

(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
2016 年 8 月 11 日 (11.08.2016)



(10) 国际公布号
WO 2016/124116 A1

- (51) 国际专利分类号:
H04L 12/58 (2006.01)
- (21) 国际申请号: PCT/CN2016/072783
- (22) 国际申请日: 2016 年 1 月 29 日 (29.01.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510058167.8 2015 年 2 月 4 日 (04.02.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属开曼群岛大开曼乔治城资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 李朝 (LI, Chao) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 王志荣 (WANG, Zhirong) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。
- (74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.);

中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100028 (CN)。

- (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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))。

(54) Title: METHOD AND DEVICE FOR DISSEMINATING INFORMATION

(54) 发明名称: 信息传播方法和装置

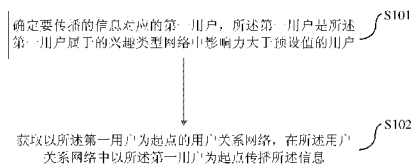


图 1

S101 First user corresponding to information to be disseminated is determined; first user is user, who has influence greater than preset value, in network of type of interest to which first user belongs
S102 User relationship network having first user as start point is obtained, and information in user relationship network is disseminated, first user being start point

(57) Abstract: The present invention provides a method and device for disseminating information, said information dissemination method comprising: determining a first user corresponding to information to be disseminated, said first user being a user whose influence in an interest-type network to which said user belongs is greater than a preset value; obtaining a user relationship network having the first user as its start point, and disseminating said information in the user relationship network, the first user being the start point. The method improves efficiency and credibility in information dissemination.

(57) 摘要: 本发明提出一种信息传播方法和装置, 该信息传播方法包括: 确定要传播的信息对应的第一用户, 所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户; 获取以所述第一用户为起点的用户关系网络, 在所述用户关系网络中以所述第一用户为起点传播所述信息。该方法能够提高信息传播的效率和可信性。



WO 2016/124116 A1

信息传播方法和装置

技术领域

本发明涉及互联网技术领域，尤其涉及一种信息传播方法和装置。

5

背景技术

随着社会信息化的发展，许多信息都需要得到有效的传播。几年来，社交网络已经成为人们获取、分享信息的主要渠道。通过社交网络传播信息，例如通过用户之间的信息分享来传播信息等，变得更容易被用户所接受。由于社交网络中的信息传播还处于初步阶段，许多信息传播的因素，例如信息传播速度、信息传播范围等参数，还处于难以预测的状态。目前，信息传播时可以采用专门的传播方式，例如，广告、营销推广等，但是，这种传播方式不容易被用户接受，效率不高。

现有技术中，可通过建立概率模型学习用户之间的信息传播概率来控制信息的传播。在传播概率学习的过程中，可以利用最大期望模型 EM(Expectation-maximization)来学习用户间的传播概率。但由于数据的稀疏性导致数据分布不均匀，EM 模型方法很容易计算得到概率为 0 或概率为 1 的极端概率情况，导致得到的传播概率往往方差比较大，实际应用后得到的传播效率仍然不高。

发明内容

本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

为此，本发明的一个目的在于提出一种信息传播方法，该方法可以提高信息传播的效率和可信性。

本发明的另一个目的在于提出一种信息传播装置。

为达到上述目的，本发明实施例提出的信息传播方法，包括：确定要传播的信息对应的第一用户，所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户；获取以所述第一用户为起点的用户关系网络，在所述用户关系网络中以所述第一用户为起点传播所述信息。

本发明实施例提出的信息传播方法，通过确定要传播的信息对应的第一用户，第一用户是影响力大于预设值的用户，并由第一用户为起点进行信息传播，可以由具有较大影响力的用户传播信息，提高信息传播的可信性，提高信息传播效率。

为达到上述目的，本发明实施例提出的信息传播装置，包括：确定模块，用于确定要

传播的信息对应的第一用户，所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户；传播模块，用于获取以所述第一用户为起点的用户关系网络，在所述用户关系网络中以所述第一用户为起点传播所述信息。

5 本发明实施例提出的信息传播装置，通过确定要传播的信息对应的第一用户，第一用户是影响力大于预设值的用户，并由第一用户为起点进行信息传播，可以由具有较大影响力的用户传播信息，提高信息传播的可信性，提高信息传播效率。

本发明附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本发明的实践了解到。

10 附图说明

本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图 1 是本发明一实施例提出的信息传播方法的流程示意图；

图 2 是本发明一实施例的兴趣类型网络的示意图；

15 图 3 是本发明一实施例的建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户的流程示意图；

图 4 是本发明一实施例的确定要传播的信息对应的第一用户的示意图；

图 5 是本发明一实施例的用户关系网络传播概率的示意图；

图 6 是本发明一实施例的获取用户之间的传播概率的流程示意图；

20 图 7 是本发明另一实施例的信息传播装置的结构示意图；

图 8 是本发明另一实施例的信息传播装置的结构示意图。

具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施

25 例是示例性的，仅用于解释本发明，而不能理解为对本发明的限制。相反，本发明的实施例包括落入所附加权利要求书的精神和内涵范围内的所有变化、修改和等同物。

下面参考附图描述根据本发明实施例的信息传播方法和装置。

图 1 是本发明一实施例提出的信息传播方法的流程示意图，该方法包括：

30 S101：确定要传播的信息对应的第一用户，所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户。

其中，要传播的信息可以是商品推广信息，也可以是其他信息，本发明对此不做

限定。要传播的信息对应的第一用户可以是一个或多个。

兴趣类型网络是基于用户的兴趣对用户进行划分后得到的类别的名称，用户的兴趣可以根据用户具有的标签确定的，用户具有的标签可以根据用户的购买或浏览历史商品信息等预先确定。

5 具体地，可以预先建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户。例如，可以预先设置多个标签，按照标签将用户划分到不同的兴趣类型网络中。如图 2 所示，兴趣类型网络包括标签，例如时尚、户外、商务、运动、旅行、电子等，每个用户都可以对应一个或多个标签。

10 第一用户是兴趣类型网络中影响力大于预设值的用户。影响力是用户的一种属性，在本实施例中，一个用户的影响力用于衡量该用户传播的信息被其他人接受的难易程度，其中，影响力大的用户传播的信息更容易被其他人接受。第一用户也可以称为达人。每个兴趣类型网络中达人可以是一个或者多个。

可选的，以要传播的信息是商品的信息为例，如图 3 所示，建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户，具体可以包括：

15 S31：根据标签传播学习算法，获取用户-标签矩阵；

具体的，根据标签传播学习算法，获取用户-标签矩阵，可以包括：

（1）计算得到商品与商品的相似度矩阵 W 。

商品与商品的相似度矩阵可以用于表示商品之间在用户行为、商品标题和商品属性上的相似度。

20 其中，用于计算相似度矩阵的商品可以是优质买家处理过的商品，处理具体可以是指购买，浏览，点击，收藏中的一项或者多项，优质买家可以根据优质买家模型确定，例如，将信用等级高或者购买次数多的买家确定为优质买家。具体的，可以获取所有买家的信息，再根据优质买家模型从所有买家中确定出优质买家，再获取优质买家处理过的商品，再根据优质买家处理过的商品中的两两商品计算相似度，得到相似
25 度矩阵 W 。

具体的，可以通过最小哈希算法对商品(pid,vid)进行哈希映射，得到商品与商品的相似度矩阵，其中，pid 是商品的 ID (Identity，身份标识)，vid 是商品属性值的 ID，pid 和 vid 通常可以从基础数据表中获取。

（2）计算得到商品-标签信息矩阵 F 。

30 其中，商品-标签信息矩阵 F 中的商品也可以具体是指优质买家处理过的商品，标签是指商品更新后的标签，在获取优质买家处理过的商品后，可以根据每个商品的初始标签，经过迭代过程计算得到商品-标签信息矩阵 F ，其中，每个商品的初始标签可

以是作为商品的一个属性预先记录在数据库中的，从而可以从数据库中获取商品的初始标签。

具体的，商品-标签信息矩阵 F 可以根据标签传播学习算法的迭代公式得到，迭代公式如下：

5 While (F 收敛)

$$F(t+1) = \alpha SF(t) + (1-\alpha) Y$$

end

其中，要计算的商品-标签信息矩阵 F 是上述公式中收敛时得到 $F(t+1)$ ， $0 \leq \alpha \leq 1$ 为预设的加权参数， S 是根据上述的商品与商品的相似度矩阵 W 计算得到的，

10 $S = D^{-1}W \in R^{n \times n}$ ， $D = \text{diag}\{D_{11}, \dots, D_{nn}\} \in R^{n \times n}$ ， $D_{ii} = \sum_{j=1}^n W_{ij} = \sum_{j=1}^n W_{ji}$ ， Y 是初始标签值， $F(t)$ 的初始值可以是根据已有的买家信息得到的商品-标签信息矩阵 F 的初始值，已有的买家信息可以是根据优质买家模型得到的。例如，根据买家的信用等级从多个买家中确定出预设个数的优质买家，再根据优质买家与优质买家对应的购买，点击或收藏的商品可以得到用户-商品信息矩阵 V ，以及，根据优质买家购买，点击或收藏的商品

