

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2021.0469

面向航空网络的机场风险传播网络

管祥民^{1,2,*}, 赵帅喆³

(1. 中国民航管理干部学院 民航通用航空运行重点实验室, 北京 100102;

2. 浙江建德通用航空研究院 浙江省通用航空运行技术研究重点实验室, 杭州 311600;

3. 北京航空航天大学 电子信息工程学院, 北京 100191)

摘 要: 随着航空网络运输量与日俱增, 机场间运行呈现出较强耦合关联。机场风险传播特性严重制约航空网络安全高效运行。目前机场运行风险量化计算及在航空网络的传播机理尚缺乏深入研究。综合考虑安全、效率等风险要素基于聚类算法提出机场运行风险耦合量化方法; 构造机场运行风险时间序列, 应用因果检验方法并基于复杂网络构建机场运行风险传播网络; 通过对比不同类型网络, 分析机场运行风险传播网络特征, 挖掘机场运行风险传播规律。结果表明: 风险传播网络度分布满足双区对数分布特点, 呈现小世界特征, 不仅具有较短的网络直径和较高的社区性, 而且可被分为若干连接密集区域, 且风险传播网络效率较低, 表明在全局传播的难度较高。

关 键 词: 航空网络; 风险耦合; 格兰杰因果检验; 风险传播; 复杂网络

中图分类号: V221+.3; TB553

文献标志码: A

文章编号: 1001-5965(2023)06-1342-10

航空网络主要包括航路网和航线网络 2 类。航线网络也称机场网络, 是以机场为节点, 航线为连边的网络模型。在航空网络中, 机场运行的主要风险由多种干扰因素构成, 包括恶劣天气条件、航班延误、大规模航班负荷及空中交通事故等。而且航空网络节点风险可以转移并通过机场之间连接进行传播扩散。中国民航业的快速发展与有限的航空网络资源之间的矛盾不断加剧机场风险传播。因此, 如何合理量化航空网络运行风险与分析风险传播特征已经成为空中交通管理领域中的重要方向。

航空网络机场运行风险传播研究可以有效揭示风险变化规律及扩散趋势, 主要传播模型包括贝叶斯网络模型、流行病毒传播模型、因果检验模型及近似延误网络-延误传播算法 (approximate network delays-delay propagation algorithms, AND-DPA) 模型等。国外方面, 文献 [1-4] 分析了单机场

的延误对国家空域系统的影响, 针对少数重要机场研究了拥挤波及传播预测。文献 [5] 首次采用 AND-DPA 模型, 研究了美国 30 多个重要机场的延误传播过程, 并针对延误传播现象采取了类似削峰填谷的优化策略, 通过流量控制减少延误传播。文献 [6] 运用流行病毒传播模型, 构建以机场为中心的网络队列模型, 运用真实数据分析了欧洲空中交通的延误特性, 并对欧洲空中交通网络的延误传播进行了研究。文献 [7-8] 提出一个航班延误级联分析模型, 以分析由地面等待和其他原因引起的航班延误级联。

国内方面, 文献 [9-10] 通过贝叶斯网络模型, 将驾驶员、管制员等人为因素与环境因素结合在一起, 分析延误关联并建立了延误传播网络。文献 [11] 基于有色 Petri 网络建立了航班延误波及链模型, 分析了航班运行与机场资源的相关性及航班延误导致的链式级联现象。文献 [12] 以饱和度为核心指

收稿日期: 2021-08-18; 录用日期: 2021-11-22; 网络出版时间: 2022-03-31 11:11

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220330.1401.003.html

基金项目: 国家自然科学基金 (U1933130, 71731001); 浙江省重点研发计划 (2022C01055); 中国科学院重点部署项目子课题 (ZDRW-KT-2020-21-2)

* 通信作者. E-mail: guanxiangmin@camic.cn

引用格式: 管祥民, 赵帅喆. 面向航空网络的机场风险传播网络 [J]. 北京航空航天大学学报, 2023, 49 (6): 1342-1351.

GUAN X M, ZHAO S Z. Airport risk propagation network oriented to aviation network [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (6): 1342-1351 (in Chinese).

标建立了机场拥挤评价方法, 基于 Elman 神经网络提出机场拥挤预测方法, 并对美国机场网络拥挤的级联传播现象进行了预测。文献 [13-14] 通过传统格兰杰因果关系检验这一统计学方法, 建立延误传播因果关联矩阵, 构建延误传播网络, 并分析复杂网络特性, 得出中小型机场对延误传播起到主要作用的结论。

目前对航空网络运行风险传播的研究聚焦于单机场、扇区等区域范围, 缺少面向全局网络的运行风险传播研究。此外, 现有研究以交通动力学模型和其他理论场景为主, 缺乏基于真实数据的运行风险传播网络研究与实证分析。同时, 风险传播研究大多体现于航班延误、航空器冲突等局部风险诱因上, 缺乏整体宏观的运行风险量化建模研究。本文根据运行风险的时空特性, 从效率和安全的角度选取局部风险因素, 将局部风险因素耦合成整体的航空网络风险系数, 建立航空网络机场节点风险量化模型, 运用真实数据结合上述风险量化方法得出机场的具体风险时间序列, 根据因果关系检验方法建立风险传播评估方法, 构建机场之间传播关联, 从而建立风险传播网络并分析其网络特性。

1 航空网络运行风险建模

航空网络运行风险是指航空网络各机场节点运行产生安全问题的危险系数, 运行风险越高, 危险系数越大。运行概念指的是航空网络各机场正常工作, 维持空中交通正常运作的状态。因此, 运行风险与航空网络安全和效率都有着密切的关系。运行风险并不能由单一风险完整表征, 是多种局部风险因素共同作用的结果。因此, 局部风险因素的选取对机场风险的量化尤为重要, 选择合适的风险因素才能有效衡量机场风险情况。本文从运行效率与运行安全 2 个角度出发, 基于时间和空间维度选取局部风险因素进行运行风险建模。

