

doi:10.19920/j.cnki.jmsc.2024.09.009

个体借贷关系网络特征与违约研究^①

——来自互联网金融平台的证据

张卫国¹, 李华², 王超^{3*}

(1. 深圳大学管理学院, 深圳 518055; 2. 华南理工大学工商管理学院, 广州 510641;
3. 广东工业大学经济学院, 广州 510520)

摘要: 本研究应用复杂网络分析方法研究了互联网金融平台个体借贷关系网络特征与借款人违约问题。提出了融资平台个体借贷关系网络的构建方法, 构建了互联网借贷平台人人贷个体借贷关系网络, 刻画了该网络的整体特征以及违约借款人群体和不同信用等级借款人群体、纯投资者群体构成的子网络特征。揭示了借款人和投资者的行为选择规律。分析了借款人网络拓扑特征与借款信息的相关性, 构建了借款人违约与借款人网络拓扑特征、借款信息的三个回归模型。研究发现: 构造的网络拓扑特征对于个体借贷关系结构具有很强的解释力; 借款人节点的网络拓扑特征与借款人的违约存在显著相关关系, 且这种关系在不同属性借款人群体中也具有稳健性; 研究成果能够用于各类互联网金融平台借款人信用风险评估及违约预测。

关键词: 个体借贷关系网络; 网络拓扑特征; 信用违约; 复杂网络

中图分类号: F83 **文献标识码:** A **文章编号:** 1007-9807(2024)09-0145-14

0 引言

近年来, 互联网技术的发展推动互联网与各行各业的融合, 使得大量微小个体得以相互连接, 为多形态的经济活动产生提供了可能。随着金融业务的线上化, 大量互联网金融独角兽先后崛起, 促使互联网环境下的个体债务市场快速发展。这一方面使得民间借贷在互联网平台中介作用下实现借贷关系扁平化和信息透明化。另一方面, 互联网环境下的信息不对称加重了债务市场固有的信用风险。在传统借贷业务中, 违约预测和信用风险预警是学术界和实践界关注的重要课题, 但已有研究讨论的多是机构投资借贷主体的风险问题。随着互联网为个体债务市场提供的技术和数据支持不断增加, 基于微观个体行为特征的信用风险研究逐渐成了学术界的关注焦点。尽管互联网借

贷在我国本土化的过程中由于存在巨大风险被清退整治, 但历史存在的交易活动为研究互联网环境下个体借贷的信用风险问题提供了详尽的个体行为数据和个体特征信息。互联网借贷平台参与者身份可以是双重的(既可以是投资者也可以是借款人), 由此, 参与者的个体借贷关系构成一个有向复杂网络。为此, 本研究从互联网借贷微观个体构成的借贷关系网络视角探究网络拓扑特征和借款人违约之间的联系。

研究互联网借贷参与者个体行为, 从借款人角度主要包括分析借款成功影响因素和预测借款人违约两类。如 Lin 等^[1] 使用 Prosper 的借贷记录, 发现借款人的在线社交关系增加了融资成功的可能性并降低了借款的利率和违约率。而 Laura 和 Yuliya^[2] 发现出借人对借款人年龄和性别的感知会影响借款人能否成功借款。在预测借款人违

① 收稿日期: 2020-03-08; 修订日期: 2021-12-11。

基金项目: 国家自然科学基金-广东省联合基金重点资助项目(U1901223); 广东省自然科学基金研究团队资助项目(2017A030312001); 广东省自然科学基金面上资助项目(2023A1515012494)。

通讯作者: 王超(1991—), 男, 陕西商洛人, 博士, 讲师。Email: sharpe_wang@163.com

约方面,包括传统的逻辑回归和线性判别分析,以及机器学习方法已被学者广泛使用^[3-11]。同时,融合非结构化数据构建借款人信用风险评估模型也是一个活跃的研究方向。如融合元水平的电话使用记录^[12]、贷款描述文本信息^[13, 14]、以及社交媒体信息^[15]和平台的参与者评论^[16]等。此外,也有学者探索了复杂网络分析方法在借款人违约预测方面的应用。如 Giudici 等^[17]基于借款人的财务特征相似性网络分析了网络的拓扑特征在提升借款人信用风险评估方面的作用。从投资者角度,学者们主要探究了互联网借贷平台中投资者是否存在羊群效应,如 Lee 和 Lee^[18]、Zhang 和 Liu^[19]。也有学者关注投资者行为异质性^[20]和双向交易者的信息价值性^[21]。更多研究可参考冯博等^[22]提供的综述工作。

从平台角度,Wei 和 Zhang^[23]以互联网借贷平台和金融中介机构为网络节点构建复杂网络模型,通过仿真模拟分析了平台间直接风险传染的一般特征。Xiao 等^[24]提出了资金流动模型用来辨别借款人异常活动从而有效识别平台风险。余桃^[25]和 Chen 等^[26]则研究了互联网借贷平台中个体借贷关系网络的生成和动态演化。利用复杂网络分析互联网借贷平台风险方面,Li 等^[27]以网络拓扑特征作为解释变量,把与违约节点连线的数量占比作为被解释变量,预测了平台的系统性风险。Tong 和 Chen^[28]探究了无线网元数据环境下基于 Spark 和复杂网络分析的 P2P 网贷违约风险问题。Xu 等^[29]从复杂系统角度来分析互联网金融的系统性风险,认为风险是通过互联网金融的内部循环向外传播。

网络分析法是结构分析策略的重要方法之一。复杂网络已经被应用于金融经济等领域^[30, 31]。小世界网络的小世界性和高聚集系数、无标度网络的小世界性和节点度分布服从幂律分布的特性都与现实世界网络十分相似^[32]。刘景卿等^[33]研究发现全球价值链贸易网络具有“小世界现象”和“无标度特性”。宫晓莉和熊熊^[34]通过构建金融市场信息溢出复杂网络分析了金融系统内部的风险传染特征。复杂网络的网络特征也被用来代理特定的金融变量,如隋聪等^[35]利用银行间借贷网络的幂律分布特征度量银行间的流动性差异。周颖刚等^[36]构造了 G20 经济体货币汇率间相

互影响的动态网络,并利用网络重要性构造了人民币国际影响力指数。Zhang 等^[37]基于复杂网络和消费者决策机制研究了互联网金融产品的扩散过程。张奇等^[38]基于复杂网络分析方法刻画了电动汽车充电桩众筹融资市场上投资者决策过程中的羊群行为,研究了社交网络负面信息与融资人的信用风险之间的联系。游运等^[39]在预测 P2P 借贷项目违约预测中使用了平台社会网络关系,发现平台社会网络关系能够提升项目违约预测准确性。