15 与商品具有的标签可以得到上述的商品-标签信息矩阵 F 的初始值。

其中，商品具有的标签可以根据统计或者 HITS (Hyperlink-Induced Topic Search, 超链诱导主题搜索) 排序算法得到。

在获取 F 的初始值后，可以根据上述迭代公式，在满足迭代收敛条件时得到最终的商品-标签信息矩阵 F 。

20 迭代收敛条件可以包括：设置最大迭代次数，当迭代次数达到最大迭代次数时满足迭代收敛条件；或者，根据迭代后的值与迭代前的值的差值，在该差值大于预设阈值时满足迭代收敛条件，例如， $\|F(t+1)-F(t)\| < \beta$ 时表明满足迭代收敛条件， $\|F(t+1)-F(t)\|$ 表示 $F(t+1)$ 与 $F(t)$ 的欧氏距离， β 表示预设阈值。

(3) 计算得到用户-标签矩阵 L 。

25 其中，用户-标签矩阵 L 中的用户也可以具体是指优质买家，标签是指用户具有的标签，用户具有的标签可以根据用户处理过的商品的更新后的标签确定。

具体的，可以采用如上所示的方式确定出优质买家，以及获取优质买家处理过的商品，以及，从数据库中获取优质买家处理过的商品的初始标签，之后可以根据优质买家处理过的商品以及上述的方式 (1) 计算得到商品与商品的相似度矩阵 W ，再根据

30 商品与商品的相似度矩阵 W 和优质买家处理过的商品具有的初始标签以及上述的方式

(2) 计算得到商品-标签信息矩阵 F ，再根据优质买家以及优质买家处理过的商品可以建立用户-商品信息矩阵 V ，之后在根据上述的 V 和 F 采用如下的方式得到用户-标签

矩阵 L 。

具体的，计算公式可以是： $L=V \cdot F$ ，其中， V 是上述得到的用户-商品信息矩阵， F 是上述得到的收敛时的最终的商品-标签信息矩阵。

5 S32：对所述用户-标签矩阵进行聚类，得到预设个数的兴趣类型网络，并获取每个兴趣类型网络中的第一用户。

在得到用户-标签矩阵 L 后，可以对该矩阵 L 进行聚类，例如，预设个数是 k 个，则可以对矩阵 L 进行双聚类得到 k 个类别，每个类别对应一个兴趣类型网络。

在对矩阵 L 进行聚类得到 k 个类别后，以每个兴趣类型网络包括一个第一用户为例，每个类别的中心点可以确定为该兴趣类型网络的第一用户。不同兴趣类型网络的第一用户可以组成列表，该列表可以称为达人列表，达人列表例如表示为：
10 $P=\{p_1, p_2, \dots, p_k\}$ ，其中， p_i ($i=1, 2, \dots, k$) 是第 i 个兴趣类型网络中的第一用户，也可以称为达人， p_i 可以由用户 ID 和该用户具有的标签组成。

在上述预先建立多个兴趣类型网络，并确定每个兴趣类型网络中的第一用户后，如上所述，可以得到由不同的兴趣类型网络中的第一用户组成的达人列表，达人列表中包括不同兴趣类型网络中的第一用户，在当前需要传播信息时，可以首先确定要传播的信息对应的第一用户。
15

可选的，确定要传播的信息对应的第一用户，包括：

获取第一标签，所述第一标签是所述要传播的信息包括的标签；

将包括所述第一标签的第一用户，确定为所述要传播的信息对应的第一用户。

20 例如，假设第一用户称为达人，如图 4 所示，达人列表中包括：服装达人，3C 达人和家居达人，则如果要传播的信息包括的标签是 3C，则该要传播的信息对应的第一用户是 3C 达人。

25 S102：获取以所述第一用户为起点的用户关系网络，在所述用户关系网络中以所述第一用户为起点传播所述信息。

其中，用户关系网络是用于描述用户与用户之间的关联关系的网络，可以直接从已有的社交网络类型的应用程序中获取用户关系网络，在社交网络类型的应用程序中，用户之间可以通过增加好友或者增加关注等方式预先建立用户关系网络。例如，可以先从第一用户的应用程序中获取到第一用户的好友包括第二用户，再在第二用户的应用程序中获取到第二用户的好友包括第三用户，则可以获取到的用户关系网络包括：
30 第一用户—>第二用户—>第三用户。

以所述第一用户为起点的用户关系网络可以从应用程序的已有数据中导入，例如，从

社交网络的应用程序中导入以确定出的第一用户为起点的用户关系网络。

例如，如图 4 所示，假设要传播的信息对应的第一用户是 3C 达人，从已有数据中获取的以 3C 达人为起点的用户关系网络是用户关系网络 41，则如图 4 所示，则可以将要传播的信息以 3C 达人为起点在用户关系网络 41 中传播。

5 可选的，所述在所述用户关系网络中以所述第一用户为起点传播所述信息，包括：

根据预设策略，在所述用户关系网络中以所述第一用户为起点传播所述信息，所述预设策略包括传播范围策略，或者，传播速度策略。

其中，传播范围策略是指优先考虑传播范围，传播速度策略是指优先考虑传播速度。

10 具体的，可以获取用户关系网络中用户与用户之间的传播概率，当采用传播范围策略时，可以不论传播概率高低都进行信息传播，当采用传播速度策略时，可以只在传播概率大于预设值的路径上进行信息传播。

例如，以传播速度策略为例，参见图 5，假设用户关系网络包括第一路径 51，第二路径 52，第三路径 53，第四路径 54 和第五路径 55，假设第一路径 51，第二路径 52 和第三路径 53 中包括的用户之间的传播概率都大于预设值，而第四路径 54 和第五路径 55 上包括的
15 的用户之间存在小于预设值的传播概率，则信息可以在第一路径 51，第二路径 52 和第三路径 53 上传播，而不在第四路径 54 和第五路径 55 上传播。

具体的，信息在用户关系网络中传播时，第一用户在初始时刻作为信息传播的种子节点，种子节点负责向其邻居节点传播信息，例如，第一用户是 3C 达人，与 3C 达人相邻的邻居节点包括第一节点和第二节点，则在初始时刻 t 设置 3C 达人是种子节点，并且由 3C
20 达人将信息传播给第一节点和第二节点，当种子节点将信息传播给邻居节点后，邻居节点在下一时间成为新的种子节点，例如，在 $t+1$ 时刻种子节点是第一节点，而不再是 3C 达人，依次类推，从初始的第一用户依次根据用户关系网络中的用户相邻关系进行信息传播，直至没有新的种子节点。另外，用户关系网络中邻居节点用户之间的传播概率是独立的，不受其他邻居节点之间的关系影响。并且，每个种子节点只有一次机会向非种子邻居节点传
25 播信息，例如，用户在 t 时刻成为种子节点，仅在 t 时刻有一次机会尝试对非种子邻居节点传播信息，如果传播成功，则该邻居节点成为 $t+1$ 时刻的种子节点，而不管该用户在 t 时刻是否传播成功，该用户再也不能在其他时刻试图传播信息给它的邻居节点。如果在同一时刻，有多个种子节点试图传播信息给同一节点，其传播的顺序可以是任意的。

可选的，所述获取所述用户关系网络中用户与用户之间的传播概率，包括：

30 根据引入传播概率方差控制因子的传播概率学习模型，获取所述用户关系网络中用户与用户之间的传播概率。

例如，传播概率学习模型可以是 EM (Expectation-maximization, 最大期望) 模型。

由于数据的稀疏性，在传播概率学习的过程中，根据 EM 模型学习到的传播概率往往方差比较大。这主要是因为 EM 模型的计算方法在稀疏数据情况下过拟合导致数据分布不均匀，很容易估计获得概率为 0 或概率为 1 的极端概率情况。

本申请实施例中，为了解决传统 EM 模型存在的上述问题，在 EM 模型中引入了传播
5 概率方差控制因子，防止 EM 模型在迭代过程中发生剧烈的波动。

可选的，所述根据引入传播概率方差控制因子的传播概率学习模型，获取所述用户关系网络中用户与用户之间的传播概率，包括：

获取所述用户关系网络，并根据所述用户关系网络和时间片断数据建立信息传播模型，所述时间片段数据是预设的信息传播扩散时间；

10 将传播概率方差控制因子引入传播概率学习模型中，得到引入传播概率方差控制因子的传播概率学习模型，并根据所述引入传播概率方差控制因子的传播概率学习模型，对所述信息传播模型进行学习，获取传播概率更新规则，所述更新规则包括第一更新规则和第二更新规则；

采用所述第一更新规则更新第一组用户之间的传播概率，采用所述第二更新规则更新
15 第二组用户之间的传播概率，所述第一组用户之间的边在所述时间片断数据内被激活，所述第二组用户之间的边在所述时间片断数据内没有被激活；

将更新后的用户之间的传播概率确定为所述用户关系网络中用户与用户之间的传播概率。

具体的，如图 6 所示，获取用户之间的传播概率的流程可以包括：

20 S61：导入用户关系网络。

例如，从已有的社交网络的应用程序中导入用户关系网络。

S62：建立独立级联模型。

独立级联模型是一种基本的传播模型，可以采用现有的方式根据用户关系网络建立。

在传播模型中，可以包括节点和边，其中，每个节点可以对应用户关系网络中的一个
25 用户，每个边是由用户关系网络中两个相邻用户组成的线段。

S63：将传播概率方差控制因子引入 EM 模型中。

EM (Expectation-maximization, 最大化期望) 模型是一种优化算法，在本实施例中，可以采用 EM 模型对独立级联模型进行学习，从而得到独立级联模型中包括的每个边的传播概率，也就是用户关系网络中用户与用户之间的传播概率。

30 传统的 EM 模型可以表示为：

$$L(\theta) = \sum_{s=1}^S \log L(q | D_s) = \sum_{s=1}^S \sum_{t=0}^{T-1} \left[\sum_{w \in D_s(t+1)} \log P_w^s + \sum_{v \in D_s(t)} \sum_{w \in Fv \setminus C(t+1)} \log(1 - k_{v,w}) \right]$$