1.1 时间维度风险因素

针对时间维度局部诱因选取, 综合考虑现实场景与理论研究要点, 本文考虑机场平均延误时间表征网络运行效率。机场 i 在 $[t, t+1]$ 时间段的平均离港延误时间 $d_i(t)$ (单位为 h) 的定义为

$$d_i(t) = \frac{y_i(t) + c_i(t) \cdot b}{n_i(t)} \quad (1)$$

式中: $y_i(t)$ 为 $[t, t+1]$ 时间段内的总延误时间; $c_i(t)$ 为 $[t, t+1]$ 时间段内取消的航班数量; $n_i(t)$ 为 $[t, t+1]$ 时间段内计划离港的航班数; b 为取消航班等价延误时间。由于在极端条件下不考虑取消航班, 可能会导致机场平均延误计算产生偏差, 取消应被视为评

估航空运输系统性能的延迟指标^[15]。因此, 根据美国联邦航空局 (FAA), 中国民用航空局 (CAAC) 和欧洲航空安全局 (EASA) 等规定, 一般等价延误时间 b 取值 180 min^[13]。

1.2 空间维度风险因素

机场饱和度可以描述网络节点机场的负载情况与运行状态, 在空间维度上侧重描述网络运行安全。机场 i 在 $[t, t+1]$ 时间段的饱和度 $s_i(t)$ 的定义为

$$s_i(t) = \frac{Q_i(t)}{C_i(t)} \quad (2)$$

式中: $Q_i(t)$ 为 $[t, t+1]$ 时间段内的进港流量; $C_i(t)$ 为 $[t, t+1]$ 时间段内的机场容量。针对机场容量, 根据民航局技术规范文件对机场时刻容量评估技术的规定^[16], 本文采用基于历史统计数据的容量评估方法, 通过对过去 3 个月航班数据的评估与分析, 确定各个机场的时刻容量 $C_i(t)$ 。同时, 根据真实国内航班数据统计进离港航班数量, 确定各机场时刻进港流量 $Q_i(t)$ 。一旦饱和度大于等于 1, 代表机场节点已超负荷运行, 机场运行瘫痪风险较大, 随之会导致机场负载分流到其他机场, 造成其他机场负荷升高, 从而出现网络节点级联失效瘫痪的问题。因此, 机场饱和度在一定程度上能够反应出机场的安全情况, 机场安全等价于网络节点安全。

1.3 机场风险耦合

针对航空网络中机场风险耦合, 文献 [17] 中选取多个局部拥挤风险指标, 运用灰色聚类模型构造拥挤评估矩阵, 采用针对不同聚类子集的白化权重函数计算各个拥挤指标的聚类权重, 通过线性加权获得拥挤风险等级。该模型的输入为多个局部因素变量, 输出为相应拥挤风险等级, 但最终还是没有给出具体量化结果。本文在聚类分级的基础上, 对机场运行风险提出了具体的量化方法。对于 $[t, t+1]$ 时间段内的平均离港延误时间与机场饱和度, 由于二者的量纲不同, 首先采用双曲正切函数对二者进行归一化处理, 之后采用 K-Means 聚类算法对所有机场的延误时间与饱和度进行聚类, 分别得到 4 个延误区间和饱和度区间。作为最经典的划分式聚类算法, K-Means 算法^[18] 是著名的基于欧氏距离的聚类算法, 在业界得到了广泛应用。本文基于 K-Means 算法对 $[t, t+1]$ 时间段内的各机场延误与饱和度进行聚类并形成各区间后, 结合式 (3) 所示的延误权重映射函数, 先将机场 k 的平均延误时间对应到相应的延误区间 i ($[\lambda_{d,i-1}, \lambda_{d,i}]$) 中, 再通过延误权重映射函数的作用, 得到 $[t, t+1]$ 时间段内的延误权重 $W_d(k, t)$ 。以此类推, 将其他所有机场在 $[t, t+1]$ 时间段内的延误权重计算出来。同理, 结合

式 (4) 所示的饱和度权重映射函数, 先将机场 k 的饱和度对应到相应的饱和度区间 $i([\lambda_{s,i-1}, \lambda_{s,i}])$ 中, 再通过饱和度权重映射函数的作用, 得到 $[t, t+1]$ 时间段内的饱和度权重 $W_s(k, t)$ 。以此类推, 将其他所有机场在 $[t, t+1]$ 时间段内的饱和度权重计算出来。最后, 根据式 (5) 求得机场 k 在 $[t, t+1]$ 时间段内的风险系数 $R_k(t)$ 。对机场 k 所在的所有时间段中都进行上述操作, 即可得出机场 k 的风险时间序列 R_k 。同理, 对其他所有机场都进行上述操作, 即可获得相应时间维度内所有机场的风险时间序列。

$$W_d(k, t) = \lambda_{d,i}^{0.5} e^{\frac{d_k(t) - \lambda_{d,i}}{\lambda_{d,i}}} + \sum_{j=1}^{i-1} \lambda_{d,j}^{0.5} \quad i = 1, 2, 3, 4 \quad (3)$$

$$W_s(k, t) = \lambda_{s,i}^{0.5} e^{\frac{s_k(t) - \lambda_{s,i}}{\lambda_{s,i}}} + \sum_{j=1}^{i-1} \lambda_{s,j}^{0.5} \quad i = 1, 2, 3, 4 \quad (4)$$

$$R_k(t) = \sqrt{W_d(k, t)^2 + W_s(k, t)^2} \quad (5)$$

式中: n 为区间总数且 $n = 4$; $\lambda_{d,i}$ 和 $\lambda_{s,i}$ 分别为延误区间 i 和饱和度区间 i 的右边界值; $d_k(t)$ 和 $s_k(t)$ 分别为机场 k 在 $[t, t+1]$ 时间段内归一化处理后的平均离港延误时间和机场饱和度的值。最终, 各时间段的风险系数组合起来就形成了机场的风险时间序列。

2 基于格兰杰检验的风险传播网络

2.1 格兰杰因果关系检验模型

2.1.1 传统检验模型

在时间序列情形下, 2 个变量 X 、 Y 之间的格兰杰因果关系定义为: 如果变量 X 有助于预测另一个变量 Y 未来的变化值, 即在 Y 关于 Y 的过去值的回归中, 添加 X 的过去值作为独立变量能够显著地增加回归的解释能力, 如式 (6) 所示, 则称 X 是 Y 的格兰杰原因。在航空网络中, 假设 2 个机场 A 和 B , 通过使用 B 的风险时间序列的过去值判断是否有助于改进 A 的风险系数时间序列的预测, 如果有助于预测, 则说明 B 的风险序列是 A 的风险序列格兰杰原因, 即可以认为机场 B 将风险传播给了机场 A 。

$$\sigma^2(Y|U^-) < \sigma^2(Y|U^- \setminus X) \quad (6)$$

式中: $\sigma^2(Y|U^-)$ 表示用所有信息预测 Y 时间序列所产生的残差平方, U^- 表示当前集合的所有信息; $\sigma^2(Y|U^- \setminus X)$ 表示用除去机场 X 之外的所有信息预测 Y 时间序列所产生的残差平方。若前者小于后者, 则说明 X 有助于 Y 的风险系数时间序列的预测, 即 X 是 Y 的格兰杰原因。

