

(19) 中华人民共和国国家知识产权局



## (12) 发明专利申请

(10) 申请公布号 CN 103838964 A

(43) 申请公布日 2014. 06. 04

(21) 申请号 201410065242. 9

(22) 申请日 2014. 02. 25

(71) 申请人 中国科学院自动化研究所

地址 100190 北京市海淀区中关村东路 95  
号

(72) 发明人 王飞跃 陈松航 朱凤华 吕宣生  
黄武陵 熊刚 孔庆杰 杨柳青

(74) 专利代理机构 中科专利商标代理有限责任  
公司 11021

代理人 宋焰琴

(51) Int. Cl.

G06F 19/00 (2011. 01)

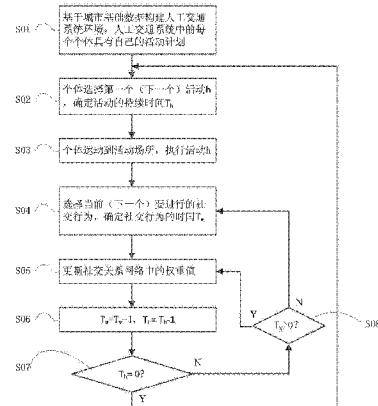
权利要求书3页 说明书6页 附图2页

### (54) 发明名称

一种基于人工交通系统的社交关系网络生成  
方法及装置

### (57) 摘要

本发明提出一种基于人工交通系统的社交关  
系网络生成方法及装置。本发明提出的方案将社  
交行为嵌入到人工交通系统中个体开展活动计划  
的过程之中，按照一定规则选择社交行为，并在开  
展过程中更新社交关系网络的权重值，从而自动  
生成社交关系网络。本发明的有益效果是，能够代  
替传统的调查问卷、电话调查、面对面调查等方  
式，通过计算机自动生成社交关系网络，实施成本  
很低，并很容易针对大规模人群进行实施。



1. 一种基于人工交通系统的社交关系网络生成方法,其特征在于,包括以下步骤:

步骤 S01 :基于城市基础数据构建人工交通系统环境,人工交通系统中的每个个体具有自己的活动计划;并构建初始的社交关系网络,其中初始的社交关系网络的各个顶点为人工交通系统中的所有个体,两个个体之间的连接表示相应两个个体的社交关系;

步骤 S02 :每个个体选择一个活动,并确定所选择活动的持续时间 Th ;

步骤 S03 :个体运动到活动场所,执行所选择的活动;

步骤 S04 :个体选择并执行一个社交行为,然后确定下一个社交行为的时间间隔 TW ;

步骤 S05 :个体执行所述社交行为,并更新社交关系网络中的权重值;

步骤 S06 :所选择社交行为与下一社交行为的时间间隔减 1,所选择活动的持续时间减 1;

步骤 S07 :如果所选择活动的持续时间为 0,则返回步骤 S02,否则执行下一步;

步骤 S08 :如果所选择社交行为与下一社交行为的时间间隔大于 0,则返回步骤 S05,否则返回步骤 S04。

2. 根据权利要求 1 的所述的社交关系网络生成方法,其特征在于,在步骤 S01 中,利用基础数据构建人工交通系统,其中所述基础数据包括实际环境中场所分布、人口分布、天气变化、经济发展、生态环境、城市建设和社会制度七个方面的调查数据。

3. 根据权利要求 1 所述的社交网络生成方法,其特征在于,在步骤 S04 中,在个体 i 选择第 n<sub>i</sub> 次社交行为时,采用以下模型选择社交行为的类型:

个体 i 所有可能行为类型的集合记为 A<sub>i</sub><sup>t</sup>,选择 a<sub>k</sub> 作为当前行为类型的概率为:

$$p_t^i(a_k) = \frac{\exp(Q_t^i(a_k)/\tau_{n_i})}{\sum_{a_k \in A_i^t} \exp(Q_t^i(a_k)/\tau_{n_i})}$$

$$\tau_{n_i} = \gamma^{n_i} + \varepsilon$$

其中, γ 是 0 到 1 之间的一个系数, ε 是一个接近 0 的小数, t 为进行社交行为的时刻,Q<sub>t</sub><sup>i</sup>(a<sub>k</sub>)是行为 a<sub>k</sub> 的效用值,其初始值为 0,在执行行为 a<sub>k</sub> 的过程中,Q<sub>t</sub><sup>i</sup>(a<sub>k</sub>)的更新方法如下:

$$Q_{t+1}^i(a_k) = (1 - r_{n_i})Q_t^i(a_k) + r_{n_i}w_{i,j}(t)$$

其中, j 是个体 i 社交行为选择的对象,r<sub>n\_i</sub> =  $\frac{1-r_{min}}{1+n_i} + r_{min}$  是学习率, r<sub>min</sub> 是最小学习率。