在引入传播概率方差控制因子后，可以根据求解过程是否收敛得到不同的引入传播概率方差控制因子的 EM 模型，采用哪种引入传播概率方差控制因子的 EM 模型可以根据实际需要确定，具体的，引入传播概率方差控制因子的 EM 模型可以是：

$$L_{\lambda}(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum [\theta_i - E(\theta_i)]^2$$

5 或者，

$$L_{\lambda}(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum |\log \theta_i - E(\log \theta_i)|$$

其中， λ 是控制因子， $k_{v,w}$ 是边(v,w)的传播概率。

S64：根据引入传播概率方差控制因子的 EM 模型获取第一更新规则和第二更新规则。

10 其中，可以先根据引入 λ 的 EM 模型确定优化方程，再对优化方程进行求解，得到第一更新规则。

具体的，如果引入 λ 的 EM 模型是：

$$L_{\lambda}(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum [\theta_i - E(\theta_i)]^2$$

其对应的优化方程是：

$$\begin{aligned} Q_{\lambda}(\theta|\theta) = & (1-\lambda) \sum_{s=1}^S \sum_{t=0}^{T-1} \sum_{v \in D_s(t)} \left(\sum_{w \in F(v) \cap D_s(t+1)} \left(\frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \frac{1}{k_{v,w}} + \left(1 - \frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \right) \frac{1}{k_{v,w} - 1} \right) + \sum_{w \in F(v) \setminus C(t+1)} \frac{1}{k_{v,w} - 1} \right) \\ & - \frac{2\lambda}{m} \sum_{(v,w) \in E} \left(k_{v,w} - E(\hat{k}_{v,w}) \right) = 0 \end{aligned}$$

15 对该优化方程进行求解后，得到第一更新规则是：

$$k_{v,w} = (1-\lambda) \frac{1}{|S_{v,w}^+| + |S_{v,w}^-|} \sum_{s \in S_{v,w}^+} \frac{\hat{k}_{v,w}}{\hat{P}_w(s)} + \lambda E(\hat{k}_{v,w})$$

其中， $|S_{v,w}^+|$ 表示 $v \in D_s(t), w \in D_s(t+1)$ ， $|S_{v,w}^-|$ 表示 $v \in D_s(t), w \notin D_s(t+1)$ ， $D_s(t)$ 表示在 t 时刻激活的点的集合， $P_w(s)$ 表示 w 被激活的概率。

对该优化方程进行求解后，得到第二更新规则是：

$$20 \quad -\frac{2\lambda}{m} \sum_{(v,w) \in E} \left(k_{v,w} - E(\hat{k}_{v,w}) \right) = 0$$

得到，

$$k_{v,w} = E(\hat{k}_{v,w})$$

如果引入 λ 的 EM 模型是:

$$L_{\lambda}(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum |\log(\theta_i) - E(\log(\theta_i))|$$

其对应的优化方程是:

$$\begin{aligned} Q_{\lambda}(\theta|\theta) = & (1-\lambda) \sum_{s=1}^S \sum_{t=0}^{T-1} \sum_{v \in D_s(t)} \left(\sum_{w \in F(v) \cap D_s(t+1)} \left(\frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \log k_{v,w} \right. \right. \\ & \left. \left. + \left(1 - \frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \right) \log(1 - k_{v,w}) \right) + \sum_{w \in F(v) \setminus C(t+1)} \log(1 - k_{v,w}) \right) \\ & - \frac{\lambda}{m} \sum_{(v,w) \in E} \left| \log k_{v,w} - E(\log \hat{k}_{v,w}) \right| \end{aligned}$$

5

对该优化方程进行求解后, 得到第一更新规则是:

$$k_{v,w} = \begin{cases} \frac{(1-\lambda) \sum_{a \in S_{v,w}^+} \frac{\hat{k}_{v,w}}{\hat{P}_w(a)} - \frac{\lambda}{m}}{(1-\lambda)(|S_{v,w}^+| + |S_{v,w}^-|) - \frac{\lambda}{m}} & \text{if } \hat{k}_{v,w} \geq \left(\prod_{(p,q) \in E} \hat{k}_{p,q} \right)^{1/|E|} \\ \frac{(1-\lambda) \sum_{a \in S_{v,w}^+} \frac{\hat{k}_{v,w}}{\hat{P}_w(a)} + \frac{\lambda}{m}}{(1-\lambda)(|S_{v,w}^+| + |S_{v,w}^-|) + \frac{\lambda}{m}} & \text{if } \hat{k}_{v,w} < \left(\prod_{(p,q) \in E} \hat{k}_{p,q} \right)^{1/|E|} \end{cases}$$

对该优化方程进行求解后, 得到第二更新规则是:

$$k_{v,w} = \left(\prod_{(p,q) \in E} \hat{k}_{p,q} \right)^{1/|E|}$$

10

S65: 判断时间片断数据是否结束, 若否, 执行 S66, 若是, 执行 S68。

其中, 时间片断数据是预设的, 用于表明信息传播扩散时间。

在得到第一更新规则和第二更新规则后, 可以在用户关系网络中选取种子节点, 然后以种子节点为起点根据用户关系网络传播预设信息, 传播时间是预设的时间片断数据。

具体的，可以得到当前时间与信息开始传播的时间之间的差值，如果该差值小于预设的时间片断数据，则确定时间片断数据没有结束，否则结束。

S66：判断要计算的边在该时间片断数据内是否被激活，若是，执行 S67，否则，重复执行 S65 及其后续步骤。

5 例如，要计算的边是用户 A 与用户 B 组成的边，在信息传播时间内，传播的信息经过用户 A 和用户 B，则可以确定用户 A 和用户 B 组成的边在该时间内被激活了，否则未激活。

S67：采用第一更新规则，对该要计算的边的传播概率进行更新，之后，执行 S69。

其中，第一更新规则的具体公式可以参见上述描述。

另外，每个边可以设置初始传播概率。

10 S68：采用第二更新规则，对整个时间片断数据内未被激活的边的传播概率进行更新，之后，执行 S69。

例如，在整个预设的时间片断数据内，用户 A 和用户 C 组成的边都没有被激活，也就是信息没有用户在用户 A 和用户 C 之间传播，则可以采用如上所示的第二更新规则对用户 A 和用户 C 组成的边的传播概率进行更新。

15 S69：将每个边更新后的传播概率写入传播概率更新库。

可以理解的是，上述以传播概率学习模型是 EM 模型为例，传播概率学习模型也可以是其他模型，例如，马尔科夫模型。

本实施例中，通过确定要传播的信息对应的第一用户，第一用户是影响力大于预设值的用户，并由第一用户为起点进行信息传播，可以由具有较大影响力的用户传播
20 信息，提高信息传播的可信性，提高信息传播效率。本实施例通过标签传播学习算法可以确定出第一用户，提高有效性。本实施例通过在传播概率学习模型中引入控制因子，可以提高传播概率的准确性。本实施例通过设置不同的传播策略，可以实现信息多样性传播。

为了实现上述实施例，本发明还提出一种信息传播装置。

25 图 7 是本发明另一实施例的信息传播装置的结构示意图。如图 7 所示，该信息传播装置包括：确定模块 100 和传播模块 200。

具体地，确定模块 100 用于确定要传播的信息对应的第一用户，所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户。其中，要传播的信息可以是商品推广信息，也可以是其他信息，本发明对此不做限定。要传播的信息对应的第一用户可
30 以是一个或多个。

兴趣类型网络可以根据兴趣类型对用户或信息进行分类标记的网络，也可以称为兴趣网络。

具体地，可以预先建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户。例如，可以预先设置多个标签，按照标签将用户划分到不同的兴趣类型网络中。如图 2 所示，兴趣类型网络包括标签，例如时尚、户外、商务、运动、旅行、电子等，每个用户都可以对应一个或多个标签。具体建立兴趣类型网络的过程将在后续实施例中介绍。

第一用户是兴趣类型网络中影响力大于预设值的用户，影响力是用户的一种属性，在本实施例中，一个用户的影响力用于衡量该用户传播的信息被其他人接受的难易程度，其中，影响力大的用户传播的信息更容易被其他人接受。

第一用户也可以称为达人。每个兴趣类型网络中达人可以是一个或者多个。

例如，假设第一用户称为达人，如图 4 所示，达人列表中包括：服装达人，3C 达人和家居达人，则如果要传播的信息包括的标签是 3C，则该要传播的信息对应的第一用户是 3C 达人。

传播模块 200 用于获取以所述第一用户为起点的用户关系网络，在所述用户关系网络中以所述第一用户为起点传播所述信息。其中，用户关系网络是用于描述用户与用户之间的关联关系的网络，可以直接从已有的社交网络类型的应用程序中获取用户关系网络，在社交网络类型的应用程序中，用户之间可以通过增加好友或者增加关注等方式预先建立用户关系网络。例如，可以先从第一用户的应用程序中获取到第一用户的好友包括第二用户，再在第二用户的应用程序中获取到第二用户的好友包括第三用户，则可以获取到的用户关系网络包括：第一用户—>第二用户—>第三用户。

以所述第一用户为起点的用户关系网络可以从应用程序的已有数据中导入，例如，从社交网络的应用程序中导入以确定出的第一用户为起点的用户关系网络。

例如，如图 4 所示，假设要传播的信息对应的第一用户是 3C 达人，从已有数据中获取的以 3C 达人为起点的用户关系网络是用户关系网络 41，则如图 4 所示，则可以将要传播的信息以 3C 达人为起点在用户关系网络 41 中传播。

