

一种面向人群疏散的高效分组方法

张建新 刘 弘 李 焱

(山东师范大学信息科学与工程学院 济南 250014)

(山东师范大学山东省分布式计算机软件新技术重点实验室 济南 250358)

摘 要 在人群疏散的过程中,个体会依据关系的亲密度产生分组现象,因此人群分组行为是人群疏散仿真中不可忽略的因素。家人、朋友、同事等会根据亲密度形成分组,在疏散过程中同组人群会聚集成簇。聚类分组时常用的 k-medoids 聚类算法对噪声敏感,容易陷入局部最优,只能发现球状簇,且对初始聚类中心点的选择敏感,在聚类准确度上不尽人意。而 DBSCAN 算法具有抗噪声能力强、可发现任意形状的簇、无须指定初始聚类中心等优点,但只能识别密度相近的簇。对此,文中提出了折半 DBSCAN 聚类算法。该算法首先对关系数据进行二分划分,将有关系的数据划分到一个网格中,然后根据每个网格的人群密度决定聚类半径 ϵ ,最后对每个网格进行 DBSCAN 聚类,因此该算法可识别密度不同的簇。人群聚类分组后,在加入同组内个体吸引力的社会力模型中驱动个体运动,并模拟关系密切程度对聚集程度的影响。实验结果表明,在考虑了现实生活中有关系的人群空间分布状况下,所提方法具有较高的聚类精度,可真实地再现现实场景中的人群疏散情况,可作为紧急情况下预测人群疏散时间和疏散状况的重要工具。

关键词 聚类算法, k-medoids, DBSCAN 聚类, 二分划分, 人群疏散仿真

中图分类号 TP319 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.06.035

Efficient Grouping Method for Crowd Evacuation

ZHANG Jian-xin LIU Hong LI Yan

(School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China)

(Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology, Shandong Normal University, Jinan 250358, China)

Abstract In the crowd evacuation process, individuals usually produce the grouping phenomenon according to the intimacy of the relationship. Therefore, the grouping behavior is a factor that can not be neglected in the evacuation simulation of the crowd. The family, friends and colleagues usually form a group according to the degree of intimacy and gather together to a cluster in the evacuation process. The commonly used k-medoids clustering algorithm is sensitive to noise and easy to fall into the local optimum. It can only find the spherical cluster and is sensitive to the selection of the initial clustering center point, which is unsatisfactory in the accuracy of clustering. The DBSCAN algorithm has the advantages of strong ability to deal with noise and can find clusters of arbitrary shape and without specifying the initial clustering center, etc. But it can only identify clusters of similar density. Therefore, this paper proposed a binary DBSCAN clustering algorithm. This algorithm firstly divides the relational data to a grid, then it determines the cluster radius ϵ according to the density of population of the grid, and finally it executes the DBSCAN clustering algorithm for each grid, so these clusters with different densities can be identified. After clustering, the individual movement is driven in the social force model which adds the individual attraction in the same group. And the influence of the intimacy degree on the aggregation degree is simulated. The experimental results show that, considering the spatial distribution of connected pedestrians in real life, this method has higher clustering accuracy. It can reappear the evacuation situation in the real scene and can be used as an important tool to predict evacuation time and evacuation situation.

Keywords Clustering algorithm, k-medoids, DBSCAN clustering, Binary partition, Crowd evacuation simulation

到稿日期:2018-07-20 返修日期:2018-10-16 本文受国家自然科学基金(61876102,61472232,61272094)资助。

张建新(1992-),女,硕士生,CCF会员,主要研究方向为群体智能计算,E-mail:1144693673@qq.com;刘弘(1955-),女,博士,教授,博士生导师,主要研究方向为分布式人工智能、软件工程及计算机辅助设计等,E-mail:lhscdn@126.com(通信作者);李焱(1980-),男,博士,讲师,主要研究方向为计算机辅助设计及智能计算。

1 引言

在建筑物密集区突发火灾、地震以及遭遇恐怖袭击等紧急事件时容易发生拥挤和踩踏事件,如2006年菲律宾马拉拉体育场外的踩踏事件、2014年上海外滩跨年夜踩踏事件、2015年沙特麦加踩踏事件等,使得社会公共安全和人民的生命财产安全面临严峻的挑战。深入挖掘突发事件下密集人群的行动规律,对可能引发的拥挤、骚乱甚至践踏等事件进行有效的预警和干预,对恐怖袭击或突发灾害后的人群疏散形成完善的硬件支持和高效的现场疏导策略具有重大意义。

为了产生真实的人群疏散行为,运用计算机仿真技术对场景建模、路径寻优以及人群运动行为建模,模拟公共场所的应急疏散现象,以在达到疏散演练效果的同时将成本最小化^[1]。

社会比较理论是由Festinger于1954年提出的一种社会心理学理论^[2],其核心思想是:人类在缺乏客观的方法来评估他们的状态时,会把他们自己的状态与其他人进行比较,然后试图纠正发现的差异。在紧急情况下的疏散过程中,社会比较更为显著,从而导致分组现象^[3]。目前,在人群疏散的模型中,比较典型的是社会力模型。该模型建模简单,能很好地模拟“出口成拱”“快即是慢”等现象,但没有考虑到人与人之间的关系亲密度对人群疏散的影响,不能模拟突发事件中依据社会关系进行的分组聚团运动^[4]。

针对这个问题,研究人员提出了依据关系分组的信息,即在紧急情况下的人群疏散过程中,亲密度高的行人总是趋于一个组,亲密度越高,人群的聚集程度越大。文献^[1]利用数据挖掘中改进的k-medoids算法,依据关系亲密度和距离对人群进行聚类;但该算法中初始聚类中心的个数和位置的选取对聚类结果的影响较大,且其只能识别球状簇,对噪音敏感且容易陷入局部最优。传统的DBSCAN聚类算法无须指定初始聚类中心的个数和位置,能识别任意形状的簇,且对噪音不敏感;但它使用全局的邻域半径 ϵ 和全局的形成高密度区域所需的最少点数 $MinPts$,因此只能发现密度相近的簇,而且聚类时需要计算任意两个粒子之间的距离,耗时较长。

