

# 数字化背景下互联网创业数据风险预警模型

苏越良, 杜芷晴\*

(华南理工大学工商管理学院, 广东 广州 510641)

**摘要:** 数字化背景下, 数据已成为互联网创业的核心资产, 在数据流上的不当行为可能对企业造成潜在损失. 选取互联网创业的典型代表共享单车为例, 提出数据流与数据风险的定义与内涵, 构建并解释数字化互联网企业数据风险体系与数据风险复杂网络. 根据其网络结构特征, 利用 PageRank 算法与负荷-容量模型, 并结合蚁群算法模拟复杂网络的相继故障, 得到最高风险传播路径, 由此搭建风险预警模型. 通过分析共享单车失败和成功案例讨论模型的实现机制和实践价值, 为互联网企业数据风险预知以及如何管理数据风险提供了直接有效的参考和判断依据, 有助于企业实现数据闭环的良性循环; 在经济上, 促进数据资产的积极流动以创造更大效益.

**关键词:** 数字化; 互联网创业; 数据风险; 风险预警模型; 复杂网络

中图分类号: TP273

文献标识码: A

文章编号: 1000-5781(2025)02-0215-17

doi: 10.13383/j.cnki.jse.2025.02.004

## Data risk warning model of Internet entrepreneurship under the background of digitalization

Su Yueliang, Du Zhiqing\*

(Faculty of Business Administration, South China University of Technology, Guangzhou 510641, China)

**Abstract:** In the digital age, data has become the core asset of Internet entrepreneurship. Misconduct in data flow can cause potential losses to the business. This paper takes shared bicycles, a typical example of Internet entrepreneurship, as a case study. It defines the concept and connotation of data flow and data risk constructs a data risk system and a data risk complex network. According to the characteristics of the network structure, the PageRank algorithm and the load-capacity model are used to simulate the successive failures of complex networks. Finally, the highest risk transmission path is obtained, and the risk early warning model is built. By analyzing the failure and success cases of shared bicycles, this paper discusses the implementation mechanism and practical value of the model. This paper provides a direct judgment basis for data risk prediction for Internet enterprises. It can help to promote the positive flow of data assets to create greater economic benefits.

**Key words:** digital; Internet entrepreneurship; data risk; risk warning model; complex networks

## 1 引言

“十四五”规划提出, 需加快数字化发展, 推动数据资源开发利用. 以大数据、物联网、移动互联网和云计算等为代表的数字技术的突破和融合发展推动了数字经济快速发展. 在数字技术创新的背景下, 涌现出以互联网创业为代表的大量创业企业, 在多个领域对商业模式、数字化产业组织形态和业态进行了创新探

收稿日期: 2021-08-21; 修订日期: 2025-01-18.

基金项目: 国家社会科学基金重点资助项目(22AZD039).

\* 通信作者

索<sup>[1]</sup>。陈晓红等<sup>[2]</sup>在创新驱动的重大创业理论与关键科学问题讨论中认为以往适用于传统创业企业的理论和实践研究在数字化时代已面临新的挑战。

目前,已有学者对此做了探究,主要就数字化背景下的创业要素<sup>[3,4]</sup>、创业产出<sup>[5,6]</sup>、创业影响因素<sup>[6-9]</sup>和创业商业模式类型<sup>[10]</sup>等展开探索性研究,但从微观的企业角度出发的研究少之又少,缺少能为创业者提供行之有效与适用的理论基础。同时,以往大多数关于互联网创业企业的研究都集中在使用正面案例探究成功决定因素上,即使当前数字化互联网创业存在一定优势,创立后3年内新创企业的死亡率仍然很高<sup>[11]</sup>,分析互联网创业失败是“一个关键但研究不足的方面”<sup>[12]</sup>。Orlandi等<sup>[11]</sup>选取3个国家层面数据库证明了数字化背景下创业,技术基础设施、人力资本和组织文化等传统情景已经被数字化并会带来新的导致创业死亡的因素。

可见,相比传统的创业企业风险管理大多从财务风险、市场风险、政策风险和技术风险等若干个维度选取指标进行定性定量相结合的研究<sup>[13]</sup>,数字化时代下互联网创业企业的风险管理被赋予了更多新的内涵<sup>[14]</sup>。此处所指的互联网创业,是指借助于数字化技术等新兴信息通信技术与其设备提供提升社会效率或为用户创造价值的新产品或服务的过程,如淘宝、滴滴出行和字节跳动等。从互联网创业企业的典型商业模式不难发现,数据已成为互联网创业企业的核心资产<sup>[15,16]</sup>。企业能够通过提供服务获取大量原始数据,清洗和结构化数据后,分析与挖掘数据的价值,为决策赋能,提升产品与服务质量,提高服务效率,最终获取更多用户及数据,形成良性的数据闭环<sup>[17]</sup>,这一过程称为“数据流”<sup>[18]</sup>。原始大数据只有通过完整的数据流,才能为创业企业的成长带来价值<sup>[19]</sup>。Raimbault<sup>[20]</sup>认为在没有适当的理论和框架的情况下,数据的滥用或使用不当存在风险,第1种可能性是会使研究方向偏向于可用的数据集(例如,许多Twitter地理信息移动性研究),有可能与理论背景脱节,而第2种可能会忽略对模型或模拟的准确性和可靠性至关重要的初步分析数据分辨率。因此,数据风险是指数据流中的不当行为可能对企业造成的潜在损失。数据风险管理旨在通过识别、预警、防范和改进措施,降低这些潜在损失发生的概率。由此看出,由于企业管理传统的物理情景部分或全部数字化<sup>[21]</sup>,数据风险管理成为了互联网创业企业风险管理的重点工作。而在经济领域,数据已经成为企业和社会的一种高价值资产。如何实现数据资产的积极流动以创造价值,有效的数据风险管理是基本保障,不仅有助于预防数据风险事件,避免对社会经济造成负面影响,还能数据风险管理过程中获得启发,以促进资产价值的流通,从而为经济创造更大、更有益的价值。

然而,目前数据风险管理更多从信息科技创新领域出发<sup>[22]</sup>,讨论信息技术层面上的风险管理<sup>[23]</sup>,或者从公共管理领域出发讨论社会治理数字化所面临的挑战<sup>[24]</sup>,主要考虑大量数据和数字社会的超连接性的各种民主提议背后的挑战、限制和后果。而数字化背景下的互联网创业是信息技术驱动下的一场业务、管理和商业模式的深度变革重构。Chen等<sup>[25]</sup>提出,在商业智能和分析(BI&A)1.0时代,数据主要是结构化的,数据管理和存储构成了数据应用的基础。而在BI&A 2.0时代,重点转向从因特网收集的大量日志数据,依赖文本挖掘(如信息提取、主题识别等)和社交网络分析等技术来开发其价值。在当前的BI&A 3.0时代,由配置各类传感器的移动设备提供的大量非结构化信息,给商业智能和分析带来了新的挑战。其中最重要的挑战之一是相关从业者不仅需要熟悉财务、管理、市场营销、物流和运营管理领域的知识,同时还应具备大数据分析、文本挖掘和网络分析等基础技术。然而,目前这一问题在学术界尚未引起足够的关注。而Frenkel<sup>[26]</sup>提出“数字风险官”的职责需要涉及IT领域之外的法律、隐私、合规、数字营销、数字供应链和数字运营等领域,不仅有足够的知识来评估和推荐处理数字化商业风险所需的技术,还要具备商业视角和管理能力,但目前缺少一套从企业的角度出发行之有效的理论,为创业者的数据风险管理提供科学的参考依据。

综上所述,目前在数字化背景下的新型风险相关的创业理论至关重要但尚较空白,主要体现在对作为互联网创业企业核心资产的数据的风险管理上,但目前学者对数据风险的研究偏向信息安全和公共管理的风险治理,而实际数据风险需要从技术与管理结合的全局视角进行管理。在2017年的Strata Data Conference<sup>[27]</sup>中,京东作为互联网企业的代表提出在每天应对海量数据的收集、储存、处理和场景下,它面临的多重现实问题与挑战:1) 京东的消费场景涵盖线上和线下领域,因此需要有效整合这两个场景

的数据,以提升消费者的体验。2) 庞大的电子商务交易量要求建立智能化的物流基础设施,而如何将数据转化为智能物流的能力成为一项关键任务。3) 随着物联网的兴起,数据采集设备不断增多,但与之相关的数据采集难度也相应上升。最后,金融科技的迅速发展为数据安全带来了巨大的挑战。因此,当前数据资产已成为企业的焦点。若不加以谨慎利用,可能会带来风险,但如果不善用这一宝贵的资产,也将导致资源的浪费。这使得数据资产的管理和应用成为了企业管理中的重要议题,需要深入研究和有效的战略规划。

此外,通过采访和一手资料的分析,发现实际企业在管理实践中,虽然察觉到问题的存在,但因缺乏理论指导和体系化管控措施,导致以下几个问题的显现,1) 数据风险缺乏系统性审视,难以对数据风险的识别进行有序的程序化处理;2) 部分企业在运营过程中偶然发现某些风险要素后,进行了重点的风险防范和改进措施,但却忽视了其他数据风险的存在。然而,数据风险通常表现为系统性特征,因此采取的风险管理措施效果受限;3) 已识别的系统性数据风险的重要性尚未得到明确定义,缺乏防范和改进的优先级次序,以及如何实现数据风险管理成本的最小化和效果的最优化。例如,2021年,滴滴因数据安全问题而遭封禁,成为负面案例;百度虽然积累了大量用户搜索数据,但由于领域间数据未能互通,无法为生态系统创造价值,未能实现数据资产的有效循环利用;相比之下,字节跳动通过整合算法、打破部门壁垒,建立了全生态服务的数据中台,从而成为了以数据为核心实现高速发展的企业。基于以上企业面临的实际问题。

综上,结合实际案例在信息通信技术创新与创业管理的交叉领域上丰富数据风险的内涵,并提出基于复杂网络的数据风险预警模型,主要解决以下问题:1) 数字化背景下互联网创业企业需要考虑哪些数据风险。2) 数据风险之间的联系有何规律。3) 企业如何预知和控制数据风险。4) 数据风险事故发生后企业如何应对?