4. 根据权利要求 1 所述的一种社交关系网络生成方法,其特征在于,在步骤 S05 中,如下更新社交关系网络中的权重值:

如果个体 i 和 j 在时刻 t 开始第一次社会交互,则个体 i 到 j 的连接权重 w<sub>i,j</sub> 按以下规则进行更新:

$$\begin{cases} w_{i,j}(t) = w_{i,j}(t-1) + f(d_{i,j})e^{-\alpha k_i} \\ \text{如果 } w_{i,j}(t) > 1, \quad w_{i,j}(t) = 1 \end{cases}$$

其中, α 是一个正的系数, k<sub>i</sub> 表示和个体 i 具有直接社交关系的个体个数, f(d<sub>i,j</sub>) 表示个体 i 到 j 的社交关系基准权重,其表示只考虑个体 i 和 j 的社会关系距离时的社会关

系权重；

已经交互过的个体 i 和 j 在时刻 t 再次交互，权重  $w_{i,j}$  按以下方法进行更新：

$$\begin{cases} w_{i,j}(t) = w_{i,j}(t-1) + f(d_{i,j})e^{-\beta w_{i,j}(t-1)} \\ \text{如果 } w_{i,j}(t) > 1, \quad w_{i,j}(t) = 1 \end{cases}$$

其中， $\beta$  是一个正的系数；

如果个体 i 和 j 在时刻 t 没有进行社会交互，权重  $w_{i,j}$  按以下方法进行更新：

$$w_{i,j}(t) = w_{i,j}(t-1)e^{-\kappa}$$

其中， $\kappa$  为减弱系数。

5. 根据权利要求 4 所述的社交网络生成方法，其特征在于，个体 i 到 j 的社交关系基准权重如下计算：

$$f(d_{i,j}) = \frac{1}{1 + d_{i,j}^\varphi}$$

$$d_{i,j}^\varphi = \sqrt{\sum_{L=1}^H w_L (h_i^L - h_j^L)^\varphi}$$

其中， $d_{i,j}^\varphi$  表示个体 i 到 j 的社会关系距离，H 为个体 i 和 j 的属性个数， $h_i^L$  和  $h_j^L$  分别为个体 i 和个体 j 的第 L 个属性， $w_L$  是个体属性的权重，满足  $\sum w_L = 1$ ， $\varphi$  是控制影响作用的系数值。

6. 根据权利要求 5 所述的社交网络生成方法，其特征在于，个体的属性包括年龄、性别、职业和性格。

7. 根据权利要求 1-6 任一项所述的社交网络生成方法，其特征在于，社交行为的类型包括：

$\alpha_0$  个体不进行社交行为；

$\alpha_1$ ：个体从当前活动场所内随机选择一个个体进行社交行为；

$\alpha_2$ ：个体和一个朋友进行社交行为；

$\alpha_3$ ：个体和一个朋友的朋友进行社交行为。

8. 根据权利要求 1-6 任一项所述的社交网络生成方法，其特征在于，步骤 S04 中，个体根据社交行为的类型，从当前所处场所上所有可能对象中随机选择一个对象，进行所选择的社交行为。

9. 一种基于人工交通系统的社交关系网络生成装置，其特征在于，包括：

构建模块，基于城市基础数据构建人工交通系统环境，人工交通系统中的每个个体具有自己的活动计划；并构建初始的社交关系网络，其中初始的社交关系网络的各个顶点为人工交通系统中的所有个体，两个个体之间的连接权重表示相应两个个体的社交关系；

活动选择模块，每个个体选择一个活动，并确定所选择活动的持续时间 Th；

活动执行模块，个体运动到活动场所，执行所选择的活动；

社交行为选择模块，个体选择并执行一个社交行为，然后确定下一个社交行为的时间

间隔 TW；

社交行为执行模块，个体执行所述社交行为，并更新社交关系网络中的权重值；

判定模块，所选择社交行为与下一社交行为的时间间隔减 1，所选择活动的持续时间减 1；如果所选择活动的持续时间为 0，则转活动选择模块；如果所选择社交行为与下一社交行为的时间间隔大于 0，则转社交行为执行模块，否则转社交行为选择模块。