针对上述问题,本文提出了折半DBSCAN聚类算法,其依据关系亲密度和距离对人群进行聚类分组,分组后在加入同组内个体吸引力的社会力模型中实现路径规划并碰撞避免。实验结果表明,折半DBSCAN聚类算法可以有效提高分组效率。聚类分组后的人群在加入同组内个体吸引力的社会力模型中聚集运动,真实地模拟了同组个体之间不同的社会关系产生的受力情况,能够更逼真地模拟人群的疏散运动。

2 相关工作

2.1 聚类算法

聚类是对物理或抽象的样本集合进行分组的过程,是将数据划分为有意义或有用的组的子集,也称为簇。聚类结果使得每个簇内部的样本之间的相关性比其他簇中样本之间的相关性更紧密,即簇内部的样本之间具有较高的相似度,而不同簇的样本之间具有较高的相异度。

聚类算法主要可以划分为:基于划分的聚类算法,基于层

次的聚类算法,基于密度的聚类算法,基于网格的聚类算法,基于模型的聚类算法,以及基于图的聚类算法。

DBSCAN聚类算法是一种基于高密度连通区域的聚类算法,其思想是通过检查数据集中每个数据的 ϵ 邻域来搜索簇,如果点 p 的 ϵ 邻域内的点数量多于 $MinPts$,则创建一个以 p 为核心对象的簇;然后通过DBSCAN聚类算法迭代地聚集从这些核心对象直接密度可达对象,这个过程可能涉及一些密度可达簇的合并。当没有新的点添加到任何簇时,该过程结束。谭颖等^[5]提出了多密度阈值的DBSCAN聚类算法,解决了DBSCAN聚类算法参数选取困难和难以发现密度相差较大的簇的问题。Xia^[6]提出了一种自适应算法sa-db-scan,该算法通过分析数据集的统计特性来自动确定输入参数,使DBSCAN聚类算法的聚类过程完全自动化。Kumar等^[7]提出了一种通过使用组的方法加速邻居搜索的快速DBSCAN聚类算法,其有效地提高了DBSCAN聚类算法的聚类速度。何熊熊等^[8]提出了一种基于密度和网格的簇心可确定聚类算法,该算法具有较高的聚类准确性和较快的执行速度。文献^[9-10]解决了如何在数据偏移的情况下实现理想的负载均衡问题。

本文在传统DBSCAN聚类算法的基础上引入二分划分的思想,对属性相近的粒子进行折半DBSCAN聚类。实验结果表明:该算法具有良好的聚类分组效果,分组的准确率较高。

2.2 人群运动仿真模型

目前,行人疏散问题的研究主要从微观和宏观两个角度进行。微观模型主要以个体行人为研究对象,分析行人的体型、移动频率、移动速度以及个体空间需求等基本参数特性^[11],研究影响密集人群疏散过程的各种因素,以及人们在面对不同情况下的心理、行为特性^[12]和个体运动过程中的视野范围、速度变化、避让障碍物等动力学特性^[13]。目前,微观模型主要有:元胞自动机模型,格子气模型,社会力模型,离心力模型,磁场力模型,以及最优速度模型等。行人流的宏观模型以整体为研究对象,将行人流看作一个可压缩的实体或连续流动的介质,对三要素的特征以及相互关系进行刻画,重点研究流量分布和时空演化特性、拥堵和群集现象产生的机理,以及自组织行为对疏散过程的影响。目前,宏观模型大多通过流体力学理论、最优化理论对密集人群的流体运动过程和路径选择行为进行刻画^[14-15]。

在社会力疏散模型方面,Hebing等^[16]对人群恐慌进行了分析,提出了社会力模型。Lakoba等^[11]对社会力模型的参数进行了研究,并对社会力模型的稳定性和精度进行了验证。Steffen等^[17]在社会力模型中加入了视野因素,研究了视野范围对行人疏散过程的影响。曹梦晓等^[18]研究了个性化的情绪感染对人群疏散的影响。Parisi等^[19]引入了减速机,运用社会力模型仿真得出常态情况下的基本图模式。纪庆革等^[20]提出利用密度场来对社会力模型进行改进,改进后的模型不仅降低了时间复杂度,而且能模拟人群自组织现象。Musse等^[21]考虑了群组关系和群组内突发行为对群组行为的影响,将群组关系和碰撞检测算法相结合,对人群行为进行了模拟。Vizzari等^[22]综合考虑了个体的运动动机和群组对

个体的影响,建立了智能体群组模型。Qiu 等^[23]提出了在行人群体中对不同群组结构进行建模的统一框架,考虑组内结构和组间关系,对群组跟随其他群组的进行了模拟,验证了群体规模、群体内部结构、群体间的关系会对群体行为产生显著影响。Fu 等^[24]将多智能体技术与元胞自动机相结合,建立了群组行为模型,体现了 agent 之间的差异,并研究了疏散人数和出口布局对疏散时间的影响。文献^[25]基于 agent 的行人群体仿真,提出了社会群组建模框架,将 agent 行为建模、群组建模和社会环境建模集成在一个分层的体系结构中,其中每一层都侧重于模拟人群的一个特定方面。Qin 等^[26]提出了一个集群运动模型来模拟双向行人流,在原始社会力模型中引入了组间吸引力和碰撞避免力。Zhang 等^[27]提出了两层社会力模型来模拟群体的聚集过程,改进的社会力模型考虑了组间吸引力和引领者对非引领者的吸引力。Liu 等^[28]提出了一个视频驱动的社会力模型来模拟人群疏散,该模型引入了适应度函数来模拟出口拥挤时行人的心理状态。Xu 等^[29]提出了一种基于物联网的灾害管理人群疏散规划系统,该系统利用移动云计算平台来开发面向行人跟踪和帮助行人疏散的优化系统,并提出了基于关系吸引的人工势场疏散规划算法。文献^[30]提出了一种新的应急疏散路径规划方法,该方法结合了扩展的社会力模型和改进的人工蜂群算法,考虑了视觉参数对行人疏散的影响,同时考虑了出口拥挤时行人的心理选择。尽管已有很多研究人员对疏散模型做出了改进,使得人群疏散模型越来越接近真实情况下的人群疏散,但是细节仍然可以进一步修改。位置属性和关系值属性特定的行人依据不同的聚类算法识别出的簇不同,而簇的精度影响着行人的疏散现象及结果,因此正确地识别簇是模拟

$$dist(i, j) = \sqrt{(particle[i]. pos_{dx} - particle[j]. pos_{dx})^2 + (particle[i]. pos_{dy} - particle[j]. pos_{dy})^2}$$

