

库, 真实数据集是从万维网上爬取了近十年应急救援新闻数据达数十万条, 包括相关案例的论文研究, 由于相关数据公开度不一致, 出现结构化数据中部分字段缺失或模糊, 但对本文研究未造成很大的影响, 本文中建立了相对完备的应急预案库, 对于数据稀疏的矩阵区域, 用人工案例库进行校正, 人工数据率为 $\mu=[0.3, 0.4]$, 消除了噪音等不合理数据, 得到的应急相关要素间的关系和发展趋势, 数据结论对应急预案、演练有极其重要的作用。

3 方法

3.1 应急执行实体和应急活动的优化

基于应急案例、应急预案数据集, 对应急执行实体和应急活动的最优化分配, 实现优化应急活动的序列整体框架如下(见图1):

首先, 对应急案例库预处理的数据清洗中, 将案例库中的非必要字段剔除, 如原采集网页中的标准等, 生成 XML 结构数据字典, 并按地域进行分类, 模糊行政区划边界(如图2), 即边界区域附近的案例互相包含。

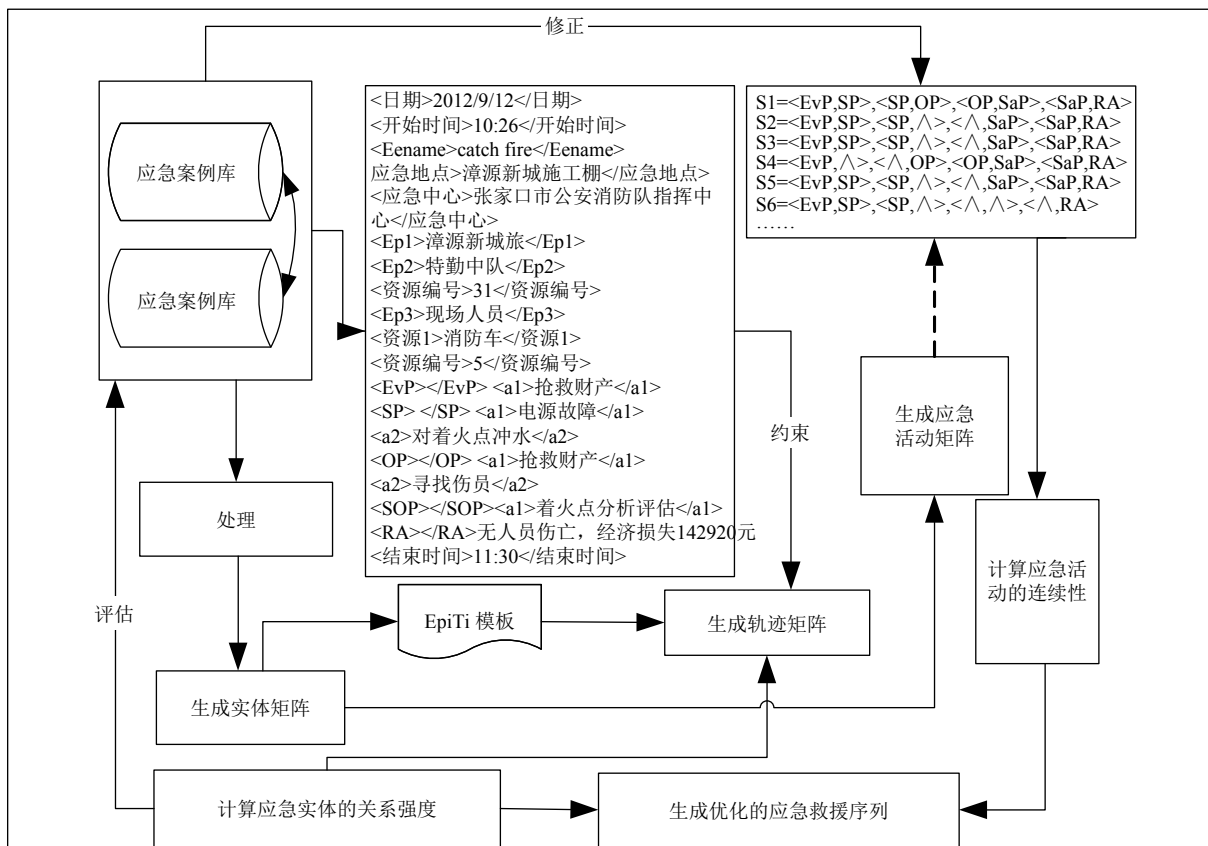


图1 基于案例数据集的实体、活动抽取框架

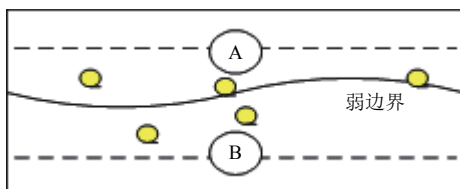


图2 边界示意图

对于地理位置分别属于 A、B 的两个位置区域中, A 和 B 为区域中心, 地理位置阈值最高, 越往边缘, 地理阈值越小, 称为弱边界, 分布在弱边界中的事件, 分

别属于两个行政区域, 这样的划分, 打破了以往的行政区划中心, 而以距离强度来分布。

将同地理区域的案例根据式(1)–式(5), 计算成应急实体轨迹矩阵和活动轨迹矩阵。稀疏矩阵中的数据元素为“^”, 则添加人工案例库数据进行修正, 修正系数 $\mu=0.3$ 。在对应急案例库和预案库过程中, 会存在出现噪音数据和冗余数据, 本文通过标签化处理, 对应急时间、属性等关键信息的标准化处理, 并剔除异常数据值, 采用经典的 Needleman-Wunsch 算法^[12]来实

现,例如,文中定义活动间的有序对元组 Set, $S = \{\langle EP, SP \rangle, \langle SP, OP \rangle, \langle OP, SaP \rangle, \langle SaP, RA \rangle\}$, 其中有 SP 缺失, 在活动字段对齐时, 用“ \wedge ”填充, 则 $S = \{\langle EP, \wedge \rangle, \langle \wedge, OP \rangle, \langle OP, SaP \rangle, \langle SaP, RA \rangle\}$, 这样可以满足概念模型的设计, 更有利于应急执行实体活动序列的优化, 有助于得到效率比较高的应急预案。

