 | 0.397436 | 1       | 0.85   |
| HOST2     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:45 | 0.503195 | 1       | 0.85   |
| HOST2     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 2:15 | 0.503145 | 1       | 0.85   |
| HOST3     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:05 | 0.348993 | 1       | 0.85   |
| HOST3     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 2:35 | 0.338028 | 1       | 0.85   |
| HOST3     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:05 | 0.248227 | 1       | 0.85   |
| HOST3     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:45 | 0.234375 | 1       | 0.85   |
| HOST3     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 2:40 | 0.252174 | 1       | 0.85   |
| HOST3     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 2:50 | 0.323944 | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:05 | 0.45     | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 2:35 | 0.639241 | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:05 | 0.831169 | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:45 | 0.165746 | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/11/1 2:35  | 0.417722 | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/11/2 1:05  | 0.463855 | 1       | 0.85   |
| HOST4     | CPUCpuUtil | DSKPercentBusy | 2015/11/3 1:05  | 0.654088 | 1       | 0.85   |
| HOST5     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 1:05 | 0.685897 | 1       | 0.85   |
| HOST5     | CPUCpuUtil | DSKPercentBusy | 2015/10/29 2:35 | 0.181287 | 1       | 0.85   |
| HOST5     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:05 | 0.801325 | 1       | 0.85   |
| HOST5     | CPUCpuUtil | DSKPercentBusy | 2015/10/30 1:45 | 0.69375  | 1       | 0.85   |
| HOST5     | CPUCpuUtil | DSKPercentBusy | 2015/11/5 2:45  | 0.75641  | 1       | 0.85   |

上表中 host\_name 代表主机名, parameter1 代表主机的某一变量, parameter2 代表主机的另一变量, windows 代表窗口期, 以 5 分钟为单位, percent 代表当前窗口期与上一窗口期规则集的相似度, S(min)代表规则集最小相似度, isalarm 代表当前规则集相似度是否小于规则集最小相似度, 1 表示小于, 0 表示大于。

通过表 4 及表 5 的检测结果对比可知, 利用本发明实施例的主机变量异常检测方法, 异常检测准确率为 98%, 检测覆盖率为 100%, 检测漏报率为 0 (传统的漏报率为 66.7%)。其实际效果要明显优于传统数据中心的异常检测。

本发明实施例还提供一种主机变量异常检测系统, 如图 2A 至图 2B 所示, 该主机变量异常检测系统主要包括: 第一关联规则集计算单元 1、第二关联规则集计算单元 2、相似度计算单元 3 及异常情况判断单元 4 等。

其中, 上述的第一关联规则集计算单元 1 用于按照第一预设时间长度的数据窗口, 通过关联规则算法, 计算待检测主机的多个变量之间的第一关联规则集; 第二关联规则集计算单元 2 用于按照第二预设时间向后滑动数据窗口, 通过关联规则算法计算按照预设时间向后滑动后的数据窗口中的变量之间的第二关联规则集; 相似度计算单元 3 用于将第二关联规则集中的每条规则分别与第一关联规则集中的每条规则进行比对, 计算第一关联规则集与第二关联规则集的相似度; 异常情况判断单元 4 用于根据相似度与规则集最小相似度进行比较, 判断待检测主机的变量是否出现异常情况。

通过上述的各个组成部分的相互协作，本发明实施例的主机变量异常检测系统，实现将滑动窗口技术与关联规则算法相结合，确定待检测主机的多个变量之间的规则集，并通过滑动窗口滑动前后数据的相似度来判断待检测主机的变量是否出现异常情况，相较于现有的异常检测技术，能够提高异常检测的覆盖率以及准确度，且此系统具有普适性，能够适用于多种数据异常检测场合。

