



(12) 发明专利申请

(10) 申请公布号 CN 116579884 A

(43) 申请公布日 2023. 08. 11

(21) 申请号 202310848427.6

(22) 申请日 2023.07.12

(71) 申请人 武汉振铭科技发展有限公司

地址 430058 湖北省武汉市经济技术开发区2MA地块办公及生产用房(东方工业园5号-6号楼)(集-HCY-5#D15)

申请人 国网湖北省电力有限公司襄阳供电公司
国网湖北省电力有限公司宜城市供电公司

(72) 发明人 叶剑锋 龚艳玲 龚廷 彭军
程航宁 苏高扬 陈俊学 付文蕊
陈鑫 杜克 张鹏超

(74) 专利代理机构 青岛海盈智专利代理事务所
(普通合伙) 37432

专利代理师 杨金凤

(51) Int. Cl.

G06Q 50/06 (2012.01)

G06Q 10/0631 (2023.01)

G06Q 10/04 (2023.01)

H02J 3/00 (2006.01)

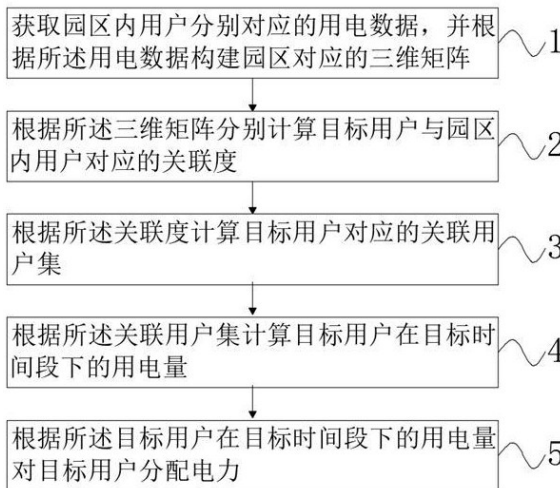
权利要求书3页 说明书8页 附图2页

(54) 发明名称

一种电力用户行为分析方法及系统

(57) 摘要

本发明属于电力控制技术领域,涉及一种电力用户行为分析方法及系统,所述方法包括以下步骤:根据所述关联度计算目标用户对应的关联用户集;根据所述关联用户集计算目标用户在目标时间段下的用电量;根据所述目标用户在目标时间段下的用电量对目标用户分配电力。本发明通过分析目标用户关联度高的用户集的用电行为,可以更精确地预测出目标用户在特定时间段的用电量;根据预测的用电量,可以实现对目标用户的精确电力分配,从而避免园区电力资源的浪费,通过精确预测和分配电力,能有效降低园区电力系统的运行压力,提高其运行效率。



1. 一种电力用户行为分析方法,其特征在于:包括以下步骤:

获取园区内用户分别对应的用电数据,并根据所述用电数据构建园区对应的三维矩阵,包括:设置时间分隔周期,所述时间分隔周期包括至少一个时间分隔区间;获取当前时间分隔周期下至少一个时间分隔区间内对应的用户用电数据;将所述用户用电数据根据用户及时间分隔区间分别对应的唯一标识进行排序和分组,得到当前时间分隔周期下园区对应的三维矩阵;

根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度,包括:根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别对应的第一数据链;根据当前时间分隔周期下至少一个时间分隔区间,获得所述第一数据链的至少一个第一子链;设置匹配阈值,计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度,若所述匹配度小于匹配阈值,则判断目标用户对应的所述第一子链与园区内用户对应的所述第一子链相匹配;将目标用户及园区内用户分别对应的第一数据链中相匹配的第一子链提取,生成第二数据链;遍历所述第二数据链中第一子链对应的时间分隔区间,并将相邻时间分割区间对应的第一子链连接,得到第二子链;将时间长度最大值对应的所述第二子链提取,得到至少一个第三数据链;

根据所述关联度计算目标用户对应的关联用户集;

根据所述关联用户集计算目标用户在目标时间段下的用电量;

根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

2. 根据权利要求1所述的电力用户行为分析方法,其特征在于:根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别对应的第一数据链,包括:

根据前时间分隔周期下园区对应的三维矩阵得到特征点 $xq_{c,p_i}^t = (a, s_{p_i}, t)$, 其中, a 为用户的类别, S_{p_i} 为用户 a 在时间分隔区间 p_i 的用电量, p_i 为第 i 个时间分隔区间, t 为用户 a 在时间分隔区间 p_i 的用电的时间戳;