为了识别出数据流上潜在的风险,建立完整客观的数据风险体系,采用案例研究与复制逻辑<sup>[28]</sup>,选取行业与企业样本的标准如下:1) 以利用互联网技术将传统线下场景数字化为业务核心的企业。2) 所在行业的创业企业兼有成功与失败案例。共享单车通过营造线上与线下互动的平台与数字化创新建立竞争优势<sup>[29]</sup>,实现精细化运营;又可通过平台聚合海量数据,强化数据处理能力,积累数字化服务企业的关键资源,塑造差异化核心能力<sup>[4]</sup>同时在行业发展过程中受到诸多争议,具有典型的失败与成功案例。因此,以共享单车创业企业为例,将行业内的典型企业相互对比,总结出各企业中数据风险的逻辑共性,确定本研究建立的数据风险体系在其他数字化背景下的互联网创业企业同样适用并得以验证。

整理归纳出风险指标体系后,传统的企业经营风险预警机制通常利用多变量模型<sup>[30]</sup>、机器学习<sup>[31]</sup>等统计学方法对评估风险大小,然后设置若干个风险等级的临界值,最后判定企业所处的风险等级。大部分研究将众多风险独立开来,从静态的角度考量,较少考虑风险的耦合作用。然而,数据风险之间相互关联,存在一定的因果关系,尤其在数据流上,信息是有向流动的,桂畅旒等<sup>[23]</sup>认为往往一个关键风险没有被及时有效控制,会导致其他风险的出现,通常存在风险逐渐扩大,走向衰落的过程,即在较长时间维度内,风险相互耦合、传播最终爆发的过程。互联网创业企业在发展成为独角兽的过程中,若无法对关键的风险进行识别与管理,由于“短板效应”的存在,不仅会阻碍企业成长与迭代,还可能会加速企业的经营恶化。因此,使用复杂网络对数据风险进行研究,能够从系统的视角出发,从企业整体而非孤立的若干个风险的角度去构建数据风险预警系统。

国内外已有不少利用复杂网络探索风险传播的研究,并且应用在不同领域中,如企业R&D网络<sup>[32]</sup>、创业产业集群<sup>[33]</sup>、交叉性金融业务<sup>[34]</sup>、房地产行业信用风险<sup>[35]</sup>和电力网络<sup>[36]</sup>等。曾维希<sup>[37]</sup>使用复杂网络模型刻画影响因素之间的内在关联关系,利用复杂网络的客观测度指标,分析评价影响因素在网络结构中的相对重要性,由此确定其权重。郭斌<sup>[38]</sup>通过层次分析与复杂网络分析,得出评价京津冀科技协同创新绩效水平的各项指标权重。胡晓峰等<sup>[39]</sup>提出,作战体系分析在理论分析方面需要重点关注复杂网络方法的应用,强调引入复杂网络的思想,构建网络化体系能力指标,并开发评估网络化体系能力的新方法。这表明,复杂网络是研究风险耦合问题和评估复杂指标能力的有效手段。在这一理论指导下,通过构建网络化模型和指标,能够从动态角度出发,更有效地防范和控制风险。

## 2 互联网创业企业数据风险复杂网络

构建数据风险复杂网络是风险预警机制的基础,构建数据风险预警机制目的是监测和掌握企业是数据流上的运营状况,及时发现异常并有效防范和控制,首先需要识别风险,明确互联网创业企业需要考虑哪些数据风险,再构建出复杂网络模型并对其特征进行分析.结合数据风险的网络特征,构建出的风险预警机制满足实用性、系统性、动态性和灵敏性的要求.

### 2.1 互联网创业企业数据分析

选取的创业企业研究对象包括美团单车、青桔单车、哈啰出行以及已破产的OFO、摩拜等企业,这些企业在互联网时代兴起,经历了高度不确定性的创业初期,目前多数已步入高速成长期.它们在资本热潮和市场筛选后,形成了行业内的多寡头垄断局面,这符合互联网创业发展的一般规律.例如,美团单车和青桔单车在经历了初期的快速扩张和市场竞争后,逐渐在市场中占据了一席之地,而OFO和摩拜则因多种因素最终未能持续运营.同时,通过对成功与失败案例的对比研究,能够进一步解释在数字化背景下数据作为核心资产和重要生产要素在互联网创业企业的成长与衰退路径上发挥何种作用.

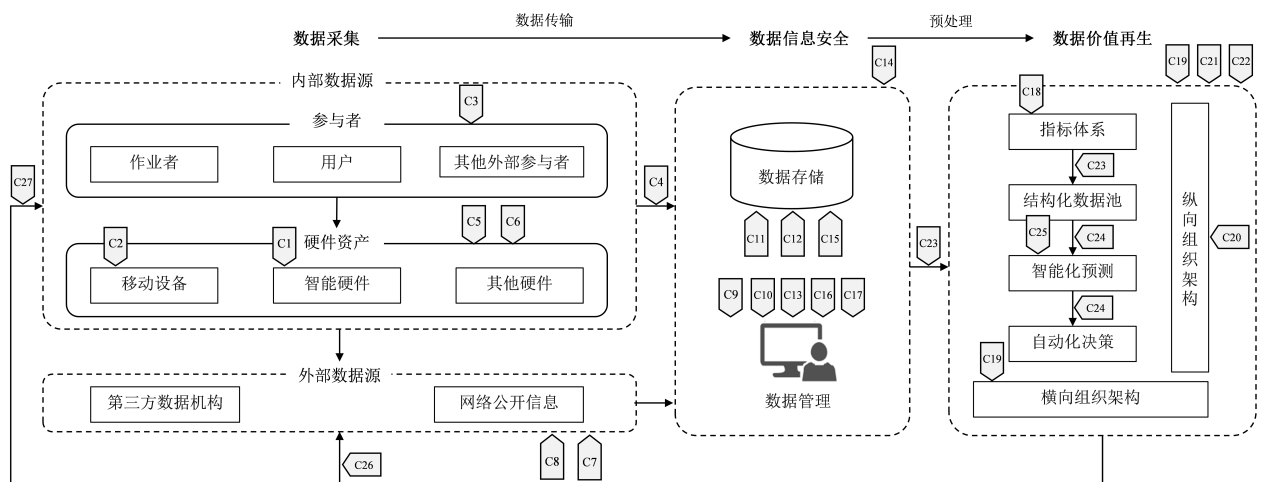


图 1 数据风险地图

Fig. 1 Data risk map

通过整理 2016 年~2021 年企业创始人及高管团队的公开讲话及访谈、相关企业公开发布资料(含年报、官方报告等)、主要媒体评论及政府官方信息、相关文献、以及与相关企业的内部员工访谈,采用编码归纳的方式,应用 ATLAS.ti 7.5.7 进行文本、视频和音频的编码,通过开放式编码获得初始概念,进而通过主轴编码找到初始概念之间的逻辑关系,从数据流上数据采集、数据信息安全和数据价值再生 3 个视角<sup>[40]</sup>对比总结出互联网创业企业 27 个数据风险,如图 1 所示,并且结合共享单车企业经营中的实际场景进行补充解释,以便读者更好理解其含义与内涵.

#### 2.1.1 数据采集

数据采集是指利用数据获取元件或应用软件,实现从数据系统外部采集并输入到系统内部的过程.从数据来源可以分为内部和外部数据源.

在内部数据源中,产生数据的参与者包括作业人员(如运维)、用户以及其他外部参与者(如关联供应商等);硬件资产包括搭载应用软件的移动设备、搭载数据获取元件的关键智能硬件(如共享单车的智能锁)及其他硬件(如蓝牙道钉).在传统创业企业中,参与者在线下场景的行为以及硬件资产的应用操作记录数字化率低,往往用少数经营结果指标以及模糊的直观感受作为调整公司战略方向与管理手段的参考依据.除内

部数据源外,对于创业企业而言,需要竞争对手、相关上下游场景等外部数据作为补充,实现更精准的智能决策。

数据采集对于新兴的互联网创业企业而言,往往是创新探索的过程,其中包含的潜在风险如下。

C1: 缺少传感器、信号收发器等数字化硬件。如 ofo 在创业初期使用物理密码锁,单车缺少中控设备,车辆的物理位置无法通过 GPS 信号定位,车况无法通过无线网络传输,企业对资产的掌控度极弱,导致大量单车失联,造成企业严重的资产损失。

C2: 通过部署数据采集点获取用户或作业者线下行为的应用软件不成熟,无法对线下参与者的行为做标准化与量化牵引。如:调度单车的运维需要在应用软件上记录线下的作业情况(如调度单车数量、调度距离等),如果缺少对应功能或部署数据采集点,则无法有效监控作业人效。

C3: 操作者未按照设计规范使用应用软件,导致数据缺失或失真。如运维鉴定车况时错误填写,没有反映真实车况,此类数据失真对数据价值再生环节产生根本性影响。

C4: 外部复杂环境导致数据无法传输至系统后台。如共享单车运用的多种定位技术的精确性与强度均受外界环境的影响,面临复杂环境,例如下雨天、障碍物遮挡等,容易定位不准确,造成大量低效运营。