文献^[33-34]将社会比较理论抽象为公式,提出了计算个体之间相异度的算法。本文依据该文献中的社会比较理论思想,同时考虑距离和关系对聚类结果的影响,来计算个体之间的相异度。距离和关系的数量级不同,需要对两个因素进行归一化。因此,改进后的相异度定义为:

$$dif = k_1 \times \frac{dist(i, j)}{rangedist} + k_2 \times \frac{1}{rangerel \times relation(i, j)} \quad (2)$$

其中, $dist(i, j)$ 表示个体 i 和个体 j 之间的几何距离, $relation(i, j)$ 表示两个个体之间的关系值, $rangedist$ 表示距离归一化因子, $rangerel$ 表示关系归一化因子。改进后的相异度既考虑了关系对分组聚类的影响,又考虑了距离对分组的影响,避免了关系较密切但距离较远的两个粒子之间相互吸引、绕远路的现象。 k_1 和 k_2 分别表示距离和关系所占的权重,本文中 $k_1 = 0.5, k_2 = 0.5$ 。

传统的 DBSCAN 聚类算法使用全局的邻域半径 ϵ , 因此难以发现密度相差较大的簇。当数据密度或簇间距离分布不均匀时,若根据较密的那些簇选取较小的 ϵ 值,则对于客观上相对较稀的那些簇中的对象,它们邻域内的数据对象数目小于 $MinPts$, 这些对象将被认为是边界对象。因此,较稀的簇会被划分成多个性质相似的簇。与此相反,若根据较稀的那些簇来选取较大的 ϵ 值,则距离较近而密度较大的那些簇将很可能被合并为同一个簇,它们之间的差异也就因为选取较

行人疏散仿真的基础。

本文采用加入同组内个体吸引力的社会力模型,该模型不仅考虑了在疏散过程中同组个体间的关系密切程度对聚集程度的影响,而且考虑了引领对每个分组的影响。

3 折半 DBSCAN 聚类算法

人与人之间存在多种社会关系,本文将这些社会关系划分为亲人、朋友、同事及陌生人 4 种,并对其量化,所有行人之间的关系值均处于 0~1 之间,关系值越大代表关系越亲密。亲人、朋友、同事关系系统称为有关系,陌生人关系则称为无关系。在疏散过程中,个体会依据关系的亲密度产生分组现象,因此在疏散过程中加入社会关系是必要的。在常态下,具有亲人、朋友、同事关系的人通常距离较近,但也存在部分具有关系的人暂时分开而后期汇合的现象,且在人群密集的区域有关系的人相距较近,在人群稀疏的区域有关系的人相距相对较远^[31]。折半 DBSCAN 聚类算法中使用二分划分的方法将有关系的群组划分到同一网格。在这个网格中根据人群密度确定邻域半径,从而避免了使用全局邻域半径无法识别不同密度群组的问题。它无须指定初始聚类中心的位置和个数,能识别任意形状的簇,且能得到全局最优解。本文将折半 DBSCAN 聚类算法与 k-medoids 算法、最大密度峰值聚类算法^[32]的聚类效果作比较。

3.1 改进的 DBSCAN 聚类算法

传统的 DBSCAN 聚类算法在计算个体之间的相异度 dif 时,仅考虑了距离这一个因素,其计算公式如下:

$$dif(i, j) = dist(i, j) \quad (1)$$

其中:

大的 ϵ 值而被忽略。显然,在上述两种情况下,其实很难选取一个合适的 ϵ 值来进行聚类且得到比较准确的聚类结果。因此,密度阈值的参数 ϵ 和 $MinPts$ 在全局唯一的情况下,只能发现密度近似的簇。

本文引入二分划分的思想,根据关系矩阵遍历所有关系值不为 0 的群组(简称有关系的群组),对每个有关系的群组进行二分划分。划分的原则是先对整个场景对半竖切,如果切完之后发现有关系的群组粒子被切分到了不同的网格,则撤销这次竖切,改为对半横切;如果切完之后发现有关系的群组粒子被切分到了相同的网格,则将群组被切到的网格作为新的场景来递归地进行上面的操作,直到新的场景中无论是横切还是竖切都会将有关系的粒子划分到不同的网格为止,此时的场景为该有关系的群组的最终场景,如图 1 所示。最终场景中可能存在与整个有关系的群组无关系的粒子。在最终的场景中遍历所有粒子,并计算每个粒子的 k - $dist$ 值。在本文中, k 取 3(3- $dist$ 意味着离某粒子 i 最近的第三个粒子的距离)。将所有粒子的 3- $dist$ 按从大到小进行排序,然后绘制排序后的 3- $dist$ 图,横坐标代表排序后的对象,纵坐标表示对应的 3- $dist$ 值。将离散点连成平滑的曲线,计算该曲线的导函数变化最快的点,该点的纵坐标即为所求的邻域半径 ϵ 。接着用 DBSCAN 聚类算法对这个最终场景中的所有粒子进行聚类。聚类后的簇称为组,即组中粒子并非全部都是有关

系的群组中的粒子。

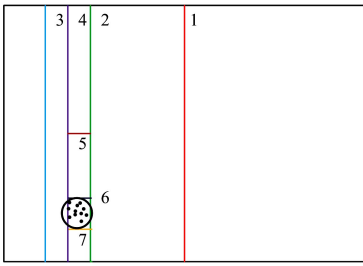


图1 二分划分过程

Fig. 1 Process of binary partition

如图1所示,假设图中圆圈圈出的为一个有关系的群组,如果在一个场景内对这个有关系的群组进行二分划分,首先需要对整个场景执行对半竖切(竖线1);该群组切完之后没有被切分到不同的网格,因此接着对群组所在的最小场景执行对半竖切(竖线2);该群组切完之后没有被切分到不同的网格,接着对群组所在的最小场景执行对半竖切(竖线3);该群组切完之后没有被切分到不同的网格,接着对群组所在的最小场景执行对半竖切(竖线4);该群组切完之后依然没有被切分到不同的网格,但是如果继续对群组所在的最小场景执行对半竖切,会将该群组划分到不同的网格,因此对群组所在的最小场景执行对半横切(横线5);该群组切完之后没有被切分到不同的网格,接着对群组所在的最小场景执行对半横切(横线6);该群组切完之后没有被切分到不同的网格,接着对群组所在的最小场景执行对半横切(横线7);该群组切完之后没有被切分到不同的网格,但是对该群组所在的最小场景无论横切还是竖切都会将群组划分到不同的网格,此时该群组所在的最小场景即为最终场景。