根据时间戳将所述特征点进行排序,分别得到目标用户及园区内用户对应的第一数据链。

3. 根据权利要求2所述的电力用户行为分析方法,其特征在于:计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度,包括:

通过公式: $f(x) = \left| \frac{s_t^u - \min(s^u)}{\max(s^u) - \min(s^u)} - \frac{s_t^w - \min(s^w)}{\max(s^w) - \min(s^w)} \right|$, 计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度 $f(x)$, 其中, s_t^u 为目标用户 u 在时间戳 t 时的用电量, s_t^w 为园区内用户 w 在时间戳 t 时的用电量, $\min(s^u)$ 和 $\max(s^u)$ 分别为目标用户 u 对应所述第一子链的用电量最小值和最大值, $\min(s^w)$ 和 $\max(s^w)$ 分别为园区内用户 w 对应所述第一子链的用电量最小值和最大值。

4. 根据权利要求3所述的电力用户行为分析方法,其特征在于:根据所述关联度计算目标用户对应的关联用户集,包括:

根据第二数据链的第一子链以及第三数据链的长度依次计算目标用户与园区内不同用户之间的同化度；

根据所述同化度将园区内不同用户进行层状式分布，得到层状式分布图；

计算选取区域，将层状式分布图对应所述选取区域的园区用户进行提取，得到关联用户集。

5. 根据权利要求4所述的电力用户行为分析方法，其特征在于：根据第二数据链的第一子链以及第三数据链的长度依次计算目标用户与园区内不同用户之间的同化度，包括：

通过公式： $sim(u, w) = pr(u, w) \times f(sim_{HH}(u, w))$ ，计算目标用户与园区内不同用户之间的同化度 $sim(u, w)$ ，其中， $pr(u, w)$ 为目标用户 u 和园区用户 w 的皮尔逊系数， $f(sim_{HH}(u, w))$ 为动态权重。

6. 根据权利要求5所述的电力用户行为分析方法，其特征在于：根据所述关联用户集计算目标用户在目标时间段下的用电量，包括：

将目标时间段与上一时间分隔周期下的时间分隔区间相对应，得到目标历史时间分隔区间；

获取所述关联用户集内各元素在所述目标历史时间分隔区间的用电量；

计算目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量；

根据所述关联用户集内各元素在所述目标历史时间分隔区间的用电量，以及目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量，确定目标用户在目标时间段下的用电量。

7. 根据权利要求6所述的电力用户行为分析方法，其特征在于：根据所述关联用户集内各元素在所述目标历史时间分隔区间的用电量，以及目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量，确定目标用户在目标时间段下的用电量，包括：

通过公式： $\tilde{s}_p^u = \sum s^u + \frac{\sum sim(u, v) \times (s_p^v - s^v)}{\sum sim(u, v)}$ ，计算目标用户在目标时间

段下的用电量，其中， \tilde{s}_p^u 为目标用户 u 在目标时间段对应当前时间分隔周期下时间分割区

间的用电量， $\sum s^u$ 为对目标用户在当前时间分隔周期对应时间分割区间的平均用电量，

s^v 为关联用户集中用户 v 在当前时间分隔周期对应时间分割区间的平均用电量， s_p^v 为关联用户集中用户 v 在所述目标历史时间分隔区间的用电量， $sim(u, v)$ 为目标用户与关联用户集中用户 v 之间的同化度。

8. 一种电力用户行为分析系统，其特征在于：所述系统包括：

构建模块(10)，用于获取园区内用户分别对应的用电数据，并根据所述用电数据构建

园区对应的三维矩阵,包括:设置时间分隔周期,所述时间分隔周期包括至少一个时间分隔区间;获取当前时间分隔周期下至少一个时间分隔区间内对应的用户用电数据;将所述用户用电数据根据用户及时间分隔区间分别对应的唯一标识进行排序和分组,得到当前时间分隔周期下园区对应的三维矩阵;

计算模块(20),用于根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度,包括:根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别对应的第一数据链;根据当前时间分隔周期下至少一个时间分隔区间,获得所述第一数据链的至少一个第一子链;设置匹配阈值,计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度,若所述匹配度小于匹配阈值,则判断目标用户对应的所述第一子链与园区内用户对应的所述第一子链相匹配;将目标用户及园区内用户分别对应的第一数据链中相匹配的第一子链提取,生成第二数据链;遍历所述第二数据链中第一子链对应的时间分隔区间,并将相邻时间分割区间对应的第一子链连接,得到第二子链;将时间长度最大值对应的所述第二子链提取,得到至少一个第三数据链;