C5: 企业内部技术与社会主流技术系统不适配。如哈啰单车、青桔单车等已全面接入主流北斗卫星定位功能,能够更准确了解每一辆车的实时位置和行动轨迹。

C6: 数据采集频率过低。如智能中控每隔一定时间发送车况信息至数据后台,如果时间间隔过长,将降低对车况的掌控强度,可能导致损坏车辆未及时回收维修,运营区外车辆未及时调度回收等情况。

C7: 必要外部数据采集渠道缺失。如各家共享单车企业的数据由于竞争关系无法互通,对于同一区域无法实现总体单车数量、骑行需求、停放管理之间的动态平衡,导致热区车辆堆积的情况,造成资源浪费与公共交通管理问题。

C8: 外部数据准确度无法保障。如对于出行需求的预测,天气作为重要影响要素,企业依赖外部数据源,数据的准确性将影响智能化决策的准确性。

### 2.1.2 数据信息安全

数据信息安全主要发生在数据储存与管理过程中。从三个视角去总结整理,数据流中的信息安全风险:一是数据威胁源,可分为内部人员主动进行破坏和数据泄露,以及企业外部的蓄意攻击与信息窃取,前者往往由于内部对人员的信息安全管理疏忽引起,能够预测与控制,后者往往是突发且不可控的;二是数据流受损的后果,可分为破坏后数量上的减少以及质量上的恶化,前者能够被察觉,而后者通常难以发现;三是数据流受损后信息的流向,可根据是否流入利益相关的外部实体分为两种情况,当内部信息被外部实体所接收时,企业将面临除数据风险以外的更严重与复杂的其他风险,若仅是数据总量的减少,企业只需加强数据风险防控。对这一环节的隐含风险进行整理,数据信息安全风险如表 1 所示。

表 1 数据信息安全风险  
Table 1 Data information security risks

信息是否流入利益相关的外部实体	数据受损后果	数据威胁源	
		内部	外部
否	数量减少	C10: 内部人为破坏	C12: 外部攻击导致数据丢失
	质量恶化	C13: 内部人为篡改数据	C14: 数据在传输、储存等过程中变形受损
是	数量减少	C9: 内部人为泄露、窃取	C11: 外部攻击导致数据外泄
	质量恶化	—	C15: 外部蓄意篡改数据

除此以外,数据权限管理作为信息安全管理的重要一环,蕴含风险如下。

C16: 企业内部与外部权限不明确. 如政府接入企业的数据系统有利于对共享单车的监管, 而企业需要明确内外网络权限设置.

C17: 业务关系人权限不明确. 如美团与滴滴除共享单车以外, 还有网约车、外卖等业务, 如业务权限不明确, 容易发生信息安全故事.

### 2.1.3 数据价值再生

数据价值再生, 是指将数据预处理后, 结合业务逻辑与成长目标, 采用数据分析、数据挖掘等技术, 客观反映业务运营现状, 发现经营异常及其原因, 利用大数据预测未来发展趋势, 使得数据最大化实现其价值的过程. 国内已有不少关于共享单车在大数据背景下如何精细化运营、专业化治理的研究<sup>[41]</sup>. 从横向纵向的组织架构与数据运用流程两个视角总结潜在风险如下.

C18: 数据指标体系化程度低, 忽视指标间内在逻辑. 如车辆调度最优化问题, 需要考虑时空颗粒度、成本、调度方式等多类型指标<sup>[42]</sup>, 各指标之间相互联系, 要实现整体效率最优化, 必须从系统的角度构建指标体系.

C19: 横向不同职能部门的数据口径不统一, 协同运营存在数据壁垒. 如管理运维薪资派发、车辆调度算法和平台定价策略等不同部门, 所用数据在对应的线下场景存在交集, 若分别使用运维应用软件端和单车端所记录的调度数据, 两者在数据采集频次, 计算逻辑缺少统一口径, 阻碍不同场景的相互协作.

C20: 纵向层级数据权限隔绝, 数据同步效率低. 如考虑信息安全与层级管理, 企业的数据权限往往自上而下存在隔绝. 从整体战略目标到各城市的线下运营, 数字鸿沟导致信息同步效率低, 因而对紧急情况反应缓慢.

C21: 不同场景间数据无法实现协同<sup>[43]</sup>. 如: 滴滴构建网约车和共享单车之间的业务连通, 积累较完整的出行数据, 能够精确分析和预判需求旺盛的地点与时间, 用相对更少的车辆满足更多用户需求.

C22: 数据缺少差异化, 如相比ofo, 滴滴构建了多场景的出行数据, 利用打车、代驾等出行数据对共享单车的数据分析进行了补充, 提升了运营效率, 实现精准营销, 从而进一步提高用户黏性.

C23: 数据预处理不当, 如: 过度的数据清洗可能导致信息的丢失, 而数据清洗不足, 可能会引入噪音和干扰, 从而导致分析中的偏差和不准确的结论; 在整合数据源时, 如果对整合的数据领域了解不足, 可能导致对数据含义的错误理解, 从而影响模型的准确性; 对数据过度降维可能使模型过于简单, 无法捕捉数据的复杂关系, 导致结果过拟合.

C24: 智能化不足, 依赖人工低效处理数据. 如哈啰基于大数据、人工智能和云计算等技术, 实时向运维人员输出调度指令, 包括去何处调度多少车辆以及最优的调度路线等, 全过程无需人工干预, 减低成本同时确保决策的准确性与实时性.

C25: 预测系统误差较大. 如: 时空需求量预测结果将直接影响车辆的供需平衡、调度效率与定价策略<sup>[44]</sup>.

C26: 形成数据孤岛, 缺席社会协同治理. 如共享单车是打造数字化城市的重要一环, 要求与产业链生态伙伴、其它公共交通工具、城市交通管理系统协同治理, 打造城市智慧交通闭环, 否则容易出现乱停乱放等痼疾.

C27: 决策下发未形成数据闭环, 策略无法通过数据反馈进行优化迭代. 如哈啰出行建立了维修标准作业模型, 优化零配件库存, 若在实施过程中, 未对维修过程与零配件库存进行数据记录, 则无法检验库存和维修效率的提升状况.

## 2.2 风险复杂网络特征分析

风险预警的传统方法, 一般是先建立树状风险体系, 然后进行单项评估, 最后再进行综合评估. 确定树状风险体系需遵循以下几个原则: 针对性、独立性、完备性、可测性、客观性和简明性等, 其中最核心的是独立性和完备性, 但这两个原则只能在系统静态或弱动态条件下才能做到, 一旦体系处于动态变化条件下, 将

难以达到预期的目标. 并且, 树状风险体系往往默认每个影响因素是独自对结果产生影响, 即二级影响因素只归属于对应的一级影响因素, 进而影响结果, 但传统的研究方法构建的指标体系不能揭示影响因素之间的相互关系, 导致不能全面分析影响因素对结果的影响. 此外, 王林尧等<sup>[45]</sup>等强调数字系统工程区别于传统系统工程的优势是模型驱动的风险分析, 即利用系统思维与数字工程原理, 量化不确定性, 以确定下一个最佳行动方案, 不仅是识别风险, 而是控制风险. 因此, 提出三个转变: 第一, 将整体风险从“简单和”转为“涌现和”, 对于复杂系统而言, 系统属性不能被简单分解, 整体风险不等于局部风险简单求和, 而应该是网络化风险的整体“涌现”效果; 第二, 抛弃风险要素独立性假设, 将“风险树”转为“风险网”, 在数据风险复杂体系评估与预警中, 既然, 在“树状思维”下难以兼具独立性和完备性, 各风险之间实质是网状结构; 第三, 舍弃结果单一性假设, 用“结果云”代替单一值, 在基于复杂网络的数据风险预警模型中, 结果不是单一值, 而是各数据风险并非孤立存在, 彼此存在因果关系, 例如数据在储存、传输过程中变形或受损, 将导致智能化预测系统与实际存在误差, 下发错误策略. 为了找到数据风险间联系的规律, 构建出有向复杂网络  $G = (X, E)$ , 其中  $X$  是风险的集合, 作为网络的节点,  $E$  是风险之间的关联集合, 其因果关系用有向边来表示, 若  $X_i$  的出现或改变会引起  $X_j$  的改变, 则  $E_{ij} = 1$ , 从而建立一个二值有向网. 利用 UCINET 软件对复杂网络进行可视化, 如图 2 所示.

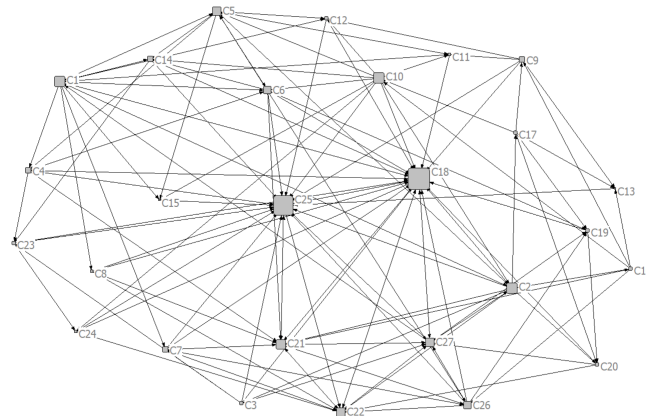


图 2 数据风险复杂网络图

Fig. 2 Complex network diagram of data risk

根据以上建立的复杂网络模型, 运用 UCINET 软件计算出风险有向复杂网络的特征值如表 2 所示.