针对 2 座机场 i 和 j , 将其风险时间序列分别设定为 Y_i 和 Y_j , 2 座机场的风险时间序列的格兰杰因果关系检验步骤如下:

步骤 1 确立原假设 H_0 : 时间序列 Y_i 不是引起 Y_j 变化的格兰杰原因, 机场 i 没有将风险传给机场 j 。

步骤 2 针对机场 j 的风险时间序列, 分别建立 2 个回归模型。

无约束回归模型为

$$y_i(t) = \sum_{m=1}^p a_m y_i(t-m) + \sum_{m=1}^p b_m y_j(t-m) + e(t) \quad (7)$$

有约束回归模型为

$$y_i(t) = \sum_{m=1}^p a_m y_i(t-m) + e(t) \quad (8)$$

式中: $y_i(t)$ 为风险时间序列 Y_i 在 t 时刻的值; $y_i(t-m)$ 为 $y_i(t)$ 的前 m 时刻的过去值; p 为回归模型的滞后期; $e(t)$ 为噪声。因此, 原假设 H_0 可以被描述为

$$b_1 = b_2 = \dots = b_p = 0 \quad (9)$$

步骤 3 计算这 2 个回归模型的残差平方 RSS_r (有约束回归的残差平方) 和 RSS_{ur} (无约束回归的残差平方), 构造 F-检验的统计量 F 为

$$F = \frac{(RSS_r - RSS_{ur})/P}{RSS_{ur}/(S - P)} \quad (10)$$

式中: S 为时间序列样本总量。通过计算 F-检验的统计量 F , 求出 F 值对应的概率 p 值, 如果 $F_p > F_\alpha$, 即 $p < \alpha$ (默认情况下 $\alpha = 5\%$), 则可抛弃原假设, 认为时间序列 Y_i 是引起 Y_j 变化的格兰杰原因, 机场 i 将风险传给机场 j 。

目前传统格兰杰检验模型在越来越多的领域用来检验因果关系, 其模型本身的不足也被逐渐发现。传统检验模型有以下 2 点不足:

1) 传统格兰杰因果检验的检验结果受检验数据平稳性的影响较大, 非平稳数据会导致伪回归的发生, 由于格兰杰检验本质为自回归分析, 因此, 非平稳数据导致的伪回归会使模型检验失误。但传统模型本身并没有对数据的处理方法做出标准说明, 需要结合其他统计学方法改善数据平稳性。

2) 传统方法直接采用 F-检验作为检验结果的唯一标准, F-检验对于数据的正态特征比较敏感, 在少数情况下会出现误差, 并且在时间序列较短的情况下容易导致伪因果关系的判定, 检验标准不够全面。

2.1.2 模型优化

针对上述传统模型的不足, 本文进行 2 方面优化:

1) 针对格兰杰因果检验对数据平稳性敏感的问题, 本文在模型数据的预处理阶段, 在现有数据的基础上, 引入 Z-score 正则化、单位根 (ADF) 检验与一阶差分模型, 保证数据平稳性, 防止伪回归。

其中, Z-score 正则化属于一种去除数据趋势化与均值的方法, 经过 Z-score 正则化操作后的数据平稳性会大幅增加, Z-score 正则化定义为

$$R'(h,t)=\frac{R(h,t)-\langle R(\cdot,t)\rangle}{\sigma(R(\cdot,t))}$$

(11)

式中: $R(h,t)$ 为 h 时间段内 t 时刻的风险值; $\langle R(\cdot,t) \rangle$ 和 $\sigma(R(\cdot,t))$ 分别为 t 时刻风险的平均值和标准差; $R'(h,t)$ 为经过 Z-score 正则化操作后的风险值。

ADF 检验^[19] 是一种检验序列平滑度的方法, 通过检测序列中是否存在单位根判断数据是否平稳。由于部分数据在经过 Z-score 正则化操作后仍存在不平稳的状态, 因此在 Z-score 正则化操作后需要再进行 ADF 检验, 如果通过检验, 则说明序列平稳。如果未通过检验, 则需要对序列继续进行平滑化操作, 并采用一阶差分方法:

$$R'(h,t)=R(h,t+1)-R(h,t)$$

(12)

2) 针对格兰杰因果检验最后的判别方法, 本文在传统检验判定中 F-检验的基础上, 引入似然比检验的思想。由于机场风险时间序列较短, 仅依靠 F-检验可能会出现伪因果关系的判定, 似然比检验可从似然的角度验证回归模型差异的显著性与拟合效果, 提高最终检验的准确率。似然比检验和其他显著性检验方法的目的, 本质上都是确定模型的约束条件是否有效, 有约束模型与无约束模型相比是否合适。但是似然比检验的通用性与有效性相比于一般的假设检验更好^[20]。因此, 本文采用似然比检验的方法与 F-检验共同作为因果关系判别的手段。

似然比检验为

$$G=2(\ln(L_{ur})-\ln(L_r))e^{\theta}$$

(13)

式中: L_{ur} 为无约束回归模型的最大似然值; L_r 为有约束回归的最大似然值; θ 为假设条件中的未知参数; G 为似然比检验的结果, 近似服从卡方分布 (χ^2)。根据 G 值得出相关概率 p 值, 与 F-检验相似, 若 $G > G_{\theta}$, 即 $p < \delta$ (默认情况下 $\delta = 5\%$), 则通过似然比检验。只有同时通过 F-检验与似然比检验, 才能抛弃原假设 H_0 , 认为二者之间存在因果关系。