以下结合具体实施例，对本发明实施例的主机变量异常检测系统做进一步说明。

上述的第一关联规则集计算单元 1，用于按照第一预设时间长度的数据窗口，通过关联规则算法，计算待检测主机的多个变量之间的第一关联规则集。

在实际应用中，待检测主机中可能包含有多种变量，如表 1 所示。

在本发明实施例中，是选取表 1 中一主机的两个变量，按照第一预设时间长度（例如是一天）的数据窗口，通过关联规则算法，计算待检测主机的两个变量之间的第一关联规则集 R1，如表 2 所示。其中规则前项、规则后项、规则前项的分箱数、规则后项的分箱数、规则最小支持度、规则最小置信度均以参数形式给出，便于后期数据验证时的调整。

需要说明的是，在此实施例中，给出的示例是针对主机中的两个变量计算关联规则的过程。但实际应用中，也可能是针对三个或三个以上的变量，本发明并不以此为限。并且，是选用 Apriori 算法作为上述的关联规则算法，但本发明并不以此为限，可根据实际情况选用不同的关联规则算法。

上述的 Apriori 算法是一种最有影响的布尔关联规则算法，因此，基于上述任务的 Apriori 算法基本思想是：首先找出所有的频繁项集，选出的项集需要满足预先设定的最小支持度。然后使用找到的频繁项集产生强关联规则，其中所产生的规则必须满足最小置信度。

上述的第二关联规则集计算单元 2，用于按照第二预设时间向后滑动数据窗口，通过关联规则算法计算按照预设时间向后滑动后的数据窗口中的变量之间的第二关联规则集。

在本发明实施例中，是将滑动窗口技术与关联规则算法相结合。例如，对于表 1 中，Host1 主机的两个变量 CPUUserTime 及 DSKPercentBus，上述的第一关联规则集即是获取上述两个变量在一天内的数据（当天 00:00 至次日 00:00）的规则集。而基于滑动窗口，按预设时间（例如是 5 分钟）向后滑动数据窗口，则是获取上述两个变量在当天 00:05 至次日 00:05 的数据，并通过 Apriori 算法计算生成第二关联规则集 R2。

上述的相似度计算单元 3，用于将第二关联规则集中的每条规则分别与第一关联规则集中的每条规则进行比对，计算第一关联规则集与第二关联规则集的相似度。

具体地，如图 3 所示，相似度计算单元 3 主要包括：相同规则数量确定模块 31，用于根据比对结果确定第一关联规则集与第二关联规则集中相同规则的数量；相似度计算模块 32，用于根据第一关联规则集中规则的总数及相同规则的数量计算相似度。

其中，上述的相同规则数量确定模块 31 是利用规则集比对算法，对第二关联规则集 R2 中的每条规则与第一关联规则集 R1 中的每条规则进行比对。其中比对内容为：规则前项、规则后项、规则最小支持度、规则最小置信度，比对方法为：规则支持度和规则置信度比较时，可以按照偏差进行设置，即第二关联规则集 R2 中的一条规则和第一关联规则集 R1 中一条规则的前后项相同，如果这两条规则的支持度偏差  $S(Dif)$  及置信度偏差  $C(Dif)$  分别在给定支持度偏差阈值范围及置信度偏差阈值范围内（ $S(Dif)$  和  $C(Dif)$  是可参数化的，根据验证结果进行调整），则这两条规则是相同的。然后，统计确定第一关联规则集 R1 与第二关联规则集 R1 中相同规则的数量  $S(Sum)$ 。

在通过相同规则数量确定模块 31 确定了第一关联规则集 R1 与第二关联规则集 R2 中相同规则的数量  $S(Sum)$  后，相似度计算模块 32 可结合第一关联规则集 R1 中规则的总数  $R1(Sum)$  来计算第一关联规则集 R1 及第二关联规则集 R2 的相似度  $S$ 。

具体地，在一实施例中，相似度计算模块 32 可通过以下公式计算第一关联规则集 R1 及第二关联规则集 R2 的相似度  $S$ ： $S = S(Sum) / R1(Sum) \times 100\%$ ，计算结果如表 3 所示。