## 一种基于人工交通系统的社交关系网络生成方法及装置

### 技术领域

[0001] 本发明涉及社交关系网络的生成方法。更具体而言，在人工交通系统对个体一天活动描述的基础上，建立个体之间社交行为模型，从而生成社交关系网络。

### 背景技术

[0002] 社交活动已经成为人们日常生活的重要组成部分，据调查，在美国、德国等发达国家，以社交活动为目的的出行已经占到总出行的 26.5% -39.5%。这个比例在发展中国家也有快速的增长，以北京为例，2011 年的比例已经达到 19.07%，比 2010 年增长了 26.8%。

[0003] 如何描述人们之间的社交活动关系，即社交关系网络，是研究社交活动影响与作用的基础。目前已经提出了大量对社交网络进行分析的方法和手段，并在传染病传播、网络舆情分析、应急事件处理中得到应用。但这些方法都是建立在社交网络已经获得的基础上，而对于如何生成社交网络还没有有效方法。目前关于一般采用调查问卷、电话咨询、面对面调查等方式获取社交关系网络，这些方面开展起来有很大难度，成本很高并且效果很不理想。加拿大多伦多市曾经开展了一次社交关系网络面对面调查，350 人中只有 87 人完成了调查。在苏黎世开展的一次社交关系网络的调查问卷，4000 人中只有 307 人完成了调查。如此低的完成比例，显然无法保证数据的有效性和可用性，如何通过较低成本生成较大人群数量的社交关系网仍是亟需解决的一个问题。

[0004] 人工交通系统 (Artificial Transportation Systems, ATS) 是近几年随着复杂系统和智能科学的深入研究而出现的新概念，旨在构建能够按照一定的指导机制，自主地进行演化并同外部进行交互的虚拟交通系统。其主要思想是通过建立个体代理模型，在计算机中营造一个虚拟的环境，利用大量代理之间的相互作用，涌现宏观交通现象。因而交通系统中的宏观交通现象并不是建模者设计进去的，而是系统自发地自下而上地涌现生成的，从而把计算机作为交通实验室，以个体为单位，以行为分析理论为基础，充分考虑个体的性别、年龄、收入、居住地等影响出行者交通模式和行为的因素，建立基于活动的交通需求模型，用于模拟人口的出行活动过程，合理预测交通需求，进而针对人口和交通需求关系问题开展分析、评估、决策等研究。

[0005] 人工交通系统能够描述个体在地理空间上的移动过程，以及各活动场所上人群分布的变化过程。

### 发明内容

[0006] 有鉴于此，本发明提出了一种基于人工交通系统的社交关系网络生成方法，以解决如何通过较低成本生成较大人群数量的社交关系网的问题。

[0007] 根据本发明一方面，其提供了一种基于人工交通系统的社交关系网络生成方法，其特征在于，包括以下步骤：

[0008] 步骤 S01：基于城市基础数据构建人工交通系统环境，人工交通系统中的每个个体具有自己的活动计划；并构建初始的社交关系网络，其中初始的社交关系网络的各个顶

点为人工交通系统中的所有个体,两个个体之间的连接权重表示相应两个个体的社交关系;

- [0009] 步骤 S02 :每个个体选择一个活动,并确定所选择活动的持续时间 Th ;
- [0010] 步骤 S03 :个体运动到活动场所,执行所选择的活动 ;
- [0011] 步骤 S04 :个体选择并执行一个社交行为,然后确定下一个社交行为的时间间隔 TW ;
- [0012] 步骤 S05 :个体执行所述社交行为,并更新社交关系网络中的权重值 ;
- [0013] 步骤 S06 :所选择社交行为与下一社交行为的时间间隔减 1 ,所选择活动的持续时间减 1 ;
- [0014] 步骤 S07 :如果所选择活动的持续时间为 0 ,则返回步骤 S02 ,否则执行下一步 ;
- [0015] 步骤 S08 :如果所选择社交行为与下一社交行为的时间间隔大于 0 ,则返回步骤 S05 ,否则返回步骤 S04 。
- [0016] 根据本发明另一方面,其提供了一种基于人工交通系统的社交关系网络生成装置,其特征在于,包括:
  - [0017] 构建模块,基于城市基础数据构建人工交通系统环境,人工交通系统中的每个个体具有自己的活动计划;并构建初始的社交关系网络,其中初始的社交关系网络的各个顶点为人工交通系统中的所有个体,两个个体之间的连接权重表示相应两个个体的社交关系;
  - [0018] 活动选择模块,每个个体选择一个活动,并确定所选择活动的持续时间 Th ;
  - [0019] 活动执行模块,个体运动到活动场所,执行所选择的活动 ;
  - [0020] 社交行为选择模块,个体选择并执行一个社交行为,然后确定下一个社交行为的时间间隔 TW ;
  - [0021] 社交行为执行模块,个体执行所述社交行为,并更新社交关系网络中的权重值 ;
  - [0022] 判定模块,所选择社交行为与下一社交行为的时间间隔减 1 ,所选择活动的持续时间减 1 ;如果所选择活动的持续时间为 0 ,则转活动选择模块;如果所选择社交行为与下一社交行为的时间间隔大于 0 ,则转社交行为执行模块,否则转社交行为选择模块。
  - [0023] 本发明的有益效果是,能够代替传统的调查问卷、电话调查、面对面调查等方式,通过计算机自动生成社交关系网络,实施成本很低,并很容易针对大规模人群进行实施。