本研究主要工作和贡献在于:现有相关研究多从借款人或投资者单一角度展开,而本研究从互联网借贷平台整体借款人和投资者借贷关系构成的有向网络角度出发,分析了个体借贷关系网络结构特征及个体行为特征,清晰刻画了借贷关系网络特征与借款人违约显著相关性,建立了借款人违约与借款人网络拓扑特征、借款信息的回归模型,为互联网融资平台借款人信用风险评估与违约预测提供了新方法。首先,提出了个体借贷关系网络构建方法,分析该借贷关系网络的网络密度、平均点度和直径以及节点间的平均距离、幂律指数等特征。其次,基于平台借贷数据,使用周期内所有投资记录构建个体借贷关系网络,并分析网络整体特征及违约借款人子网络、不同信用等级借款人子网络、不同信用等级借款人子网络与纯投资者子网络的特征,以揭示借款人和投资者的行为选择规律。再次,计算借款人节点入度中心度、近邻威望和中介中心度,并检验这些节点网络拓扑特征与信用等级、借款利率、借款金额等借款信息的相关性。然后,对低信用等级的借款人群体分别建立借款人违约与节点网络拓扑特征、借款信息的 Logistic 回归模型,并通过 Probit 回归验证了稳健性。最后,针对互联网融资平台借贷风险的有效监管提出管理对策和建议。

1 个体借贷关系网络构建及整体特征描述

1.1 借贷关系网络构建

以借贷平台参与者(投资者或借款人)为节点,以借贷关系作为连接边,可建立一个有向借贷关系网络 $G = (V, E)$, 这里 $V = \{v_i, i = 1, 2, \dots,$

$n\}$; $E = \{e_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, n\}$, n 是网络中节点个数, e_{ij} 是从节点 v_i 指向 v_j 的边。

为清晰说明网络结构,图1给出了互联网平台个体借贷关系网络的最简子图。

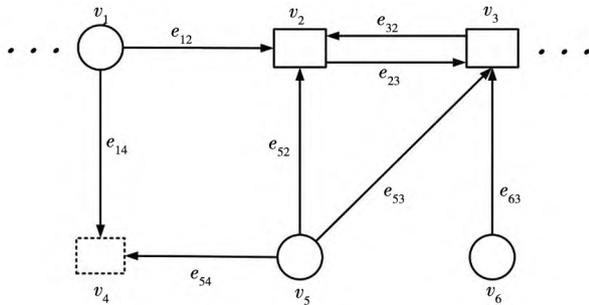


图1 个体借贷关系网络的最简子图

Fig.1 A simple graph of individual lending relationship network

图1所示的借贷关系网络中,集合 V 包含了三类节点:1)纯投资者,即在网络中只有出度的节点,图中用实线圆圈表示;2)纯借款人,即在网络中只有入度的节点,图中用虚线方框表示;3)中介参与者,即在网络中有出度也有入度的节点,图中用实线方框表示.集合 E 表示投资者与借款人之间的所有借贷关系.元素 e_{ij} 代表投资者 v_i 投资了借款人 v_j 发起的借款申请,代表投资选择关系,同时也包含了资金流动方向.该有向网络存在多重边但不包含环,即同一个投资者对同一个借款人发起的信用标投资了多次,代表了重复的投资选择关系,但不存在借款人对自己发起的标的进行投资的行为。

1.2 借贷关系网络的整体特征描述

网络密度和平均点度、网络直径和平均距离、节点度分布是复杂网络的整体特征.其中,网络密度刻画了网络连接的总体特征.节点度描述个体节点的连接情况,由于是无标度网络,则其度分布服从幂律分布^[27].网络直径和平均最短路径体现网络信息流的传递效率。

1)网络密度. $Density = |E| / n(n-1)$, E 为网络边集合, $|E|$ 表示网络中边的数量; $n(n-1)$ 是有向网络的最大可能边数.根据 Nier 等^[40]关于网络结构与金融系统稳定性的研究,在未超过一定阈值下,较低的网络连通性意味着较低的风险传染性。

2)平均点度. $avr_k = \frac{1}{n} \sum_{i=1}^n (k_{v_i-in} + k_{v_i-out})$,其

中节点 v_i 的入度 $k_{v_i-in} = \sum_{k \in V} 1_{e_{ki}} > 0, 1_{e_{ki}} > 0$ 表示存在从 v_k 指向 v_i 的边,节点 v_j 的出度 $k_{v_j-out} = \sum_{j \in V} 1_{e_{ij}} > 0, 1_{e_{ij}} > 0$ 表示存在从 v_i 指向 v_j 的边.一个节点的度等于其入度与出度之和.网络平均点度代表平均意义下一个参与者参与借贷活动的次数。

3)网络直径和平均距离.节点间的距离 $Dist_{ij}$ 定义为两个可达节点所有路径中的最短路径长度.有向网络中节点的可达性考虑了边的方向性,因此在不同方向上,同样两个节点的距离可能不同.网络直径为所有可达节点距离的最大值,即 $diameter = \text{Max}(Dist_{ij})$.网络直径刻画平台中最长的投资关系共经过多少位中介参与者的连接构成;网络平均距离表示平均意义下一个参与者与多少其他参与者有借贷联系。

4)幂律指数.网络中边的数量通常服从分布函数为 $P(x) = cx^{-\alpha}$ 的幂律分布^[41].其中 α 是幂律指数,描述了网络中特征变量的变化速度.幂律分布反映了网络通常情况下表现为少数节点与大量其他节点相连,而多数节点的连接边很少的特征.现实世界中的多数网络都满足幂律特征,这类网络被称为无标度网络^[42].节点的入度在投资关系中表示有多少投资者向该节点代表的借款人进行投资,对该节点来说,描述的是借款行为.节点的出度在投资关系中表示该节点所代表的投资者投资了多少个借款人发布的借款,对于该节点来说,描述的是投资行为。

1.3 人人贷平台借贷关系网络的整体特征描述

人人贷平台成立于2010年,是我国早期的网络借贷信息服务中介之一.经过多年稳健发展,成为我国互联网借贷行业的领军者.本研究选择2010年11月到2016年2月人人贷平台上所有信用标的的借贷记录构建借贷关系网络.所建立的网络共有71499个参与者节点,包括49666个(69.46%)纯投资者节点,20959个(29.32%)纯借款人节点和873个(1.22%)中介参与者节点.有775858条边,其中99140条为多重边.通过合并相同指向边(合并后的边权为原边权之和)对原始网络简化处理,最终得到包含676718条边的简化网络。

所构建的网络密度为 1.3×10^{-4} .网络平均点度约为19,表明在平均意义下,一个参与者参

与借贷活动大约 19 次. 网络直径为 10, 说明平台中最长的投资关系共经过 9 位中介参与者连接构成. 网络平均距离为 3.64, 表示在平均意义下, 1 个参与者只与平台中其他 2 位参与者有借贷联系. 这些结果表明该网络边的关系是稀疏的, 存在很多可以挖掘的潜在业务机会. 另外, 节点的入度和出度经过验证都符合幂律分布, 对应的幂律指数分别是 2.45 和 2.94. 从幂律指数来看, 节点出度的幂律指数比入度的幂律指数大, 说明平台交易中投资次数变化速度更快.