2.2 风险传播网络建模

根据 2.1.1 节的模型优化, 本节介绍风险传播网络的具体搭建方法, 如图 1 所示。首先对原始的风险时间序列进行数据预处理, 经过 Z-score 正则化对数据进行去趋势化与去均值操作。之后, 对数据进行 ADF 检验, 若通过检验, 则可以继续进行 F-检验和似然比检验, 否则需要进行一阶差分操作, 直至通过 ADF 检验。经过 F-检验和似然比检验后, 可以得出二者之间的因果关系, 即风险传播关系。

用同样的方法对航空网络中所有的机场两两之间互相进行相同的因果检验操作, 得出网络内全部的机场风险传播关系。最后, 以机场为节点, 机场间风险传播关系为有向的连边, 构建风险传播网络。

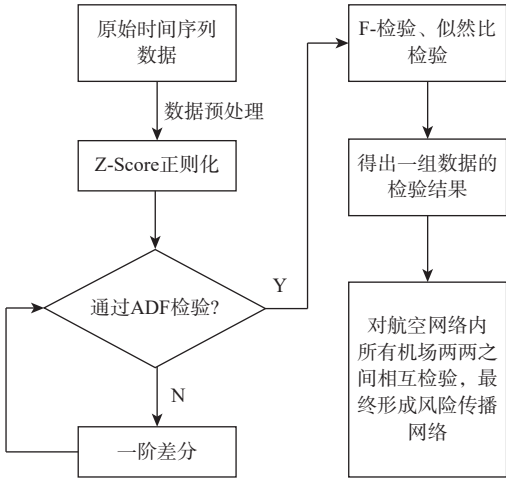


图 1 运行风险传播网络构建流程

Fig. 1 Flowchart of risk propagation network construction

3 风险传播网络特征分析

3.1 网络分析指标

为了深入挖掘风险传播网络的特点, 本节基于复杂网络理论对 1 天的风险传播网络进行特征分析。首先选取 1 天的航班数据, 计算每个机场每个小时的风险系数, 形成一个长度为 24 的时间序列, 并对每个机场的时间序列彼此间做格兰杰因果关系检验, 最终得到 1 天的运行风险传播网络。根据所得到的风险传播网络, 基于复杂网络理论进行基本特性分析。网络的基本属性定义如下: 规定 $A = [a_{ij}]_{N \times N}$ 为网络的邻接矩阵, N 为网络节点数量。其中 $a_{ij} = 1$ 表示存在节点 i 到节点 j 的连边, 否则 $a_{ij} = 0$ 。网络节点数即网络中的个体点数, 连边数即节点之间的连线的个数。无向网络中, $a_{ij} = a_{ji}$ 。在此引入以下复杂网络指标。

1) 网络节点度数的计算。在风险传播网络中, 节点的出度代表机场传播风险到其他机场的数量, 入度则代表机场接收来自其他机场传播的风险的数量。网络节点 i 的出度 k_i^{out} 与入度 k_i^{in} 的计算公式为

$$k_i^{\text{out}}=\sum_{j=1}^Na_{i,j}$$

(14)

$$k_i^{\text{in}}=\sum_{j=1}^Na_{j,i}$$

(15)

在风险传播网络中, 引入出入度比 r_i , 其定义为

$$r_i=\frac{k_i^{\text{in}}+1}{k_i^{\text{out}}+1}$$

(16)

2) 网络密度 $L_D^{[21]}$ 。用来描述网络中连边的疏密程度,在风险传播网络中描述的是机场风险传播的密度,定义如式(17)所示。网络中的连边越多,风险传播越密集,网络密度越大。

$$L_D = \frac{\sum_{j=1}^N a_{i,j}}{N(N-1)} \tag{17}$$

3) 网络聚类系数 $C^{[22]}$ 。用来表示网络内节点结成集团的程度,即节点的固有聚类趋势,也是网络中三角形区域占比大小。在风险传播网络中,机场节点的聚类系数是指具有直接风险传播连边(即网络中的三角形数量)的相邻机场对(具有风险传播连边的机场)的比例。网络的平均聚类系数计算方法为

$$C = \frac{1}{2N} \sum_{i=1}^n \frac{\sum_j \sum_h (a_{i,j} + a_{ji}) (a_{i,h} + a_{hi}) (a_{h,j} + a_{jh})}{d_{single} (d_{single} - 1) - 2d_{double}} \tag{18}$$

式中:

$$d_{single} = \sum_{j \neq i} a_{i,j} + \sum_{j \neq i} a_{ji} \tag{19}$$

$$d_{double} = \sum_{j \neq i} a_{i,j} a_{ji} \tag{20}$$

4) 最大连通子图 $LCC^{[23]}$ 。用来表示网络中所有节点都连接在一起的子图中节点最多的子图,通常被用于描述网络的规模,最大连通子图尺度为 N_{LCC} 。在风险传播网络中,最大连通子图可以直接描述风险的传播范围。

5) 互异性 $R_C^{[24]}$ 。用来表示网络节点之间互相连接的程度,即网络中所有节点对之间为双向连接的比率,定义如式(21)所示。在风险传播网络中,满足互异性的一对机场都有着相互影响,彼此间互相传播风险的特点。

$$R_C = \frac{\sum_{i \neq j}^N (a_{i,j} - \bar{a}) (a_{ji} - \bar{a})}{\sum_{i \neq j}^N (a_{i,j} - \bar{a})^2} \tag{21}$$

式中:

$$\bar{a} = \frac{\sum_{i \neq j}^N a_{i,j}}{N(N-1)} \tag{22}$$

6) 社区性 $Q_C^{[25]}$ 。用来评估网络的模块化特点。在风险传播网络中,社区性可用于评估机场之间的风险传播是否可以分为几个子区域,其中机场

的每个子区域内部都有密集的风险传播链路,而与网络其余部分的链路稀疏。社区性强弱的评估方法为^[26]

$$Q_C = \frac{1}{M} \sum_{i,j} \left(a_{i,j} - \frac{k_i^{out} k_j^{out}}{M} \right) \delta(c_i, c_j) \tag{23}$$

式中: M 为网络中的边数; $\delta(c_i, c_j)$ 表示如果节点 i 和 j 在同一个社区中,则返回 1, 否则返回 0。

7) 网络效率 $E^{[27]}$ 。用来描述网络中一个节点到达另一节点的难易程度,其定义为

$$E = \frac{1}{N(N-1)} \sum_{i,j \neq i} \frac{1}{d_{i,j}} \tag{24}$$

式中: $d_{i,j}$ 为节点 i 和 j 之间的距离(经过节点的数量)。在风险传播网络中,网络效率描述一个机场的风险传播到另一个机场的难易程度,即风险传播的容易度。一个节点到达另一个节点所需的路程距离越大,说明到达越困难。