### 附图说明

- [0024] 图 1 是本发明提出的生成社交关系网络方法的流程图。
- [0025] 图 2 是应用本发明生成的社交关系网络的连接密度 (Tie density) 图。
- [0026] 图 3 是应用本发明生成的社交关系网络的集群系数 (Clustering coefficient) 图。

### 具体实施方式

- [0027] 为使本发明的目的、技术方案和优点更加清楚明白,以下结合具体实施例,并参照附图,对本发明作进一步的详细说明。
- [0028] 本发明在人工交通系统软件的基础上进行实施。人工交通系统起源于人工社会的

理念,采用自底向上的方法生成虚拟交通系统。人工交通系统的核心包括三部分。首先,实际社会中的每个个体被描述成具有一定年龄、性别、家庭位置等属性的智能体,智能体本身具有一定智能性和主观能动性,其全体就构成了一个城市的人工人口。其次,采用基于活动生成出行的方法,为每个智能体生成一天的活动计划。最后,每个智能体根据活动计划完成一天的出行,采用跟车模型、换道模型等微观模型描述智能体的微观运动过程,通过大量简单个体运动过程,生成复杂的宏观交通现象,在此基础上可以开展交通管理方案的实验、评估与优化等功能。

[0029] 在具体实施过程中,本发明提出的社交关系网络生成方法可以作为一个插件,运行在人工交通系统软件的基础之上。

[0030] 本发明中的社交关系网络定义为一个无向图  $G=(V, E, W)$ , 其中  $V$  是顶点集合,  $E$  是连接顶点的边的集合。每个顶点代表一个个体,个体的属性可以用一个  $H$  维的向量  $\vec{h}_i = (h_i^1, h_i^2, \dots, h_i^H)$  表示,其中每一维表示一个属性,如年龄、性别、职业、性格等。 $W=\{w_{i,j} \in [0,1]\}$  是一个表示个体之间社交关系的权重矩阵,  $w_{i,j}$  接近 0 表示个体  $i$  到  $j$  的关系比较疏远,接近 1 表示个体  $i$  到  $j$  的关系比较密切。如果  $w_{i,j}$  大于预设的阈值(常用的值包括 0.05, 0.25, 0.50 和 0.75),表示个体  $i$  到  $j$  有社交关系,否则没有社交关系。

[0031] 图 1 是本发明提出的社交关系网络生成方法的实现流程图,其执行过程包括以下步骤:

[0032] 步骤 S01: 基于城市基础数据构建人工交通系统环境,人工交通系统中的每个个体具有自己的活动计划;并构建初始的社交关系网络,其中初始的社交关系网络的各个顶点为人工交通系统中的所有个体,连接每两个顶点边表示相应两个个体的社交关系;

[0033] 这一步完成人工交通系统的初始化过程,使用的基础数据包括场所分布、人口分布、天气变化、经济发展、生态环境、城市建设、规章制度七个方面的调查数据。在此基础上,生成与实际交通系统相对应的人工交通系统。其中的每个个体被描述成具有一定智能性的代理,个体具有自己的活动计划,根据活动计划的需求生成交通出行需求。

[0034] 步骤 S02: 每个个体选择第一个(下一个)活动  $h$ ,确定活动的持续时间  $T_h$ ;

[0035] 步骤 S03: 个体运动到活动场所,执行活动  $h$ ;