关联模块(30),用于根据所述关联度计算目标用户对应的关联用户集;

预测模块(40),用于根据所述关联用户集计算目标用户在目标时间段下的用电量;

分配模块(50),用于根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

一种电力用户行为分析方法及系统

技术领域

[0001] 本发明涉及电力控制技术领域,尤其涉及一种电力用户行为分析方法及系统。

背景技术

[0002] 园区电力系统主要是指在园区内,例如工业园区、科技园区、大学城等,为园区内的企业或者机构提供的电力供应服务;园区电力的分配通常需要考虑园区内各个用户的电力需求和消费行为,以实现电力资源的高效利用。

[0003] 在传统的电力分配系统中,对用户的电力需求进行预测和分配通常基于历史数据和经验规则,这种方法在一定程度上可以满足电力系统的运行需求,但这种方法在处理大规模用户数据时,往往需要大量的计算资源和时间,而且难以处理用户之间的相互影响和依赖关系,无法实时反应电力系统的运行状态和用户的用电需求以应对电力需求的突然变化,这不仅降低了电力系统的运行效率,也可能导致电力资源的分配不公平。

[0004] 因此,如何准确预测园区内用户的电力需求,并根据预测结果实现精准的电力分配,是在园区电力控制中亟待解决的问题。

发明内容

[0005] 为了现有技术中:无法准确预测园区内用户的电力需求,并根据预测结果实现精准的电力分配,本发明提供了一种电力用户行为分析方法及系统,可以有效解决背景技术中的问题。

[0006] 为了解决上述技术问题,本发明提供的技术方案具体如下:

第一方面,一种电力用户行为分析方法,包括以下步骤:

获取园区内用户分别对应的用电数据,并根据所述用电数据构建园区对应的三维矩阵;

根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度;

根据所述关联度计算目标用户对应的关联用户集;

根据所述关联用户集计算目标用户在目标时间段下的用电量;

根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

[0007] 在上述任一方案中优选的是,获取园区内用户分别对应的用电数据,并根据所述用电数据构建园区对应的三维矩阵,包括:

设置时间分隔周期,所述时间分隔周期包括至少一个时间分隔区间;

获取当前时间分隔周期下至少一个时间分隔区间内对应的用户用电数据;

将所述用户用电数据根据用户及时间分隔区间分别对应的唯一标识进行排序和分组,得到当前时间分隔周期下园区对应的三维矩阵。

[0008] 在上述任一方案中优选的是,根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度,包括:

根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别

对应的第一数据链；

根据当前时间分隔周期下至少一个时间分隔区间，获得所述第一数据链的至少一个第一子链；

设置匹配阈值，计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度，若所述匹配度小于匹配阈值，则判断目标用户对应的所述第一子链与园区内用户对应的所述第一子链相匹配；

将目标用户及园区内用户分别对应的第一数据链中相匹配的第一子链提取，生成第二数据链；

遍历所述第二数据链中第一子链对应的时间分隔区间，并将相邻时间分割区间对应的第一子链连接，得到第二子链；

将时间长度最大值对应的所述第二子链提取，得到至少一个第三数据链。

[0009] 在上述任一方案中优选的是，根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别对应的第一数据链，包括：

根据前时间分隔周期下园区对应的三维矩阵得到特征点 $xq_{c,p_i}^t = (a, s_{p_i}, t)$ ，

其中， a 为用户的类别， s_{p_i} 为用户 a 在时间分隔区间 p_i 的用电量， p_i 为第 i 个时间分隔区间， t 为用户 a 在时间分隔区间 p_i 的用电的时间戳；

根据时间戳将所述特征点进行排序，分别得到目标用户及园区内用户对应的第一数据链。

[0010] 在上述任一方案中优选的是，计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度，包括：

通过公式： $f(x) = \left| \frac{s_t^u - \min(s^u)}{\max(s^u) - \min(s^u)} - \frac{s_t^w - \min(s^w)}{\max(s^w) - \min(s^w)} \right|$ ，计算目标用户及园区

内用户分别对应的所述第一子链之间的匹配度 $f(x)$ ，其中， s_t^u 为目标用户 u 在时间戳 t 时的用电量， s_t^w 为园区内用户 w 在时间戳 t 时的用电量， $\min(s^u)$ 和 $\max(s^u)$ 分别为目标用户 u 对应所述第一子链的用电量最小值和最大值， $\min(s^w)$ 和 $\max(s^w)$ 分别为园区内用户 w 对应所述第一子链的用电量最小值和最大值。