本实施例中，通过确定要传播的信息对应的第一用户，第一用户是影响力大于预设值的用户，并由第一用户为起点进行信息传播，可以由具有较大影响力的用户传播信息，提高信息传播的可信性，提高信息传播效率。

图 8 是本发明另一实施例的信息传播装置的结构示意图。如图 8 所示，该信息传播装置包括：确定模块 100、第二获取子模块 110、第一确定子模块 120、传播模块 200、第三获取子模块 210、获取单元 211、建模单元 212、更新单元 213、确定单元 214、第二确定子模块 220、建立模块 300、第一获取子模块 310 和聚类子模块 320。其中，建立模块 300 包括第一获取子模块 310 和聚类子模块 320；确定模块 100 包括第二获取子模块 110 和

第一确定子模块 120; 传播模块 200 包括第三获取子模块 210 和第二确定子模块 220; 第三获取子模块 210 包括获取单元 211、建模单元 212、更新单元 213 和确定单元 214。

具体地, 建立模块 300 用于建立预设个数的兴趣类型网络, 并在每个兴趣类型网络中, 确定对应的第一用户。以要传播的信息是商品的信息为例, 建立模块 300 具体可以包括:

5 第一获取子模块 310, 用于根据标签传播学习算法, 获取用户-标签矩阵。具体可以包括:

(1) 计算得到商品与商品的相似度矩阵 W 。

商品与商品的相似度矩阵可以用于表示商品之间在用户行为、商品标题和商品属性上的相似度。

10 其中, 用于计算相似度矩阵的商品可以是优质买家处理过的商品, 处理具体可以是指购买, 浏览, 点击, 收藏中的一项或者多项, 优质买家可以根据优质买家模型确定, 例如, 将信用等级高或者购买次数多的买家确定为优质买家。具体的, 可以获取所有买家的信息, 再根据优质买家模型从所有买家中确定出优质买家, 再获取优质买家处理过的商品, 再根据优质买家处理过的商品中的两两商品计算相似度, 得到相似

15 度矩阵 W 。
更具体的, 第一获取子模块 310 可以通过最小哈希算法对商品(pid,vid)进行哈希映射, 得到商品与商品的相似度矩阵, 其中, pid 是商品的 ID (Identity, 身份标识), vid 是商品属性值的 ID, pid 和 vid 通常可以从基础数据表中获取。

(2) 计算得到商品-标签信息矩阵 F 。

20 其中, 商品-标签信息矩阵 F 中的商品也可以具体是指优质买家处理过的商品, 标签是指商品更新后的标签, 在获取优质买家处理过的商品后, 可以根据每个商品的初始标签, 经过迭代过程计算得到商品-标签信息矩阵 F , 其中, 每个商品的初始标签可以是作为商品的一个属性预先记录在数据库中的, 从而可以从数据库中获取商品的初始标签。

25 更具体的, 商品-标签信息矩阵 F 可以根据标签传播学习算法的迭代公式得到, 迭代公式如下:

While (F 收敛)

$$F(t+1) = \alpha SF(t) + (1-\alpha) Y$$

end

30 其中, 要计算的商品-标签信息矩阵 F 是上述公式中收敛时得到 $F(t+1)$, $0 \leq \alpha \leq 1$ 为预设的加权参数, S 是根据上述的商品与商品的相似度矩阵 W 计算得到的, $S = D^{-1}W \in R^{n \times n}$, $D = \text{diag}\{D_{11}, \dots, D_{nn}\} \in R^{n \times n}$, $D_{ii} = \sum_{j=1}^n W_{ij} = \sum_{j=1}^n W_{ji}$, Y 是初始标签值,

$F(t)$ 的初始值可以是根据已有的买家信息得到的商品-标签信息矩阵 F 的初始值, 已有的买家信息可以是根据优质买家模型得到的。例如, 根据买家的信用等级从多个买家中确定出预设个数的优质买家, 再根据优质买家与优质买家对应的购买, 点击或收藏的商品可以得到用户-商品信息矩阵 V , 以及, 根据优质买家购买, 点击或收藏的商品与商品具有的标签可以得到上述的商品-标签信息矩阵 F 的初始值。

其中, 商品具有的标签可以根据统计或者 HITS (Hyperlink-Induced Topic Search, 超链诱导主题搜索) 排序算法得到。

在获取 F 的初始值后, 可以根据上述迭代公式, 在满足迭代收敛条件时得到最终的商品-标签信息矩阵 F 。

迭代收敛条件可以包括: 设置最大迭代次数, 当迭代次数达到最大迭代次数时满足迭代收敛条件; 或者, 根据迭代后的值与迭代前的值的差值, 在该差值大于预设阈值时满足迭代收敛条件, 例如, $\|F(t+1)-F(t)\| < \beta$ 时表明满足迭代收敛条件, $\|F(t+1)-F(t)\|$ 表示 $F(t+1)$ 与 $F(t)$ 的欧氏距离, β 表示预设阈值。

(3) 计算得到用户-标签矩阵 L 。

其中, 用户-标签矩阵 L 中的用户也可以具体是指优质买家, 标签是指用户具有的标签, 用户具有的标签可以根据用户处理过的商品的更新后的标签确定。

具体的, 可以采用如上所示的方式确定出优质买家, 以及获取优质买家处理过的商品, 以及, 从数据库中获取优质买家处理过的商品的初始标签, 之后可以根据优质买家处理过的商品以及上述的方式 (1) 计算得到商品与商品的相似度矩阵 W , 再根据商品与商品的相似度矩阵 W 和优质买家处理过的商品具有的初始标签以及上述的方式 (2) 计算得到商品-标签信息矩阵 F , 再根据优质买家以及优质买家处理过的商品可以建立用户-商品信息矩阵 V , 之后在根据上述的 V 和 F 采用如下的方式得到用户-标签矩阵 L 。

更具体的, 计算公式可以是: $L=V \cdot F$, 其中 V 是上述得到的用户-商品信息矩阵, F 是上述得到的收敛时的最终的商品-标签信息矩阵。

聚类子模块 320, 用于对所述用户-标签矩阵进行聚类, 得到预设个数的兴趣类型网络, 并获取每个兴趣类型网络中的第一用户。在得到用户-标签矩阵 L 后, 可以对该矩阵 L 进行聚类, 例如, 预设个数是 k 个, 则可以对矩阵 L 进行双聚类得到 k 个类别, 每个类别对应一个兴趣类型网络。

在对矩阵 L 进行聚类得到 k 个类别后, 以每个兴趣类型网络包括一个第一用户为例, 每个类别的中心点可以确定为该兴趣类型网络的第一用户。不同兴趣类型网络的

第一用户可以组成列表，该列表可以称为达人列表，达人列表例如表示为： $P=\{p_1, p_2, \dots, p_k\}$ ，其中， p_i ($i=1, 2, \dots, k$) 是第 i 个兴趣类型网络中的第一用户，也可以称为达人， p_i 可以由用户 ID 和该用户具有的标签组成。

在上述预先建立多个兴趣类型网络，并确定每个兴趣类型网络中的第一用户后，
5 如上所述，可以得到由不同的兴趣类型网络中的第一用户组成的达人列表，达人列表中包括不同兴趣类型网络中的第一用户，在当前需要传播信息时，可以首先确定要传播的信息对应的第一用户。

所述确定模块 100 具体包括：

第二获取子模块 110 用于获取第一标签，所述第一标签是所述要传播的信息包括的标
10 签；

第一确定子模块 120 用于将包括所述第一标签的第一用户，确定为所述要传播的信息对应的第一用户。

例如，假设第一用户称为达人，如图 4 所示，达人列表中包括：服装达人，3C 达人和家居达人，则如果第二获取子模块 110 获取到要传播的信息包括的标签是 3C，则第一确定
15 子模块 120 确定该要传播的信息对应的第一用户是 3C 达人。

传播模块 200 还用于根据预设策略，在所述用户关系网络中以所述第一用户为起点传播所述信息，所述预设策略包括传播范围策略，或者，传播速度策略。其中，传播范围策略是指优先考虑传播范围，传播速度策略是指优先考虑传播速度。

更具体的，第三获取子模块 210 可以获取用户关系网络中用户与用户之间的传播概率，
20 当采用传播范围策略时，可以不论传播概率高低都进行信息传播，当采用传播速度策略时，可以只在传播概率大于预设值的路径上进行信息传播。例如，以传播速度策略为例，参见图 5，假设用户关系网络包括第一路径 51，第二路径 52，第三路径 53，第四路径 54 和第五路径 55，假设第一路径 51，第二路径 52 和第三路径 53 中包括的用户之间的传播概率都大于预设值，而第四路径 54 和第五路径 55 上包括的用户之间的存在小于预设值的传播概
25 率，则信息可以在第一路径 51，第二路径 52 和第三路径 53 上传播，而不在第四路径 54 和第五路径 55 上传播。

更具体的，信息在用户关系网络中传播时，第一用户在初始时刻作为信息传播的种子节点，种子节点负责向其邻居节点传播信息，例如，第一用户是 3C 达人，与 3C 达人相邻的邻居节点包括第一节点和第二节点，则在初始时刻 t 设置 3C 达人是种子节点，并且由 3C
30 达人将信息传播给第一节点和第二节点，当种子节点将信息传播给邻居节点后，邻居节点在下一时间成为新的种子节点，例如，在 $t+1$ 时刻种子节点是第一节点，而不再是 3C 达人，