[0036] 在该步骤,个体首先将活动  $h$  的活动场所设为目的地,根据个人偏好、出行习惯、当前拥堵情况等生成由当前位置到目的地的出行方式和出行路径,按运动模型(包括步行、公共交通、私家车)运动到目的地,到达目的地后开始执行活动  $h$ 。

[0037] 步骤 S04: 个体选择并执行社交行为  $w$ ,然后确定下一个社交行为的时间间隔  $T_W$ ;

[0038] 在该步骤,个体首先采用 Softmax 动作选择模型选择社交行为的类型,个体之间的社交行为可以有以下四种类型:

[0039]  $\alpha_0$ : 个体不进行社交行为;

[0040]  $\alpha_1$ : 个体从当前活动场所内随机选择一个个体进行社交行为;

[0041]  $\alpha_2$ : 个体和一个朋友进行社交行为;

[0042]  $\alpha_3$ : 个体和一个朋友的朋友进行社交行为;

[0043] 将个体  $i$  所有可能行为类型的集合记为  $A_i^t$ ,选择  $\alpha_k$  作为当前行为类型的概率为:

$$[0044] p_t^i(\alpha_k) = \frac{\exp(Q_t^i(\alpha_k)/\tau_{n_i})}{\sum_{\alpha_k \in A_t^i} \exp(Q_t^i(\alpha_k)/\tau_{n_i})}$$

$$[0045] \tau_{n_i} = \gamma^{n_i} + \varepsilon$$

[0046] 其中,  $n_i$  是个体  $i$  进行选择的累计次数 (初始为 0),  $\gamma$  是 0 到 1 之间的一个系数, 缺省为 0.9,  $\varepsilon$  是一个接近 0 的小数, 其缺省值为 0.1, 避免当  $n_i$  趋近于无穷大时,  $\tau_{n_i}$  趋近于 0。 $Q_t^i(\alpha_k)$  是行为  $\alpha_k$  的效用值, 其初始值为 0, 在执行行为  $\alpha_k$  的过程中,  $Q_t^i(\alpha_k)$  的更新方法如下:

$$[0047] Q_{t+1}^i(\alpha_k) = (1 - r_{n_i})Q_t^i(\alpha_k) + r_{n_i}w_{i,j}(t)$$

[0048] 其中,  $j$  是个体  $i$  社交行为选择的对象,  $r_{n_i} = \frac{1-r_{min}}{1+n_i} + r_{min}$  是学习率, 当  $n_i$  趋近于无穷大时,  $r_{n_i}$  趋近于  $r_{min}$ 。 $r_{min}$  是最小学习率, 其缺省值为 0.1。

[0049] 然后, 个体选择社交行为的对象。个体根据行为类型  $\alpha_k$ , 从当前处于同一地点上的所有可能对象中随机选择一个对象  $j$ , 将当前地点上所有可能对象的集合记为  $\Omega$ , 则个体  $j$  被选中的概率为:

$$[0050] P\{j = k | \forall k \in \Omega\} = \frac{1}{|\Omega|}$$

[0051] 最后, 个体确定社交行为的时间。假设个体  $i$  的两次社交行为之间的行为间隔  $T_w$  服从参数为  $\frac{1}{\lambda_i}$  的负指数分布, 即  $T_w$  的概率分布为

$$[0052] F(t) = P\{T_w < t\} = 1 - e^{-\lambda_i t}$$

[0053] 那么, 个体  $i$  的累积交互次数  $X^i(t)$  就形成一个参数为  $\lambda_i$  的泊松过程。可推导得  $X^i(t)$  的期望为  $\lambda_i t$ , 所以通过调整  $\lambda_i$  就可以控制代理的交互频率, 从而间接地加速或减慢社交网络的生长速度。

[0054] 步骤 S05: 个体执行社交行为, 更新社交关系网络中的权重值;

[0055] 如果个体  $i$  和  $j$  在当前时刻开始第一次社会交互, 则个体  $i$  和其行为对象  $j$  之间的连接权重  $w_{i,j}$  按以下规则进行更新:

[0056]

$$\begin{cases} w_{i,j}(t) = w_{i,j}(t-1) + f(d_{i,j})e^{-\alpha k_i} \\ \text{如果 } w_{i,j}(t) > 1, \quad w_{i,j}(t) = 1 \end{cases}$$

[0057] 其中,  $\alpha$  是一个正的系数,  $k_i$  表示和个体  $i$  具有直接社交关系的个体个数,  $f(d_{i,j})$  表示个体  $i$  到  $j$  的社交关系基准权重, 其表示只考虑个体  $i$  和  $j$  的社会关系距离时的社会关系权重; 计算方法如下