[0011] 在上述任一方案中优选的是，根据所述关联度计算目标用户对应的关联用户集，包括：

根据第二数据链的第一子链以及第三数据链的长度依次计算目标用户与园区内不同用户之间的同化度；

根据所述同化度将园区内不同用户进行层状式分布，得到层状式分布图；

计算选取区域，将层状式分布图对应所述选取区域的园区用户进行提取，得到关联用户集。

[0012] 在上述任一方案中优选的是，根据第二数据链的第一子链以及第三数据链的长度依次计算目标用户与园区内不同用户之间的同化度，包括：

通过公式： $sim(u, w) = pr(u, w) \times f(sim_{HH}(u, w))$ ，计算目标用户与园区内不同用户之间的同化度 $sim(u, w)$ ，其中， $pr(u, w)$ 为目标用户 u 和园区用户 w 的皮尔逊系数， $f(sim_{HH}(u, w))$ 为动态权重。

[0013] 在上述任一方案中优选的是，根据所述关联用户集计算目标用户在目标时间段下的用电量，包括：

将目标时间段与上一时间分隔周期下的时间分隔区间相对应，得到目标历史时间分隔区间；

获取所述关联用户集内各元素在所述目标历史时间分隔区间的用电量；

计算目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量；

根据所述关联用户集内各元素在所述目标历史时间分隔区间的用电量，以及目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量，确定目标用户在目标时间段下的用电量。

[0014] 在上述任一方案中优选的是，根据所述关联用户集内各元素在所述目标历史时间分隔区间的用电量，以及目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量，确定目标用户在目标时间段下的用电量，包括：

通过公式： $\tilde{s}_p^u = \sum \bar{s}^u + \frac{\sum sim(u, v) \times (s_p^v - \bar{s}^v)}{\sum sim(u, v)}$ ，计算目标用户在目标时间段下的用电量，其中， \tilde{s}_p^u 为目标用户 u 在目标时间段对应当前时间分隔周期下时间分割区间的用电量， $\sum \bar{s}^u$ 为对目标用户在当前时间分隔周期对应时间分割区间的平均用电量， \bar{s}^v 为关联用户集中用户 v 在当前时间分隔周期对应时间分割区间的平均用电量， s_p^v 为关联用户集中用户 v 在所述目标历史时间分隔区间的用电量， $sim(u, v)$ 为目标用户与关联用户集中用户 v 之间的同化度。

[0015] 第二方面，一种电力用户行为分析系统，所述系统包括：
构建模块，用于获取园区内用户分别对应的用电数据，并根据所述用电数据构建园区对应的三维矩阵；
计算模块，用于根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度；
关联模块，用于根据所述关联度计算目标用户对应的关联用户集；
预测模块，用于根据所述关联用户集计算目标用户在目标时间段下的用电量；
分配模块，用于根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

[0015] 第二方面，一种电力用户行为分析系统，所述系统包括：

构建模块，用于获取园区内用户分别对应的用电数据，并根据所述用电数据构建园区对应的三维矩阵；

计算模块，用于根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度；

关联模块，用于根据所述关联度计算目标用户对应的关联用户集；

预测模块，用于根据所述关联用户集计算目标用户在目标时间段下的用电量；

分配模块，用于根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

[0016] 与现有技术相比，本发明的有益效果：

本发明通过分析目标用户关联度高的用户集的用电行为，可以更精确地预测出目标用户在特定时间段的用电量；根据预测的用电量，可以实现对目标用户的精确电力分配，从而避免园区电力资源的浪费，通过精确预测和分配电力，能有效降低园区电力系统的

运行压力,提高其运行效率。

附图说明

[0017] 附图用于对本发明的进一步理解,与本发明的实施例一起用于解释本发明,并不构成对本发明的限制。

[0018] 图1是本发明电力用户行为分析方法的流程示意图;

图2是本发明电力用户行为分析系统的模块示意图。

[0019] 图中标号说明:

10、构建模块;20、计算模块;30、关联模块;40、预测模块;50、分配模块。

具体实施方式

[0020] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,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。

[0021] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征。在本发明的描述中,“多个”的含义是两个或两个以上,除非另有明确具体的限定。

[0022] 为了更好地理解上述技术方案,下面将结合说明书附图及具体实施方式对本发明技术方案进行详细说明。

[0023] 如图1所示,一种电力用户行为分析方法,包括以下步骤:

步骤1,获取园区内用户分别对应的用电数据,并根据所述用电数据构建园区对应的三维矩阵;

步骤2,根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度;