表 2 数据风险有向复杂网络测度  
Table 2 Data risk oriented complex network measure

特征值名称	统计值
聚类系数	0.305
介数	6.199
平均路径长度	2.55
网络的关联度	0.491
互惠性	41.06 %

当网络的聚类系数为 0, 即网络中所有的点均为孤立节点; 若网络中每个节点都有边直接相连, 即全局耦合, 则聚类系数为 1. 可见数据风险之间存在一定的耦合作用, 但是并非全局耦合, 所以基于其耦合关系研究风险的传播路径是可行且有意义的. 任意两个节点之间的平均路径长度较短, 仅为 2.550. 与各节点直接相连的边的数量为各节点的度, 统计数据风险复杂网络各节点度, 如图 3 所示.

因此所构建的数据风险网络为 BA 无标度网络, 其具有以下两个特性, 符合数据风险网络的特征:

1) 增长特性: 网络的规模并非固定的, 而是不断扩大的. 数据风险随着技术进步、环境变化, 在数量上必然会不断增长.

2) 优先连接特性: 新加入的节点更倾向于与已经具有较高连接度的重要节点相连接. 数据风险之间的重要性显然会因为其所处数据流的位置而有所不同, 重要的节点往往随着技术与环境变化会衍生更多新数据风险.

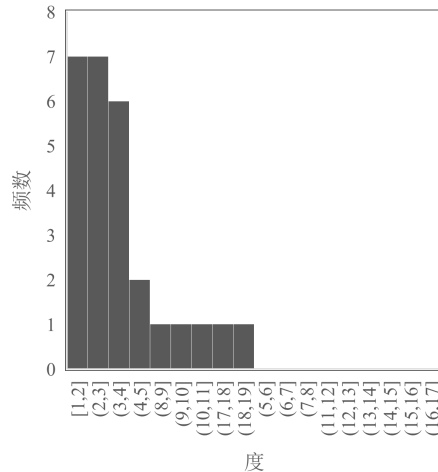


图 3 数据风险复杂网络度分布

Fig. 3 Distribution of complex network degree of data risk

从网络介数可见节点信息传播的控制能力较强, 一个风险出现, 后续存在 6 种以上风险传播的可能性, 若同时进行排查与控制, 无疑是低效且盲目的, 因此希望构建能够帮助企业有的放矢地对其中一种可能性进行预防与控制的模型. 对于有向网络而言, 若任何节点之间都能够建立联系, 则称为关联网. 可见数据风险网络的节点之间的可达性不强, 并非所有节点之间都能够建立联系, 说明风险传播是会终止的, 并不会传播至整个网络. 互惠性记为双向连接的边占所有边的比例, 可见各风险展出的因果性比双向关联性更强, 也就是说风险的传播是有方向的, 研究传播路径是有意义的.

### 3 基于相继故障理论的风险传播模型

在实际网络中, 少数节点或边发生内部故障或被蓄意攻击破坏时, 会通过节点间的耦合关系引起其他节点发生故障, 发生连锁效应, 最终导致相当一部分节点甚至整个网络的崩溃, 称为相继故障. 显然, 在数据风险网络中, 由于信息和数据本身的流动性与传递性, 符合复杂网络的相继故障理论. 风险传播模型根据复杂网络结构特征将数据风险间的传播特性进行量化, 满足风险预警机制动态地监测与诊断的功能要求, 企业由此可预知和控制数据风险. 由于数据风险复杂网络是 BA 无标度网络, 具有优先连接特性, 通过量化节点重要性得到各要素的风险传播能力是风险传播模型的基础. 利用复杂网络研究其他风险传播的研究往往仅考虑节点或边其中一个, 然而数据流上信息传递环环相扣, 不仅要考虑节点的风险传播能力, 连接节点的边的风险传播强度也是数据风险传播的重要影响因子, 需要将两者紧密结合.

#### 3.1 节点风险传播能力

每个数据风险的影响力显然是不同的, 若用节点所连接节点数量来评估节点的重要性, 考虑维度相对单一. 目前, 对节点重要性的研究受到广泛关注, 研究学者提出多种评估指标<sup>[46]</sup>, 如度中心性、介数中心性与特征向量中心性等, 但以上指标主要针对无向复杂网络. 而数据风险复杂网络是一个具有增长性的有向网络, 因此, 选择使用 PageRank 算法<sup>[47]</sup>计算风险的重要性, 其算法思想是从网络的拓扑结构出发, 在度量邻居节点网络结构与重要性均纳入算法中. PageRank 算法通过迭代求解, 对于一个包含  $n$  个节点构成的网络, 初始时, 给定网络中每个节点一个初始的 PageRank 值  $\theta_i(0)$ , 使得  $\sum_{i=1}^n \theta_i(0) = 1$ . 因此, 令  $\theta_i(0) = \frac{1}{n}$  后进行

迭代, 各节点将当前的  $\theta$  值平分给它所指向的节点, 即节点在第  $k$  步迭代后的  $\theta$  为

$$\theta_i(k) = s \sum_{j=1}^n a_{ij} \frac{\theta_j(k-1)}{k_j^o} + \frac{(1-s)}{n}, \quad (1)$$

其中  $a_{ij}$  为网络邻接矩阵  $A$  的元素, 若存在从节点  $j$  指向节点  $i$  的有向边, 则  $a_{ij} = 1$ , 否则  $a_{ij} = 0$ .  $k_j^o$  为节点  $j$  的出度.

各节点得到稳定的  $\theta$  值停止迭代, 即  $\theta_i(k) = \theta_i(k-1)$ , 此时可以得到网络中每个节点的  $\theta$  值,  $\theta$  值越高, 节点越重要. 在迭代过程中, 始终有  $\sum_{i=1}^n \theta_i(k) = 1$ . 若存在一个只有入度没有出度的节点, 它将不断吸收  $\theta$  值. 为了解决此问题, 将指向节点的  $\theta$  值之和分成两部分, 先将其用阻尼因子  $s \in (0, 1)$  进行缩减, 再把  $1-s$  的部分平均分给每个节点, 在本模型中意味着某一风险节点崩溃后风险传播至下一风险节点的可能性, 选取  $s = 0.5$ , 完成算法的修正.

计算出数据风险节点  $i$  的  $\theta_i$  值, 定义风险的风险传播能力  $S_i$  为  $\theta_i$  值归一化的结果, 表示风险出现或对应风险事故发生时, 对其它风险产生影响的能力, 即

$$S_i = \theta_i \left( \sum_{i=1}^n \theta_i \right)^{-1}. \quad (2)$$

### 3.2 网络风险传播强度

数据风险的出现并不意味着数据风险的故障, 在实际中往往对风险具有一定的包容性, 因此选用负荷-容量模型<sup>[48]</sup>行建模, 当数据风险负荷超过一定阈值时, 才认为发生数据风险事故. 在数据风险复杂网络中, 数据风险节点在初始状态都承担着一定的负荷. 每个数据风险节点的初始负荷均小于其最大负荷, 整个网络处于稳定状态, 当对其中某节点发生故障, 即发生该节点代表的风险事故, 该节点的负荷就会分配到其相邻节点上, 当相邻节点收到所分配的负荷后再加上自身初始负荷, 若负荷总和超出其能承受最大负荷时便会导致该节点也发生故障, 从而导致负荷的再次重新分配从而引发数据风险网络中的相继故障发生. 而这个过程, 风险是通过数据风险节点之间的有向边上根据一定策略传播的, 风险的传播存在优先级风险的传播存在优先级<sup>[48]</sup>.

传统模型中通过节点自身的中心度  $k_i$  以及其指向的相邻节点的中心度  $k_m$  计算得到每个节点的初始负荷  $L_i$ . 然而度中心度仅考虑节点连接其它节点的数量, 维度过于单一, 因此选取上文定义的风险传播能力  $S$  作为初始负荷的自变量, 即

$$L_i = \left( S_i \left( 1 + \sum_{m \in \tau_i} S_m - S_i \right) \right)^\alpha, \quad (3)$$

其中  $\tau_i$  是一个点集, 代表节点  $i$  所指向的相邻点集合.  $\alpha$  为调节负荷大小的控制参数, 此处取  $\alpha = 1$ .

节点  $i$  故障崩溃后, 负荷按照一定策略分配到其相邻节点  $j$  上, 节点  $i$  指向的相邻节点  $j$  将接受负荷  $\Delta L_j$ .

$$\Delta L_j = L_i \Pi_m = L_i \frac{L_j}{\sum_{m \in \tau_j} L_m} = \frac{L_i \left( S_j \left( 1 + \sum_{m \in \tau_j} S_m \right) - S_j \right)^\alpha}{\sum_{m \in \tau_j} \left( S_m \left( 1 + \sum_{h \in \tau_j} S_h \right) - S_m \right)^\alpha}. \quad (4)$$

由此可知,  $\Delta L_j$  结合了风险传播能力  $S$  以及风险网络对风险的包容性表示风险复杂网络中边的强度, 通过定义在风险节点发生故障后, 各条有向边所传播的风险负荷量  $\Delta L_j$  来表征风险传播强度  $I_{ij}$ , 风险传播强度越大, 优先级越高. 由上式可知, 风险传播强度  $I_{ij}$  和每个节点的初始负荷  $L_i$  以及递减的分配策略  $\Pi_m$  有关, 其中  $\Pi_m$  是节点  $j$  的初始负荷在  $i$  指向的所有节点负荷之和的占比, 如果节点  $j$  在和与节点  $i$  相邻的所有节点  $\tau_i$  上, 初始负荷占比越多, 所接受的负荷越多, 符合复杂网络优先连接特性, 这种分配策略

实际上会降低数据风险复杂网络的鲁棒性,可以使得风险预警模型在较为苛刻的条件下得到最高风险传播路径;同时,初始负荷  $L_i$  越大,在故障时将有更多的风险负荷传播给相邻的节点,边所传播的风险负荷量  $\Delta L_j$  相应增加,而根据(3)可知,初始负荷  $L_i$  与节点本身风险传播能力  $S_i$  以及跟相邻节点风险传播能力之和  $\sum_{m \in \tau_i} S_m$  的差值相关.因此每条有向边所传播的风险负荷不仅与所连接的数据风险节点本身风险传播能力  $S_i$  有关,而且基于节点的局域特征关注了数据风险节点相邻节点风险传播能力  $\sum_{m \in \tau_i} S_m$  对节点负荷的影响.