[0058]

$$f(d_{i,j}) = \frac{1}{1 + d_{i,j}^\varphi}$$

[0059]

$$d_{ij}^\varphi = \sqrt{\sum_{L=1}^H w_L (h_i^L - h_j^L)^\varphi}$$

[0060]  $d_{ij}^\varphi$  表示个体 i 到 j 的社会关系距离, 所述社会关系包括两个个体的属性如年龄、性别、职业、性格等之间的关系,  $w_L$  是个体属性的权重, 满足  $\sum w_L = 1$ ,  $\varphi$  是控制影响作用的系数值, 缺省为 2。

[0061] 如果已经交互过的个体 i 和 j 在当前时刻再次交互, 个体 i 和其活动对象 j 之间的连接权重  $w_{i,j}$  按以下规则进行更新

[0062]

$$\begin{cases} w_{i,j}(t) = w_{i,j}(t-1) + f(d_{i,j}) e^{-\beta w_{i,j}(t-1)} \\ \text{如果 } w_{i,j}(t) > 1, w_{i,j}(t) = 1 \end{cases}$$

[0063] 其中  $\beta$  是一个正的系数, 缺省取值 1。

[0064] 如果个体 i 和 j 之间在当前时刻没有发生社交行为, 二者之间的连接权重  $w_{i,j}$  按以下方法进行更新:

[0065]  $w_{i,j}(t) = w_{i,j}(t-1) e^{-\kappa}$

[0066] 其中  $\kappa$  为减弱系数, 缺省值为 0.5。从上面的公式可以看出, 当 i 和 j 之间没有活动关系时, 个体 i 到 j 的关系会逐渐减弱。

[0067] 步骤 S06 : TW=TW-1, Th=Th-1 ;

[0068] 步骤 S07 : 如果 Th=0 (当前活动结束), 返回步骤 S02, 执行下一活动, 否则执行下一步;

[0069] 步骤 S08 : 如果  $T_w > 0$  (当前社交行为还未结束), 返回步骤 S05, 继续执行当前社交行为, 否则返回步骤 S04, 执行下一社交行为。

[0070] 所生成的社交网络将保存为开放格式的 Pajek NET 文件, 大部分的社交网络分析工具 (如 Pajek, UCINET, Gephi 等) 都支持这种格式, 使用者也可以根据 Pajek Net 格式的定义自己的解析文件, 从而获得建立社交网络所需的数据。生成的社交网络除了可以用于人工交通系统本身的社交出行研究外, 还可以用于其它领域的多项研究。下面举出两个典型应用的例子:

[0071] (1) 研究疾病 (信息) 的传播过程。

[0072] 社交活动是传染性疾病的传播的主要渠道之一。社交网络中的边表示个体之间的一次活动, 即传染病疾病传播的一个渠道。社交网络中的节点状态表示个体是否感染疾病。通过社交网络的演化过程, 可以研究传染性疾病的传播过程。这里的疾病还可以进一步推广为能通过社交活动传播的信息, 如商品价格、股市波动等信息。

[0073] (2) 研究网络节点的影响力

[0074] 在网络安全、网络舆情分析、疾病传播的过程中, 不同的个体、团体之间影响力差别非常大。及时发现影响力大, 可能造成严重影响的个体和团体对这类分析非常关键。以社交网络中的节点代表个体, 边代表个体之间的影响关系, 通过观察社交网络在不同初始条件下的演化过程, 可以找出对整个网络有严重影响的个体和团体。

[0075] 为了验证模型的有效性,我们以北京市海淀区中关村区域作为物理世界参考模型,建立了人工交通系统,模拟该区域内 20000 人口 62 天(两个月)的活动过程,将每一天结束后得到社交关系网络记为  $G^T$ ,其中上标 T 为运行的天数。将阈值分别设为 0.05,0.25,0.50 和 0.75,过滤掉低于阈值的边,分别得到四组社交关系网络,记为  $G_{0.05}^T, G_{0.25}^T, G_{0.50}^T, G_{0.75}^T$ 。我们对生成的社交关系网络做了详尽分析,证明了本发明方法的有效性。下面我们以连接密度和集群系数两个分析社交关系网络时的常用指标为例进行具体说明。