步骤3,根据所述关联度计算目标用户对应的关联用户集;

步骤4,根据所述关联用户集计算目标用户在目标时间段下的用电量;

步骤5,根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

[0024] 应当指出的是,上述步骤仅是优选的实施顺序,在具体实施过程中,在不影响整体实施效果的前提下,部分步骤可以调换,为更加清晰的对本申请的技术方案进行阐述,下述内容以一个优选的方式对本方案进行解释。

[0025] 在本发明另一可选的实施例中,上述步骤1还包括:

步骤11,设置时间分隔周期,所述时间分隔周期包括至少一个时间分隔区间;

步骤12,获取当前时间分隔周期下至少一个时间分隔区间内对应的用户用电数据;

步骤13,将所述用户用电数据根据用户及时间分隔区间分别对应的唯一标识进行排序和分组,得到当前时间分隔周期下园区对应的三维矩阵。

[0026] 在该实施例中,对于上述步骤11中涉及到确定时间分隔周期而言,时间分隔周期可以是按照小时、按照天、按照周或者按照月等来划分。例如按照一天作为时间分隔周期,那么一天中的每个小时就构成了至少一个时间分隔区间,以便于更好的分析和预测用户在

不同时间周期内的用电情况;对于上述步骤13而言,在获取了用户用电数据后,需要将用户用电数据进行排序和分组,排序和分组的依据是用户和时间分隔区间的唯一标识,比如用户ID和日期等,在排序后,可以直接得到园区内的用户在各个时间区间的用电情况;分组后,可以看到园区内各个用户在同一时间区间的用电情况,以便于比较和分析;最后将排序和分组的数据转换为三维矩阵,在这个矩阵中,一个维度代表用户,一个维度代表时间分隔区间,另一个维度代表用电量。

[0027] 在本发明另一可选的实施例中,上述步骤2还包括:

步骤21,根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别对应的第一数据链;

步骤22,根据当前时间分隔周期下至少一个时间分隔区间,获得所述第一数据链的至少一个第一子链;

步骤23,设置匹配阈值,计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度,若所述匹配度小于匹配阈值,则判断目标用户对应的所述第一子链与园区内用户对应的所述第一子链相匹配;

步骤24,将目标用户及园区内用户分别对应的第一数据链中相匹配的第一子链提取,生成第二数据链;

步骤25,遍历所述第二数据链中第一子链对应的时间分隔区间,并将相邻时间分割区间对应的第一子链连接,得到第二子链;

步骤26,将时间长度最大值对应的所述第二子链提取,得到至少一个第三数据链。

[0028] 在该实施例中,通过根据当前时间分隔周期下园区对应的三维矩阵,可以为每个用户(包括目标用户和园区内的其他用户)计算一个第一数据链;所述第一数据链为在一定时间分隔周期下,用户的用电行为序列;对于第一数据链,可以根据时间分隔周期划分出多个第一子链。例如,如果数据链是按照每小时记录的用电量,那么可以将每天的24小时作为一个子链;进而设定一个匹配阈值,然后通过计算目标用户和园区内其他用户的第一数据链中第一子链的匹配度,找到相匹配的第一子链,将它们从各自的第一数据链中提取出来,组成仅包含有目标用户和园区内其他用户相同用电行为的第二数据链,进而再通过上述处理得到最长即时间线连贯度最大的第三数据链,可以更便于判断两用户之间的用电行为相似度。

[0029] 可选的,上述步骤21中,根据当前时间分隔周期下园区对应的三维矩阵计算目标用户及园区内用户分别对应的第一数据链,包括:

根据前时间分隔周期下园区对应的三维矩阵得到特征点 $xq_{c,p_i}^t = (a, s_{p_i}, t)$,

其中, a 为用户的类别, s_{p_i} 为用户 a 在时间分隔区间 p_i 的用电量, p_i 为第 i 个时间分隔区间, t 为用户 a 在时间分隔区间 p_i 的用电的时间戳;

根据时间戳将所述特征点进行排序,分别得到目标用户及园区内用户对应的第一数据链。

[0030] 在该实施例中,通过上述内容,可以把原始的三维矩阵数据转换为更方便处理和数据分析的数据链,提高数据处理的效率。

[0031] 可选的,上述步骤23中,计算目标用户及园区内用户分别对应的所述第一子链之间的匹配度,包括:

通过公式: $f(x) = \left| \frac{s_t^u - \min(s^u)}{\max(s^u) - \min(s^u)} - \frac{s_t^w - \min(s^w)}{\max(s^w) - \min(s^w)} \right|$, 计算目标用户及园区

