arXiv:2611.32230v1 [physics.soc-ph] 24 Nov 2026

《老友记》的互动动力学

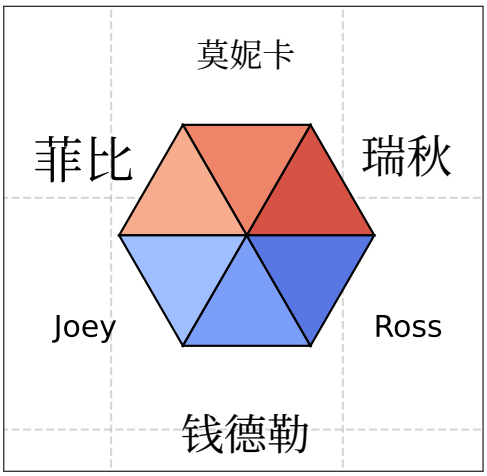
Shun Zhang[1*] **Tabia Tanzin Prama**[1,2], **ChristopherM. Danforth**[1,3], **Peter SheridanDodds**[1,2,4,5†]

[1] 计算故事实验室、佛蒙特复杂系统研究所、MassMutual复杂系统与数据科学卓越中心、佛蒙特高级计算中心，佛蒙特大学，美国佛蒙特州伯灵顿，VT05405[2] 佛蒙特大学计算机科学系，美国佛蒙特州伯灵顿，VT 05405[3] 佛蒙特大学数学与统计学系，美国佛蒙特州伯灵顿，VT 05405[4] 圣菲研究所，美国新墨西哥州圣菲市海德公园路1399号，NM 87501[5] 复杂性科学中心，奥地利维也纳Metternichgasse8，1030

2026年11月25日

故事梗概

我们将 Ousiometry 的力量—危险框架应用于群像情景喜剧《老友记》，以揭示对话中的情感意义如何在单个角色之间演变、区分不同互动语境，并在群体叙事中形成反复出现的时间模式。



*shun.zhang@uvm.edu
†peter.dodds@uvm.edu

摘要

计算叙事分析日益寻求能够捕捉语言如何反映不断变化的社会与互动语境的表征。我们将 Ousiometry 这一围绕力量、危险和结构组织起来的基于词典的框架应用于《老友记》的完整对话语料（236 集，$\sim$750K 个词元），考察六位主要角色的时间变化、群体层面定位和成对协变。我们发现，角色轨迹在不同剧集之间存在显著波动，PELT 检测出的变点在描述层面上与主要叙事事件相吻合。在群体层面，角色组合分布于权力–危险空间的不同区域；权力的离散度高于危险，尽管这些差异经邦费罗尼校正后未能保留统计显著性，且轨迹聚类显示出较弱的结构（轮廓系数 $= 0.27$）。两两皮尔逊相关显示角色之间存在程度有限的协变：经校正后，权力维度有三个显著配对，危险维度有两个，结构维度有六个。偏相关显著减弱了这些关联，仅在权力维度保留一个显著配对、在危险维度保留一个显著配对，这表明所观察到的大部分协变反映的是群体整体动态，而非配对特异性动态。总体而言，我们的研究结果表明，单集级 Ousiometry 能够捕捉角色对话中有意义的时间变化；与此同时，关于稳定的群体层面结构和配对特异性协变的证据仍然有限。因此，《老友记》中的语义变化更适合被刻画为连续且依赖语境的变化。

关键词

Ousiometry、计算叙事分析、语义词典、权力–危险–结构空间、角色互动动态、群像情景喜剧、《老友记》

目录

1 引言 4

2 文献综述 4

2.1 叙事中的计算角色分析.2.2 本体计量学.2.3 情景喜剧对话.2.4 研究空白. 4 4 5 5

3 数据与方法 5

3.1 数据收集与预处理. . . . . 53.2 本体计量学框架. 63.3 RQ1：主要角色的语义轨迹.63.4 RQ2：角色组合中的情感模式.73.5 RQ3：角色的语义画像是否会跨剧集系统性地共同变化？ 11

4 结果与讨论 12

4.1 RQ1：主要角色的情感弧线 。 12

4.1.1 总体权力—危险度概况. . . . 12

4.1.2 词汇差异：Word 转变分析 12

4.1.3 主要角色的情感弧线 14

4.2 RQ2：角色组合中的情感模式 .14

4.2.1 角色组合的空间分布.14

4.2.2 自助法置信椭圆 . . 15

4.2.3 角色组合的轨迹聚类.16

4.3 RQ3：角色的语义画像是否会在不同剧集中系统性地共同变化？ 16

4.3.1 角色层面的皮尔逊相关性164.3.2 折半信度. . . . . 164.3.3 偏相关 16

5 结论 17

6 局限性和未来工作 18

参考文献 19

A 附录 A1

A1保留的字符组合 A1

A2每个字符的变化点 A4

A3用于轨迹聚类的角色组合 A6

A4分半信度 A6

1 介绍

情感分析已成为从文本中提取情感意义的标准工具 [1], ，广泛应用于社会科学、公共卫生和计算叙事分析领域 [1, 2]。大多数情感分析系统都依赖语义词典运行：即由词语及其对应的情感标注构成的结构化词汇表 [3, 4]。系统通过对给定文本中各词语的情感分数取平均值，估算读者所感受到的文本整体情感倾向 [5]。