基于上文中数据, 计算应急活动的连续性 $contin_{a_i, a_j}$ 和应急实体的关系强度 RS 等, 进一步计算结合实体的技能贡献度得到实体协作能力, 进一步挖掘出优化的应急活动序列。

3.2 挖掘应急实体的高协作度算法

为了挖掘应急案例库中具有高协作度的实体, 基于应急实体轨迹、实体技能贡献度等实现算法 1, 首先计算所有的实体轨迹的矩阵, 再基于此矩阵, 通过最小地域置信度阈值和技能贡献度等过滤得到具有高协作度的实体即高协作模式。

算法 1. 高协作度关联实体挖掘算法

输入: 实体轨迹矩阵; 初始化 i, j, U_{Epi} /实体技能贡献度
 Min_Ept_conf /最小地域置信度
输出: 关联实体的协作度表

1. 依据式 (2) 构建实体关系强度矩阵 $RS[]$
2. 依据式 (5) 构建实体间协作能力矩阵 $Coop[]$
3. FOR $i=1$ to n DO
4. 活动数组 $Ea[i] \leftarrow$ 应急活动矩阵中第 i 行非空下标
5. END FOR
6. FOR $i=1$ to n DO //根据活动的非空下标得到活动序列
7. $ActivitySeqMap[i] \leftarrow$ getActivitySeq(活动数组 $Ea[i]$)
8. If $ActivitySeqMap[i] ==$ 稀疏 Then
9. $ActivitySeqMap[i] \leftarrow$ getSimActivitySeq()
10. END FOR
11. FOR $i=1$ to n DO
12. $entitySeqMap[i] \leftarrow RS[i] \times Coop[i] \times U_{Epi}$
13. END FOR
14. FOR each $entitySeqMap[i]$ of $ActivitySeqMap[i]$ DO
15. If $ActivitySeqMap[i] \in Min_Ept_conf \times ActivitySeqMap[i]$ Then
16. 高协作实体 $HCEntity[i] \leftarrow entitySeqMap[i]$
17. END FOR

基于算法得到实体间协作能力的一般矩阵和有高协作能力的实体矩阵。

3.3 挖掘高连续性活动

基于应急案例库, 挖掘应急实体活动的轨迹挖掘同区域活动的协作度。首先初始化应急活动轨迹矩阵, 计算一般活动的连续度 $contin$, 根据整个案例库中所有活动为基数, 计算某一个活动出现的频度 F , 产生 θ 调优值, 可以根据不同的值得到几组结果, 进行迭代。