内用户分别对应的所述第一子链之间的匹配度 $f(x)$,其中, s_t^u 为目标用户 u 在时间戳 t 时的用电量, s_t^w 为园区内用户 w 在时间戳 t 时的用电量, $\min(s^u)$ 和 $\max(s^u)$ 分别为目标用户 u 对应所述第一子链的用电量最小值和最大值, $\min(s^w)$ 和 $\max(s^w)$ 分别为园区内用户 w 对应所述第一子链的用电量最小值和最大值。

[0032] 在该实施例中,因为不同用户的用电量可能存在较大的差异,直接比较可能无法反映出其实际的匹配程度,通过上述公式可以对目标用户及园区内用户的用电量进行归一化处理,使得不同用户在相同时间戳下的用电量可以在相同标准下进行对比;并且可以计算出目标用户与园区内其他用户在同一时间段的用电模式的相似度,从而可以识别出具有相似用电模式的用户。

[0033] 在本发明另一可选的实施例中,上述步骤3可以包括:

步骤31,根据第二数据链的第一子链以及第三数据链的长度依次计算目标用户与园区内不同用户之间的同化度;

步骤32,根据所述同化度将园区内不同用户进行层状式分布,得到层状式分布图;

步骤33,计算选取区域,将层状式分布图对应所述选取区域的园区用户进行提取,得到关联用户集。

[0034] 可选的,上述步骤31中,根据第二数据链的第一子链以及第三数据链的长度依次计算目标用户与园区内不同用户之间的同化度,包括:

通过公式: $sim(u, w) = pr(u, w) \times f(sim_{HH}(u, w))$, 计算目标用户与园区内不同用户之间的同化度 $sim(u, w)$,其中, $pr(u, w)$ 为目标用户 u 和园区用户 w 的皮尔逊系数, $f(sim_{HH}(u, w))$ 为动态权重。

[0035] 在该实施例中,通过公式:

$$f(sim_{HH}(u, w)) = \lambda \cdot e^{\frac{|ct(u, w)|}{\sqrt{|ct(u, u)| \cdot |ct(w, w)|}}} + (1 - \lambda) \cdot e^{\frac{|l(u, w)|}{\sqrt{|l(u, u)| \cdot |l(w, w)|}}},$$

计算动态权重 $f(sim_{HH}(u, w))$,其中, λ 为常数, $|ct(u, w)|$ 为目标用户 u 与园区内用户 w 的第二数据链中第一子链的个数, $|ct(u, u)|$ 为目标用户 u 的第一数据链中第一子链的个数, $|ct(w, w)|$ 为园区内用户 w 的第一数据链中第一子链的个数, $|l(u, w)|$ 为目标用户 u 与园区内用户 w 的第三数据链的长度, $|l(u, u)|$ 为目标用户 u 的第一数据链的长度, $|l(w, w)|$ 为目标用户 w 的第一数据链的长度;在本实施例中,所述第二数据链中的第一子链的数量越

多,则反映目标用户与园区内用户的相似度越高,所述第三数据链的长度越长,则也能反映目标用户与园区内用户的相似度越高,因此,可通过上述公式,将所述第二数据链和所述第三数据链的特征进行融合,从而得到一个相较于一般相似度算法更为准确的计算结果,进而通过将融合后的相似度计算结果集成到皮尔逊系数,可以得到最终同化度计算结果,能够充分准确计算出目标用户与园区内用户的用电行为的相同性。

[0036] 可选的,上述步骤32中,根据所述同化度将园区内不同用户进行层状式分布,得到层状式分布图,包括:

设置同化度阈值,获取满足同化度阈值的园区内用户数量 k ;

取 k 的平方根值,并将 k 的平方根值四舍五入至最近整数 m ;

将判断 k/m 是否为整数,若非整数则将 k/m 四舍五入至最近整数 Y ,并设置 m 为层状式分布的层级数量, Y 为每层分布的用户数量;

将用户根据用户的同化度从高到低进行排序,并根据每层的用户数量 Y ,得到每层的用户同化度范围;

根据每层的用户同化度范围,将用户分配到相应的层级上,得到层状式分布图。

[0037] 可选的,上述步骤33中,计算选取区域,将层状式分布图对应所述选取区域的园区用户进行提取,得到关联用户集,包括:

获取层状式分布图中用户同化度范围最小值满足同化度阈值的最大层号 U ;

通过公式: $ST = b_1 \cdot U + (U - 1)b_2$, 计算选取区域的面积 ST ,其中, b_1 为层宽,

b_2 为层间距;