3.2 算法伪代码

折半 DBSCAN 聚类算法的伪代码如算法1所示。

算法1 折半 DBSCAN 聚类算法

输入:数据集 D , 参数 $\text{MinPts}=3$

输出:簇集合、噪声

1. 遍历数据 D , 找到所有有关系的群组 group ;
2. for 数据集 D 中每个有关系的群组 group do
3. while $i=1$
4. if 竖切之后有关系的群组被分到同一个网格
5. 执行竖切 $\text{verticalcut}()$;
6. else
7. if 横切之后有关系的群组被分到同一个网格
8. 执行横切 $\text{crosscut}()$;
9. else
10. 得到最终场景;
11. break;
12. end if
13. end while
14. end while
15. 针对该群组的最终场景,计算其 $k\text{-dist}$ 函数,根据 $k\text{-dist}$ 函数确定其 ϵ , 其中 $k=3$;
16. 将该群组的最终场景中的数据集 $d(d \subset D)$ 中的所有对象标记为未处理状态;
17. for 数据集 d 中的每个对象 p do

18. if p 已被标记 then
19. continue;
20. else
21. 将与对象 p 同一区域的 ϵ 邻域内的对象存入 $p.\text{tmplist}$;
22. if $p.\text{tmplist.size}()$ 大于或等于 MinPts then
23. $p.\text{tmplist}$ 的点标记为已分组,并存入 resultList 中;
24. end if
25. end if
26. end for
27. 合并
28. end for
29. 输出聚类结果.

3.3 聚类效率比较与分析

$k\text{-medioids}$ 算法是比较经典的基于划分的聚类算法,研究人员多用该算法对人群进行聚类分组。最大密度峰值聚类算法是2014年发表在 *Science* 上的聚类算法,其思想新颖、精度较高,因此受到了广泛的关注。为了验证本文提出的折半 DBSCAN 聚类算法的有效性,对比其与 $k\text{-medoids}$ 算法、最大密度峰值聚类算法对多种规模下的人群进行聚类的效果。因为人群中的社会关系是客观存在的,所以仿真程序会事先给定人群的关系矩阵。为了更接近现实,所有个体间并非人人都有关系,而是分成人数不等的多个朋友圈和一部分孤立的个人,并且这些朋友圈间两两没有交集,每个朋友圈就是一个事先给定的群组,单个孤立的个人不属于任何朋友圈,但在聚类的过程中可能由于距离某个朋友圈较近而与他们归为一组。由于在常态下人群中有关系的人通常距离较近,但也存在部分有关系的人群暂时分开而后期汇合的情况,因此在初始化关系时为多数距离较近的人群赋予非0值,为少数距离较远的人群赋予非0值。本文中定义粒子总个数为 n , 有关系的簇的个数为 \sqrt{n} , 由于人群中存在孤立粒子,因此每个有关系的簇中的人数少于 \sqrt{n} 。

3.3.1 聚类精度的比较

轮廓系数是一种评价聚类效果的方法,最早由 Peter 于1986年提出。其基本流程如下(本节中的距离均指式(2)中定义的相异度 dif):

Step1 计算样本 i 到同簇其他样本的平均距离 a_i , 并将 a_i 称为样本 i 的簇内不相似度。

Step2 计算样本 i 到其他簇 c_j 中所有样本的平均距离 b_{ij} , 并将 b_{ij} 定义为样本 i 的簇间不相似度 $b_i = \min\{b_{i1}, b_{i2}, \dots, b_{ik}\}$ 。

Step3 根据样本 i 的簇内不相似度 a_i 和簇间不相似度 b_i 计算样本 i 的轮廓系数 $s(i)$, 公式如下:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3)$$

Step4 根据 $s(i)$ 的值,判断样本聚类是否合理:

- 1) 若 $s(i)$ 接近1, 则说明样本 i 聚类合理;
- 2) 若 $s(i)$ 接近-1, 则说明样本 i 更应该被分类到另外的簇;
- 3) 若 $s(i)$ 近似为0, 则说明样本 i 在两个簇的边界上。

Step5 所有样本的 $s(i)$ 的均值称为聚类结果的轮廓系数,用于度量聚类分组是否合理且有效。

本文实验采用的机器配置为奔腾双核处理器,主频为

3.40 GHz,内存为 16 GB。根据轮廓系数函数对各人群规模做 50 组实验,得到折半 DBSCAN 聚类算法、最大密度峰值聚类算法、k-medoids 算法的聚类精度,对其取均值,得到如图 2 所示的折线图。

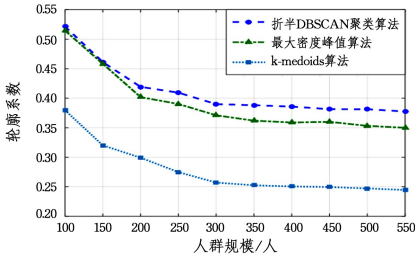


图 2 不同人群规模下 3 种算法的聚类效果对比

Fig. 2 Comparison of clustering effect of three algorithms in different population sizes

从图 2 中可以看出,k-medoids 轮廓系数最低。这是因为该算法的聚类结果对初始聚类中心的选择和噪声较为敏感,容易陷入局部最优,且该算法只能发现球状簇;但它能识别不同密度的簇。也可以看出,折半 DBSCAN 聚类算法的轮廓系数最高。这是因为该算法能识别任意形状的簇,不需要指定初始聚类中心的个数和位置,对噪声不敏感,划分网格后能识别不同密度的簇。最大密度峰值聚类算法的轮廓系数处于前两者之间。虽然最大密度峰值聚类算法能识别任意形状的簇和不同密度的簇,但在数据集中的噪声数据太多或者簇间存在明显噪声干扰的情况下,本文所提算法的聚类效果明显更优。粒子个数较少的情况下,轮廓系数相对较大,原因是粒子个数较少,簇与簇之间的距离较远, $b(i)$ 与 $a(i)$ 相差较大,轮廓系数中分子 $b(i)-a(i)$ 较大,因此 $s(i)$ 较大。随着粒子个数的增加,簇与簇之间的距离逐渐缩小, $b(i)$ 与 $a(i)$ 的差异逐渐缩小。当粒子个数特别多时, $b(i)$ 与 $a(i)$ 的距离差异越来越小,它们之间的差异几乎只体现在关系上。轮廓系数中分子 $b(i)-a(i)$ 较小,因此 $s(i)$ 较小。