2 不同属性借款人群体及纯投资者群体的行为特征

将互联网平台个体借贷网络上的节点对应到其所代表参与者的属性特征, 可将参与者划分为违约借款人、无违约借款人、不同信用等级借款人和纯投资者. 将节点按照不同的属性特征进行分

类, 并将同属性特征的参与者节点及与其相连的边提取出来, 构成违约借款人子网络、无违约借款人子网络、不同信用等级借款人子网络和纯投资者子网络. 通过考察各个属性子网络的内部结构及群体间行为特征, 研究不同属性特征参与者的差异及联系.

2.1 违约借款人群体的行为特征

对于违约借款人子网络, 根据违约借款人群体占全部借款人比率, 可以得出网络借款违约问题是否突出. 根据违约节点之间边的连接情况, 可以判断违约是否为直接传染所导致的. 就人人贷平台个体借贷关系网络来看, 本研究提取了 4 131 个违约节点组成违约子网络, 具体情况如表 1 所示.

表 1 显示人人贷平台有违约借款人和无违约借款人分别占 5.77% 和 24.76%. 纯投资者在全部参与者中占多数 (69.46%), 为借款人总数的 2.27 倍. 有违约记录的借款人占全部借款人的 18.92%, 说明借款人违约问题比较突出.

表 1 节点分布

Table 1 Distribution of node

属性	频数	占参与者比率	占借款人比率
违约	4 131	5.77%	18.92%
无违约	17 702	24.76%	81.08%
纯投资者	49 666	69.46%	227.48%

在局部视角下, 将违约借款人节点和它们相连的边从网络中提取出来形成局部违约子网络, 来观察违约借款人之间的内部联系. 在互联网借贷平台的个体借贷网络中, 风险的传染通过点与点之间的资金链进行. 由于本研究提取的违约子网络只有一条边, 是从借款人 ID 为 81 316 的节点指向借款 ID 为 117 068 的节点, 对应的边权为 100. 如果违约节点之间没有边相连, 便不具备违约直接传染的基础. 因此, 该违约子网络结构表明人人贷平台在 2010 年 11 月至 2016 年 2 月期间借款人发生的贷款违约不是由其他借款人的违约传染所导致的.

2.2 不同信用等级借款人群体的行为特征

根据借款人的信用等级, 可以将借款人分为不同信用等级的借款人群体, 由此提取不同信用等级的借款人子网络. 根据各信用等级子网络中违约借款人占全部借款人比率, 便于了解各个信

用等级借款人群体违约情况是否突出. 也可以了解不同信用等级的借款人群体是否有显著差异, 进一步分析信用等级子网络特征, 清晰刻画不同信用等级子网络下借款人之间的投资行为选择.

具体就人人贷平台借贷关系网络来看, 将建立的借贷关系网络的节点对应到信用等级属性特征, 可以将借款人节点分为七种, 分别是信用等级为 AA、A、B、C、D、E 和 HR 对应的类别. 由于同一个借款人在一段时间内发起的借款行为可能是多次的, 而每一次借款都可能对其重新评级, 因此同一个借款人在不同时期可能有不同的信用等级. 在违约预测问题中, 把一个违约借款人错判为非违约借款人造成的损失远大于把一个非违约借款人错判为违约借款人造成的损失. 同时, 投资者进行投资决策时对有违约记录的借款人也更加谨慎和敏感. 在这种情况下, 取最低的信用等级与该借款人对应. 信用等级及违约分布见表 2.

表2 按借款人信用等级属性划分的借款人频率分布及违约分布

Table 2 Distribution of borrowers and defaults grouped by credit level

信用等级	频数	占借款人比率	违约率
AA	92	0.421%	0.005%
A	38	0.174%	0.000%
B	307	1.406%	0.009%
C	567	2.597%	0.032%
D	3 470	15.893%	1.287%
E	3 622	16.590%	1.296%
HR	13 737	62.919%	16.310%

表2显示人人贷平台上,信用等级为AA和

表3 信用等级子网络的网络特征

Table 3 Features of credit level sub-network

信用等级	节点数	边数	多重连线	密度	平均度
HR	13 737	3 368	654	0.000 017 85	0.49
E	3 622	182	11	0.000 013 88	0.10
D	3 470	728	84	0.000 060 48	0.42
C	567	585	182	0.001 822 87	2.06
B	307	202	75	0.002 150 26	1.32
A	38	6	1	0.004 267 43	0.32
AA	92	1 340	782	0.160 057 33	29.13

表3表明信用等级为AA级的借款人之间投资行为相对于其他信用等级借款人之间的投资行为有更紧密的“联系”。评级为HR和E的借款人内部借贷关系的网络密度明显小于其他的子网络,这类借款人的信用水平低,他们也不愿意对自己同级的其他借款人进行投资。

网络中的中介参与者是值得关注的,如果他们所投标的发生违约,将会影响他们的资产水平,从而可能影响自身的还款行为。因此,从盈利的角度来看,高信用等级且有比较活跃投资行为的借款人更值得平台关注。

2.3 不同信用等级借款人群体及纯投资者群体间的行为特征

为研究不同信用等级借款人群体及纯投资者群体间的行为特征,将纯投资者的信用属性标签对应为N。基于整体视角,将借贷关系网络中同一属性的节点看成一个节点,指向该节点的边的权重为原网络中指向该类节点的所有边的权重之和。如果边的起点和终点属于同一属性类别,则在

A的参与者占比很低,分别为0.42%和0.17%。从A到HR,信用等级越低的借款人在全部借款人中占比越高,特别是信用等级为HR的借款人在全部借款人中占比高达62.92%,表明绝大多数借款人是高风险借款人。同时,信用等级越低的借款人违约率越高,其中信用等级为HR的借款人违约率高达16.31%,表明平台绝大多数借款人信用等级过低而导致整体违约率过高的特征。

根据节点对应的借款人信用等级属性分别提取出7个信用等级子网络,观察同信用评级中参与者之间的投资选择。信用等级子网络的网络特征如表3所示。

新的网络中形成环。整体视角的网络中只有8个节点,除了标记为N的节点只有出度没有入度外,其他节点在该网络中都有一条出边、一条入边和环。通过整体视角,可以看到不同信用等级借款人之间的投资关系,由边的权重可直观了解到不同属性参与者在投资金额上的差异。

具体就人人贷平台借贷关系网络来看,表4列出了该网络下所有边的权重,表示不同信用等级借款人群体对其他信用等级借款人群体投资的金额及纯投资者群体对各信用等级借款人群体投资的金额。

信用等级为A的借款人投资金额相对低于其他中介参与者,但该类参与者在平台中获得的投资是最高的。由于信用等级为A的借款人群体违约率最低,这说明该平台信用等级为A的借款人更受欢迎也更可信。信用等级AA的借款人对处于任何一个信用等级的借款人投资的金额都是最高的。另外,AA级借款人群体的违约率只有0.005%,与高信用等级代表的高还款能力相匹

配,他们具有较强的投资能力. 只有投资行为的投资者群体对 HR 到 AA 各信用等级借款人群体

投资的金额基本保持不断增加,其中对于 A 级信用借款人投资额最大.