[0076] 图 2 是四组社交关系网络中的连接密度 (Tie density) 图。网络中的连接密度定义为:

$$[0077] \rho = \frac{2e}{N(N - 1)}$$

[0078] 其中,e 是网络中边的数量,N 是网络中节点的数量。从图中可以看出,四组社交关系网络的连接密度都趋向与一个稳定的值。

[0079] 图 3 是四组社交关系网络的集群系数 (Clustering coefficient) 图。一个节点的集群系数定义为:

$$[0080] C_i = \begin{cases} \frac{2m_i}{k_i(k_i - 1)} & \text{if } k_i > 1 \\ 0 & \text{if } k_i \leq 1 \end{cases}$$

[0081] 其中  $k_j$  表示和节点连接的边的数量, $m_j$  是和这个节点连接的三角形的数量。一个网络中的集群系数定义为网络中所有节点集群系数的平均值。从结果中可以看出,随着时间的推移,四组社交关系网络的连接密度都趋近于稳定。根据实际调查结果,社交网络的集群系数范围一般是 0.13 ~ 0.88,图 3 中的结果也是符合的。

[0082] 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,应理解的是,以上所述仅为本发明的具体实施例而已,并不用于限制本发明,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。

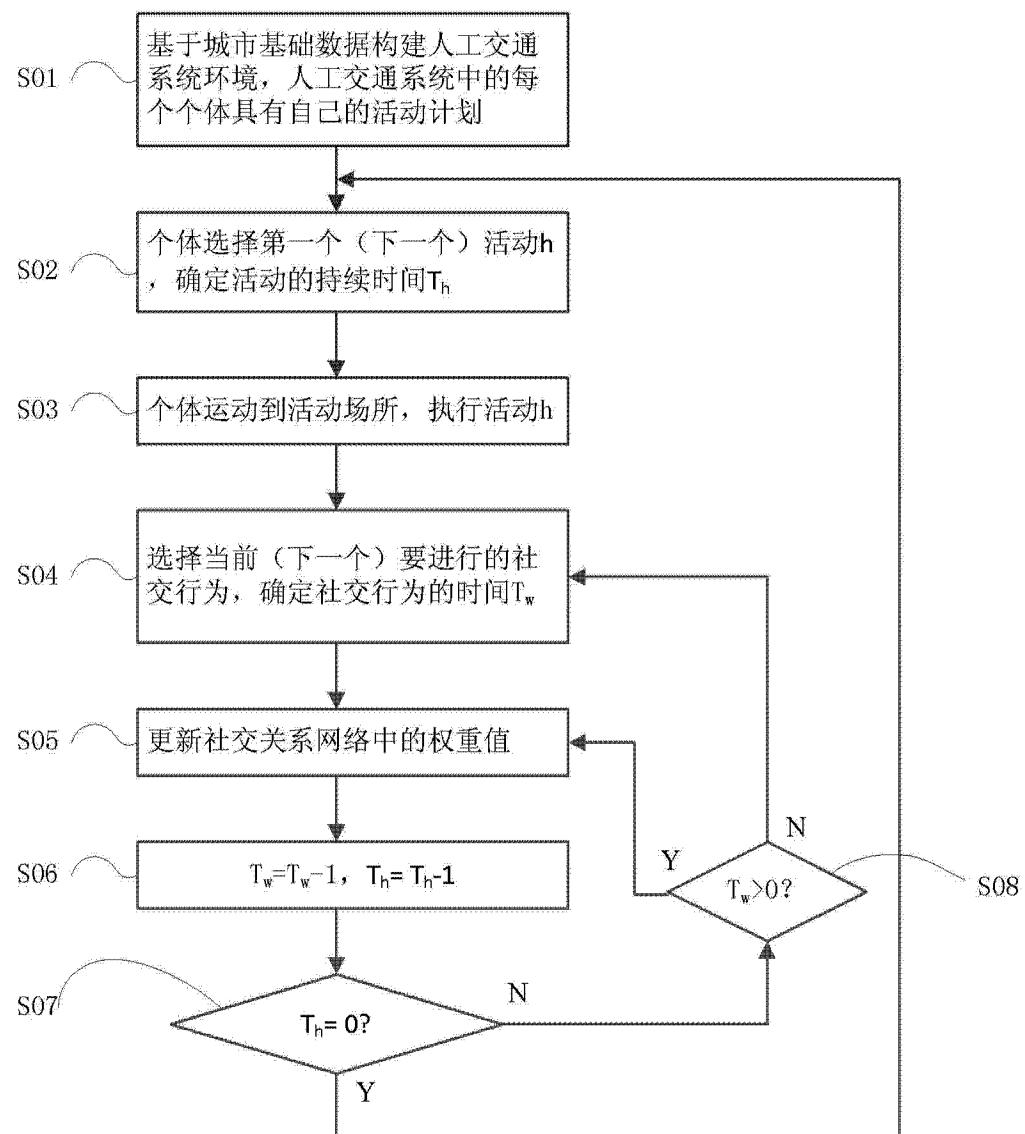


图 1

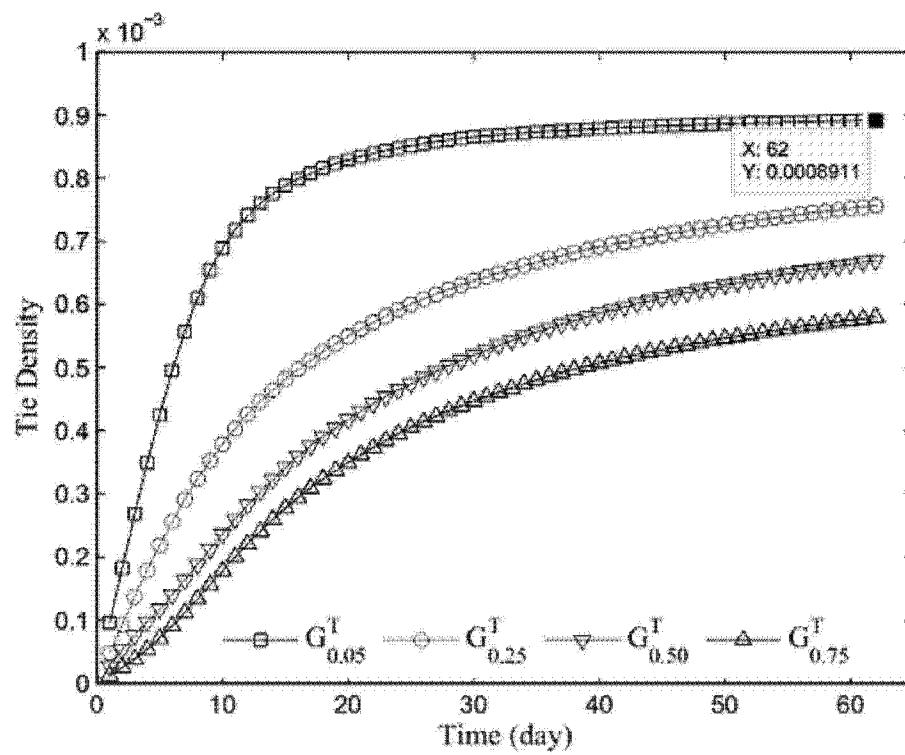


图 2

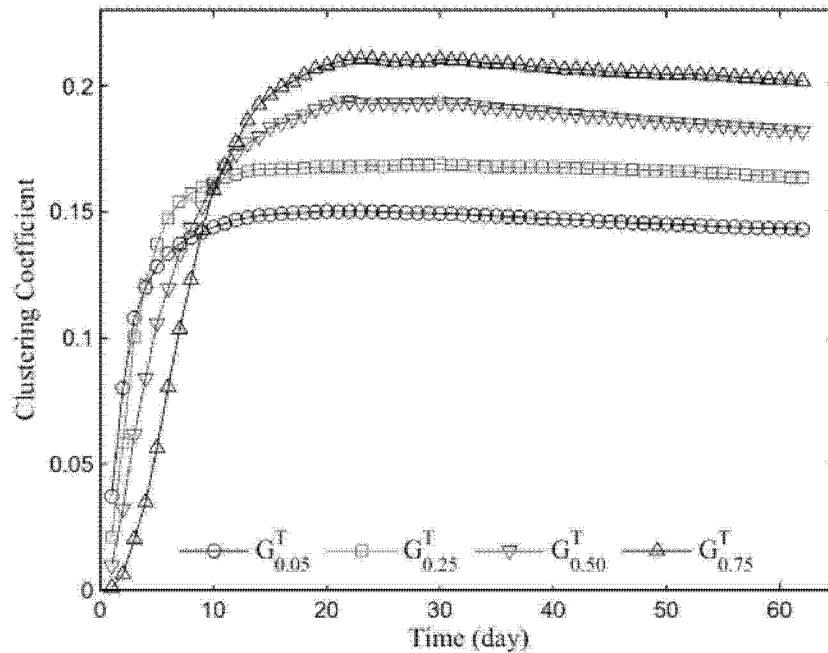


图 3