总体而言,3 种聚类算法的轮廓系数值都不是很高,改进的 DBSCAN 聚类算法大概在 0.38 左右。这表示这个数据集中,簇内的数据不是很密集(即簇内数据比较疏松),或簇与簇之间的差异较小。

3.3.2 聚类时间对比

对不同人群规模进行聚类分组,比较折半 DBSCAN 聚类算法、最大密度峰值聚类算法、k-medoids 算法的聚类时间,并绘制折线图,如图 3 所示。

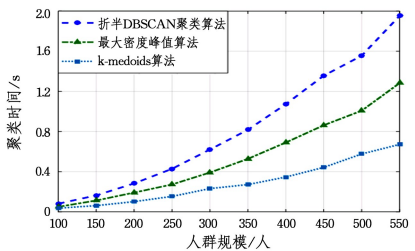


图 3 不同人群规模下 3 种算法的聚类时间对比

Fig. 3 Comparison of clustering time of three algorithms in different population sizes

从图 3 中可以看出,3 种算法的运行时间都随着人群规模的增大而增加。k-medoids 算法的运行时间最短。首先随机选择 k 个粒子作为中心点,本文中 k 取 \sqrt{n} ,将 n 个对象赋给这 k 个点所代表的簇,时间复杂度为 $O(nk)$ 。计算目标函数及新的中心点的时间复杂度为 $O(n)$,迭代上述过程。因此,整个算法的时间复杂度为 $O(tkn)$,其中 t 为迭代次数。最大密度峰值聚类算法的运行时间居中,它需要计算每个粒子的局部密度,时间复杂度为 $O(n^2)$;计算某粒子最相近的、局部密度比该粒子大的粒子之间的距离的时间复杂度为 $O(n^2)$;选择 \sqrt{n} 个粒子作为聚类中心并将其余粒子归并到最近的聚类中心所在的簇所需的时间复杂度为 $O(n \times \sqrt{n})$,即 $O(\sqrt{n^3})$ 。因此,整个算法的时间复杂度为 $O(n^2)$ 。折半 DBSCAN 聚类算法的运行时间最长,首先遍历所有粒子以找到所有有关系的群组,时间复杂度为 $O(n^2)$;对每个有关系的组进行多次二划分,直到再继续划分会把有关系人群划分到不同的网格为止,最坏情况下的时间复杂度为 $O(\sqrt{n} \log n)$;接着依据人群密度求各个有关系的小区域的邻域半径 ϵ ,这个过程需要计算各个小区域上任意两个粒子之间的相异度,时间复杂度为 $O(\sqrt{n} \times (\sqrt{n})^2)$,即 $O(\sqrt{n^3})$;最后以所求的 ϵ 值对小区域中的所有粒子执行 DBSCAN 聚类算法,所需的时间复杂度为 $O(\sqrt{n} \times (\sqrt{n})^2)$,即 $O(\sqrt{n^3})$ 。因此,折半 DBSCAN 聚类算法的时间复杂度为 $O(n^2)$ 。

4 仿真实验

4.1 聚类后的人群疏散

在原始社会力模型中,个体的运动受到 3 个力的共同作用:个体自身产生的向目标运动的驱动力,与其他个体的排斥力,以及障碍物对个体产生的排斥力。其中,驱动力用来引导个体向目标运动,排斥力用来实现个体与个体之间或个体与障碍物之间的碰撞避免。

由牛顿第二定律可知,行人所受的合力用式(4)表示:

$$m_i \frac{d\vec{v}_i(t)}{dt} = \vec{f}_i^0 + \sum_{j(\neq i)} \vec{f}_{ij}^r + \sum_w \vec{f}_{iw}^r \quad (4)$$

本文在原始社会力模型的基础上加入了同组内个体吸引力^[35],同时同组中考虑引领的作用,对个体自身产生的向目标运动的驱动力 \vec{f}_i^0 做出调整。每个组中距离出口最近的粒子被定义为引领,该出口设为这个引领的目标,而这个引领设为组内其他粒子的目标,以使得每个组在引领的带领下移动。孤立粒子将就近出口作为其目标。改进之后的社会力模型如式(5)所示:

$$m_i \frac{d\vec{v}_i}{dt} = \vec{f}_i^0 + \sum_{j(\neq i)} \vec{f}_{ij}^r + \sum_w \vec{f}_{iw}^r + \sum_{c \in C} \vec{f}_{ic} \quad (5)$$

其中:

$$\vec{f}_{ic} = R_{ij} G_i \exp[(r_{ic} - d_{ij})/H_i] \vec{n}_{ic} \quad (6)$$

其中,如果行人 i 和行人 j 都不是引领,那么 R_{ij} 表示 i 与 j 之间的社会关系值;如果 i 和 j 中有一个是引领,那么 R_{ij} 被定义为 1,如表 1 所列。通过这种方式定义 R_{ij} ,实现了同组内非引领粒子根据社会关系值相互吸引以及引领粒子对非引领粒子的吸引力最大的目标。 G_i 与 H_i 为吸引力参数, G_i 表示

吸引力强度, H_i 表示行人之间的安全距离。通过关系值对吸引力加权来划分不同关系值的聚合速度和紧密程度。个体间的吸引力越大, 聚合速度越快、越紧密。

表 1 同组个体吸引力参数 R_{ij} 的取值

Table 1 R_{ij} of individual attraction parameters in the same group

关系类型	关系值	
引领关系	1	
社会关系	亲人关系	(0.7, 1)
	朋友关系	(0.4, 0.7]
	同事关系	(0, 0.4]
	陌生人关系	0

本文在区域大小为 $300 \times 200 \text{ pix}$ 的房间进行疏散仿真, 粒子个数为 300, 用折半 DBSCAN 聚类算法对加入关系的人群进行聚类。聚类后, 在加入同组内个体吸引力的社会力平台上模拟人群疏散的过程。具体框架如图 4 所示。