依次类推，从初始的第一用户依次根据用户关系网络中的用户相邻关系进行信息传播，直至没有新的种子节点。另外，用户关系网络中邻居节点用户之间的传播概率是独立的，不受其他邻居节点之间的关系影响。并且，每个种子节点只有一次机会向非种子邻居节点传播信息，例如，用户在 t 时刻成为种子节点，仅在 t 时刻有一次机会尝试对非种子邻居节点传播信息，如果传播成功，则该邻居节点成为 $t+1$ 时刻的种子节点，而不管该用户在 t 时刻是否传播成功，该用户再也不能在其他时刻试图传播信息给它的邻居节点。如果在同一时刻，有多个种子节点试图传播信息给同一节点，其传播的顺序可以是任意的。

可选地，第三获取子模块 210 还用于根据引入传播概率方差控制因子的传播概率学习模型，获取所述用户关系网络中用户与用户之间的传播概率。例如，传播概率学习模型可以是 EM (Expectation-maximization, 最大化期望) 模型。由于数据的稀疏性，在传播概率学习的过程中，根据 EM 模型学习到的传播概率往往方差比较大。这主要是因为 EM 模型的计算方法在稀疏数据情况下过拟合导致数据分布不均匀，很容易估计获得概率为 0 或概率为 1 的极端概率情况。

本申请实施例中，为了解决传统 EM 模型存在的上述问题，在 EM 模型中引入了传播概率方差控制因子，防止 EM 模型在迭代过程中发生剧烈的波动。

可选的，所述第三获取子模块 210，包括：

获取单元 211 用于获取所述用户关系网络，例如，从已有的社交网络的应用程序中导入用户关系网络，并根据所述用户关系网络和时间片断数据建立信息传播模型，例如，可以建立独立级联模型。独立级联模型是一种基本的传播模型，可以采用现有的方式根据用户关系网络建立。

其中，时间片断数据是预设的，用于表明信息传播扩散时间。

在传播模型中，可以包括节点和边，其中，每个节点可以对应用户关系网络中的一个用户，每个边是由用户关系网络中两个相邻用户组成的线段。

建模单元 212 用于将传播概率方差控制因子引入传播概率学习模型中，得到引入传播概率方差控制因子的传播概率学习模型，并根据所述引入传播概率方差控制因子的传播概率学习模型，对所述信息传播模型进行学习，获取传播概率更新规则，所述更新规则包括第一更新规则和第二更新规则。

EM (Expectation-maximization, 最大化期望) 模型是一种优化算法，在本实施例中，可以采用 EM 模型对独立级联模型进行学习，从而得到独立级联模型中包括的每个边的传播概率，也就是用户关系网络中用户与用户之间的传播概率。

传统的 EM 模型可以表示为：

$$L(\theta) = \sum_{s=1}^S \log L(q | D_s) = \sum_{s=1}^S \sum_{t=0}^{T-1} \left[\sum_{w \in D_s(t+1)} \log P_w^s + \sum_{v \in D_s(t)} \sum_{w \in Fv \setminus C(t+1)} \log(1 - k_{v,w}) \right]$$

在引入传播概率方差控制因子后，可以根据求解过程是否收敛得到不同的引入传播概率方差控制因子的 EM 模型，采用哪种引入传播概率方差控制因子的 EM 模型可以根据实际需要确定，具体的，引入传播概率方差控制因子的 EM 模型可以是：

$$L_\lambda(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum [\theta_i - E(\theta_i)]^2$$

或者，

$$L_\lambda(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum |\log \theta_i - E(\log \theta_i)|$$

其中， λ 是控制因子， $k_{v,w}$ 是边(v,w)的传播概率。

其中，可以先根据引入 λ 的 EM 模型确定优化方程，再对优化方程进行求解，得到第一更新规则。

具体的，如果引入 λ 的 EM 模型是：

$$L_\lambda(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum [\theta_i - E(\theta_i)]^2$$

其对应的优化方程是：

$$\begin{aligned} Q_\lambda(\theta | \theta) = & (1-\lambda) \sum_{s=1}^S \sum_{t=0}^{T-1} \sum_{v \in D_s(t)} \left(\sum_{w \in F(v) \cap D_s(t+1)} \left(\frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \frac{1}{k_{v,w}} + \left(1 - \frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \right) \frac{1}{k_{v,w} - 1} \right) + \sum_{w \in F(v) \setminus C(t+1)} \frac{1}{k_{v,w} - 1} \right) \\ & - \frac{2\lambda}{m} \sum_{(v,w) \in E} \left(k_{v,w} - E(\hat{k}_{v,w}) \right) = 0 \end{aligned}$$

对该优化方程进行求解后，得到第一更新规则是：

$$k_{v,w} = (1-\lambda) \frac{1}{|S_{v,w}^+| + |S_{v,w}^-|} \sum_{s \in S_{v,w}^+} \frac{\hat{k}_{v,w}}{\hat{P}_w(s)} + \lambda E(\hat{k}_{v,w})$$

其中， $|S_{v,w}^+|$ 表示 $v \in D_s(t), w \in D_s(t+1)$ ， $|S_{v,w}^-|$ 表示 $v \in D_s(t), w \notin D_s(t+1)$ ， $D_s(t)$ 表示在 t 时刻激活的点的集合， $P_w(s)$ 表示 w 被激活的概率。

对该优化方程进行求解后，得到第二更新规则是：

$$-\frac{2\lambda}{m} \sum_{(v,w) \in E} \left(k_{v,w} - E(\hat{k}_{v,w}) \right) = 0$$

得到,

$$k_{v,w} = E(\hat{k}_{v,w})$$

如果引入 λ 的 EM 模型是:

$$L_\lambda(\theta) = (1-\lambda)L(\theta) - \frac{\lambda}{m} \sum |\log(\theta_i) - E(\log(\theta_i))|$$

5 其对应的优化方程是:

$$\begin{aligned} Q_\lambda(\theta|\theta) = & (1-\lambda) \sum_{s=1}^S \sum_{t=0}^{T-1} \sum_{v \in D_s(t)} \left(\sum_{w \in F(v) \cap D_s(t+1)} \left(\frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \log k_{v,w} \right. \right. \\ & \left. \left. + \left(1 - \frac{\hat{k}_{v,w}}{\hat{P}_w(s)} \right) \log(1-k_{v,w}) \right) + \sum_{w \in F(v) \setminus C(t+1)} \log(1-k_{v,w}) \right) \\ & - \frac{\lambda}{m} \sum_{(v,w) \in E} \left| \log k_{v,w} - E(\log \hat{k}_{v,w}) \right| \end{aligned}$$

对该优化方程进行求解后, 得到第一更新规则是:

$$k_{v,w} = \begin{cases} \frac{(1-\lambda) \sum_{a \in S_{v,w}^+} \frac{\hat{k}_{v,w}}{\hat{P}_w(a)} - \frac{\lambda}{m}}{(1-\lambda)(|S_{v,w}^+| + |S_{v,w}^-|) - \frac{\lambda}{m}} & \text{if } \hat{k}_{v,w} \geq \left(\prod_{(p,q) \in E} \hat{k}_{p,q} \right)^{1/|E|} \\ \frac{(1-\lambda) \sum_{a \in S_{v,w}^+} \frac{\hat{k}_{v,w}}{\hat{P}_w(a)} + \frac{\lambda}{m}}{(1-\lambda)(|S_{v,w}^+| + |S_{v,w}^-|) + \frac{\lambda}{m}} & \text{if } \hat{k}_{v,w} < \left(\prod_{(p,q) \in E} \hat{k}_{p,q} \right)^{1/|E|} \end{cases}$$

对该优化方程进行求解后, 得到第二更新规则是:

$$k_{v,w} = \left(\prod_{(p,q) \in E} \hat{k}_{p,q} \right)^{1/|E|}$$

10

更新单元 213 用于采用所述第一更新规则更新第一组用户之间的传播概率, 采用所述第二更新规则更新第二组用户之间的传播概率, 所述第一组用户之间的边在所述时间片断

数据内被激活，所述第二组用户之间的边在所述时间片断数据内没有被激活。在得到第一更新规则和第二更新规则后，可以在用户关系网络中选取种子节点，然后以种子节点为起点根据用户关系网络传播预设信息，传播时间是预设的时间片断数据。更具体地，可以判断时间片断数据是否结束，例如，可以获取当前时间与信息开始传播的时间之间的差值，

5 如果该差值小于预设的时间片断数据，则确定时间片断数据没有结束，否则结束。

若时间片断数据没有结束，则可以判断要计算的边在该时间片断数据内是否被激活，例如，要计算的边是用户 A 与用户 B 组成的边，在信息传播时间内，传播的信息经过用户 A 和用户 B，则可以确定用户 A 和用户 B 组成的边在该时间内被激活了，否则未激活。如果被激活，则采用第一更新规则，对该要计算的边的传播概率进行更新，之后，将每个边更新后的传播概率写入传播概率更新库。如果未被激活，则返回继续判断时间片断数据是否结束。