表 4 整体视角网络下所有边的权重(单位:万元)

Table 4 Weights of all edges of overall-view-network(Unit: 10 000 Yuan)

From \ To	HR	E	D	C	B	A	AA
HR	84	28	188	773	944	2 746	2 482
E	32	31	100	263	490	953	592
D	46	32	125	336	746	1 555	1 108
C	51	12	92	544	793	1 875	828
B	27	9	64	377	780	1 241	966
A	3	0	27	25	23	30	8
AA	403	113	610	3 056	6 868	12 520	12 053
N	21 350	27 251	47 561	137 214	172 712	407 588	225 924

3 借款人网络拓扑特征与借款信息相关性分析

通常参与者在平台上进行投资或者借款的行为时间或者周期依赖性是不同的. 在确定划分网络的时间周期时,首先,要考虑在多长时间形成的借贷关系会对借款人当前违约情况造成影响. 其次,当前借款人的违约是在其借款到期后才形成的状态,这个状态并不受其还款日之后的投融资行为的影响. 综合考虑借贷关系会对借款人当前违约情况造成影响需要周期积累性(纳入周期内更多的投资者投资行为的影响)以及投资者投资行为的时间频繁性因素,因此应该选取尽量短的时间周期.

具体就人人贷平台借贷关系网络来看,通过对 2010 年 11 月至 2016 年 1 月人人贷平台上所有信用标的的借款周期进行统计分析,发现只有 8 个标的的借款期限是小于 3 个月的. 另外,选择 3 个月为周期时在相邻两个周期的借贷关系网络中,有重复出现的投资者节点或借款人节点. 但其中借款人节点相对比较少,体现了投资者或者借款者在平台上的行为有不同时间或周期依赖性. 从而选择 3 个月期限内发生的投资关系来构建借贷关系网络,得到的网络特征最合适. 基于此建立了 21 个借贷关系网络. 周期内的每一个借款人的借贷关系网络特征,由该借款人所在周期对应的

网络计算得到.

3.1 借款人网络拓扑特征的计算

节点中心性用于衡量网络的可达性,如果一个节点的中心程度越高,则意味着有越多的其他节点与此节点有相连. 在社会网络中,中心程度越高的节点代表的个体具有越大的“威望”或社会权利. 本研究建立的个体借贷关系网络是一个有向网络,在这里主要关注借款人的入度中心度(Input Degree, ID),即节点 v_i 的入度代表从节点 v_j 指向 v_i 的边数. 从选择关系来看,入度越大意味着选择把钱借给此借款人的投资者就越多. 从资金链的关系来看,入度越大的节点与更多的投资者有负债关系.

节点的“域”是评价节点重要性的一种方法. 一个节点的“域”是指与该节点存在途径的其他节点的数量,节点 v_i 的入域用 N_{v_i} 表示. 网络中任意两个可达节点 v_i 和 v_j 的距离是指它们之间的最短路径长度,用 $g(v_i, v_j)$ 表示. 由于在网络借贷中,平台参与者的资产负债水平受投资关系影响. 间接投资关系中是否违约也会间接影响到该参与者的资产情况. 同时,这种影响随着距离的增加而变弱,即直接投资关系对被观测节点中心程度影响比间接投资关系对其中心性的影响更大. 距离越短,节点就越容易通达网络中的其他点,越容易接收到网络中流传的信息,传输效率越高. 入域中的节点到达被观测节点的距离存在不同,因此在衡量被观测节点的中心性程度时,需要考虑入域

内所有节点到达被观测节点的距离。“近距威望”(proximity prestige, PP)可以很好的衡量这一特征,其计算公式为

$$PP_i = \frac{1}{N-1} \sum_{v_j \in N_{v_i}} g(v_i, v_j) \quad (1)$$

较大的入域规模使节点代表的借款人具有较高的近距威望,这是因为有更多的投资者直接或间接地选择了该借款人。较小的平均距离也会增大借款人的近距威望,这是因为有更多的近邻节点选择了该借款人。

节点的中介性是另外一种评价节点重要性的方法。如果网络中任意其他两个节点的最短路径经过该节点,那么该节点就作为这两个节点信息传播路径上的中介。如果该节点作为中介的次数越多,那么该节点的中介中心度(Betweenness Centrality, BC)

就越高,其在流通中起到的作用就越重要。节点中介中心度的计算公式为

$$c_B(v) = \sum_{i \neq v \neq j} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \quad (2)$$

其中 $\sigma_{ij}(v)$ 表示节点 v_i 到节点 v_j 的最短路径中经过节点 v 的数量, σ_{ij} 表示节点 v_i 到节点 v_j 的最短路径数量。在个体借贷关系网络中,节点的中介中心度描述了该节点在多大程度上参与了他人的投资关系。从节点作为中介的定义来看,作为中介的节点必然既有出度也有入度,因此该节点所代表的参与者具有投资者和借款人的双重身份。中介程度越高,反映的是该节点对应的参与者在平台中的投资借贷行为越活跃。

通过整理发现人人贷平台21 833名借款人中只有4%的中介参与者。对应的借款人入度中心度、近距威望和中介中心度特征如表5所示。

表5 借款人网络拓扑特征统计描述

Table 5 Summary of topology characteristics of borrower network

网络拓扑特征	平均值	标准差	最小值	中位数	最大值
入度中心度(ID)	26.434 3	0.269 7	1.000 0	16.000 0	1 037.000 0
近距威望(PP)	0.015 6	0.000 2	4.91E-05	0.001 5	0.294 0
中介中心度(BC)	4.25E-05	5.47E-06	0.000 0	0.000 0	0.066 4

表5显示借款人的入度中心度平均值达到26.434 3,最大值达到1 037。由于网络中的节点数量大,因此计算得到的节点入域小,导致近距威望数值上较小。节点中介中心度的最大值为0.066 4,与只有4%的借款人是中介节点有关,但节点的中介性特征仍不可忽视。

3.2 借款人网络拓扑特征与借款信息相关性检验

下面检验三个节点网络特征与借款信息(借款人信用等级、借款金额和借款利率)之间的相关关系。之所以选择借款信息去检验相关性,是因

为在计算借款人的网络拓扑特征是基于每笔投资记录建立的借贷关系网络,只有借款信息才是与每笔记录相对应的。这些指标中除了信用等级是有序的分类变量外,其他指标都是连续变量。在检验节点网络特征与借款利率和借款金额的相关关系时,计算两个变量间的皮尔逊相关系数。在检验节点网络特征与信用等级的相关关系时,计算两个变量的斯皮尔曼等级相关系数。

具体就人人贷平台借贷关系网络来看,借款人网络特征与借款人信用等级、借款利率和借款金额的相关关系见表6。

表6 借款人网络特征与借款特征相关性表

Table 6 Correlation table of characteristics of borrower network and loan features

网络拓扑特征	信用等级	借款利率	借款金额
入度中心度(ID)	0.184 **	-0.007	0.641 **
近距威望(PP)	0.086 **	0.540 **	-0.004
中介中心度(BC)	0.288 **	0.055	0.176 **

注:** 指示在1%的显著性水平下显著。

表6显示入度中心度与信用等级相关系数为0.184,且在1%显著性水平下显著.说明借款人在网络中的中心程度与其信用等级是有联系的.出借人会更倾向于把钱出借给高信用等级借款人,相应的其在网络中的入度中心度会越大.另外,借款人在平台中总的借款金额与直接被投资的次数(入度中心度)有较强正相关关系.但是入度中心度与借款利率间关系不显著,可以解释为入度中心度越大说明该节点所代表的借款人获得越多投资者投资,对应借款金额越大,而低借款利率与投资者投资的次数无明显相关关系.