通过表 3 可以看出，在此实施例中，是以每 5 分钟作为预设时间，向后移动滑动窗口，从而获取第二关联规则集 R2。

然后，即可通过异常情况判断单元 4，将计算得到的第一关联规则集 R1 及第二关联规则集 R2 的相似度  $S$  与规则集最小相似度进行比较，判断待检测主机的变量是否出现异常情况。

其中，利用历史数据验证，调整相关参数，得到不同规则集的最小相似度  $S(Min)$ 。例如是，观察历史数据，在实际生产环境中没有发生故障和事件时，其相应的  $S(Dif)$  为 0.006， $C(Dif)$  为 0.03， $S(Min)$  为 85%。

将相似度  $S$  与规则集最小相似度  $S(Min)$  进行比较，如果  $S \geq S(Min)$ ，则判定待检测主机的变量正常；如果  $S < S(Min)$ ，则判定待检测主机的变量出现异常情况。

结合表 3 中的数据可以看出，对于规则集最小相似度  $S(\text{Min})$  为 85% 时，HOST1 的变量 CPUUserTime 及 DSKPercentBus 的异常情况出现在 2015/10/29 1:05，其余时间均为正常。

需要说明的是，上述实施例中所列举的具体时间、数值及百分比等数据均为举例说明，而并非用以限制本发明。在具体实施时，上述的具体时间、数值及百分比等数据均可根据实际情况和实际需要进行调整。

并且，上文中所提及的变量也仅是实例说明，在实际应用中，主机中可包含多种变量，针对不同的变量之间进行异常检测，均可通过本发明实施例的主机变量异常检测系统的上述过程实现，在此不再赘述。

以上即是本发明实施例的主机变量异常检测系统在进行异常检测的主要过程，通过上述过程，融合关联分析和滑动窗口技术，在实际应用中，根据验证结果，调整相应的参数，得出不同滑动窗口规则集之间的相似度，然后与最小相似度进行比较，判断变量的异常情况。能够有效提高异常检测准确率及检测覆盖率，并减小检测漏报率。

在一实施例中，本发明实施例的主机变量异常检测系统还可包括异常情况处理单元，用于当确定待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据。

如果出现异常情况，异常情况处理单元用上一个没有异常的滑动窗口期数据替换本窗口期的数据，然后重新计算规则集。用最新的窗口期规则集和上一个窗口期的规则集对比，形成当前窗口期的规则集相似度。

例如，对于表 3 中的数据而言，当 2015/10/29 1:05 的数据出现异常时，用 2015/10/29 1:00 的数据替换 2015/10/29 1:05 的数据，并用替换后的 2015/10/29 1:05 的数据与 2015/10/29 1:10 的数据再进行相似度计算。

在一实施例中，还提供一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述方法。其中，所述处理器执行方法还包含：当确定所述待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据。在上述实施例中，将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度，具体包括：根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量；根据所述第一关联规则集中规则的总数及所述相同规则的数量计算所述相似度。

在一实施例中，所述处理器执行判断所述待检测主机的变量是否出现异常情况，具体包括：当所述相似度大于或等于所述规则集最小相似度时，则判定所述待检测主机的变量正常；当所述相似度小于所述规则集最小相似度时，则判定所述待检测主机的变量出现异常情况。

- 5        在一实施例中，本发明还提供一种计算机可读存储介质，所述计算机可读存储介质存储有执行上述方法的计算机程序。其中，当确定所述待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度，具体包括：
- 10    根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量；根据所述第一关联规则集中规则的总数及所述相同规则的数量计算所述相似度。

- 在上述实施例中，所述计算机程序执行判断所述待检测主机的变量是否出现异常情况，具体包括：当所述相似度大于或等于所述规则集最小相似度时，则判定所述待检测主机的变量正常；当所述相似度小于所述规则集最小相似度时，则判定所述待检测主机的变量出现异常情况。
- 15