3.2 传播网络基本属性

本文采用 2019 年前 5 个月的国内航班数据,根据这 5 个月的国内航班数据得到的日均国内航班量约为 8 844 架次,涵盖 223 座机场。其中,2019 年的春节,2 月 5 日的国内航班量最少,共计 6 952 架次,而作为传统假日,5 月 1 日的航班量为这几个月中最多,共计 9 421 架次。3 月 2 日的国内航班量为 8 846 架次,最接近平均水平。因此,本节采用 2019 年 3 月 2 日的航班数据进行分析。

本文构建风险传播网络涉及数据信息包括航空公司名称、航班代码、计划起飞时间、实际起飞时间、计划到达时间、实际到达时间、始发机场、目的机场、航班状态以及飞行日期。根据机场运行风险建模方法,对各机场一天内每个小时的运行风险进行量化建模,最终对每个机场形成长度为 24 的风险时间序列。根据各机场的风险时间序列,通过风险传播识别方法检验机场间的风险传播关系,从而研究当天整体的风险传播情况。

图 2 为 2019 年 3 月 2 日的整体风险传播情况可视化结果,不同机场节点代表着它们的风险吸收系数 r_i 不同。 $r_i > 1$ 意味着机场 i 在风险传播过程中,吸收来自其他机场的风险数量比自身传播出去的风险数量多,对风险的传播有着阻滞作用。同理, $r_i < 1$ 意味着机场 i 吸收来自其他机场的风险数量比自身传播出去的风险数量少,对风险的传播有着加剧作用。从图 2 可知多数的大机场倾向于吸收风险的传播,而多数小机场倾向于加剧风险的传播。

根据 2019 年前 5 个月的数据,国内有航班运行的机场共计 223 座,因此网络节点数为 223。为了

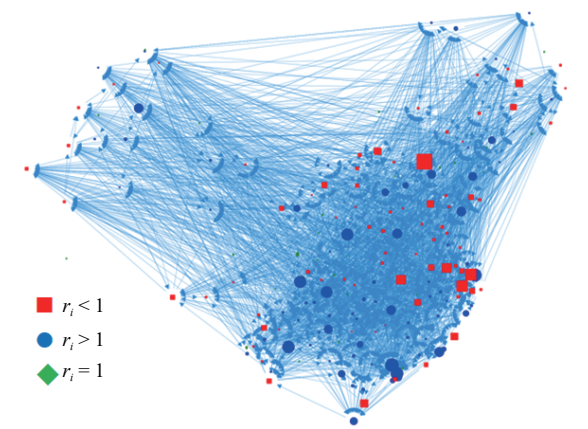


图 2 2019 年 3 月 2 日整体风险传播网络视图
Fig. 2 General risk propagation network view for 2 March 2019

更清楚地了解风险传播网络的网络指标, 本节同时对具有和传播网络相同节点数与边数的有向随机

网络与同一天的航线网络也进行了相应的指标检测。本部分涉及的随机网络结果均是生成 20 次随机网络平均后的结果。

表 1 为风险传播网络的基本属性, 结果表明, 3 种网络的密度均较小、网络直径小, 具有小世界特点。与相同数量边和节点的随机网络相比, 风险传播网络显现出较高的群簇性、互异性与社区性, 说明风险传播网络中的机场节点存在传播聚集的特点。其中, 风险传播网络的社区性相比随机网络和航线网络又具有比较明显的差别, 风险传播网络中社区性体现了机场之间的风险传播是否可以分为几个子区域, 社区性较高可能是由于风险在传播过程中机场本身的调节能力产生了部分的风险吸收, 致使区域机场产生的风险在传播中逐渐减小甚至消弭, 区域风险的传播范围始终有限。

表 1 风险传播网络的基本属性										
Table 1 Basic attributes of risk propagation network										
网络	节点数	边数	平均出入度	网络密度	平均聚类系数	最大连通子图	互异性	网络直径	社区性	网络效率
风险传播网络	223	2 772	12.43	0.06	0.11	160	0.04	3	0.23	0.31
随机网络	223	2 772	12.43	0.06	0.06	223	0.02	3	0.17	0.55
航线网络	223	1 656	14.85	0.07	0.51	179	1	4	0.14	0.33

此外, 传播网络的效率低于随机网络, 且最大连通子图的尺寸较小, 说明相比于网络分布较均匀的随机网络, 风险传播网络则更集中于某个区域, 而部分区域并不明显, 风险在全局网络的传播容易度较低。与航线网络相比, 风险传播网络的群簇性较小, 且航线网络的互异性为 1, 这是因为航线网络是双向网络, 相比单向的风险传播网络具有更强的群簇性, 而传播网络的社区性相较其他 2 种网络较大, 说明传播网络倾向于分成几个子区域。此外, 传播网络的最大连通子图尺寸也小于航线网络, 说明风险传播的规模并没有航线网络的规模大。由于风险的发生存在某些诱因, 而这些诱因仅可能影响航线网络中的少部分机场, 而少部分机场产生的风险辐射范围有限, 并不能影响到整个网络, 故风险传播网络的规模要比航线网络小。网络聚类系数表示网络内节点结成集团的程度, 风险传播网络的聚类系数要小于航线网络, 这是因为机场在对其航线相连机场进行风险传播时, 并非所有与其相连的机场都会受到风险的波及, 部分机场由于自身周转能力较好而抵御住了风险, 故风险传播网络的社区性虽较高但聚类系数相比航线网络低。网络效率描述一个机场的风险传播到另一个机场的难易程度, 即风险传播的容易度, 风险传播网络的效率比航线网络的低可能也是由于这个因素。

3.3 传播网络度相关性分析

通过对风险传播网络的基本分析, 可以看出风险传播网络密度小、具有社区性和小世界网络特点, 风险传播区域较为集中, 网络效率较低, 风险传播连通性较弱。除了这些基本的网络参数外, 网络度数相关的分析对于了解网络属性也尤为重要。因此, 本节将对 2019 年 3 月 2 日的风险传播网络的度数进行相关研究。