将层状式分布图对应所述面积 ST 的园区用户进行提取,得到关联用户集。

[0038] 在该实施例中,通过上述内容,可以在园区用户数量非常多的情况下,更高效地处理大规模的用户数据;通过根据用户的同化度进行层状分布,每一层的用户同化度都有一个明确的范围,可以更精确地反映用户之间的差异,可以更快速的找到与目标用户更加相似的用户群体,节约计算资源和时间。

[0039] 在本发明另一可选的实施例中,上述步骤4,包括:

步骤41,将目标时间段与上一时间分隔周期下的时间分隔区间相对应,得到目标历史时间分隔区间;

步骤42,获取所述关联用户集内各元素在所述目标历史时间分隔区间的用电量;

步骤43,计算目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量;

步骤44,根据所述关联用户集内各元素在所述目标历史时间分隔区间的用电量,以及目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量,确定目标用户在目标时间段下的用电量。

[0040] 进一步的,步骤44,根据所述关联用户集内各元素在所述目标历史时间分隔区间的用电量,以及目标用户与所述关联用户集内各元素对于当前分割周期的平均用电量,确定目标用户在目标时间段下的用电量,包括:

通过公式： $\tilde{s}_p^u = \sum s^u + \frac{\sum \text{sim}(u,v) \times (s_p^v - s^v)}{\sum \text{sim}(u,v)}$ ，计算目标用户在目标时间段下的用电量，其中， \tilde{s}_p^u 为目标用户u在目标时间段对应当前时间分隔周期下时间分割区间的用电量， $\sum s^u$ 为对目标用户在当前时间分隔周期对应时间分割区间的平均用电量， s^v 为关联用户集中用户v在当前时间分隔周期对应时间分割区间的平均用电量， s_p^v 为关联用户集中用户v在所述目标历史时间分隔区间的用电量， $\text{sim}(u,v)$ 为目标用户与关联用户集中用户v之间的同化度。

[0041] 在该实施例中，通过上述公式可以将目标用户的历史用电量和关联用户集的同化度与用电量进行集合，能够提高对目标用户在目标时间段的用电量预测准确率，进而根据上述步骤5，在目标时间段对目标用户分配合理电力；可以根据实际应用场景进行动态调整，扩大了整体算法的应用场景。

[0042] 如图2所示，本发明还提供了一种电力用户行为分析系统，所述系统包括：

构建模块10，用于获取园区内用户分别对应的用电数据，并根据所述用电数据构建园区对应的三维矩阵；

计算模块20，用于根据所述三维矩阵分别计算目标用户与园区内用户对应的关联度；

关联模块30，用于根据所述关联度计算目标用户对应的关联用户集；

预测模块40，用于根据所述关联用户集计算目标用户在目标时间段下的用电量；

分配模块50，用于根据所述目标用户在目标时间段下的用电量对目标用户分配电力。

[0043] 以上仅为本发明的优选实施例而已，并不用于限制本发明，尽管参照前述实施例对本发明进行了详细的说明，对于本领域的技术人员来说，其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