本领域普通技术人员当可知，以上主机变量异常检测方法及系统中各实施例中记载的技术方案及其任意组合方案均可通过上述计算机可读存储介质所运行的程序以及上述计算机设备中处理器执行，本发明在此就不再一一举例说明。

- 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分步骤可以通过程序来指令相关的硬件来完成，该程序可以存储于一计算机可读取存储介质中，比如ROM/RAM、磁碟、光盘等。
- 20

- 以上所述的具体实施例，对本发明的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本发明的具体实施例而已，并不用于限定本发明的保护范围，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。
- 25

## 权利要求书

1. 一种主机变量异常检测方法，其特征在于，所述的主机变量异常检测方法包括：

按照第一预设时间长度的数据窗口，通过关联规则算法，计算待检测主机的多个变量之间的第一关联规则集；

5       按照第二预设时间向后滑动所述的数据窗口，通过关联规则算法计算按照预设时间向后滑动后的数据窗口中所述变量之间的第二关联规则集；

将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度；

10       根据所述相似度与规则集最小相似度进行比较，判断所述待检测主机的变量是否出现异常情况。

2. 根据权利要求 1 所述的主机变量异常检测方法，其特征在于，当确定所述待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据。

15       3. 根据权利要求 1 所述的主机变量异常检测方法，其特征在于，将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度，具体包括：

根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量；

根据所述第一关联规则集中规则的总数及所述相同规则的数量计算所述相似度。

20       4. 根据权利要求 3 所述的主机变量异常检测方法，其特征在于，根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量，具体包括：

比对所述第一关联规则集与第二关联规则集的每条规则之间，是否满足以下条件：

所述第一关联规则集与第二关联规则集的规则前项及规则后项是否相同；

所述第一关联规则集与第二关联规则集的最小支持度的支持度偏差在支持度偏差阈值范围内；

25       所述第一关联规则集与第二关联规则集的最小置信度的置信度偏差在置信度偏差阈值范围内；

如果均满足所述条件，则判定所述第一关联规则集与第二关联规则集中的规则属于相同规则；

统计所述相同规则的数量。

5. 根据权利要求 3 所述的主机变量异常检测方法, 其特征在于, 通过以下公式计算所述相似度:

$$S = S(\text{Sum}) / R1(\text{Sum}) \times 100\%,$$

其中, S 为所述相似度; S(Sum)为所述相同规则的数量; R1(Sum)为所述第一关联规则集中规则的总数。

6. 根据权利要求 1 所述的主机变量异常检测方法, 其特征在于, 判断所述待检测主机的变量是否出现异常情况, 具体包括:

当所述相似度大于或等于所述规则集最小相似度时, 则判定所述待检测主机的变量正常;

10 当所述相似度小于所述规则集最小相似度时, 则判定所述待检测主机的变量出现异常情况。

7. 一种主机变量异常检测系统, 其特征在于, 所述的主机变量异常检测系统包括:

第一关联规则集计算单元, 用于按照第一预设时间长度的数据窗口, 通过关联规则算法, 计算待检测主机的多个变量之间的第一关联规则集;

15 第二关联规则集计算单元, 用于按照第二预设时间向后滑动所述的数据窗口, 通过关联规则算法计算按照预设时间向后滑动后的数据窗口中所述变量之间的第二关联规则集;

相似度计算单元, 用于将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对, 计算所述第一关联规则集与第二关联规则集的相似度;

20 异常情况判断单元, 用于根据所述相似度与规则集最小相似度进行比较, 判断所述待检测主机的变量是否出现异常情况。

8. 根据权利要求 7 所述的主机变量异常检测系统, 其特征在于, 所述主机变量异常检测系统还包括:

25 异常情况处理单元, 用于当确定所述待检测主机的变量出现异常情况时, 用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据。