首先对表 1 中的风险传播网络、随机网络与航线网络的出、入度的累积分布进行分析, 结果如图 3 和图 4 所示。图 3 为风险传播网络的出、入度分布结果, 横坐标 k_{out} 表示出度, k_{in} 表示入度。纵坐标为节点度数大于等于 k_{out} 或 k_{in} 的概率。结果显示风险传播网络的出、入度均服从双区对数分布的结果即 $P(X \geq k) = \lambda \ln k$, 图中用 $P(X \geq k_{out})$ 及 $P(X \geq k_{in})$ 表示, λ 为对数分布系数。并且出、入度的结果均显示 $\lambda_1 > \lambda_2$, 说明不论是出度还是入度, 多数节点都倾向于具有较大的度数, 而只有少数节点的度数较小。这表示网络中的大部分机场都参与到了风险传播的过程中, 且传播度数较大。

图 4 为与风险传播网络同等规模的随机网络的出、入度累积分布情况。由图 4 可知, 随机网络的度分布首位相对平缓而中间部分相对线性且均匀, 整体近似满足正态分布特点。这说明网络中各

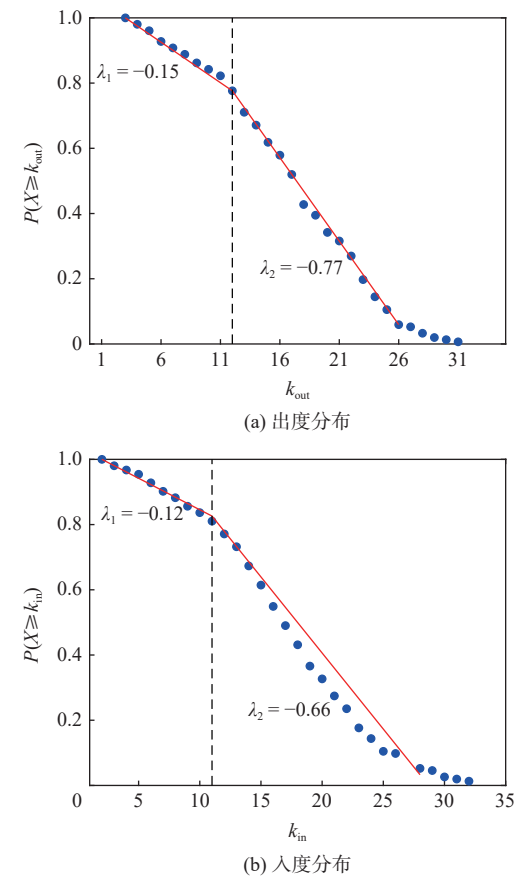


图3 风险传播网络的出、入度分布

Fig. 3 Out-degree and in-degree distribution of risk propagation network

节点的度数分布聚集于整个网络中间的水平, 度数过大或过小的节点数量都相对较少, 中间度数的节点较多, 且分布均匀。这些都符合随机网络随机生成节点之间连边的特点。

图5为2019年3月2日航线网络的度分布结果, 横坐标 k 表示度数, 纵坐标为节点度数 $X \geq k$ 的概率。由于航线网络为双向网络, 出、入度结果一致, 因此统一为度分布。航线网络的度分布满足指数分布的特点, 即 $P(X \geq k) \propto e^{-\lambda k}$ 。这表明航线网络的度分布特点与风险传播网络相反, 航线网络的度分布倾向于少数节点拥有大的度数而多数节点的度数较小。这也显示了航线网络的异质性, 符合目前航空网络现状, 即少数大型机场为空中交通枢纽, 连接大部分中、小型机场的特点。

除了网络的出、入度分布, 为了研究机场的风险传播复杂度与风险集群的关系, 本节还将研究网络的度簇相关性特点, 即网络节点的节点度数与群簇性(倾向于聚类的程度)的关系。研究对象仍然是上述的3种不同的网络, 最终结果如图6所示, 横坐标为度数, 纵坐标为聚类系数 C 。由于随机网络与风险传播网络的出度和入度的度簇相关性基本一致, 因此, 图6仅展示二者的出度度簇相关

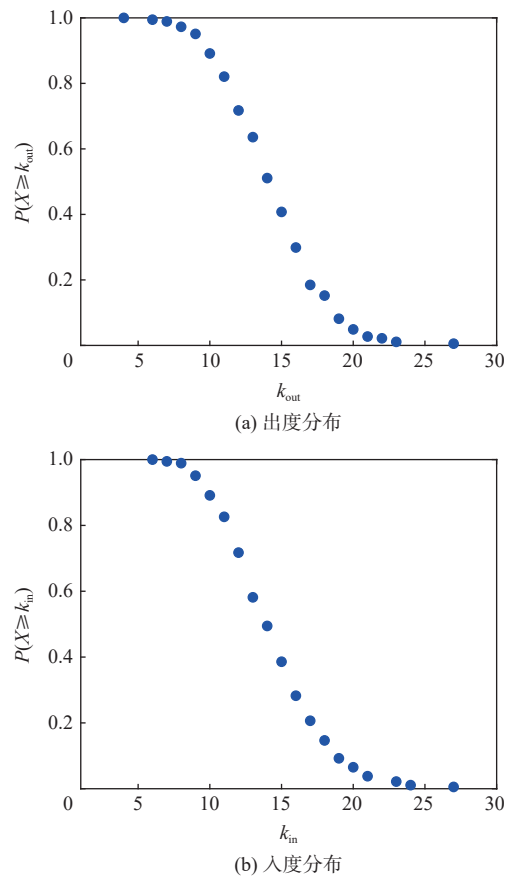


图4 随机网络的出、入度分布

Fig. 4 Out-degree and in-degree distribution of random network

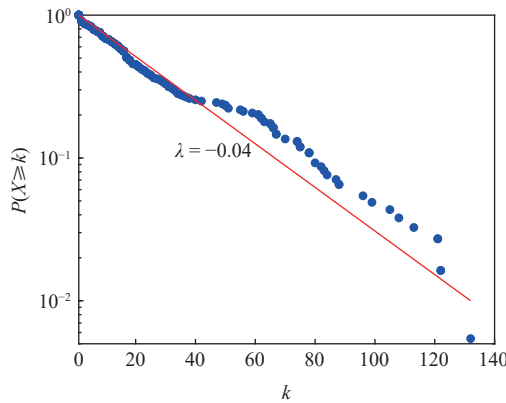


图5 航线网络的度分布

Fig. 5 Distribution of airline network degree

性。由图6可知, 随机网络与风险传播网络的度簇相关性并不明显, 风险传播网络出度的度簇相关系数为0.30(入度的度簇相关系数 R 为0.28), 而随机网络入度的簇度相关系数 R 仅为-0.07(入度的度簇相关系数 R 为-0.04), 表明这2种网络节点的群簇性与节点的度数并没有明显联系。这说明无论风险传播关系复杂还是简单的机场, 并不意味着在其周围会产生传播集群。由3.2节可知, 风险传播可以被分为几个连接密集的区域, 但是这些区域并不是以传播关系复杂或者简单的机场为中心的。因