若时间片断数据已结束，则采用第二更新规则，对整个时间片断数据内未被激活的边的传播概率进行更新，之后，将每个边更新后的传播概率写入传播概率更新库。

可以理解的是，上述以传播概率学习模型是 EM 模型为例，传播概率学习模型也可以是其他模型，例如，马尔科夫模型。

确定单元 214 用于将更新后的用户之间的传播概率确定为所述用户关系网络中用户与用户之间的传播概率。

第二确定子模块 220 用于将所述传播概率大于预设值的路径确定为传播路径，并根据所述传播路径传播所述信息，以实现最大的传播速度。

20 本实施例中，通过确定要传播的信息对应的第一用户，第一用户是影响力大于预设值的用户，并由第一用户为起点进行信息传播，可以由具有较大影响力的用户传播信息，提高信息传播的可信性，提高信息传播效率。本实施例通过标签传播学习算法可以确定出第一用户，提高有效性。本实施例通过在传播概率学习模型中引入控制因子，可以提高传播概率的准确性。本实施例通过设置不同的传播策略，可以实现信息多样性传播。

需要说明的是，在本发明的描述中，术语“第一”、“第二”等仅用于描述目的，而不能理解为指示或暗示相对重要性。此外，在本发明的描述中，除非另有说明，“多个”的含义是两个或两个以上。

30 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分，并且本发明的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能，这应被本发明的

实施例所属技术领域的技术人员所理解。

应当理解，本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如，如果用硬件来实现，和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列（PGA），现场可编程门阵列（FPGA）等。

本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，在本发明各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读存储介质中。

上述提到的存储介质可以是只读存储器，磁盘或光盘等。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不一定指的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任何一个或多个实施例或示例中以合适的方式结合。

尽管上面已经示出和描述了本发明的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本发明的限制，本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

权 利 要 求 书

1、一种信息传播方法，其特征在于，包括：

确定要传播的信息对应的第一用户，所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户；

5 获取以所述第一用户为起点的用户关系网络，在所述用户关系网络中以所述第一用户为起点传播所述信息。

2、根据权利要求1所述的方法，其特征在于，还包括：

建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户；

所述建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户，

10 包括：

根据标签传播学习算法，获取用户-标签矩阵；

对所述用户-标签矩阵进行聚类，得到预设个数的兴趣类型网络，并获取每个兴趣类型网络中的第一用户。

3、根据权利要求2所述的方法，其特征在于，所述第一用户的标识信息包括：用户标

15 识和标签，兴趣类型网络包括标签，所述确定要传播的信息对应的第一用户，包括：

获取第一标签，所述第一标签是所述要传播的信息包括的标签；

将包括所述第一标签的第一用户，确定为所述要传播的信息对应的第一用户。

4、根据权利要求1所述的方法，其特征在于，所述在所述用户关系网络中以所述第一用户为起点传播所述信息，包括：

20 根据预设策略，在所述用户关系网络中以所述第一用户为起点传播所述信息，所述预设策略包括传播范围策略，或者，传播速度策略。

5、根据权利要求4所述的方法，其特征在于，当所述预设策略是传播速度策略时，所述根据预设策略，在所述用户关系网络中以所述第一用户为起点传播所述信息，包括：

获取所述用户关系网络中用户与用户之间的传播概率；

25 将所述传播概率大于预设值的路径确定为传播路径，并根据所述传播路径传播所述信息。

6、根据权利要求5所述的方法，其特征在于，所述获取所述用户关系网络中用户与用户之间的传播概率，包括：

根据引入传播概率方差控制因子的传播概率学习模型，获取所述用户关系网络中用户

30 与用户之间的传播概率。

7、根据权利要求6所述的方法，其特征在于，所述根据引入传播概率方差控制因子的

传播概率学习模型，获取所述用户关系网络中用户与用户之间的传播概率，包括：

获取所述用户关系网络，并根据所述用户关系网络和时间片断数据建立信息传播模型，所述时间片段数据是预设的信息传播扩散时间；

5 将传播概率方差控制因子引入传播概率学习模型中，得到引入传播概率方差控制因子的传播概率学习模型，并根据所述引入传播概率方差控制因子的传播概率学习模型，对所述信息传播模型进行学习，获取传播概率更新规则，所述更新规则包括第一更新规则和第二更新规则；

采用所述第一更新规则更新第一组用户之间的传播概率，采用所述第二更新规则更新第二组用户之间的传播概率，所述第一组用户之间的边在所述时间片断数据内被激活，所述第二组用户之间的边在所述时间片断数据内没有被激活；

10 将更新后的用户之间的传播概率确定为所述用户关系网络中用户与用户之间的传播概率。

8、一种信息传播装置，其特征在于，包括：

确定模块，用于确定要传播的信息对应的第一用户，所述第一用户是所述第一用户属于的兴趣类型网络中影响力大于预设值的用户；

15 传播模块，用于获取以所述第一用户为起点的用户关系网络，在所述用户关系网络中以所述第一用户为起点传播所述信息。

9、根据权利要求 8 所述的装置，其特征在于，还包括：

20 建立模块，用于建立预设个数的兴趣类型网络，并在每个兴趣类型网络中，确定对应的第一用户；

所述建立模块，包括：

第一获取子模块，用于根据标签传播学习算法，获取用户-标签矩阵；

聚类子模块，用于对所述用户-标签矩阵进行聚类，得到预设个数的兴趣类型网络，并获取每个兴趣类型网络中的第一用户。

25 10、根据权利要求 9 所述的装置，其特征在于，所述第一用户的标识信息包括：用户标识和标签，兴趣类型网络包括标签，所述确定模块，包括：

第二获取子模块，用于获取第一标签，所述第一标签是所述要传播的信息包括的标签；

第一确定子模块，用于将包括所述第一标签的第一用户，确定为所述要传播的信息对应的第一用户。

30 11、根据权利要求 8 所述的装置，其特征在于，所述传播模块还用于根据预设策略，在所述用户关系网络中以所述第一用户为起点传播所述信息，所述预设策略包括传播范围策略，或者，传播速度策略。

12、根据权利要求 11 所述的装置，其特征在于，当所述预设策略是传播速度策略时，所述传播模块，包括：

第三获取子模块，用于获取所述用户关系网络中用户与用户之间的传播概率；

5 第二确定子模块，用于将所述传播概率大于预设值的路径确定为传播路径，并根据所述传播路径传播所述信息。

13、根据权利要求 12 所述的装置，其特征在于，所述第三获取子模块还用于根据引入传播概率方差控制因子的传播概率学习模型，获取所述用户关系网络中用户与用户之间的传播概率。

14、根据权利要求 13 所述的装置，其特征在于，所述第三获取子模块，包括：

10 获取单元，用于获取所述用户关系网络，并根据所述用户关系网络和时间片断数据建立信息传播模型，所述时间片段数据是预设的信息传播扩散时间；

建模单元，用于将传播概率方差控制因子引入传播概率学习模型中，得到引入传播概率方差控制因子的传播概率学习模型，并根据所述引入传播概率方差控制因子的传播概率学习模型，对所述信息传播模型进行学习，获取传播概率更新规则，所述更新规则包括第
15 一更新规则和第二更新规则；

更新单元，用于采用所述第一更新规则更新第一组用户之间的传播概率，采用所述第二更新规则更新第二组用户之间的传播概率，所述第一组用户之间的边在所述时间片断数据内被激活，所述第二组用户之间的边在所述时间片断数据内没有被激活；