#### 4 基于 ACO 的数据风险预警模型

对于数据风险预警,传统方法会提供一个“单一值”结果,企业通过评估自身现阶段整体风险是否超过阈值作为调整风险管理策略的依据.但希望结合复杂网络相继故障理论为创业企业提供一个有效避免数据风险“涌现”的明确风险控制方向,因此所构建的数据风险预警模型即对最高风险传播路径的搜索,风险节点崩溃后将自身的负荷传播至相邻节点,每个风险节点指向下一个节点的路径并非单一的,本质上是 TSP 问题,即对所有可能的遍历且只经过一次的路径进行排序选出最短路径.在初始节点确定情况下,容易证明该问题是 NP 难问题,运用穷举法计算复杂度过高,为  $O(2^n n^2)$  且容易陷入局部最优,所得到的最高风险传播路径并非全局最优;蚁群算法(Ant Colony Optimization, ACO)是一种群智能算法,通过蚂蚁在数据风险复杂网络各路径上释放信息素,后续的蚂蚁将会选择积累信息素较多的路径,最终信息素浓度最高的路径就是蚁群在多次试验中发现的最优路径,每只蚂蚁同一节点只能途经一次,与相继故障理论的规律一致,启发式算法可增加找到全局最优解的概率.

$\tau_{ij}$  代表某一风险节点在  $t$  时刻的信息素浓度,  $\rho$  为信息素挥发因子,用于调节信息素的挥发比例,目的是避免残留信息素过多而淹没启发信息.蚂蚁每次经过该节点时,将会释放信息素  $\Delta\tau_{ij}$ ,因此,该节点下一时刻的信息素浓度等于残留信息素与新增信息素浓度之和,即

$$\tau_{ij} = (1 - \rho) \tau_{ij} + \Delta\tau_{ij}, \quad (5)$$

其中  $\Delta\tau_{ij} = \sum_{k=1}^m \Delta\tau_{ij}^k$ ,  $m$  为蚂蚁数量.

当数据风险节点  $i$  崩溃时,即蚂蚁  $k = 1, 2, \dots, m$  在起始各节点  $i$  上面临多条可选路径时,此时结合启发信息  $\eta_{ij}^\beta(t)$  与信息素浓度决定其选择下一个转移节点的概率.启发信息  $\eta_{ij}^\beta(t)$  在传统 ACO 算法中设定为节点间距离的倒数.选定风险传播强度  $I_{ij}$  作为启发信息,即  $\eta_{ij}^\beta(t) = I_{ij}$ ,如上文所述,风险在通过网络中有向边的传播是存在优先级的,  $I_{ij}$ ,而在数据风险网络中,风险传播强度  $I_{ij}$  表示数据风险节点在故障时的风险外溢,即既考虑数据风险节点本身风险传播能力  $S_i$  节点强度,也考虑关注了边传播强度相邻节点的局部结构,能够完整描述数据风险,网络的初始特征,  $I_{ij}$  作为蚂蚁收到的启发信息  $\eta_{ij}^\beta(t)$  越大,蚂蚁选择该路径的概率  $P_{ij}^k$  越大,因此选定  $I_{ij}$  作为启发信息,即  $\eta_{ij}^\beta(t) = I_{ij}$ .定义  $t$  时刻,第  $k$  只蚂蚁在数据风险节点  $i$  选择前往数据风险节点  $j$  的概率  $P_{ij}^k$ ,决定数据风险节点  $i$  故障后蚂蚁将风险传播至节点  $j$  的可能性,

$$P_{ij}^k = \begin{cases} \tau_{ij}^\alpha(t) I_{ij}(t) \left( \sum_{s \in N_k} \tau_{is}^\alpha(t) I_{is}(t) \right)^{-1}, & j \in N_k \\ 0, & \text{其它,} \end{cases} \quad (6)$$

其中  $N_k$  为蚂蚁  $k$  下一步被允许访问的数据风险节点集合,在数据风险复杂网络中  $i$  所指向的节点集合.

#### 5 案例应用分析与结论

ofo 单车和摩拜单车在同一时期兴起, ofo 在 2019 年开始濒临破产,摩拜在 2018 年 4 月被美团收购,两者的商业模式基本一致,但数据运营理念却在早期便展示出诸多差异,因此结合共享单车失败案例和成功

案例研究两者对于数据风险管理的差异, 讨论基于复杂网络的互联网创业企业数据风险管理的实现机制和实践价值, 回答数据风险事故发生后企业如何应对这一问题。

首先根据式(1)和式(2)计算出风险传播能力  $S$  如表 3 所示. 可见 C18 (数据指标体系化程度低, 忽视指标间内在逻辑) 与 C2 (应用软件不成熟, 无法对线下参与者的行为量化) 的风险传播能力最高. 在实际中, 应用软件已经成为互联网创业的标配, 应用软件的产品设计直接决定了对商品或服务的标准化程度, 成熟的软件产品设计能够将线下多种多样的非标准行为和服务进行标签化、结构化, 为后续的数据价值再生打下坚实的基础. 而数据指标体系反过来也会对智能硬件和软件的设计起推进作用, 是形成数据闭环的关键。

表 3 数据风险节点风险传播能力  $S$   
Table 3 Risk transmission capability of data risk nodes  $S$

风险节点	$S$	风险节点	$S$
C1	0.047 1	C15	0.020 2
C2	0.052 7	C16	0.023 1
C3	0.031 6	C17	0.019 1
C4	0.035 8	C18	0.080 3
C5	0.037 5	C19	0.025 8
C6	0.042 9	C20	0.027 1
C7	0.036 9	C21	0.050 7
C8	0.029 6	C22	0.046 4
C9	0.026 2	C23	0.028 1
C10	0.046 3	C24	0.027 2
C11	0.023 7	C25	0.072 2
C12	0.030 9	C26	0.041 9
C13	0.014 0	C27	0.047 4
C14	0.035 3		

计算出各个节点的传播能力  $S$  后, 根据式(3)和式(4)可算出每个节点的初始负荷  $L_i$  以及节点故障后向邻近节点传播的负荷  $\Delta L_i$  后作为启发信息集合利用蚁群算法模拟风险网络相继故障, 设置蚁群算法参数如表 4.

表 4 蚁群算法参数  
Table 4 Ant colony algorithm parameters

$\alpha$	$\beta$	$\rho$	$Q$	$m$	迭代次数
1	3	0.1	0.5	54	50

2014 年到 2017 年, ofo 早期在高校内试运行, 获得市场热烈的回应. 为了快速占领市场, ofo 早期采用造价低廉的物理锁, 车身没有 GPS, 用户换车必须打散密码盘, 每辆车的开锁密码不更新, 用户可以跳过应用软件直接开锁, 导致车况与定位失控(C1). 同期, 摩拜为了减少后期维修运营成本采用电子锁, 用新颖的外观设计增加用户对产品的认同感, 但受技术限制, 定位漂移、开锁成功率低、开锁时间长和耗电等问题存在(C1). 对于此风险, 两者表现出不同的风险管理态度, 因此选取 C1 作为识别出来具有破坏性的数据风险, 通过蚁群算法模拟最高风险传播路径, 得到风险传播路径如图 4 所示。

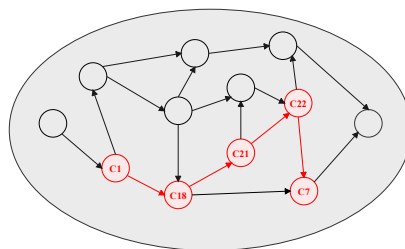


图 4 风险最高传播路径

Fig. 4 Highest risk propagation path

数据风险并不是一触即发,而是一个缓慢累积的过程.因此,在风险事故发生之前,企业可以从已显著的风险因素中感知到未来一段时间内可能发生的风险事故,提前采取相应的风险控制措施,从而降低自身爆发风险的可能性.这种企业对于未来的不确定性,可能发生的对自身产生不利影响事件的察觉,称为风险感知<sup>[51]</sup>(risk perception).“预知数据风险”本质上是提高对可能即将发生的风险感知,运用基于复杂网络的互联网创业企业数据风险预警模型,能够帮助企业提高风险感知能力以降低数据风险事故发生的可能性.因此,从风险感知的角度,可以将数据风险分为两种状态:已被感知风险(Y)与未被感知风险(N).

因此可以得到风险间最高传播路径 C1-C18-C21-C22-C7,即在 C1 发生时, C18, C21, C22 和 C7 这 4 个风险最大概率将会随之发生.因此,在风险控制时需要前置地对这四个风险进行评估与控制.以下结合 ofo 单车和摩拜单车两个对比案例来阐述在不同的风险感知下,企业采用的不同风险管理方式以及数据风险事故发生的不同结果,从而回答“企业如何预知和控制数据风险”这一问题.