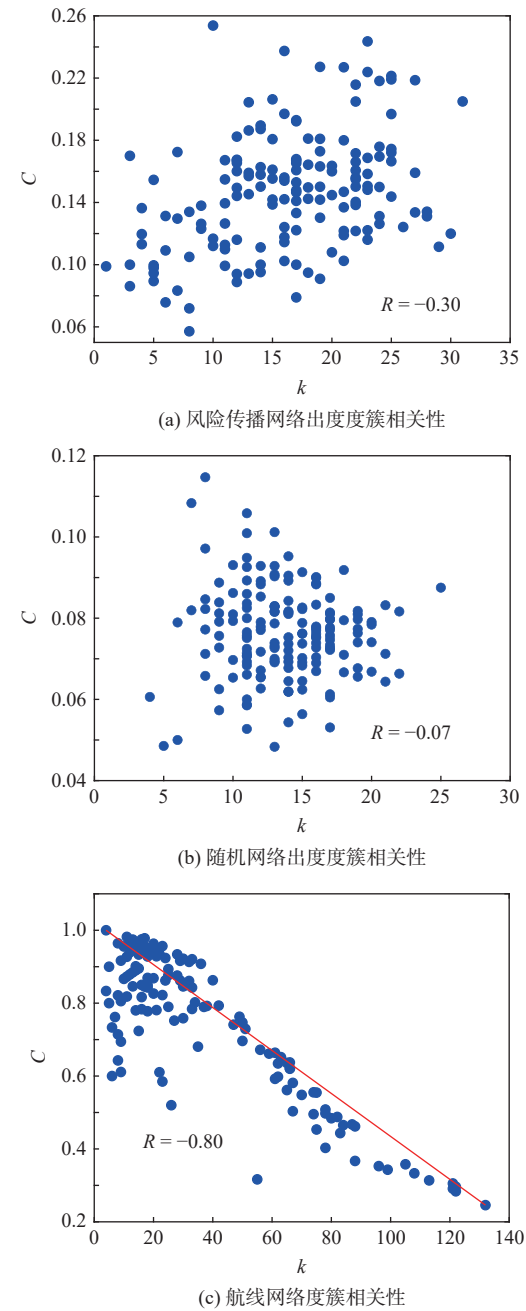


图 6 3 种网络的度簇相关性

Fig. 6 Degree cluster correlation of three types of networks

此,在传播集群问题上,并不是预料中的传播关系复杂的机场所主导的。而航线网络的节点度数与节点群簇性有着明显的关系(度簇相关系数 R 为 -0.80),其度簇相关性呈现负相关关系,即度数越大的机场节点,其群簇系数就越小,机场节点的聚类趋势越弱,邻居的邻居也是邻居的可能性越小。这是因为航线网络具有较强的异质性,度数小的机场更倾向于连接度数大的机场。而大机场和众多小机场都有连边,因此,小机场的邻居(小机场)的邻居(大机场)也是这个小机场邻居的概率大,度数小的机场的聚类系数相对较高。而大机场的邻居(小机场)的邻居(小机场)也是其邻居的概率小,受城

市发展原因影响,一般三线及以下的城市之间的直接的航线很少,小机场之间彼此相连的概率较小,因此,度数大的机场的聚类系数就相对较小。

4 结 论

- 1) 风险传播网络具有较短的网络直径和较高的社区性,说明传播网络具有小世界特点,且风险传播可以被分为几个连接密集的区域。而风险传播网络的网络效率较低,说明风险在全局传播的难度较高。
- 2) 通过对风险传播网络的度数相关特点分析,发现网络的度分布满足双区对数分布的特点,说明多数机场的度数都较大,参与到风险传播中,小度数的机场较少。
- 3) 航线网络由于其双向连边与地域发展程度不同的特点,具有较高的聚类系数与相对较大的网络直径,然而其网络效率与社区性都较低。此外,航线网络的节点具有较明显的度簇负相关性,度数大的机场节点群簇性弱,而度数小的机场具有较强群簇性。

参考文献 (References)

[1] CAI K Q, ZHANG J, DU W B, et al. Analysis of the Chinese air route network as a complex network[J]. Chinese Physics B, 2012, 21(2): 028903.

[2] ZHANG J, CAO X B, DU W B, et al. Evolution of Chinese airport network[J]. Physica A:Statistical Mechanics and Its Applications, 2010, 389(18): 3922-3931.

[3] KAFLE N. Modeling flight delay propagation: A new analytical-econometric approach[J]. Transportation Research Part B:Methodological, 2016, 93: 520-542.

[4] NAYAK N, ZHANG Y. Estimation and comparison of impact of single airport delay on national airspace system with multivariate simultaneous models[J]. Transportation Research Record:Journal of the Transportation Research Board, 2011, 2206(1): 52-60.

[5] PYRGIOTIS N, MALONE K M, ODONI A. Modelling delay propagation within an airport network[J]. Transportation Research Part C:Emerging Technologies, 2013, 27: 60-75.

[6] BASPINAR B, URE N K, KOYUNCU E, et al. Analysis of delay characteristics of European air traffic through a data-driven airport-centric queuing network model[J]. IFAC-PapersOnLine, 2016, 49(3): 359-364.

[7] ABDELGHANY A, EKOLLU G, NARASIMHAN R, et al. A proactive crew recovery decision support tool for commercial airlines during irregular operations[J]. Annals of Operations Research, 2004, 127(1): 309-331.

[8] ABDELGHANY A, ABDELGHANY K, EKOLLU G. A genetic algorithm approach for ground delay program management: The airlines' side of the problem[J]. Air Traffic Control Quarterly, 2004, 12(1): 53-74.