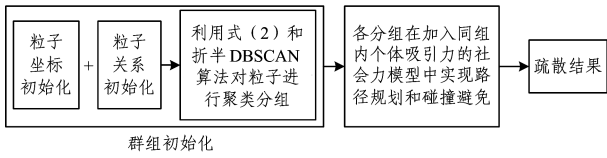


图 4 人群疏散框架图

Fig. 4 Framework map of crowd evacuation

社会力参数设置如下: 个体半径 $r = 0.25 \text{ m}$, 个体质量 $m = 80 \text{ kg}$, 期望速度 $v_i^0 = 0.8 \text{ m} \cdot \text{s}^{-1}$, 社会力模型常数 $k = 1.2 \times 10^5 \text{ kg} \cdot \text{s}^{-2}$, $\kappa = 2.4 \times 10^5 \text{ kg} \cdot \text{m}^{-1} \cdot \text{s}^{-1}$, $A = 2000 \text{ N}$, $B = 0.08 \text{ m}$, $G = 2000 \text{ N}$, $H = 0.05 \text{ m}$ 。用 MATLAB2016a 平台展示聚类及疏散过程, 如图 5 所示。

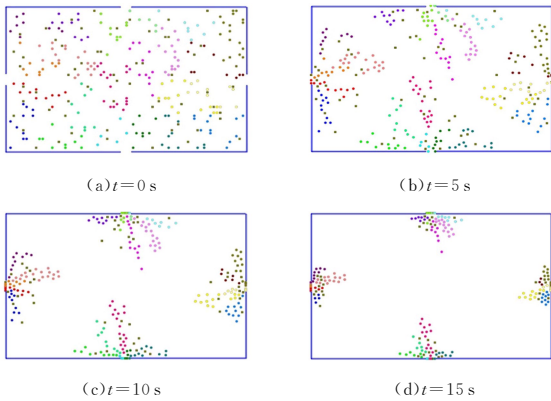


图 5 聚类分组后不同时刻的人群疏散情况

Fig. 5 Crowd evacuation at different times after clustering

图 5(a) 是聚类分组后的结果, 聚类后相同颜色的粒子为一组。小方块为孤立粒子, 它不属于任何组。图 5(b) — 图 5(d) 分别表示在社会力模型的基础上粒子在不同时刻的疏散状态。可以看出: 在疏散过程中, 相同颜色的人群在疏散过程中逐渐聚集在一起, 即同组内的个体在不断向出口运动的同时也渐渐聚集, 并从同一个门口逃生; 孤立粒子选择就近的门口逃生; 同时在出口处出现了“出口成拱”“快即是慢”的效应。改进后的疏散模型能更真实地模拟人群疏散的过程。

4.2 人群疏散仿真对比

为了验证折半 DBSCAN 聚类算法在人群疏散过程中的有效性, 本文以 Visual Studio 2012+OSG 为开发平台, 在自

行研发的加入同组内个体吸引力的社会力人群疏散仿真系统中进行实验。区域大小为 $300 \times 200 \text{ pix}$ 。实验以 300 个粒子为例, 仿真结果如图 6—图 8 所示。

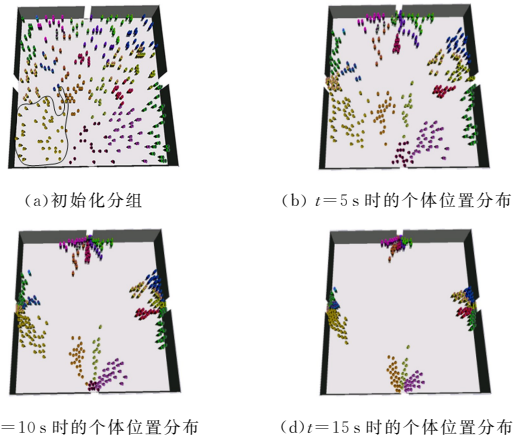


图 6 人群分组后的疏散模拟(k-medoids 算法)

Fig. 6 Evacuation simulation after crowd grouping (k-medoids algorithm)

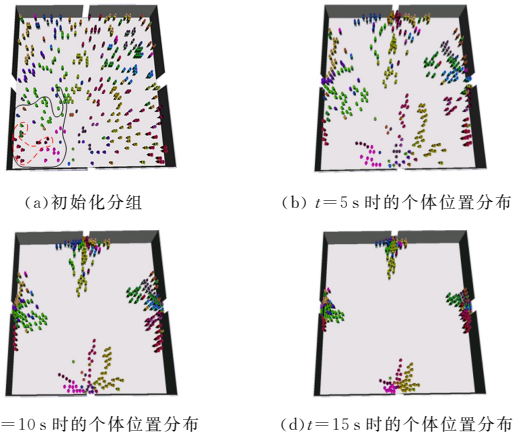


图 7 人群分组后的疏散模拟(最大密度峰值聚类算法)

Fig. 7 Evacuation simulation after crowd grouping(maximum density peak clustering algorithm)

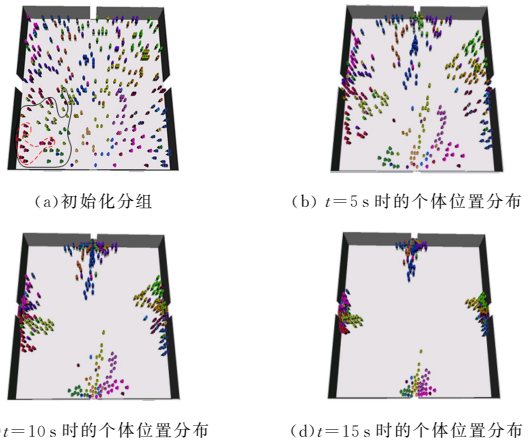


图 8 人群分组后的疏散模拟(折半 DBSCAN 聚类算法)

Fig. 8 Evacuation simulation after crowd grouping(binary DBSCAN clustering algorithm)

图 8 为折半 DBSCAN 聚类算法在加入同组内个体吸引

力的社会力模型中的应用。初始化结果为距离较近、关系较大的个体被归为一组,同组粒子颜色一致,组内不同颜色的粒子为孤立粒子,如图 8(a)所示。通过该仿真结果可以看出,在加入同组内个体吸引力的社会力模型上的疏散过程中,同组的个体逐渐靠拢,选择从同一门口逃生,如图 8(b)所示。在图 8(c)中,组内个体的聚集程度更明显,门口出现拥堵现象。在图 8(d)中,部分分组已从同一门成功逃离,部分分组已逃离了部分个体,此时组内未成功逃离的个体聚集程度更高,拥堵现象更明显,出现了“出口成拱”“快即是慢”的现象。