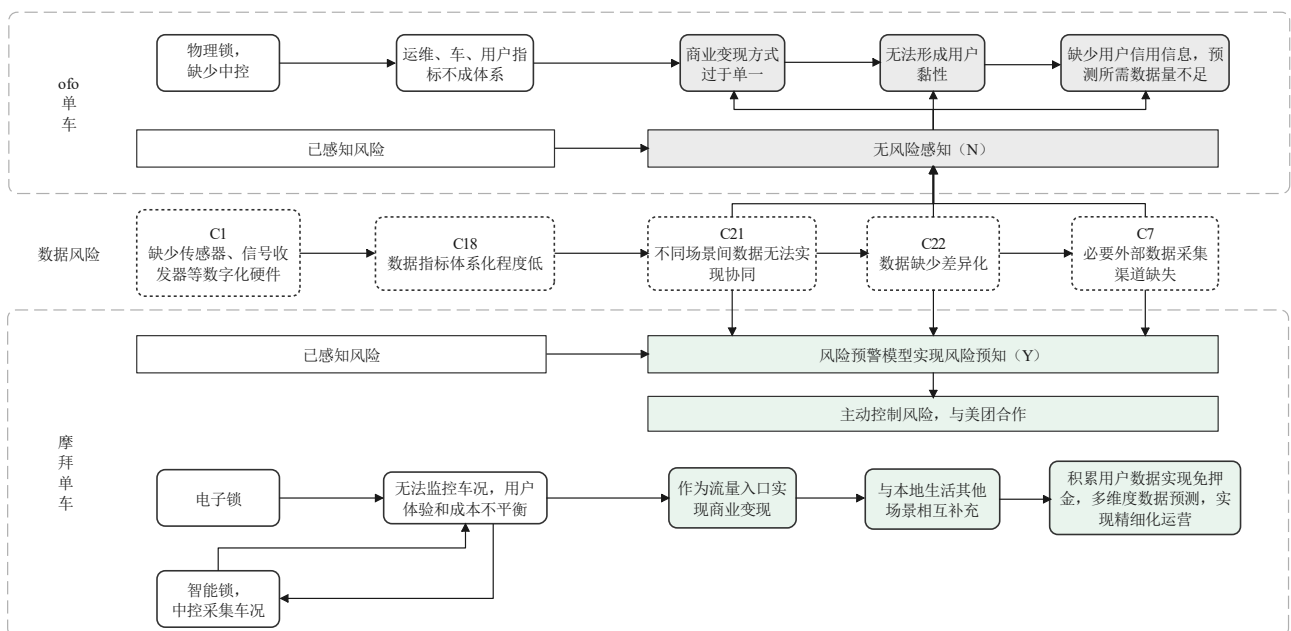


图5 共享单车实现数据风险预知与控制

Fig. 5 Realization path of shared bike data risk warning

在共享单车发展初期, ofo 和摩拜单车均已感知数据风险 C1 和 C18, 早期 ofo 采用物理锁(C1)导致用户、运维、单车三个数据产生端口的时空数据均不完整且无法成体系, 忽视内在逻辑测算收入和毛利. 根据风险投资机构和高管公开谈话, ofo 用每辆车日均被骑行次数(车日均周转)乘订单收入来测算每日收入, 按照此逻辑 ofo 在高校外推行也能获得正向收益, 但 ofo 忽视了数据指标之间的内在逻辑在场景转换时也发生变化, 在校园封闭场景下车日均周转远高于外部开放场景, 同时运营面积增大导致运营成本和车辆丢失损失的概率迅速增大, 最终导致入不敷出的低效运营与大量资产损坏损失(C18); 而摩拜投放的第一批车辆采用电子锁, 但受技术限制和 ofo 同样面临定位漂移、开锁成功率低、开锁时间长和耗电等问题(C1). 在提升用户用车体验同时缺少对成本类指标的考虑, 第一代产品的成本高达 3 000 元/台, 车身总量达 25 kg, 导致调度和全国投调难度极大增加, 在全国市场布局远远落后与 ofo(C18). 在感知相同数据风险后, 两者对最高风险传播路径上的数据风险在风险感知上存在差异, ofo 单车在对 C21, C22 和 C7 无风险感知的情况下, 仅控制对数据风险 C1 与 C18, 将物理锁逐步替换为智能锁, 逐步搭建数字化运营系统, 但未对由 C1, C18 引发的数据风险进行防御控制, 因此后续 ofo 除了用户的骑行数据外, 没有用户其他任何类型的数据, 因此需要通过收取押金保证用户用车行为的规范性; 商业模式本质上为简单的分时租赁, 变现模式和为用户提供的场景过于单一(C21). 以上风险对用户造成的伤害, 最终难以形成用户黏性, 在失去价格优势时便失去大批用户(C22). 此时, 企业所积累用车热点、用车高频时间、用户信用信息等必要的运营数据随着用户量减

少数数据量与可分析数据维度也在减少; 在竞争者免押金时 ofo 由于信用记录数据的缺失, 仍然需要收取押金, 最终导致数字化竞争优势逐步消失(C7); 而摩拜单车能够利用风险预警模型实现风险预知, 在感知数据风险 C21, C22 和 C7 的情况下, 在 2017 年 8 月, 摩拜正式推行第一代智能锁, 为后续精细化运营打下基础; 2018 年 4 月, 摩拜被美团收购, 管理团队留任. 场景过于单一的问题在被美团收购得到解决, 摩拜作为重要流量入口能够与其它业务产生正向协同作用, 美团能够通过原有的外卖与团购业务获取丰富的热门商区与生活区信息, 作为骑行用车热点的补充, 同时骑行场景能够进一步完善美团在本地生活的布局, 为其他业务引入更多流量完成变现(C21). 在为用户提供骑行服务的同时, 能够协同其他业务为用户提供完整的本地生活场景, 用户需求能够连贯且精准被满足, 更容易形成使用黏性(C22). 在此基础上, 摩拜通过竞争优势与平台其他业务积累的大量用户数据, 能够为用户提供免押金服务, 并且在用车时空预测上减少对外部数据的依赖(C7), 形成数据的良性自闭环. 通过案例对比可知, 相比风险感知的缺失, 在风险预知的情况下通过前置对最高风险传播路径上的数据风险进行控制, 能够有效避免数据风险事故的发生.

通过两个案例分析, 证明基于复杂网络的互联网创业企业数据风险预警模型和管理机制符合实际, 具有可行性和实践意义. 总结创业企业数据风险管理机制(见图 6), 总结了创业企业如何控制数据风险以及应对数据风险事故, 依据数据风险地图和数据风险体系进行风险识别, 结合数据风险传播能力, 利用基于蚁群算法的风险预警模型进行风险分析, 而不仅仅对已发生的数据风险做“亡羊补牢”式的改进, 更要针对性地对风险最高传播路径上的数据风险“未雨绸缪”地强化风险感知, 通过优化风险最高传播路径上的规章制度、管理流程、组织架构、指标设置、产品工具和日常运营做到防患于未然, 结合数据风险复杂网络对数据风险实现动态监控, 防微杜渐.

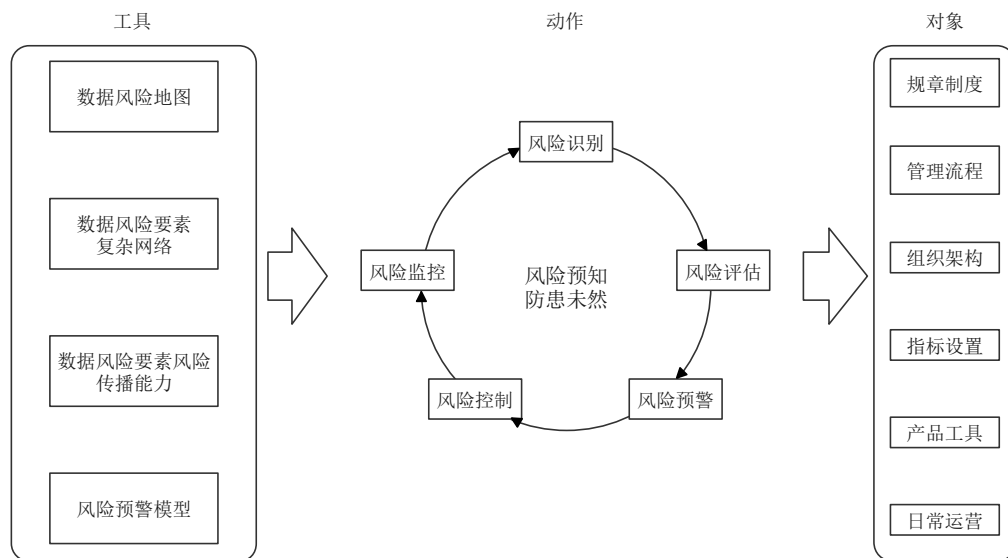


图 6 创业企业数据风险管理机制

Fig. 6 Data risk management mechanism of entrepreneurial enterprises

### 6 拓展

在前面的模型的 PageRank 算法中, 假设了阻尼因子  $s = 0.5$ , 在这一节放宽这一假设, 观察  $s$  对风险风险传播能力  $S$  和风险传播强度  $I$  的分布. 不改变其它条件, 分别计算出  $s = 0.2, s = 0.3, s = 0.4, s = 0.5, s = 0.6, s = 0.7, s = 0.8, s = 0.9$  时各个风险节点的  $S$  和  $I$ , 分布见图 7 与图 8. 可见随着阻尼因子  $s$  的变化, 各节点的风险传播能力  $S$  和风险传播强度  $I$  基本保持不变, 不影响风险预警模型的结果.

观察图 9 可以发现, 随着  $s$  的增大, 各节点  $S$  的方差  $\sigma^2(S)$  和  $I$  的方差  $\sigma^2(I)$  增大, 即各节点往下传播的概率越大, 各节点的风险传播能力  $S$  和风险传播强度  $I$  的分化更明显, 并且呈现非线性关系, 这一结果符合直觉.