近距威望与信用等级的相关系数为0.086,且在1%显著性水平下显著.由于在计算借款人近距威望时考虑了距离,即在计算一个借款人的近距威望时,同时考虑了其他借款人对自身的选择和其他投资者与该借款人在借贷关系网络中的距离.这可视作对中心程度不同的节点赋予了不同的权重.可以理解为,如果一个借款人近距威望越高代表他被更多中心程度高的参与者节点选择.由借款人近距威望与借款利率、借款金额的相关系数表明:借款人近距威望和借款人平均借款利率有较强的正相关关系,而与借款金额无显著相关关系.

借款人中介中心度与信用等级的相关系数为

0.288,且在1%显著性水平下显著,即借款人中介中心度与信用等级之间具有正相关关系.中介中心度在一定程度上代表了借款人借贷行为的活跃程度,说明越活跃的借款人在平台中一般具有较高的信用等级.由借款人中介中心度与借款利率、借款金额的相关系数可知,借款人的中介性与自身的借款利率水平无显著相关关系,与自身的借款金额有较弱的正相关关系.说明平台中参与越多借贷关系的中介借款人,其自身在平台中也有越高的借款金额.

4 借款人网络拓扑特征、借款信息与借款人违约的回归分析

为了方便论述关于借款人网络拓扑特征、借款信息与借款人违约关系的研究,不失一般性,本节针对人人贷平台借贷关系网络进行具体的回归分析.

4.1 模型变量描述

本研究基于 Logistic 模型探究借贷网络节点特征与借款人违约的关系.相关变量的描述见表7,网络特征变量与借款信息特征变量的相关关系见表8.

表7 变量描述

Table 7 Variable description

变量	变量描述
违约情况 (Default)	分类变量, Default = 1 表示借款人发生违约, 否则未违约
入度中心度 (Input Degree)	网络中某节点被多少其他节点指向. 表示选择把钱借给该节点代表的借款人的投资者数量, 以及该借款人与多少投资者有负债关系
近距威望 (Proximity Prestige)	距离加权的节点连通关系强度, 计算方法见式(1). 该特征给予直接投资关系和间接投资关系影响以不同的权重, 以有效衡量节点影响力
中介中心性 (Betweenness Centrality)	网络中某节点出现在其他任意两个节点的最短路径中的次数, 计算方法见式(2). 中介中心性越高, 反映该节点在流通中起到越重要的作用
信用等级 (Credit)	借款人的信用等级, 分类变量, 包括 AA、A、B、C、D、E、HR 七个等级, 从 AA 到 HR 信用等级逐渐降低
借款利率 (Interest)	借款人发起借款项目的账面支付利率
借款金额 (Amount)	借款人发起借款项目的借款金额

注: 实证分析中对网络拓扑特征和借款利率以及借款金额进行自然对数变化.

表8 变量相关性

Table 8 Correlation

变量	入度中心度	近距威望	中介中心度	借款利率	借款金额
近距威望	0.478 **				
中介中心度	0.395 **	0.795 **			
借款利率	-0.007	0.540 **	0.005		
借款金额	0.641 **	-0.004	0.176 **	-0.262 **	
信用等级	0.184 **	0.086 **	0.288 **	-0.171 **	0.237 **

注：** 指示在1%显著性水平下显著。

4.2 回归模型与结果

本节采用 Logistic 回归分析网络结构特征与借款人违约情况之间的联系。其中因变量为借款人的还款情况，用是否违约来表示， $default = 1$ 代表借款人违约， $default = 0$ 代表借款人没有发生违约。Logistic 回归的基本模型为

$$\text{Logit}(p) = \ln \left(\frac{P}{1-p} \right) = \beta_0 + \sum_i^n \beta_i x_i \quad (3)$$

其中 p 为借款人违约的概率， $1-p$ 为借款人无违约的概率。

此外，为说明分析结果不受模型选择的影响，也采用 Probit 回归模型替换 Logistic 回归模型以降低由于模型选择而带来的分析结果偏差。Probit 回归的基本模型为

$$\text{Probit}(p) = \Phi^{-1}(P) = \beta_0 + \sum_i^n \beta_i x_i \quad (4)$$

其中 $\Phi(x)$ 为标准正态分布的概率分布函数。

上文检验结果表明三个节点的网络特征指标（入度中心度、近距威望、中介中心度）与节点所代表的借款人借款特征（借款等级、借款利率、借款金额）存在显著相关。据此，本节建立如下的三个回归模型。

模型1： $F(x) = f(ID, Interest, Credit)$

模型2： $F(x) = f(PP, Amount, Credit)$

模型3： $F(x) = f(BC, Amount, interest, Credit)$

首先使用 Logistic 回归分析方法拟合上述三个模型。回归分析结果显示每个模型的 Omnibus 检验 P 值都小于0.01，表明所以模型在1%显著性水平下显著整体有效。相应的回归结果如表9所示。

表9 Logistic 回归分析结果

Table 9 Regression results of Logistic analysis

模型	模型1		模型2		模型3	
	系数	标准误	系数	标准误	系数	标准误
<i>Input Degree</i>	0.109 ***	(0.019)				
<i>Proximity Prestige</i>			-0.473 ***	(0.015)		
<i>Betweenness Centrality</i>					-0.242 ***	(0.077)
<i>Interest</i>	-1.316 ***	(0.084)			-1.237	(1.375)
<i>Amount</i>			0.786 ***	(0.022)	1.318 ***	(0.249)
<i>Credit (HR)</i>	3.649 ***	(0.451)	3.302 ***	(0.454)	19.774	(2 891.763)
<i>Credit (E)</i>	2.083 ***	(0.454)	1.486 ***	(0.457)	0.362	(5 224.395)
<i>Credit (D)</i>	2.038 ***	(0.453)	1.531 ***	(0.456)	16.443	(2 891.763)
<i>Credit (C)</i>	1.159 **	(0.477)	0.715	(0.48)	0.444	(4 573.478)
<i>Credit (B)</i>	0.491	(0.536)	-0.006	(0.539)	0.085	(6 124.531)
<i>Credit (A)</i>	1.393 **	(0.594)	1.054 *	(0.599)	-0.402	(8 502.935)
常数项	-1.595 ***	(0.498)	-14.888 ***	(0.533)	-34.393	(2 891.767)
样本容量	29 276		29 276		29 276	