图 6 和图 7 分别为 k-medoids 算法、最大密度峰值聚类算法在加入同组内个体吸引力的社会力模型中的应用。可以看出,初始时期分组的不同会导致后期疏散过程的不同。

图 6(a)、图 7(a)、图 8(a)分别为 3 种聚类算法的聚类结果。在图 6(a)中几乎所有的簇都是凸状的。例如,用实线圈出的粒子是用 k-medoids 算法聚类形成的一个簇,可以看出其外观呈凸状。在最大密度峰值聚类及折半 DBSCAN 聚类结果中圈出同样的人数,发现这两个算法将圈出的人分成了多个簇和孤立的粒子,而且在外观上其形状不都是凸状。对实线圈出的所有粒子分别计算 3 种聚类算法的轮廓系数,得出折半 DBSCAN 聚类算法最大,最大密度峰值聚类算法次之,k-medoids 算法最小。这说明折半 DBSCAN 聚类算法的聚类效果比 k-medoids 算法的好,其能识别凹状簇。图 8(a)中用虚线圈出的粒子是用折半 DBSCAN 聚类算法得出的簇,在图 7(a)中却将其分成了多个簇。对这个虚线圈出的所有粒子计算轮廓系数,得出折半 DBSCAN 聚类算法的轮廓系数比最大密度峰值聚类算法的大,这说明折半 DBSCAN 聚类算法识别任意密度簇的能力比最大密度峰值聚类算法强。因此,仿真实验结果证明,本文提出的折半 DBSCAN 聚类算法能够精确地聚类人群中关系值比较大、距离较近的粒子,也能精确地显示孤立粒子,提高了人群疏散仿真的准确度。加入同组吸引力的社会力模型能逼真地模拟同组人群依据关系亲密度产生的聚集、引领及跟随、向出口逃生、“出口成拱”、“快即是慢”等现象,也能逼真地展现孤立粒子的运动过程,为人群疏散演变规律提供了新思路。

结束语 k-medoids 算法中初始聚类中心的个数和位置都是人为设定的,且该算法对噪音敏感,只能识别球状簇,容易陷入局部最优。传统的 DBSCAN 聚类算法虽无须指定初始聚类中心的个数和位置,能识别任意形状的簇,对噪音不敏感,但它只能识别密度相近的簇,且需要计算任意两个粒子之间的距离。本文提出的折半 DBSCAN 聚类算法不需要人为确定聚类个数和初始聚类中心点的位置,能识别任意形状和不同密度的簇,且能找到全局最优解,因此它能准确地将关系和距离值在一定范围内的人聚集在一起,并且能将孤立粒子分离出来。与 k-medoids 算法相比,所提算法具有更高的准确性,聚类后在加入同组吸引力的社会力模型上实现疏散运动,并引入同组个体吸引力参数 R_{ij} 模拟关系亲密程度对疏散过程中聚集程度的影响,更为逼真地模拟了人群疏散时聚集成团的现象,为进一步研究人群疏散行为奠定了基础。但是折半 DBSCAN 聚类算法引入二分划分将粒子划分到同一网格中,并用式(2)聚类,缺少了偶然性。例如,某个人和与其有

关系的人群相距较远,则按照式(2)聚类划分不到一个分组中,但某人可能在突发事件下心理归属感比较强,一定要和与其有关系的群组汇合到一起逃生;或某个人和与其有关系的群组关系较为密切且距离较近,则按照式(2)聚类会划分到一个分组中,但是该人在突发事件下可能未跟随本群组而是独自选择他所熟悉的路径逃生了;或者某个陌生人与一个群组较近,突发事件下可能跟随这个群组逃生,但是由于二分划分时并未将此人与这个群组划分到同一网格,因此其并未跟随该群组一起逃生。但是,本文在考虑不同社会关系间行人的分布状况的前提下,使用折半 DBSCAN 聚类算法对行人聚类分组,分组后在引入同组内吸引力的社会力模型中实现人群疏散。这对一般情况下的人群疏散预测具有一定的指导意义和参考价值,可以作为预测疏散时间和评价场景建筑分布合理状况的重要工具,能为利用人群分组现象引导人群疏散以提高疏散效率的研究提供更多的新思路。

本文考虑了距离和关系在突发事件下对人群聚类分组现象的影响,并考虑了组内作用力及引领对人群疏散的影响。但在实际情况下人群疏散还存在诸多因素的影响,比如视野范围、行人体型、障碍物的位置及形状、个体自身状况、情绪感染状况、场景出口个数、大小及位置。因此在未来的工作中将进一步研究以上因素对疏散效率的影响;同时也将研究如何对人群聚集成簇现象加以引导和利用,制定相应的人群疏散机制来提高人群的疏散效率。

参 考 文 献

- [1] LIU G P, LIU H, LV L, et al. Relationship-integrated crowds simulation[J]. Journal of Chinese Mini-Micro Computer Systems, 2016, 37(8): 1735-1740. (in Chinese)
柳广鹏, 刘弘, 吕蕾, 等. 融入关系分组的人群运动仿真[J]. 小型微型计算机系统, 2016, 37(8): 1735-1740.
- [2] FESTINGER L. A theory of social comparison processes[J]. Human Relations, 1954, 7(7): 117-140.
- [3] TSAI J, FRIDMAN N, BOWRING E, et al. ESCAPES- Evacuation Simulation with Children, Authorities, Parents, Emotions, and Social comparison[J]. Copyright, 2013, 50(6): 563-564.
- [4] LI Y, LIU H, LIU G P, ET A L. A grouping method based on grid density and relationship for crowd evacuation simulation [J]. Physica A Statistical Mechanics & Its Applications, 2017, 473(1): 319-336.
- [5] TAN Y, HU R F, YIN G F. Adapted DBSCAN with multi-threshold[J]. Journal of Computer Applications, 2008, 28(3): 745-748. (in Chinese)
谭颖, 胡瑞飞, 殷国富. 多密度阈值的 DBSCAN 改进算法[J]. 计算机应用, 2008, 28(3): 745-748.
- [6] XIA L N. SA-DBSCAN: A self-adaptive density-based clustering algorithm[J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2009, 26(4): 530-538.
- [7] KUMAR K M, REDDY A R M. A fast DBSCAN clustering algorithm by accelerating neighbor searching using Groups method [J]. Pattern Recognition, 2016, 58: 39-48.
- [8] HE X X, GUAN J Y, YE X Z, et al. A density-based and grid-based cluster centers determination clustering algorithm [J].