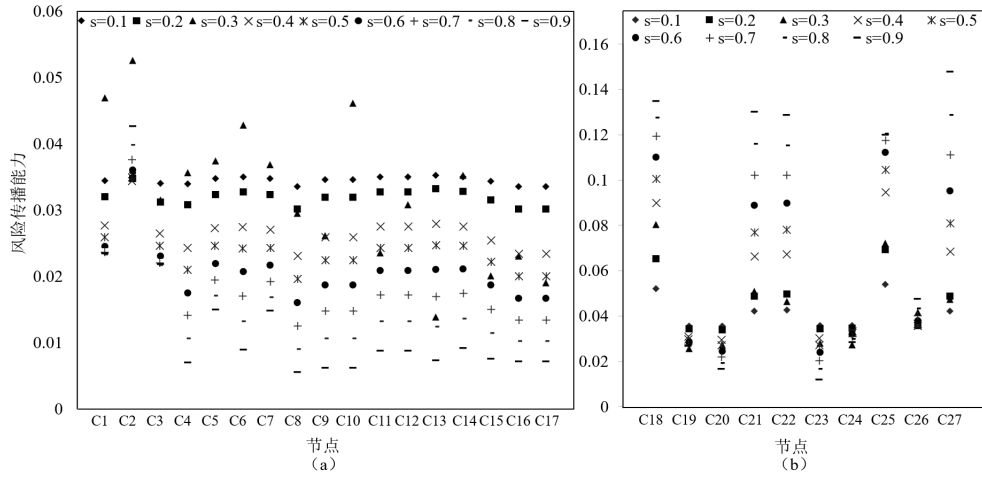


图7 不同阻尼因子  $s$  下各节点的风险传播能力  $S$

Fig. 7 Risk transmission capability  $S$  of each node under different damping factors  $s$

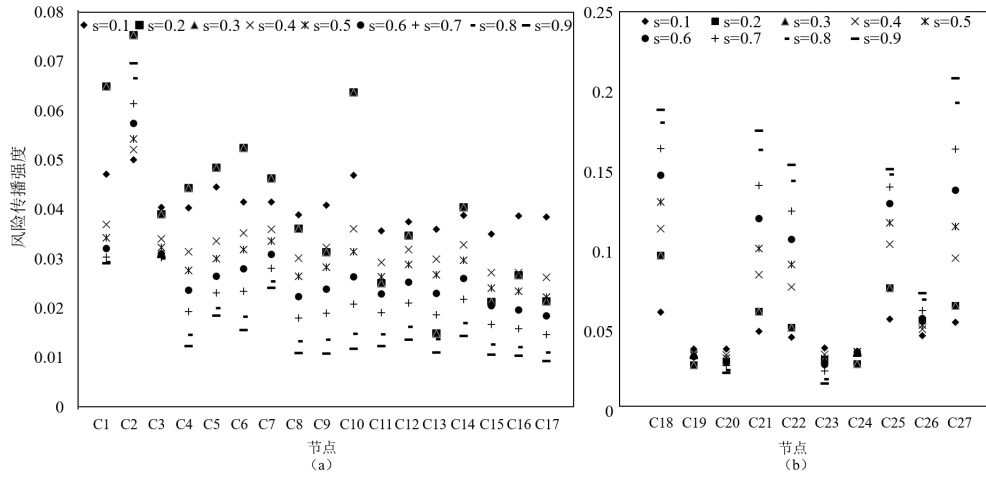


图8 不同阻尼因子  $s$  下各节点的风险传播强度  $I$

Fig. 8 Risk propagation intensity  $I$  of each node with different damping factors  $s$

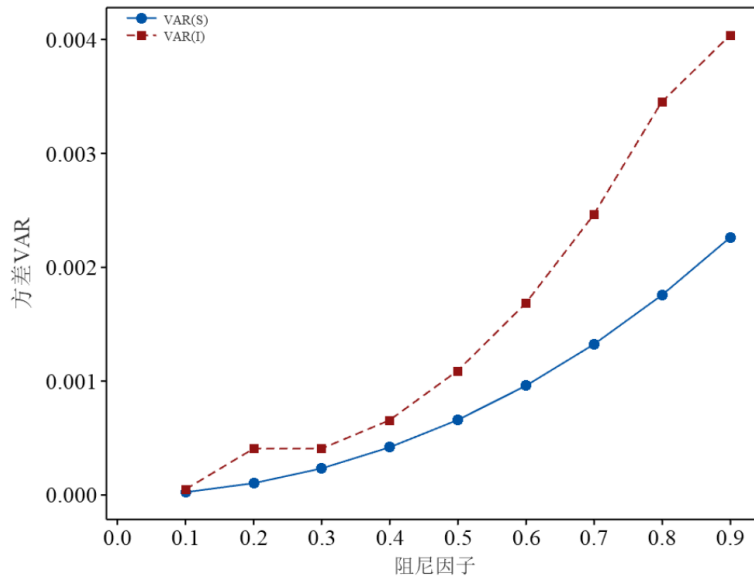


图9  $S$  和  $I$  的方差随阻尼因子  $s$  的变化

Fig. 9 Variation of variance of  $S$  and  $I$  with damping factor  $s$

## 7 结束语

选取共享单车这一典型数字化背景下互联网创业代表为例,对数据风险进行了定义与内涵解释,基于数据流构建了互联网企业数据风险体系,在信息通信技术创新与创业管理的交叉领域上丰富数据风险的内涵,考虑信息的流动性与数据风险的耦合性,将数据风险根据因果关系构建出有向复杂网络,发现数据风险网络符合 BA 无标度网络的特性.根据增长性和优先连接特性,选取 PageRank 算法来衡量风险的传播能力  $S$ .在此基础上,利用负荷-容量模型对复杂网络的相继故障进行探索,将故障后节点负荷的转移作为蚁群算法的启发信息,利用蚁群算法对风险复杂网络的相继故障进行模拟,通过得到最高风险传播路径搭建风险预警模型.最后,通过对共享单车风险管理机制实现路径的分析说明数据风险预警机制的可行性和实践价值.互联网创业企业能够利用数据风险体系与风险传播路径模型,对风险进行识别并且结合风险传播路径对应加强风险预防措施,有助于企业搭建有效完善的数据风险预警系统,提高风险消除效率,在数据化背景下形成良性的数据闭环.

本研究工作还需要对数据风险传播模型进一步细化,包括风险要素之间相互影响的强度、阻尼因子的设定.数据风险要素随时代发展会发生变化,如何客观地识别更新数据风险要素将会是未来一个重要的研究方向.

### 参考文献:

- [1] Sussan F, Acs Z J. The digital entrepreneurial ecosystem. *Small Business Economics*, 2017, 49(5): 1–19.
- [2] 陈晓红, 蔡 莉, 王重鸣, 等. 创新驱动的重大创业理论与关键科学问题. *中国科学基金*, 2020, 34(2): 108–116.  
Chen X H, Cai L, Wang C M, et al. An empirical study of the relationship between entrepreneurship and innovation. *Science Foundation of China*, 2020, 34(2): 108–116. (in Chinese)
- [3] Steininger D M. Linking information systems and entrepreneurship: A review and agenda for IT-associated and digital entrepreneurship research. *Information Systems Journal*, 2019, 29(2): 363–407.
- [4] Nambisan S. Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory & Practice*, 2017, 32(1): 18–34.
- [5] He X, Koveos P. Digital entrepreneurship solution to rural poverty: Theory, practice and policy implications. *Journal of Developmental Entrepreneurship*, 2019, 11: 13–19.
- [6] Song A K. The Digital entrepreneurial ecosystem: A critique and reconfiguration. *Small Business Economics*, 2019, 2019, 53(3): 569–590.
- [7] Hu H, Huang T, Zeng Q, et al. The role of institutional entrepreneurship in building digital ecosystem: A case study of Red Collar Group (RCG). *International Journal of Information Management*, 2016, 36(3): 496–499.
- [8] Martinez D A, Martin L, Marlow S. Emancipation through digital entrepreneurship: A critical realist analysis. *Organization*, 2018, 25(5): 585–560.
- [9] Hansen B. The digital revolution: Digital entrepreneurship and transformation in Beijing. *Small Enterprise Research*, 2019, 26(1): 36–54.
- [10] Hair N, Wetsch L R, Hull C E, et al. Market orientation in digital entrepreneurship: Advantages and challenges in a Web 2.0 networked world. *International Journal of Innovation and Technology Management*, 2013, 12: 209–213.
- [11] Orlandi L B, Zardini A, Rossignoli C. Highway to hell: Cultural propensity and digital infrastructure gap as recipe to entrepreneurial death. *Journal of Business Research*, 2021, 123: 188–195.
- [12] Coad A. Death is not a success: Reflections on business exit. *International Small Business Journal*, 2012, 32(7): 721–732.
- [13] 侯合银. 高新技术创业企业风险的系统分析: 辨识与规避. *科技管理研究*, 2008, 28(10): 4–16.  
Hou H Y. System analysis of high-tech entrepreneurial risk: Identification and avoidance. *Science and Technology Management Research*, 2008, 28(10): 4–16. (in Chinese)
- [14] Amit R, Han X. Value creation through novel resource configurations in a digitally enabled world. *Strategic Entrepreneurship Journal*, 2017, 11(3): 228–242.