注：***, **, * 指示分析系数分别在1%, 5%和10%的显著性水平下显著。

表 9 显示借款人节点的网络特征变量与借款人的贷款违约情况显著相关. 其中, 借款人节点的入度中心度与借款人违约正相关, 表明与借款人产生投资关系的投资者越多则对应的借款人更容易发生违约. 在贷款数额一定的情况下, 投资者越多, 反映的是大量投资者在投资时仅使用小额资金进行投资, 从侧面反映了投资者认为当前借款人可能存在较高的违约风险. 而借款人节点的近距威望和中介中心度与借款人违约负相关. 表明借款人在平台越活跃其发生违约的可能性越小, 这与现实情况相符. 中介参与者既作为借款人发布借款申请, 也愿意作为投资者去投资其他借款人的借款. 这一定程度上反映了其对这种借贷模式

的认可, 从而在作为借款人的情况下, 其越活跃越不倾向于违约自己的借款. 此外, 关于其他借款申请的特征变量分析结果与既有文献保持一致.

为避免模型选择差异导致分析结果存在偏差, 使用 Probit 回归方法重新拟合上述三个模型. 对应的回归结果展示在表 10 中.

Probit 拟合的三个模型经过卡方检验皆在 10% 的显著性水平下显著, 说明构建的三个模型整体上显著有效. 表 10 对应的分析结果与表 9 的 Logistic 回归分析结果保持一致. 说明本研究关于借款人借贷关系网络特征变量与借款人贷款违约的分析结果不受回归模型选择的影响, 分析结果具有一定的稳健性.

表 10 Probit 回归分析结果

Table 10 Regression results of Probit analysis

模型 变量	模型 1		模型 2		模型 3	
	系数	标准误	系数	标准误	系数	标准误
<i>Input Degree</i>	0.053 ***	(0.011)				
<i>Proximity Prestige</i>			-0.267 ***	(0.008)		
<i>Betweenness Centrality</i>					-0.096 ***	(0.033)
<i>Interest</i>	-0.765 ***	(0.049)			-0.641	(0.591)
<i>Amount</i>			0.441 ***	(0.012)	0.557 ***	(0.113)
<i>Credit (HR)</i>	1.765 ***	(0.172)	1.641 ***	(0.186)	2.055 ***	(0.676)
<i>Credit (E)</i>	0.918 ***	(0.173)	0.637 ***	(0.188)	0.244	(1.045)
<i>Credit (D)</i>	0.900 ***	(0.173)	0.657 ***	(0.188)	0.661	(0.741)
<i>Credit (C)</i>	0.488 ***	(0.185)	0.260	(0.199)	0.394	(0.875)
<i>Credit (B)</i>	0.216	(0.206)	-0.049	(0.222)	0.078	(1.308)
<i>Credit (A)</i>	0.581 **	(0.246)	0.407	(0.267)	-0.154	(2.054)
常数项	-0.580 ***	(0.211)	-8.188 ***	(0.241)	-8.103 ***	(2.14)
样本容量	29 276		29 276		29 276	

注: ***, **, * 指示分析系数分别在 1%, 5% 和 10% 的显著性水平下显著.

在前面关于不同信用等级的人群体分析中, 发现信用等级为 D、E、HR 的三类借款人群体违约率高于 1%, 其他信用等级的借款人群体违约率均低于 0.033%. 由于发生违约的多是低信用等级(高信用风险)群体, 本研究随后将重点聚焦在低信用等级的借款人群体子样本上. 下面分别对信用等级为 D、E、HR 的三类借款人群体关于其对应的网络拓扑特征、借款信息与贷款违约进行 Logistic 回归分析, 结果如表 11 所示.

表 11 显示, 关于信用等级为 HR 的借款人子样本回归的结果与表 9 的结果一致. 该群体的样本量占借款人整体的 56.3%, 对表 9 所示整体回归结果影响较大. 对于信用等级为 E 的借款人群体来讲, 违约借款人所对应的中介中心度除了一个节点非零, 其他节点的中介中心度均为 0, 即 E 信用等级的借款人多是纯借款人. 由于可用的数据匮乏, 本研究放弃关于 E 信用等级借款人群体的模型 3 分析. 表 11 的三个子样本分析结果显示

借款人子网络的节点入度中心度、近距威望和中介中心度三个网络特征与借款人的违约显著相关。其中,借款人节点的入度中心度与借款人违约呈

正相关,借款人节点的近距威望和中介中心度与借款人违约呈负相关。这些子样本分析结果与上文分析保持一致,进一步表明本研究分析结果的稳健性。

表 11 低信用等级群体下的回归分析的结果

Table 11 Regression results of low credit groups

信用等级	HR			E		D		
	模型 1	模型 2	模型 3	模型 1	模型 2	模型 1	模型 2	模型 3
变量								
<i>Input Degree</i>	0.184 *** (0.022)			0.105 *** (0.07)		0.326 *** (0.055)		
<i>Proximity Prestige</i>		-0.44 *** (0.015)			-0.171 *** (0.063)		-0.738 *** (0.057)	
<i>Betweenness Centrality</i>			-0.233 *** (0.077)					-0.729 *** (0.772)
<i>Interest</i>	-1.726 *** (0.093)		-1.346 (1.401)	2.875 *** (0.333)		1.132 *** (0.286)		3.636 (8.294)
<i>Amount</i>		0.825 *** (0.025)	1.322 *** (0.259)		0.647 *** (0.083)		0.76 *** (0.079)	1.669 *** (1.409)
常数项	2.911 *** (0.246)	-11.744 *** (0.296)	-14.258 (4.796)	-9.955 *** (0.876)	-9.964 *** (1.078)	-4.501 *** (0.737)	-14.847 *** (1.066)	-40.951 (38.121)
样本容量	16 483	16 483	16 483	4 708	4 708	5 483	5 483	5 483

注: ***, **, * 指示分析系数分别在 1%, 5% 和 10% 的显著性水平下显著。括号中报告估计系数的标准误。

5 结束语

目前已有较多研究基于社交网络理论,结合网络拓扑特征对个体行为进行解释,但基于互联网金融平台的个体债务市场的相关研究较为欠缺。本研究为互联网借贷平台的参与者个体行为研究提供了一个新视角:基于个体借贷关系网络,考虑了参与者个体在网络中的位置(中心性和中介性),即在研究参与者行为的同时考虑了平台中其他参与者关联性影响。这些特征能为互联网借贷市场上的投资行为提供新的解释。在以往的研究中,很少学者关注中介参与者的投资行为,但这一类参与者在个体借贷关系网络中是值得关注的。因为中介参与者使借贷网络的连接关系更紧密。在违约风险的传染中,如果他们所投资的标的发生违约,将会影响他们的资产水平,从而影响其自身还款行为。本研究弥补了目前研究中关于借款人投资行为的研究空缺。此外,还检验了借贷关系网络特征与借款人违约之间的关系,为借款人信用风险评估提供了新思路。