9. 根据权利要求 7 所述的主机变量异常检测系统, 其特征在于, 所述的相似度计算单元包括:

30 相同规则数量确定模块, 用于根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量;

相似度计算模块，用于根据所述第一关联规则集中规则的总数及所述相同规则的数量计算所述相似度。

10. 根据权利要求 7 所述的主机变量异常检测系统，其特征在于，所述异常情况判断单元具体用于：

5 当所述相似度大于或等于所述规则集最小相似度时，则判定所述待检测主机的变量正常；

当所述相似度小于规则集最小相似度时，则判定所述待检测主机的变量出现异常情况。

11. 一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的  
10 计算机程序，其特征在于，所述处理器执行所述计算机程序时实现权利要求 1 所述方法。

12. 根据权利要求 11 所述的计算机设备，其特征在于，所述处理器执行方法还包含：当确定所述待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据。

15 13. 根据权利要求 11 所述的计算机设备，其特征在于，所述处理器执行将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度，具体包括：

根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量；

根据所述第一关联规则集中规则的总数及所述相同规则的数量计算所述相似度。

20 14. 根据权利要求 11 所述的计算机设备，其特征在于，所述处理器执行判断所述待检测主机的变量是否出现异常情况，具体包括：

当所述相似度大于或等于所述规则集最小相似度时，则判定所述待检测主机的变量正常；

25 当所述相似度小于所述规则集最小相似度时，则判定所述待检测主机的变量出现异常情况。

15. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有执行权利要求 1 所述方法的计算机程序。

16. 根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述计算机程序执行方法还包含：当确定所述待检测主机的变量出现异常情况时，用前一数据窗口中的所述变量的数据替换所述按照预设时间向后滑动后的数据窗口中所述变量的数据。  
30

17. 根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述计算机程序执行将所述第二关联规则集中的每条规则分别与所述第一关联规则集中的每条规则进行比对，计算所述第一关联规则集与第二关联规则集的相似度，具体包括：

根据比对结果确定所述第一关联规则集与第二关联规则集中相同规则的数量；

5 根据所述第一关联规则集中规则的总数及所述相同规则的数量计算所述相似度。

18. 根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述计算机程序执行判断所述待检测主机的变量是否出现异常情况，具体包括：