算法 2. 基于活动轨迹和活动连续度挖掘高连续性活动

输入: 活动轨迹矩阵; 初始化 i, j ; 活动连续度 $contin$;
 F 频度; θ 调优参数
输出: 强相关连续性高的活动表

1. 初始化 θ
2. 依据式 (1) 构计算活动连续度表 $contin[]$
3. FOR $i=1$ to n DO
4. $ActivityMax[] \leftarrow$ Max(a_i, a_j 连接次数)
5. $H\ Activity = contin[] \times ActivityMax[]$
6. END FOR

4 结果

图 3 中曲线 a 可以看出某城市中, 不同应急消防团队实体在协作次数和协作能力上的表现, 随着合作次数的增加稳步提升, 可以超过指数 1, 在尾部也表现为增幅放缓, 跟合作多次之后, 采用经验和旧案例判断有关, 需要适应性更新案例库和预案库。

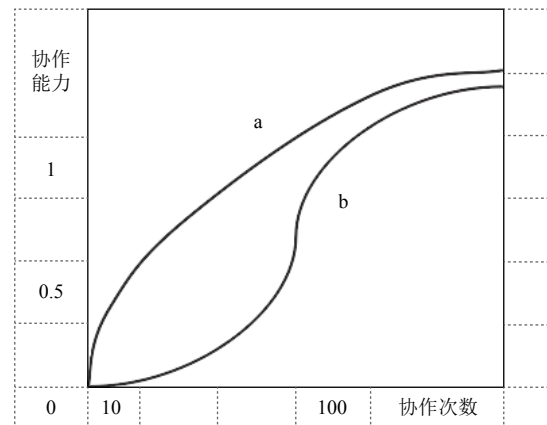


图 3 应急实体协作次数与协作能力关系

曲线 b 可以看出, 在同属同一位置区域中, 位置强度关系显著的应急实体间的协作能力, 医护实体和消防实体之间随着协作次数的增加, 协作能力稳步增强, 增长期间的会有波动, 跟参与的活动有关, 在协作一定次数之后, 反而协作能力会出现下降, 因为在多次协同之后, 经验占据整个决策的主导, 导致经验主义的误判, 在应急预案的生成中, 要重点规划合作次数较少的和合作有一定次数的。

图 4 中曲线 a 可以看出边界消防实体和中心消防实体之间的关系, 随着协作的次数, 它们的协作能力稳步提升, 跟消防团队的调配和管理方式相关, 消防实体在应急中任务明确, 平时应急演练标准化程度高的作用体现。

曲线 b 可以看出在同区域的医护实体实体和中心

位置的医护实体之间的协调度随着协作次数的增长而增长,呈现泛指数分布,中间反而会出现协作能力的上升,协调能力可以达到0.5,跟实体的技能度相关,也反应了在边界区域的实体之间的协作能力也有表现强的.在应急预案中不能仅仅关注行政区域中心的应急实体,这种现象在大规模应急事件中会有所体现,往往调度一个区域的医护实体不够的时候,需要调度多个医护实体才能应对的突发事件,需要考虑此种情形.

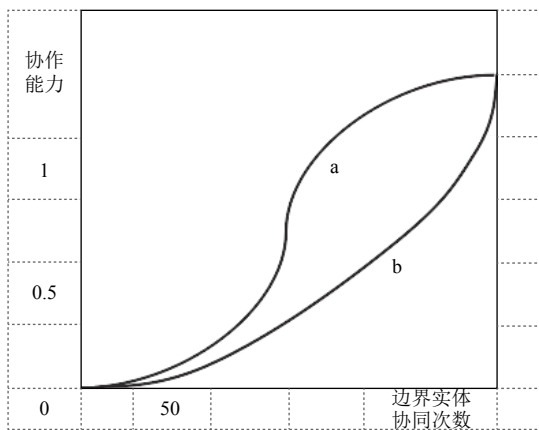


图4 边界实体协同次数与协作能力的关系

从图5曲线上可以看出,应急活动的连续性与应急案例库中的应急频率直接成正比.曲线不平滑,分段明显.应急预案库包括应急预案和旧预案.基于某一特定领域的案例库的完整性,针对某一特定应急活动的演练越多,应急活动的连续性就越强.很强,反映了应急预案的重要性.事实上,应急计划通常非常相似,但事件的情况是独一无二的.同类突发事件发生频率较高,具有预警作用.同样,数据库中相应应急活动的频率也下降了.如图所示,当活动连续性为零时,数据中的活动频率不是零.数据库中存在单一的活动关系,不一定没有活动连续性,这与实际的应急指挥一致.有时根据现场情况,临时产生应急动员方案.