以人人贷为例的互联网平台个体借贷关系网络密度低,度分布服从许多复杂网络都具有的幂律分布特征。在检验了网络中节点的拓扑特征与借款人信用特征的相关性之后,发现入度中心度、近距威望和中介中心度与信用等级之间具有正相关关系。这表明出借人更倾向于把钱出借给高信用等级的借款人。另外,平台中活跃的中介参与者一般具有较高的信用等级,他们自身的借款金额也越高,同时在投资选择中会倾向选择高利率的借款人。这表明还款能力强的借款人同时具备比其他信用等级借款人更高的投资能力。通过借贷关系网络特征与借款人违约之间的关系的检验,可以得到更高的信息传输效率对借款人违约具有负向影响。

对于互联网个体债务市场的平台运营者来说,从盈利角度,要促进平台高信用等级借款人与其他参与者的借贷联系。如对平台中高信用等级的借款人给予利率优惠或分期付款的政策倾斜,同时使用运营手段鼓励他们在平台中进行投资。从风险角度,本研究验证了借款人网络拓扑特征与借款人违约之间的联系。结论表明在互联网借

贷平台的借款人信用风险问题中加入借款人所在平台的个体借贷关系网络拓扑结构特征对借款人信用风险进行评估和防范是一种新的思路和方法. 另一方面, 现实中市场参与者会在不同平台上

进行投资和借款, 从而相关监管机构可以使用本研究提出的方法构建跨平台的广义借贷网络模型, 来监控借款人的违约风险, 促进我国互联网平台个体债务市场的健康发展.

参考文献:

- [1] Lin M F, Prabhala N R, Viswanathan S. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending[J]. *Management Science*, 2013, 59(1): 17–35.
- [2] Laura G, Yuliya K L. When can a photo increase credit? The impact of lender and borrower profiles on online peer-to-peer loans[J]. *Journal of Behavioral and Experimental Finance*, 2014, 2(6): 44–58.
- [3] 熊正德, 刘臻焯, 熊一鹏. 基于有序 logistic 模型的互联网金融客户违约风险研究[J]. *系统工程*, 2017, 35(8): 29–38.
Xiong Zhengde, Liu Zhenxuan, Xiong Yipeng. Predicting borrowers' default activity in lending through internet platform based on ordinal logistic model[J]. *Systems Engineering*, 2017, 35(8): 29–38. (in Chinese)
- [4] Wiginton J C. A note on the comparison of logit and discriminant models of consumer credit behavior[J]. *Journal of Financial & Quantitative Analysis*, 1980, 15(3): 757–770.
- [5] 张卫国, 卢媛媛, 刘勇军. 基于非均衡模糊近似支持向量机的 P2P 网贷借款人信用风险评估及应用[J]. *系统工程理论与实践*, 2018, 38(10): 2466–2478.
Zhang Weiguo, Lu Yuanyuan, Liu Yongjun. The borrowers' credit risk assessment in P2P platform based on fuzzy proximal support vector machine and its application[J]. *Systems Engineering: Theory & Practice*, 2018, 38(10): 2466–2478. (in Chinese)
- [6] Meko M C S, Lyn C T. Modelling the profitability of credit cards by Markov decision processes[J]. *European Journal of Operational Research*, 2011, 212(1): 123–130.
- [7] 张 维, 李玉霜, 王春峰. 递归分类树在信用风险分析中的应用[J]. *系统工程理论与实践*, 2000, 20(3): 50–55.
Zhang Wei, Li Yushuang, Wang Chunfeng. Application of classification tree in credit risk analysis[J]. *Systems Engineering: Theory & Practice*, 2000, 20(3): 50–55. (in Chinese)
- [8] Huang Z, Chen H, Hsu C J, et al. Credit level analysis with support vector machines and neural networks: A market comparative study[J]. *Decision Support Systems*, 2004, 37(4): 543–558.
- [9] Kim H S, Sohn S Y. Support vector machines for default prediction of SMEs based on technology credit[J]. *European Journal of Operational Research*, 2010, 201(3): 838–846.
- [10] Xia Y, Liu C, Liu N. Cost-sensitive boosted tree for loan evaluation in peer-to-peer lending[J]. *Electronic Commerce Research and Applications*, 2017, 24: 30–49.
- [11] Niu K, Zhang Z, Liu Y, et al. Resampling ensemble model based on data distribution for imbalanced credit risk evaluation in P2P lending[J]. *Information Sciences*, 2020, 536: 120–134.
- [12] Ma L, Zhao X, Zhou Z, et al. A new aspect on P2P online lending default prediction using meta-level phone usage data in China[J]. *Decision Support Systems*, 2018, 111: 60–71.
- [13] Zhang W, Wang C, Zhang Y, et al. Credit risk evaluation model with textual features from loan descriptions for P2P lending[J]. *Electronic Commerce Research and Applications*, 2020, 42: 100989.
- [14] Jiang C, Wang Z, Wang R, et al. Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending[J]. *Annals of Operations Research*, 2018, 266(1): 511–529.
- [15] Ge R, Feng J, Gu B, et al. Predicting and deterring default with social media information in peer-to-peer lending[J]. *Journal of Management Information Systems*, 2017, 34(2): 401–424.
- [16] Wang C, Zhang Y, Zhang W, et al. Textual sentiment of comments and collapse of P2P platforms: Evidence from China's P2P market[J]. *Research in International Business and Finance*, 2021: 101448.
- [17] Giudici P, Hadji-Misheva B, Spelta A. Network based scoring models to improve credit risk management in peer to peer lending platforms[J]. *Frontiers in Artificial Intelligence*, 2019, 2: 3.