- [15] 张新香, 胡立君. 商业模式动态演化机制: 基于互联网业的多案例内容分析. 科研管理, 2018, 39(3): 110–121.  
Zhang X X, Hu L J. Dynamic evolution of business models: A multi-case content analysis based on Internet Industry. Research Management, 2018, 39(3): 110–121. (in Chinese)
- [16] 邢小强, 周平录, 张竹, 等. 数字技术、BOP商业模式创新与包容性市场构建. 管理世界, 2019, 35(12): 122–142.  
Xing X Q, Zhou P L, Zhang Z, et al. Digital technology, BOP business model innovation and inclusive market construction. Management World, 2019, 35(12): 122–142. (in Chinese)
- [17] Ansari S S, Garud R, Kumaraswamy A. The disruptor's dilemma: TiVo and the U.S. television ecosystem. Strategic Management Journal, 2016, 37(9): 1829–1853.
- [18] 赵强, 单炜. 大数据政府创新: 基于数据流的公共价值创造. 中国科技论坛, 2014(12): 23–27.  
Zhao Q, Shan W. Big data government innovation: Public value creation based on data flow. China Science and Technology Forum, 2014(12): 23–27. (in Chinese)
- [19] Von B F, Davidsson P, Recker J. Digital technologies as external enablers of new venture creation in the IT hardware sector. Entrepreneurship Theory and Practice, 2018, 42(1): 47–69.
- [20] Raimbault J. For a cautious use of big data and computation. Royal Geographical Society, 2016, 12: 48–58.
- [21] Elia G, Margherita A, Passiante G. Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. Technological Forecasting and Social Change, 2020, 150: 119–126.
- [22] 黄煜. 企业数据风险处置中的方法研究. 信息通信, 2015, 4: 114–115.  
Huang Y. Research on methods of enterprise data risk management. Information Communication, 2015, 4: 114–115. (in Chinese)
- [23] 桂畅旒, 李维杰, 穆琳. 网络社会数据风险识别与应对. 北京邮电大学学报: 社会科学版, 2018, 20(6): 9–13.  
Gui C N, Li W J, Mu L. Risk identification and response of online social data. Journal of Beijing University of Posts and Telecommunications: Social Sciences, 2018, 20(6): 9–13. (in Chinese)
- [24] 林奇富, 贺竞超. 大数据与公共生活: 冲击、风险及治理反思. 探索与争鸣, 2015, 7: 86–90.  
Lin Q F, He J C. Big data and public life: Impact, risk and governance reflection. Exploration and Debate 2016, 7: 86–90. (in Chinese)
- [25] Chen H, Chiang R H L, Storey V C. business intelligence and analytics: From big data to big impact. MIS Quarterly, 2012, 36: 1165–1188.
- [26] Frenkel K A. What you need to know about digital risk officers. Cio Insight, 2014, 23(6): 10–16.
- [27] Lee T. JD.com security intelligence and analytics: From big data to big impact // The Strata Data Conference in Singapore, Singapore: 2017, 12: 128–133.
- [28] Eisenhardt K M. Better stories and better constructs: The case for rigor and comparative logic. Academy of Management Review, 1991, 16(3): 620–627.
- [29] 周文辉, 周依芳, 任胜钢. 互联网环境下的创业决策, 价值共创与创业绩效. 管理学报, 2017, 14(8): 1105–1113.  
Zhou W H, Zhou Y F, Ren S G. Entrepreneurial decision-making, value co-creation and entrepreneurial performance under the Internet environment. Chinese Journal of Management, 2017, 14(8): 1105–1113. (in Chinese)
- [30] 刘伟, 张振国, 汪志波. 企业经营风险预警指标体系与预警系统模型设计. 科技进步与对策, 2005, 7: 42–44.  
Liu W, Zhang Z G, Wang Z B. Enterprise management risk early warning index system and early warning system model design. Science and Technology Progress and Countermeasures, 2005, 7: 42–44. (in Chinese)
- [31] 肖毅, 熊凯伦, 张希. 基于TEI@I方法论的企业财务风险预警模型研究. 管理评论, 2020, 32(7): 226–235.  
Xiao Y, Xiong K L, Zhang X. Research on early warning model of enterprise financial risk based on TEI@I methodology. Management Review, 2020, 32(7): 226–235. (in Chinese)
- [32] 刘慧, 杨乃定, 张延禄, 等. 考虑风险感知和项目关联的R&D网络风险传播模型. 系统工程学报, 2019, 34(3): 11–21.  
Liu H, Yang N D, Zhang Y L, et al. Risk propagation model of R&D network considering risk perception and project association. Journal of Systems Engineering, 2019, 34(3): 11–21. (in Chinese)
- [33] 刘凤, 吴件, 唐静. 差异性与动态性并存: 大学生创业失败学习内容多案例研究. 科技进步与对策, 2018, 445(9): 151–157.  
Liu F, Wu C, Tang J. Diversity and dynamic coexistence: A multi-case study of college students' entrepreneurial failure learning content. Science and Technology Progress and Countermeasures, 2018, 445(9): 151–157.
- [34] 吴田, 胡海青, 张丹, 等. 基于复杂网络的交叉性金融业务风险传染仿真. 系统工程, 2018, 36(1): 22–30.  
Wu T, Hu H Q, Zhang D, et al. Risk contagion simulation of cross-financial business based on complex network. Systems Engineering, 2018, 36(1): 22–30. (in Chinese)

- [35] 龙剑友, 凌毓秀, 谢 赤. 基于复杂网络理论的房地产行业信用风险传染研究. 系统工程学报, 2022, 37(3): 289–302.  
Long J Y, Ling Y X, Xie C. Research on credit risk contagion in real estate industry based on complex network theory. Journal of Systems Engineering, 2022, 37(3): 289–302. (in Chinese)
- [36] 伍志韬, 杜 伟, 刘蕾蕾, 等. 恶意攻击下的电力耦合网络风险传播模型研究. 电网技术, 2020, 44(6): 42–49.  
Wu Z T, Du W, Liu L L, et al. Research on risk propagation model of power coupling network under malicious attack. Power Grid Technology, 2020, 44(6): 42–49. (in Chinese)
- [37] 曾维希. 基于复杂网络模型构建人事测评指标体系. 心理学探新, 2013, 1: 4–9.  
Zeng W X. Building personnel evaluation index system based on complex network model. Psychological Exploration, 2013, 1: 4–9. (in Chinese)
- [38] 郭 斌. 京津冀科技协同创新绩效体系重构: 基于文献编码的复杂网络分析. 中央财经大学学报, 2016, 6: 87–96.  
Guo B. Reconstruction of Beijing-Tianjin-Hebei collaborative innovation performance system: Complex network analysis based on document coding. Journal of Central University of Finance and Economics, 2016, 6: 87–96. (in Chinese)
- [39] 胡晓峰, 张 昱, 李仁见, 等. 网络化体系能力评估问题. 系统工程理论与实践, 2015(5): 7–16.  
Hu X F, Zhang Y, Li R J, et al. Research on the capability evaluation of networked systems. Systems Engineering: Theory & Practice, 2015, 5: 7–16. (in Chinese)
- [40] 杨 玮. 企业级数据中心数据流处理方案设计. 中国科技信息, 2007, 15: 111–123.  
Yang W. Design of data flow processing scheme for enterprise-level data center. China Science and Technology Information, 2007, 15: 111–123. (in Chinese)
- [41] 郝雅立, 温志强. 共建共治共享: 大数据支持下共享单车智能化治理路径. 管理评论, 2019, 31(1): 251–256.  
Hao Y L, Wen Z Q. Co-construction, co-governance and sharing: Intelligent governance path of shared bikes supported by big data. Management Review, 2019, 31(1): 251–256. (in Chinese)
- [42] 鄢章华, 刘 蕾. 考虑服务水平与动态转移规律的共享单车投放策略研究. 中国管理科学, 2019, 17(9): 195–204.  
Yan Z H, Liu L. Research on sharing bicycle placement strategy considering service level and dynamic transfer law. Chinese Journal of Management Science, 2019, 17(9): 195–204. (in Chinese)
- [43] Maciuliene M, Skarzauskiene A. Evaluation of co-creation perspective in networked collaboration platforms. Journal of Business Research, 2016, 69(11): 4826–4830.
- [44] 种颖珊, 韩晓明. 基于随机森林与时空聚类的共享单车站点需求量预测. 科学技术与工程, 2018, 18(32): 94–99.  
Chong Y S, Han X M. Prediction of demand for bike-sharing stations based on random forest and spatial-temporal clustering. Science Technology and Engineering, 2018, 18(32): 94–99. (in Chinese)
- [45] 王林尧, 赵 滢, 张仁杰. 数字工程研究综述. 系统工程学报, 2023, 38(2): 265–274.  
Wang L Y, Zhao Y, Zhang R J. Digital engineering research overview. Journal of Systems Engineering, 2023, 38(2): 265–274. (in Chinese)
- [46] 任晓龙, 吕琳媛. 网络重要节点排序方法综述. 科学通报, 2014, 59: 1175–1197.  
Ren Xiaolong, Lu Linyuan. An overview of sorting important nodes in network. Chinese Science Bulletin, 2014, 59: 1175–1197. (in Chinese)
- [47] Page L. The PageRank citation ranking: Bringing order to the Web. The Web Conference, 1998, 39: 117–119.
- [48] Motter A E, Lai Y C. Cascade-based attacks on complex networks. Physical Review E: Statistical Nonlinear & Soft Matter Physics, 2002, 66(6): 65–102.
- [49] 王建伟, 荣莉莉. 面向相继故障的复杂网络上边袭击策略研究. 系统工程学报, 2011, 26(1): 8–23.  
Wang J W, Rong L L. Research on sequential fault-oriented top-down attack strategy for complex networks. Journal of Systems Engineering, 2011, 26(1): 8–23. (in Chinese)
- [50] Das T K, Bing S T. A risk perception model of alliance structuring. Journal of International Management, 2001, 7(1): 1–29.

### 作者简介:

苏越良(1969—), 男, 湖南平江人, 博士, 副教授, 研究方向: 创新与创业, 风险管理, 价值评估, Email: syxcg@scut.edu.cn;  
杜芷晴(1997—), 女, 广东佛山山人, 硕士生, 研究方向为创新与创业, 风险管理, Email: zhiqingdu@foxmail.com.