在图6中,横轴表示活动的连续性,纵轴表示实体协作的程度.曲线表现出相似的二项分布,反映了应急活动的连续性与应急实体协作程度的关系.在紧急事件中,应急活动的连续性和参与应急活动的实体的协作程度相互影响.应急活动的连续性越高,应急实体的协作程度越高.同样,应急协作程度越高,应急活动的连续性越高,应急活动越活跃.当动态连续性超过指数1时,由于应急实体协调次数增加,活动与活动的优先

关系设计,曲线中后部的增长速度加快.当活动连续性值接近1.8时,不再上升,与实际应急处置环节一致.每个突发事件的处置过程都有其自身的特点.在实际的应急指挥中,需要通过应急实体的协调和应急活动的连续性,直接确定应急处置的时间.时间是紧急处置中最昂贵的成本之一,首先应该考虑到时间,提高应急响应的满意度.

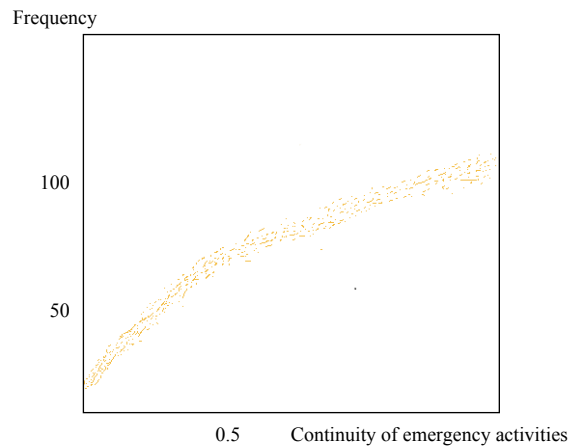


图5 频率和应急活动的关系

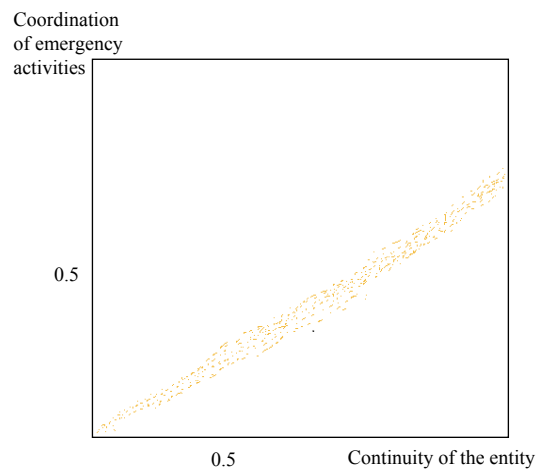


图6 活动连续性与实体协作度的关系

5 基于应急案例库的应急预案产生的框架

如图7,基于案例库和预案库进行数据处理,其中的数据有实时从外部感知的数据,这样可以动态的调整数据,在数据分析的基础上建立预测模型,为应急演练、仿真分析等作准备,应急预警、应急预案和应急决策是建立在对数据逻辑推理的基础上进行的,可以提供辅助决策的准确度,应急个体的属性包括应急执