- [18] Lee E, Lee B. Herding behavior in online P2P lending: An empirical investigation[J]. *Electronic Commerce Research and Applications*, 2012, 11(5): 495–503.
- [19] Zhang J, Liu P. Rational herding in microloan markets[J]. *Management Science*, 2012, 58(5): 892–912.
- [20] Puro L, Teich J E, Wallenius H, et al. Bidding strategies for real-life small loan auctions[J]. *Decision Support Systems*, 2011, 51(1): 31–41.
- [21] 丁杰, 李悦雷, 曾燕, 等. P2P网贷中双向交易者的双重信息价值及信息传递[J]. *南开管理评论*, 2018, 21(2): 4–15.
Ding Jie, Li Yuelei, Zeng Yan, et al. Dual information value and information transmission of two-way traders in P2P lending[J]. *Nankai Business Review*, 2018, 21(2): 4–15. (in Chinese)
- [22] 冯博, 叶绮文, 陈冬宇. P2P网络借贷研究进展及我国问题研究展望[J]. *管理科学学报*, 2017, 20(4): 113–126.
Feng Bo, Ye Qiwen, Chen Dongyu. Review on P2P online lending and new research opportunities for China's case[J]. *Journal of Management Sciences in China*, 2017, 20(4): 113–126. (in Chinese)
- [23] Wei Q, Zhang Q. P2P lending risk contagion analysis based on a complex network model[J]. *Discrete Dynamics in Nature and Society*, 2016, 6: 1–8.
- [24] Xiao Z, Li Y L, Zhang K. Visual analysis of risks in peer-to-peer lending market[J]. *Personal and Ubiquitous Computing*, 2018, 22(4): 825–838.
- [25] 余桃. 基于复杂网络结构的P2P平台内部信用风险传染研究——以“人人贷”平台为实例[D]. 成都: 电子科技大学, 2017.
She Tao. Research of credit risk contagion in P2P lending platform based on the complex network topology: Evidence from “RenrenDai.com”[D]. Chengdu: University of Electronic Science and Technology of China, 2017. (in Chinese)
- [26] Chen L, Li P, Li Q. The evolving networks of debtor-creditor relationships with addition and deletion of nodes: A case of P2P lending[D]. Ithaca: Cornell University, 2018.
- [27] Li Y L, Hao A T, Zhang X T, et al. Network topology and systemic risk in peer-to-peer lending market[J]. *Physica A: Statistical Mechanics and its Applications*, 2018, 508: 118–130.
- [28] Tong Z, Chen X. P2P net loan default risk based on Spark and complex network analysis based on wireless network element data environment[J]. *EURASIP Journal on Wireless Communications and Networking*, 2019, 2019(1): 1–7.
- [29] Xu R, Mi C, Mierziak R, et al. Complex network construction of Internet finance risk[J]. *Physica A: Statistical Mechanics and its Applications*, 2020, 540: 122930.
- [30] 杨晓兰, 王伟超, 高媚. 股市政策对股票市场的影响——基于投资者社会互动的视角[J]. *管理科学学报*, 2020, 23(1): 15–32.
Yang Xiaolan, Wang Weichao, Gao Mei. The impact of stock market policies on stock market: From the perspective of investor social interaction[J]. *Journal of Management Sciences in China*, 2020, 23(1): 15–32. (in Chinese)
- [31] 周开国, 季苏楠, 杨海生. 系统性金融风险跨市场传染机制研究——基于金融协调监管视角[J]. *管理科学学报*, 2021, 24(7): 1–20.
Zhou Kaiguo, Ji Sunan, Yang Haisheng. Cross-market contagion mechanism of systemic risk from the perspective of coordinated supervision[J]. *Journal of Management Sciences in China*, 2021, 24(7): 1–20. (in Chinese)
- [32] Albert R, Barabási A-L. Statistical mechanics of complex networks[J]. *Reviews of Modern Physics*, 2002, 74(1): 47–97.
- [33] 刘景卿, 车维汉, 夏方杰. 全球价值链贸易网络分析与国际风险传导应对[J]. *管理科学学报*, 2021, 24(3): 1–17.
Liu Jingqing, Che Weihang, Xia Fangjie. Network analysis of global value chain and coping with international risk transmission[J]. *Journal of Management Sciences in China*, 2021, 24(3): 1–17. (in Chinese)
- [34] 宫晓莉, 熊熊. 波动溢出网络视角的金融风险传染研究[J]. *金融研究*, 2020, (5): 39–58.
Gong Xiaoli, Xiong Xiong. A study of financial risk contagion from the volatility spillover network perspective[J]. *Journal of Financial Research*, 2020, (5): 39–58. (in Chinese)
- [35] 隋聪, 王宪峰, 王宗尧. 银行间债务网络流动性差异对风险传染的影响[J]. *管理科学学报*, 2020, 23(3): 65–72.

- Sui Cong, Wang Xianfeng, Wang Zongyao. The impacts of interbank debt network liquidity differences on risk contagion [J]. *Journal of Management Sciences in China*, 2020, 23(3): 65–72. (in Chinese)
- [36] 周颖刚, 程欣, 王艺明. 为什么人民币越来越重要呢? ——基于网络分析方法的汇率证据[J]. *管理科学学报*, 2019, 22(9): 12–38.
- Zhou Yinggang, Cheng Xin, Wang Yiming. Why is CNY increasingly important? Network evidence from exchange rates [J]. *Journal of Management Sciences in China*, 2019, 22(9): 12–38. (in Chinese)
- [37] Zhang J M, Zhu S Z, Yan W, et al. The construction and simulation of internet financial product diffusion model based on complex network and consumer decision-making mechanism[J]. *Information Systems and E-Business Management*, 2020, 18(4): 545–555.
- [38] 张奇, 李彦, 王歌, 等. 基于复杂网络的电动汽车充电桩众筹市场信用风险建模与分析[J]. *中国管理科学*, 2019, 27(8): 66–74.
- Zhang Qi, Li Yan, Wang Ge, et al. Credit risk modeling and analysis for crowdfunding market of electric vehicle charging pile based on complex network[J]. *Chinese Journal of Management Science*, 2019, 27(8): 66–74. (in Chinese)
- [39] 游运, 万常选, 江腾蛟. 考虑社会网络关系的 P2P 借贷项目违约风险预测[J]. *计算机工程与应用*, 2021, 57(13): 239–245.
- You Yun, Wan Changxuan, Jiang Tengjiao. Project default risk prediction considering social network in P2P lending[J]. *Computer Engineering and Applications*, 2021, 57(13): 239–245. (in Chinese)
- [40] Nier E, Yang J, Yorulmazer T, et al. Network models and financial stability[J]. *Journal of Economic Dynamics & Control*, 2007, 31(6): 2033–2060.
- [41] Clauset A, Shalizi C R, Newman M E J. Power-law distributions in empirical data[J]. *SIAM Review*, 2009, 4(51): 661–703.
- [42] Newman M E J. The structure and function of complex networks[J]. *SIAM Review*, 2003, 45(2): 167–256.

Characteristics of individual lending relationship network and borrowers' default: Evidence from Internet financing platforms

ZHANG Wei-guo¹, LI Hua², WANG Chao^{3*}

1. School of Management, Shenzhen University, Shenzhen 518055, China;
2. School of Business Administration, South China University of Technology, Guangzhou 510641, China;
3. School of Economics, Guangdong University of Technology, Guangzhou 510520, China

Abstract: This paper applies the complex network analysis method to study the individual lending network characteristics on Internet financing platforms and borrower's default problem. The paper proposes a method to establish the individual lending network, constructs the lending network of Renrendai.com, and depicts the features of the overall network and sub-networks that are separated by borrowers and pure investors according to their default status and credit ratings. Based on this, the behavior rules of borrowers and investors are discussed and the correlation between the network topology characteristics of borrowers and loan information is analyzed. Then three regression models for borrowers' network topology characteristics, loan information and their defaults are constructed. The results show that the network topology has a strong explanatory power for the structure of lending-borrowing relationship. There is a significant correlation between the network topology characteristics of borrower nodes and the borrowers' default, and this correlation is also robust among groups of borrowers with different credit ratings. These findings can be applied to borrower credit risk assessment and default prediction on Internet financing platforms.

Key words: individual lending relationship network; topology characteristics; credit default; complex network