[9] WU C L, LAW K. Modelling the delay propagation effects of mul-

- multiple resource connections in an airline network using a Bayesian network model[J]. *Transportation Research Part E:Logistics and Transportation Review*, 2019, 122: 62-77.
- [10] 李俊生, 丁建立. 基于贝叶斯网络的航班延误传播分析[J]. *航空学报*, 2008, 29(6): 1598-1604.
- LI J S, DING J L. Analysis of flight delay propagation using Bayesian networks[J]. *Acta Aeronautica et Astronautica Sinica*, 2008, 29(6): 1598-1604(in Chinese).
- [11] 王珊珊, 王建东, 丁建立. 航班延误波及链的有色出现网模型[J]. *计算机科学*, 2009, 36(2): 241-244.
- WANG S S, WANG J D, DING J L. Colored occurrence net model of flight delay propagation chain[J]. *Computer Science*, 2009, 36(2): 241-244(in Chinese).
- [12] 李善梅, 徐肖豪, 孟令航. 基于聚类神经网络的机场拥挤等级预测[J]. *计算机工程与应用*, 2013, 49(17): 254-257.
- LI S M, XU X H, MENG L H. Forecasting of airport congestion level based on cluster and neural network algorithms[J]. *Computer Engineering and Applications*, 2013, 49(17): 254-257(in Chinese).
- [13] DU W B, ZHANG M Y, ZHANG Y, et al. Delay causality network in air transport systems[J]. *Transportation Research Part E:Logistics and Transportation Review*, 2018, 118: 466-476.
- [14] ZANIN M, BELKOURA S. Network analysis of Chinese air transport delay propagation[J]. *Chinese Journal of Aeronautics*, 2017, 30(2): 491-499.
- [15] XIONG J, HANSEN M. Value of flight cancellation and cancellation decision modeling[J]. *Transportation Research Record:Journal of the Transportation Research Board*, 2009, 2106(1): 83-89.
- [16] 中国民用航空局空管行业管理办公室. 机场时刻容量评估技术规范: AP-93-TM-2017-01[S]. 北京: 中国民用航空局, 2017: 3-4.
- Air Traffic Management Industry Management Office of Civil Aviation Administration of China. Technical specification for airport time capacity evaluation: AP-93-TM-2017-01[S]. Beijing: Civil Aviation Administration of China, 2017: 3-4 (in Chinese).
- [17] 李善梅. 空中交通拥挤的识别与预测方法研究[D]. 天津: 天津大学, 2014.
- LI S M. Research on identification and prediction methods of air traffic congestion[D]. Tianjin: Tianjin University, 2014 (in Chinese).
- [18] JAIN A K. Data clustering: 50 years beyond K-means[J]. *Pattern Recognition Letters*, 2010, 31(8): 651-666.
- [19] CHEUNG Y W, LAI K S. Lag order and critical values of the augmented dickey-Fuller test[J]. *Journal of Business & Economic Statistics*, 1995, 13(3): 277-280.
- [20] 沃尔夫冈·哈德勒, 利奥波德·西马. 应用多元统计分析[M]. 陈诗一, 译. 北京: 北京大学出版社, 2011: 157-183.
- WOLFGANG H, LEOPOLD S. Applied multivariate statistical analysis[M]. CHEN S Y, translated. Beijing: Peking University Press, 2011: 157-183 (in Chinese).
- [21] DE LAAT M, LALLY V, LIPPONEN L, et al. Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis[J]. *International Journal of Computer-Supported Collaborative Learning*, 2007, 2(1): 87-103.
- [22] FAGIOLO G. Clustering in complex directed networks[J]. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 2007, 76(2 Pt 2): 026107.
- [23] BRODER A, KUMAR R, MAGHOUL F, et al. Graph structure in the web[J]. *Computer Networks*, 2000, 33(1-6): 309-320.
- [24] WATTS D J, STROGATZ S H. Collective dynamics of ‘small-world’ networks[J]. *Nature*, 1998, 393(6684): 440-442.
- [25] ARENAS A, DUCH J, FERNÁNDEZ A, et al. Size reduction of complex networks preserving modularity[J]. *New Journal of Physics*, 2007, 9(6): 176.
- [26] NEWMAN M E J. Modularity and community structure in networks[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(23): 8577-8582.
- [27] LATORA V, MARCHIORI M. Efficient behavior of small-world networks[J]. *Physical Review Letters*, 2001, 87(19): 198701.

Airport risk propagation network oriented to aviation network

GUAN Xiangmin^{1, 2, *}, ZHAO Shuaizhe³

(1. CAAC Key Laboratory of General Aviation Operation, Civil Aviation Management Institute of China, Beijing 100102, China;

2. Zhejiang Key Laboratory of General Aviation Operation Technology, General Aviation

Research Institute of Zhejiang Jiande, Hangzhou 311600, China;

3. School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

Abstract: Aviation network traffic has increased day by day, and operations between airports are closely coupled. Airport risk diffusion has grown to be a major issue that affects their ability to operate safely and effectively. However, the spread mechanism of airport risk at the network level has not yet been fully understood. This paper first proposes an airport risk coupling quantification method based on clustering algorithm and taking into account multiple risk factors, then constructs risk time series and applies causality testing methods to construct airport risk propagation network in order to better study the mechanism of airport risk propagation at the network level. By comparing the performance of different types of networks, analyzing the characteristics of risk propagation networks, and studying the overall characteristics and laws of risk propagation. The findings demonstrate that the degree distribution of the risk propagation network meets the requirements for dual-zone logarithmic distribution, exhibits small-world properties, has a small network width and high community, and can be separated into a number of densely connected areas. The low network efficiency of the risk propagation network indicates that it is more difficult for the risk to spread globally.

Keywords: aviation network; risk coupling; Granger causality test; risk propagation; complex network

Received: 2021-08-18; **Accepted:** 2021-11-22; **Published Online:** 2022-03-31 11:11

URL: kns.cnki.net/kcms/detail/11.2625.V.20220330.1401.003.html

Foundation items: National Natural Science Foundation of China (U1933130,71731001); Zhejiang Key Research and Development Plan (2022C01055); Sub-Project of Chinese Academy of Sciences key Deployment Project (ZDRW-KT-2020-21-2)

* **Corresponding author.** E-mail: guanxiangmin@camic.cn