当所述相似度大于或等于所述规则集最小相似度时，则判定所述待检测主机的变量正常；

10 当所述相似度小于所述规则集最小相似度时，则判定所述待检测主机的变量出现异常情况。

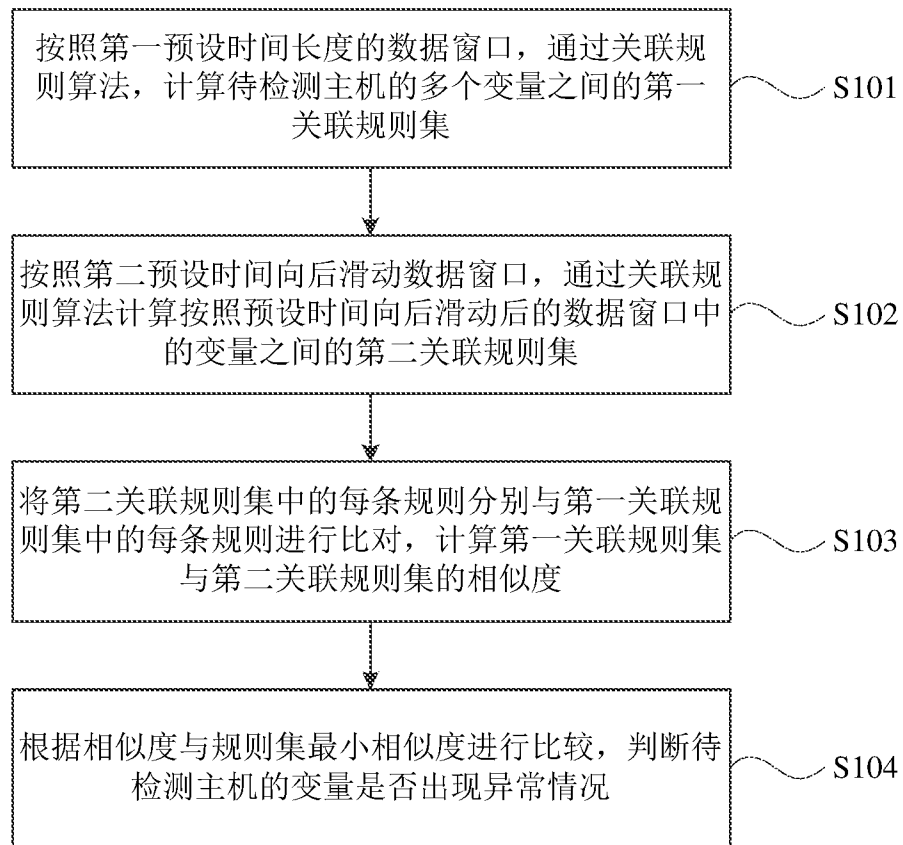


图1

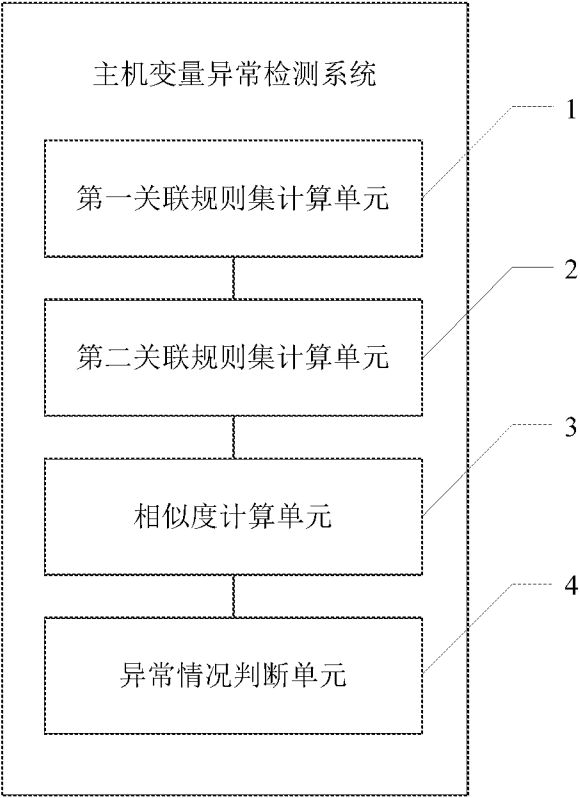


图2A

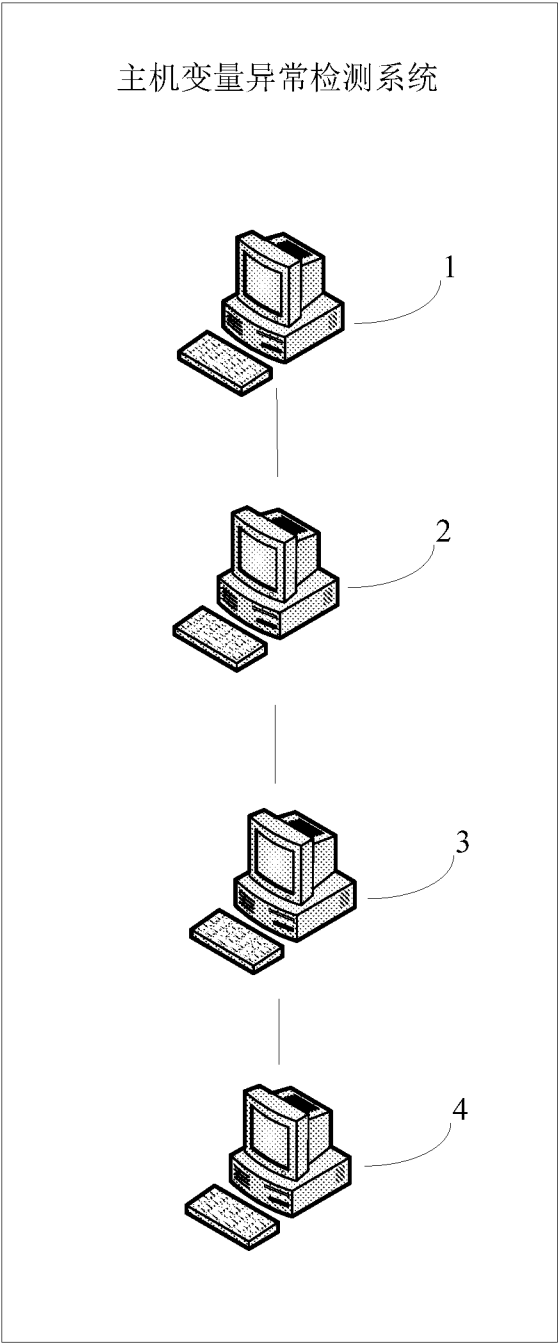


图2B

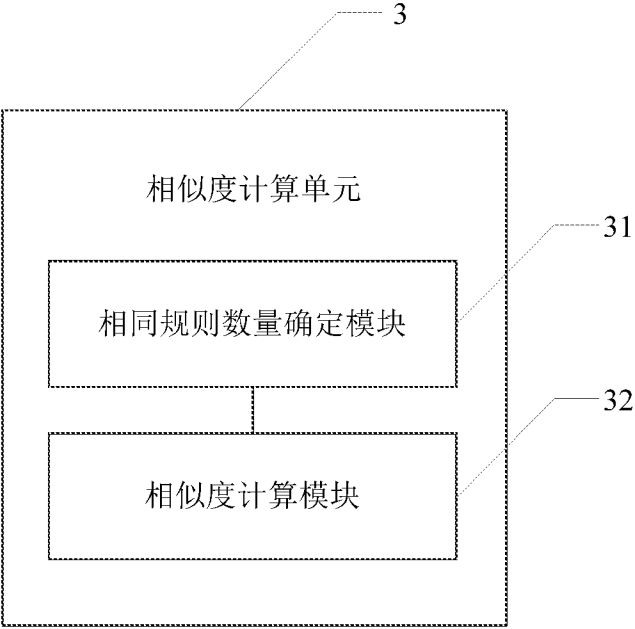


图3

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/CN2017/107407

## A. CLASSIFICATION OF SUBJECT MATTER

G06F 11/07 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DWPI, CNTXT, SIPOABS, CNKI: 数据库, 关联规则, 数据窗口, 相似度, 异常, 变量, database, associat+ rules, data window, similarity

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                     | Relevant to claim No. |
|-----------|------------------------------------------------------------------------------------------------------------------------|-----------------------|
| PX        | CN 106528313 A (CHINA UNIONPAY CO., LTD.), 22 March 2017 (22.03.2017), entire document                                 | 1-18                  |
| A         | CN 105099759 A (SHANGHAI HUAWEI TECHNOLOGIES CO., LTD.), 25 November 2015 (25.11.2015), entire document                | 1-18                  |
| A         | CN 101853277 A (NANJING UNIVERSITY OF INFORMATION SCIENCE & TECHNOLOGY), 06 October 2010 (06.10.2010), entire document | 1-18                  |
| A         | US 9043354 B1 (IQOR US INC.), 26 May 2015 (26.05.2015), entire document                                                | 1-18                  |