- Control and Decision, 2017, 32(5):913-919. (in Chinese)
何熊熊, 管俊轶, 叶宣佐, 等. 一种基于密度和网格的簇心可确定聚类算法[J]. 控制与决策, 2017, 32(5):913-919.
- [9] PATWARY M M A, PALSETIA D, AGRAWAL A, et al. A new scalable parallel DBSCAN algorithm using the disjoint-set data structure[C]// High Performance Computing, Networking, Storage and Analysis. IEEE, 2013:1-11.
- [10] HE Y, TAN H, LUO W, et al. MR-DBSCAN: a scalable MapReduce-based DBSCAN algorithm for heavily skewed data[J]. Frontiers of Computer Science, 2014, 8(1):83-99.
- [11] LAKOBA T I, KAUP D J, FINKELSTEIN N M. Modifications of the Helbing-Molnar-Farkas-Vicsek social force model for pedestrian evolution[J]. Simulation, 2005, 81(5):339-352.
- [12] OSARAGI T. Modeling of pedestrian behavior and its applications to spatial evaluation[C]// Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2. IEEE Computer Society, 2004: 836-843.
- [13] MOUSSAÏD M, HELBING D, THERAULAZ G. How simple rules determine pedestrian behavior and crowd disasters[J]. Proceedings of the National Academy of Sciences, 2011, 108(17):6884-6888.
- [14] HUGHES R L. A continuum theory for the flow of pedestrians [J]. Transportation Research Part B: Methodological, 2002, 36(6):507-535.
- [15] HOOGENDOORN S, BOVY P. Gas-kinetic modeling and simulation of pedestrian flows [J]. Transportation Research Record: Journal of the Transportation Research Board, 2000, 1710(1): 28-36.
- [16] HELBING D, JOHANSSON A, AL-ABIDEEN H Z. Dynamics of crowd disasters: An empirical study[J]. Physical Review E, 2007, 75(4):046109.
- [17] STEFFEN B. A modification of the social force model by foresight[M]// Pedestrian and Evacuation Dynamics 2008. Berlin: Springer, 2010:677-682.
- [18] CAO M X, ZHANG G J, HUANG L J, et al. Crowd Animation Generation Method Based on Personalized Emotional Contagion [J]. Computer Science, 2017, 44(6):306-311. (in Chinese)
曹梦晓, 张桂娟, 黄丽君, 等. 基于个性化情绪感染的人群动画生成方法[J]. 计算机科学, 2017, 44(6):306-311.
- [19] PARISI D R, GILMAN M, MOLDOVAN H. A modification of the Social Force Model can reproduce experimental data of pedestrian flows in normal conditions [J]. Physica A Statistical Mechanics and Its Applications, 2012, 388(17):3600-3608.
- [20] JI Q G, HE H, WANG F C. Social Force Model for Crowd Simulation Using Density Field[J]. Computer Science, 2015, 42(6): 12-17. (in Chinese)
纪庆革, 何浩, 王福川. 密度场下的短程社会力模型[J]. 计算机科学, 2015, 42(6):12-17.
- [21] MUSSE S R, THALMANN D. A model of human crowd behavior: Group inter-relationship and collision detection analysis [J]. Computer Animation and Simulation, 1997, 97(9):39-51.
- [22] VIZZARI G, MANENTI L, CROCIANI L. Adaptive pedestrian behaviour for the preservation of group cohesion[J]. Complex Adaptive Systems Modeling, 2013, 1(1):1-29.
- [23] QIU F, HU X. Modeling group structures in pedestrian crowd simulation[J]. Simulation Modelling Practice and Theory, 2010, 18(2):190-205.
- [24] FU Y W, LIANG J H, LIU Q P, et al. Crowd Simulation for Evacuation Behaviors Based on Multi-agent System and Cellular Automaton[C]// International Conference on Virtual Reality and Visualization. IEEE, 2015:103-109.
- [25] QIU F, HU X. A Framework for Modeling Social Groups in Agent-Based Pedestrian Crowd Simulations [J]. International Journal of Agent Technologies and Systems, 2012, 4(1):39-58.
- [26] QIN X, LIU H, ZHANG H, et al. A collective motion model based on two-layer relationship mechanism for bi-direction pedestrian flow simulation[J]. Simulation Modelling Practice and Theory, 2018, 84:268-285.
- [27] ZHANG H, LIU H, QIN X, et al. Modified two-layer social force model for emergency earthquake evacuation[J]. Physica A Statistical Mechanics and Its Applications, 2018, 492: 1107-1119.
- [28] LIU B, LIU H, ZHANG H, et al. A social force evacuation model driven by video data[J]. Simulation Modelling Practice and Theory, 2018, 84:190-203.
- [29] XU X, ZHANG L, SOTIRIADIS S, et al. CLOTHO: A Large-Scale Internet of Things based Crowd Evacuation Planning System for Disaster Management [J]. IEEE Internet of Things Journal, 2018, 99:1-1.
- [30] LIU H, XU B, LU D, et al. A Path Planning Approach for Crowd Evacuation in Buildings Based on Improved Artificial Bee Colony Algorithm[J]. Applied Soft Computing, 2018, 68: 360-376.
- [31] MOUSSAÏD M, PEROZO N, GARNIER S, et al. The walking behaviour of pedestrian social groups and its impact on crowd dynamics[J]. Plos One, 2010, 5(4):e10047.
- [32] RODRIGUEZ A, LAIO A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191):1492-1496.
- [33] FRIDMAN N, KAMINKA G A. Comparing human and synthetic group behaviors: A model based on social psychology[C]// International Conference on Cognitive Modeling (ICCM-09). 2009.
- [34] FRIDMAN N, KAMINKA G A. Towards a cognitive model of crowd behavior based on social comparison theory[C]// National Conference on Artificial Intelligence. AAAI Press, 2007: 731-737.
- [35] WANG L, CAI Y, XU Q. Modifications to Social Force Model [J]. Journal of Nanjing University of Science and Technology (Natural Science), 2011, 35(1):144-149. (in Chinese)
汪蕾, 蔡云, 徐青. 社会力模型的改进研究[J]. 南京理工大学学报, 2011, 35(1):144-149.