行实体的各种态势, 应急活动的各项指标, 结合实际感知到突发事件的数据, 为应急决策提供理论基础。

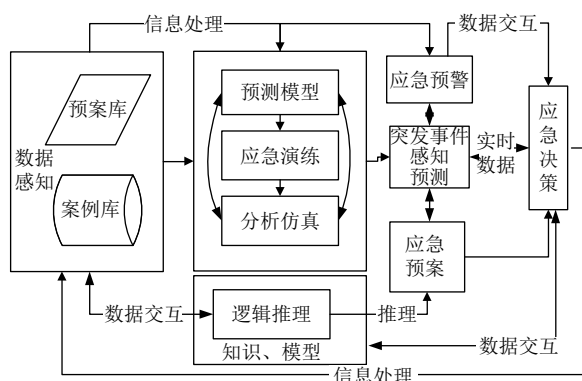


图7 基于应急案例库的应急预案产生的框架

6 结论

本文以应急案例库和预案库为基础, 对应急实体和应急活动的必要属性进行定义, 对应急执行实体的技能贡献度, 应急实体的关系强度、应急活动的连续度等做了细致分析, 用实验数据得出同一区域同一类型应急实体之间的协作能力和协作次数之间的关系, 不同实体在同一区域的协作能力的趋势, 位置关系强度较弱的实体之间的协作能力关系, 在此基础上产生的应急预案, 有效性和针对性大幅提高, 更合理的选择利用经验知识, 缩小误判率, 文中考虑的因素相对比较理想化, 影响应急事件演变因素很多, 需要根据确定的某一事件做细化, 今后工作中, 将从下面3点不断深入应急知识库中的知识表示和知识学习, 提高知识库和应急要素对接的精准度, 提高产生应急预案的准确度和可行度, 其一, 实验的数据量直接影响了结果, 由于案例库的不足, 预案库支持率上升, 使得有些实验数据跟现实情况有差值, 还需要不断的爬取新的数据量, 使案例库不断完备。其二, 影响应急救援结果的因素, 不止文中着重研究的应急实体和活动等要素, 还有诸多突发的要素, 文中就这两点从团队组合的角度去考虑, 有许多救援其实跟资源有关, 文中这点比较薄弱。其三, 在应急区域边界的划分上, 各个区域有所不同, 文中采用了模糊化的手段, 需要分区域进行考虑, 严格划分位置区域, 本文是理想位置状态。

参考文献

1 Jiao PF, Cai F, Feng YD, *et al.* Link predication based on

matrix factorization by fusion of multi class organizations of the network. *Scientific Reports*, 2017, 7: 8937. [doi: 10.1038/s41598-017-09081-9]

2 Liu X, Wang WJ, He DX, *et al.* Semi-supervised community detection based on non-negative matrix factorization with node popularity. *Information Sciences*, 2017, 381: 304–321. [doi: 10.1016/j.ins.2016.11.028]

3 Liu D, Liu X, Wang WJ, *et al.* Semi-supervised community detection based on discrete potential theory. *Physica A: Statistical Mechanics and Its Applications*, 2014, 416: 173–182. [doi: 10.1016/j.physa.2014.08.051]

4 Ly LT, Rinderle S, Dadam P, *et al.* Mining staff assignment rules from event-based data. In: Bussler CJ, Haller A, eds. *Business Process Management Workshops*. Berlin: Springer, 2006. 177–190.

5 Liu YB, Wang JM, Yang Y, *et al.* A semi-automatic approach for workflow staff assignment. *Computers in Industry*, 2008, 59(5): 463–476. [doi: 10.1016/j.compind.2007.12.002]

6 Xu RB, Liu X, Xie Y, *et al.* A Gaussian fields based mining method for semi-automating staff assignment in workflow application. *Proceedings of 2014 International Conference on Software and System Process*. Nanjing, China. 2014. 178–182.

7 许荣斌, 鲍广华, 杨培全, 等. 基于最大依赖度及最小冗余度的员工协作优化策略. *计算机集成制造系统*, 2017, 23(5): 1014–1019.

8 Meddah I, Khaled B. Discovering patterns using process mining. *International Journal of Rough Sets and Data Analysis*, 2016, 3(4): 21–31. [doi: 10.4018/IJRSDA.2016100102]

9 Lappas T, Liu K, Terzi E. A survey of algorithms and systems for expert location in social networks. In: Aggarwal CC, ed. *Social Network Data Analytics*. Boston: Springer, 2011. 215–241.

10 Kargar M, An AJ. Discovering top-k teams of experts with/without a leader in social networks. *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. Glasgow, UK. 2011. 985–994.

11 俞东进, 王娇娇, 柳诚飞. 基于协作模式的工作流最优员工分配方法. *软件学报*, 2018, 29(11): 3340–3354. [doi: 10.13328/j.cnki.jos.005483]

12 Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 1970, 48(3): 443–453. [doi: 10.1016/0022-2836(70)90057-4]