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                | “T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention                                              |
| “A” document defining the general state of the art which is not considered to be of particular relevance                                                                | “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone                                                                     |
| “E” earlier application or patent but published on or after the international filing date                                                                               | “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | “&” document member of the same patent family                                                                                                                                                                                                    |
| “O” document referring to an oral disclosure, use, exhibition or other means                                                                                            |                                                                                                                                                                                                                                                  |
| “P” document published prior to the international filing date but later than the priority date claimed                                                                  |                                                                                                                                                                                                                                                  |

|                                                                                                                                                                                                               |                                                                       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Date of the actual completion of the international search<br>09 January 2018                                                                                                                                  | Date of mailing of the international search report<br>24 January 2018 |
| Name and mailing address of the ISA<br>State Intellectual Property Office of the P. R. China<br>No. 6, Xitucheng Road, Jimenqiao<br>Haidian District, Beijing 100088, China<br>Facsimile No. (86-10) 62019451 | Authorized officer<br>YIN, Jianfeng<br>Telephone No. (86-10) 62089526 |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
PCT/CN2017/107407

| Patent Documents referred<br>in the Report | Publication Date | Patent Family | Publication Date |
|--------------------------------------------|------------------|---------------|------------------|
| CN 106528313 A                             | 22 March 2017    | None          |                  |
| CN 105099759 A                             | 25 November 2015 | None          |                  |
| CN 101853277 A                             | 06 October 2010  | None          |                  |
| US 9043354 B1                              | 26 May 2015      | US 8489872 B1 | 16 July 2013     |
|                                            |                  | US 9405814 B1 | 02 August 2016   |
|                                            |                  | US 8812537 B1 | 19 August 2014   |
|                                            |                  | US 9390132 B1 | 12 July 2016     |

## 国际检索报告

国际申请号

PCT/CN2017/107407

## A. 主题的分类

G06F 11/07(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

## B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

DWPI, CNTXT, SIPOABS, CNKI, 数据库, 关联规则, 数据窗口, 相似度, 异常, 变量, database, associat+ rules, data window, similarity

## C. 相关文件

| 类 型* | 引用文件, 必要时, 指明相关段落                                                | 相关的权利要求 |
|------|------------------------------------------------------------------|---------|
| PX   | CN 106528313 A (中国银联股份有限公司) 2017年 3月 22日 (2017 - 03 - 22)<br>全文  | 1-18    |
| A    | CN 105099759 A (上海华为技术有限公司) 2015年 11月 25日 (2015 - 11 - 25)<br>全文 | 1-18    |
| A    | CN 101853277 A (南京信息工程大学) 2010年 10月 6日 (2010 - 10 - 06)<br>全文    | 1-18    |
| A    | US 9043354 B1 (IQOR US INC) 2015年 5月 26日 (2015 - 05 - 26)<br>全文  | 1-18    |

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

\* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&amp;” 同族专利的文件

国际检索实际完成的日期

2018年 1月 9日

国际检索报告邮寄日期

2018年 1月 24日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)  
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

尹剑峰

电话号码 (86-10)62089526

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2017/107407

| 检索报告引用的专利文件 |           |    | 公布日<br>(年/月/日) | 同族专利 |         |    | 公布日<br>(年/月/日) |
|-------------|-----------|----|----------------|------|---------|----|----------------|
| CN          | 106528313 | A  | 2017年 3月 22日   | 无    |         |    |                |
| CN          | 105099759 | A  | 2015年 11月 25日  | 无    |         |    |                |
| CN          | 101853277 | A  | 2010年 10月 6日   | 无    |         |    |                |
| US          | 9043354   | B1 | 2015年 5月 26日   | US   | 8489872 | B1 | 2013年 7月 16日   |
|             |           |    |                | US   | 9405814 | B1 | 2016年 8月 2日    |
|             |           |    |                | US   | 8812537 | B1 | 2014年 8月 19日   |
|             |           |    |                | US   | 9390132 | B1 | 2016年 7月 12日   |

表 PCT/ISA/210 (同族专利附件) (2009年7月)