20 确定单元，用于将更新后的用户之间的传播概率确定为所述用户关系网络中用户与用户之间的传播概率。

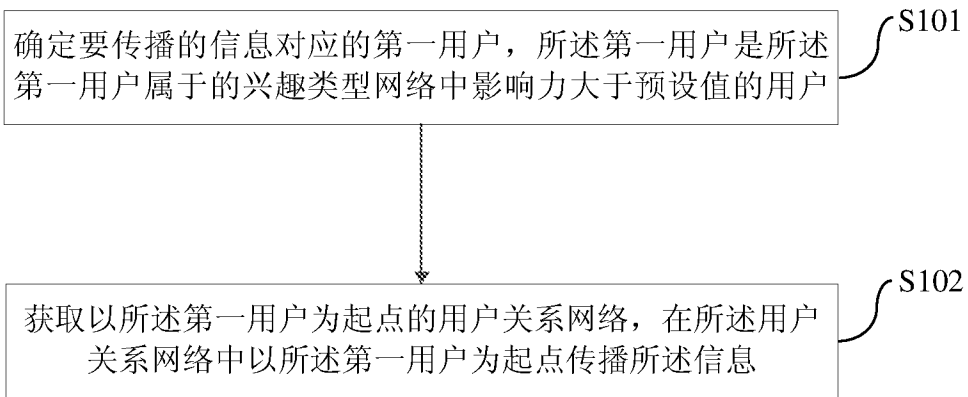


图 1

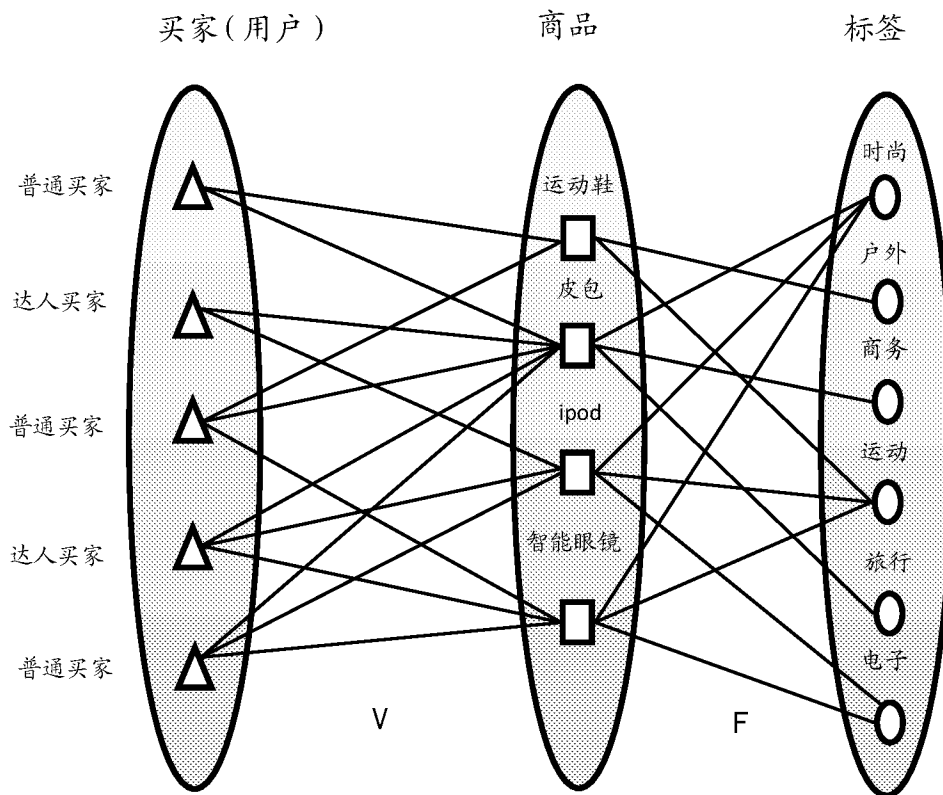


图 2

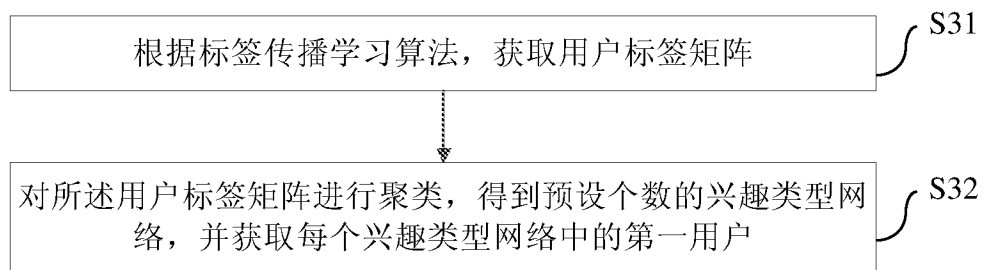


图 3

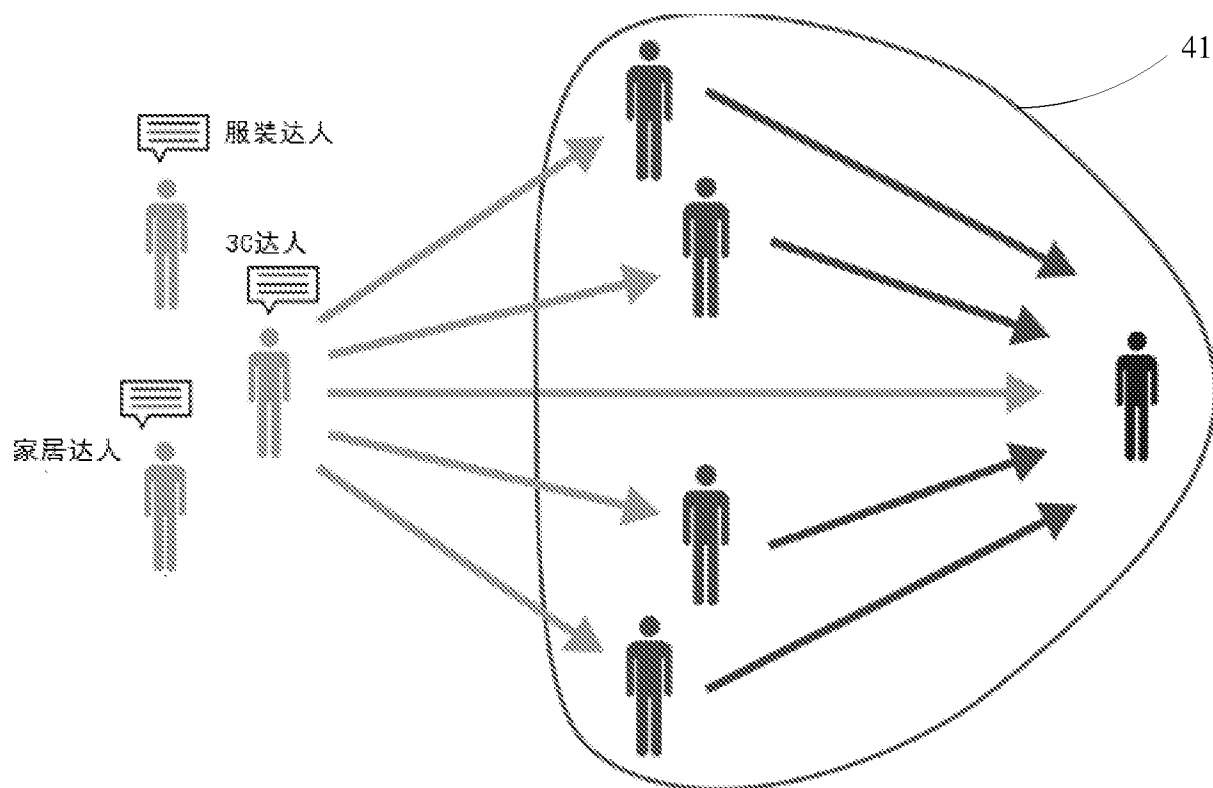


图 4

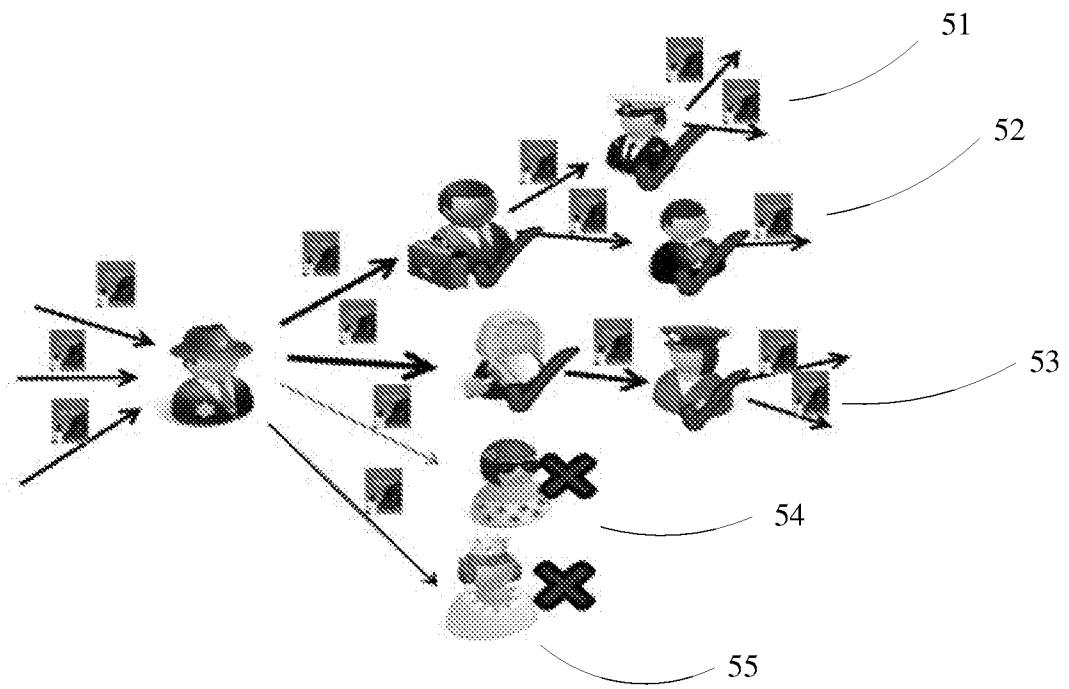


图 5

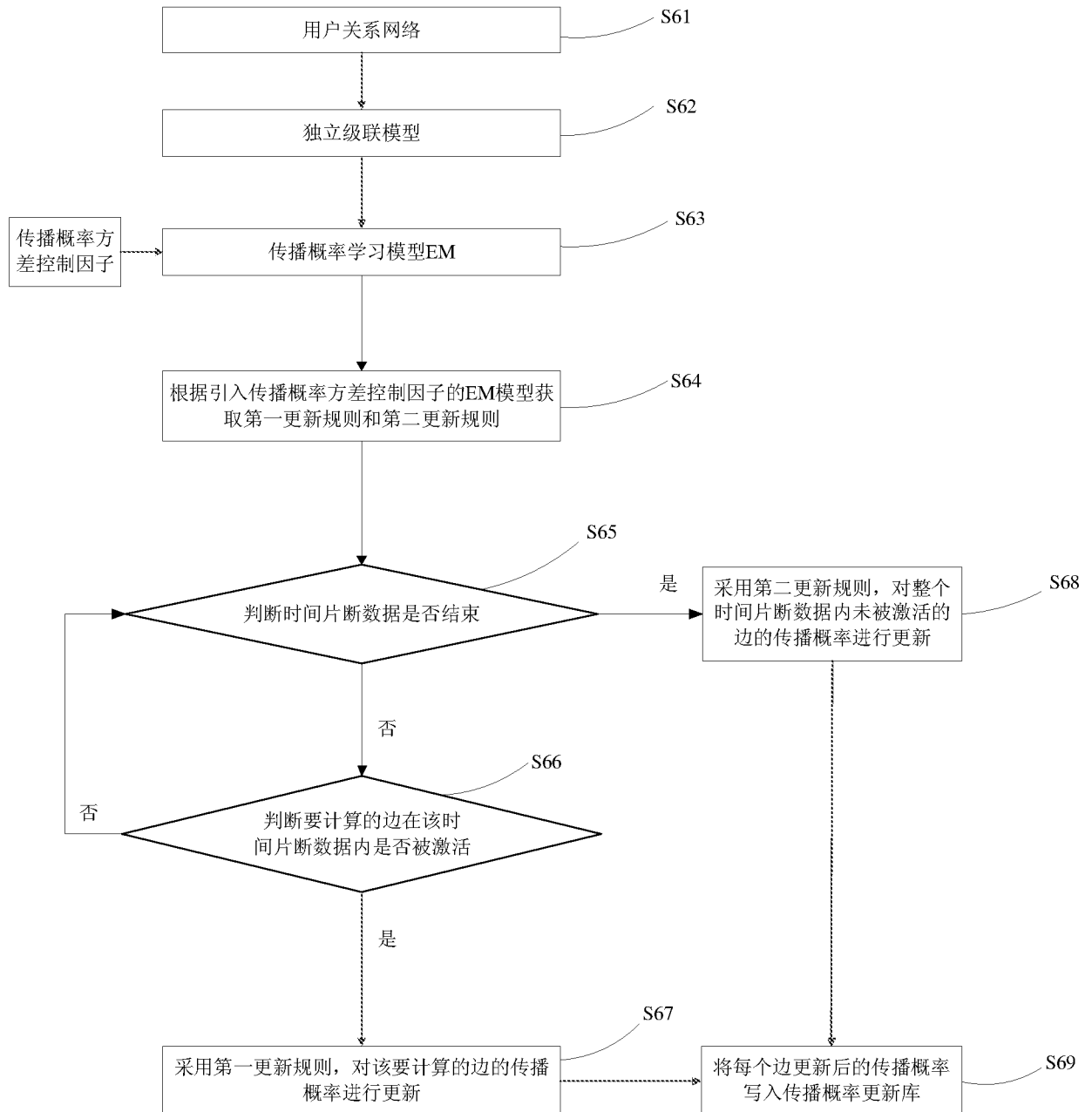


图 6

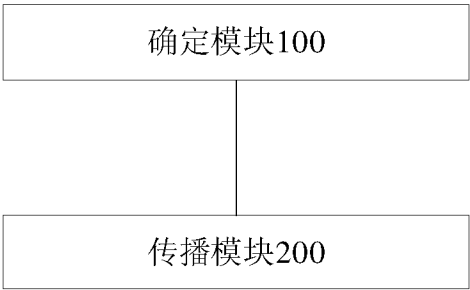


图 7

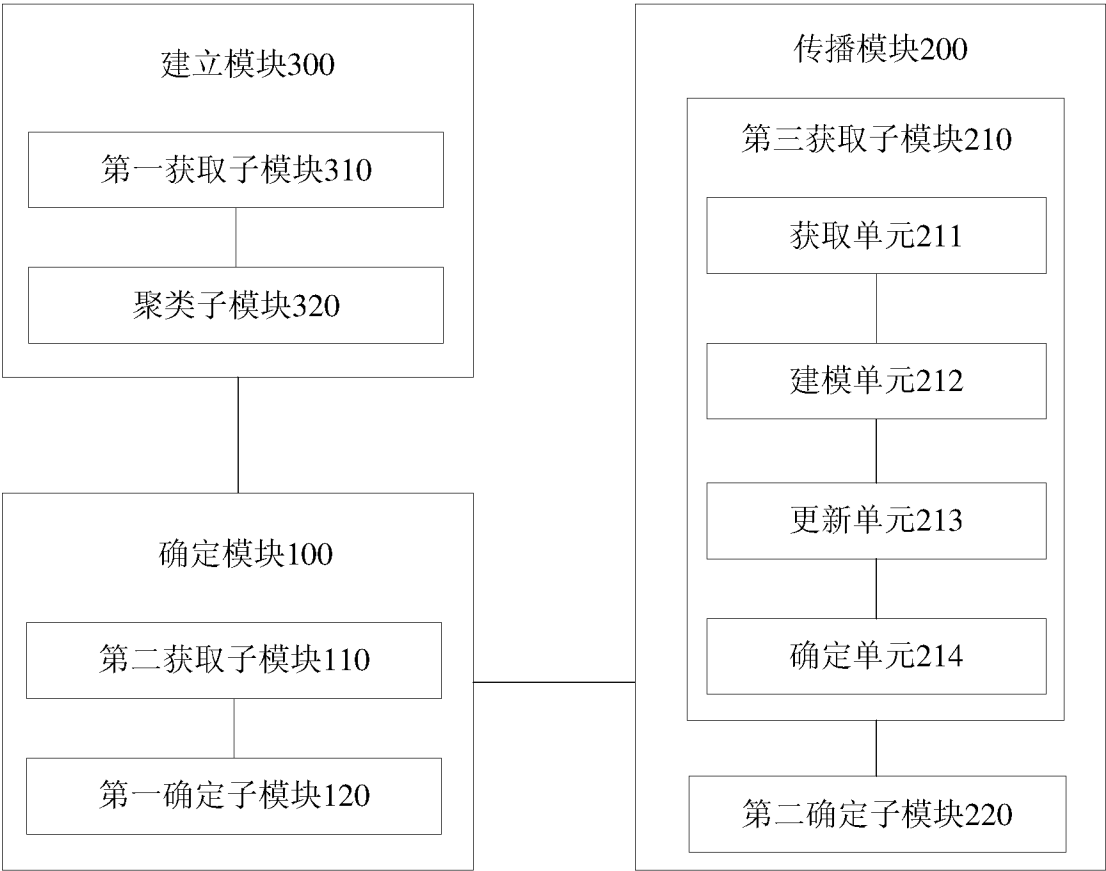


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/072783

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/58 (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, H04L

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)

CPRSABS, CNTXT, CNKI, VEN: marketing, diffuse, publish, advertisement, information, push, distribute, affect, influence, interest, social network

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103279512 A (HUZHOU TEACHERS COLLEGE), 04 September 2013 (04.09.2013), description, paragraphs [0018], [0044] and [0064]	1, 4, 8, 11
A	CN 103279512 A (HUZHOU TEACHERS COLLEGE), 04 September 2013 (04.09.2013) , the whole description	2, 3, 5-7, 9, 10, 12-14
A	CN 103412872 A (XI'AN JIAOTONG UNIVERSITY), 27 November 2013 (27.11.2013) , the whole description	1-14
A	CN 103177382 A (WUHAN UNIVERSITY), 26 June 2013 (26.06.2013) , the whole description	1-14

☐ 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 21 March 2016 (21.03.2016)	Date of mailing of the international search report 22 April 2016 (22.04.2016)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer WANG, Jing Telephone No.: (86-10) 62088433

International application No.
PCT/CN2016/072783

Form PCT/ISA/210 (patent family annex) (July 2009)