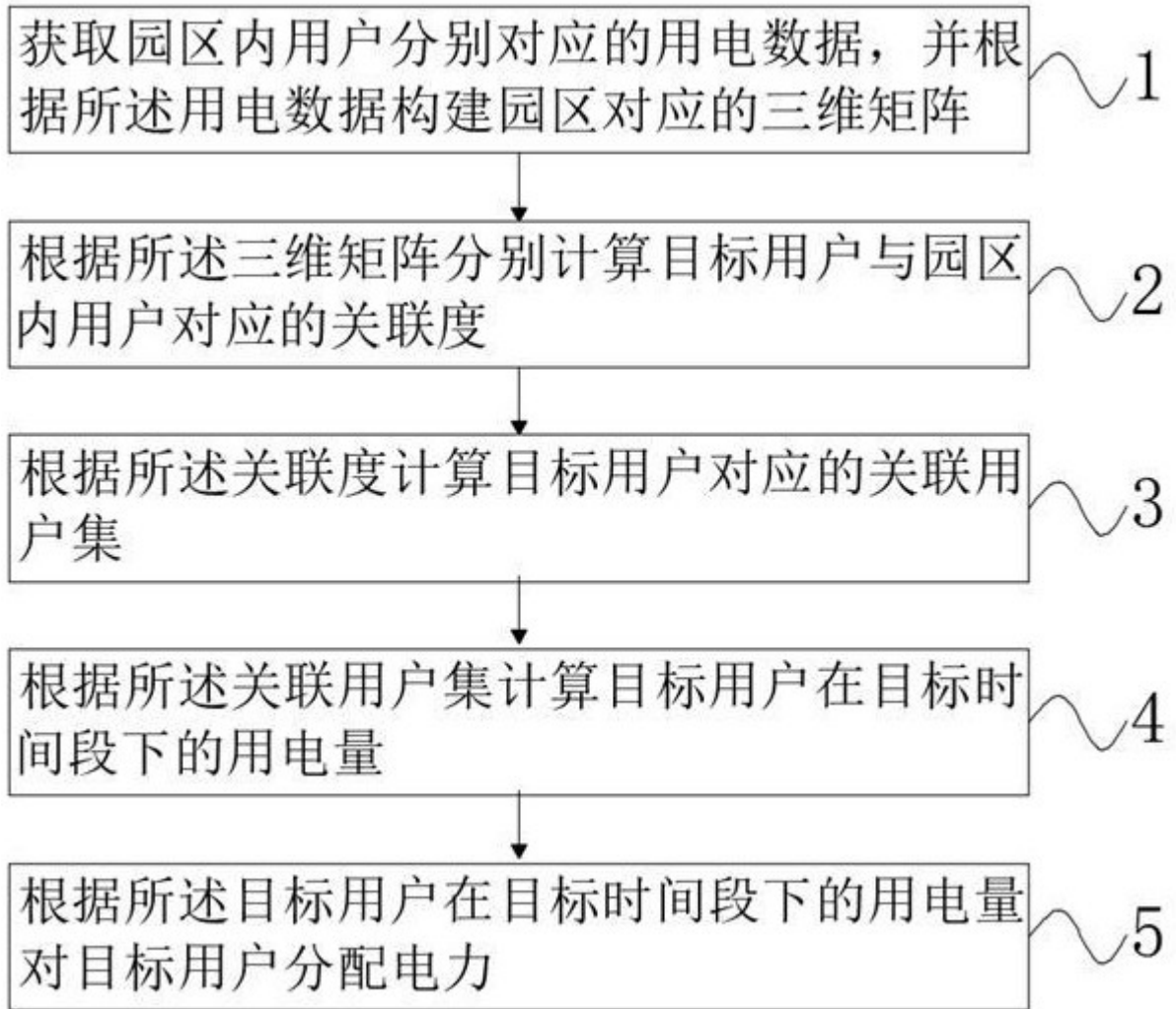


图 1

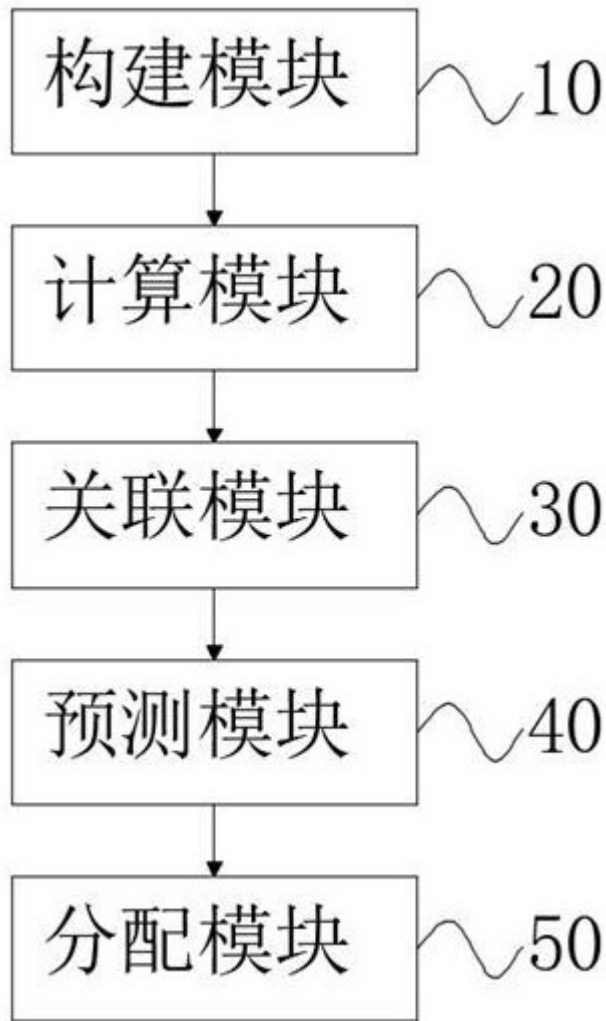


图 2