A. 主题的分类 H04L 12/58 (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, H04L 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)) CPRSABS, CNTXT, CNKI, VEN: 信息, 广告, 营销, 推送, 传播, 扩散, 发布, 影响, 兴趣, 社交网络, advertisement, information, push, distribute, affect, influence, interest, social network,																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 103279512 A (湖州师范学院) 2013年 9月 4日 (2013 - 09 - 04) 说明书第[0018]、[0044]、[0064]段</td> <td>1, 4, 8, 11</td> </tr> <tr> <td>A</td> <td>CN 103279512 A (湖州师范学院) 2013年 9月 4日 (2013 - 09 - 04) 说明书全文</td> <td>2, 3, 5-7, 9, 10, 12-14</td> </tr> <tr> <td>A</td> <td>CN 103412872 A (西安交通大学) 2013年 11月 27日 (2013 - 11 - 27) 说明书全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 103177382 A (武汉大学) 2013年 6月 26日 (2013 - 06 - 26) 说明书全文</td> <td>1-14</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 103279512 A (湖州师范学院) 2013年 9月 4日 (2013 - 09 - 04) 说明书第[0018]、[0044]、[0064]段	1, 4, 8, 11	A	CN 103279512 A (湖州师范学院) 2013年 9月 4日 (2013 - 09 - 04) 说明书全文	2, 3, 5-7, 9, 10, 12-14	A	CN 103412872 A (西安交通大学) 2013年 11月 27日 (2013 - 11 - 27) 说明书全文	1-14	A	CN 103177382 A (武汉大学) 2013年 6月 26日 (2013 - 06 - 26) 说明书全文	1-14
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 103279512 A (湖州师范学院) 2013年 9月 4日 (2013 - 09 - 04) 说明书第[0018]、[0044]、[0064]段	1, 4, 8, 11															
A	CN 103279512 A (湖州师范学院) 2013年 9月 4日 (2013 - 09 - 04) 说明书全文	2, 3, 5-7, 9, 10, 12-14															
A	CN 103412872 A (西安交通大学) 2013年 11月 27日 (2013 - 11 - 27) 说明书全文	1-14															
A	CN 103177382 A (武汉大学) 2013年 6月 26日 (2013 - 06 - 26) 说明书全文	1-14															
☐ 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 “&” 同族专利的文件																	
国际检索实际完成的日期 2016年 3月 21日		国际检索报告邮寄日期 2016年 4月 22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		授权官员 王静 电话号码 (86-10) 62088433															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/072783

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	103279512	A	2013年 9月 4日	无	
CN	103412872	A	2013年 11月 27日	无	
CN	103177382	A	2013年 6月 26日	无	