情感分析的基础是本体计量学：对本质意义的测量。在近期重新审视 Osgood 等人长期以来提出的评价-力量-活动（EPA）框架（及其在情感反应中的具体体现——效价-唤醒-支配）的研究中，研究表明，本质意义实际上应由一个三维空间来描述，其正交的语义差异轴分别是权力（{weak ⇔powerful}）、危险（{safe ⇔ dangerous}）和结构（{rigid ⇔loose}） [6]。在由权力和危险构成的主平面上，两个中间方位维度分别是善性（权力 - 危险，{bad ⇔good}）和攻击性（权力 ＋危险，{gentle ⇔ aggressive}）。我们将完整的环形模型称为GPADS框架。

在本研究中，我们使用这一权力—危险—结构框架分析对话中的互动动力学。我们主要关注权力和危险这两个维度。

我们以电视情景喜剧《老友记》为案例研究 [7],，该剧提供了丰富的语料库，可用于考察情绪与社会结构的交汇 [7]。该剧于1994年至2004年播出，共十季、236集，讲述了六位生活在纽约市的亲密朋友（瑞秋、莫妮卡、菲比、乔伊、钱德勒和罗斯）的生活 [7]。《老友记》的叙事结构围绕反复出现的角色组合展开，这使我们能够考察同一角色与不同伙伴互动时语义特征如何发生变化。

本研究围绕三个具体的研究问题展开：

1. **个体层面：** 六位主要角色各自的权力和危险得分如何贯穿十季不断变化？这些轨迹又如何与关键叙事事件相吻合？

   在个体层面，我们追踪每位说话者的权力—危险语义画像如何随时间演变，从而揭示与叙事发展相对应的关键意义弧线。

2. **组合层面：** 不同的组合是否占据权力-危险空间中的不同区域？

   随后，我们将这一框架扩展到不同角色组合的比较，考察不同的角色分组如何占据语义空间中的不同区域。

3. **两两层面：** 角色的语义画像是否会在不同剧集中呈现系统性的协变？

   我们还使用两两皮尔逊相关来考察角色轨迹是否表现出非均匀的协变。

2 文献综述

2.1 叙事中的计算角色分析

角色是叙事结构的基础，将事件联系起来，并在故事随时间发展时维持连贯性 [8, 9]。因此，计算叙事研究发展出了一系列用于大规模表示虚构角色的方法。这些方法通过角色的叙事角色、人格特质、对话和互动模式来刻画角色，从而互补地呈现角色在故事中的功能 [10–12]。

2.2 本体计量学

本体计量学提供了一个定量框架，用于沿着可解释的语义维度表示词语的本质含义 [6]。GPADS框架从善性、力量、攻击性、危险和结构这几个维度刻画词语，从而对语义变化进行多维表示 [6]。具体而言，力量和危险构成两个基本语义轴，而善性和攻击性分别对应力量—危险平面的两条正交平分线 [6]。这一几何结构提供了两种描述同一二维语义空间的互补坐标系：词语可以根据力量和危险进行定位，也可以等价地沿旋转后的善性和攻击性方向进行解读。结构则提供了一个额外的语义组织维度。

本体计量学的一项重要实证发现，涉及语言在语义空间中的分布。在现实世界语料库中对词语进行加权后可以发现，语言在力量—危险平面中的分布并不对称，而是明显偏向低危险、也就是安全的区域 [6]。这一发现从几何角度重新诠释了人类语言中广为人知的积极性偏差，这种偏差通常与

波莉安娜原则 [13, 14]相关。由于 {bad ⇔good} 轴是力量—危险平面中的一条平分线，因此可以将观察到的积极性偏差理解为更广泛的安全偏差在善性维度上的一维投影 [6]。

本体计量维度可以在文本的连续单元中进行汇总，以考察语义位置如何随叙事时间变化。Fudolig 等人 [15] 将本体计量测量应用于书籍，并将语义波动表示为累积词时上的轨迹，展示了语义维度的变化如何揭示叙事结构的大尺度变异。

2.3 情景喜剧对话

情景喜剧对话为研究社会互动中的语义特征提供了独特优势。这类对话以一群稳定且反复出现的角色为特色，他们之间的关系会在较长时期内逐渐发展；在此期间，同一批人物会在不同的社会配置中反复互动 [10, 16]。这类长篇叙事尤其适合此类分析，因为反复出现的角色群体会贯穿许多剧集，使研究者能够在较长的叙事序列中反复观察角色行为 [17]。

在群像情景喜剧中，《老友记》尤其适合作为语料库 [7, 18]。整整十季中，该剧始终围绕六位核心角色（瑞秋、莫妮卡、菲比、乔伊、钱德勒和罗斯）展开 [7]，每一集都以同一社交群体成员之间的互动为核心。瑞秋、莫妮卡、菲比、乔伊、钱德勒和罗斯反复形成的各种组合，营造出丰富多样的会话场景，从一对一交谈到全体讨论皆有涵盖。该语料库包含约 750,000 个对话词元，足以支持定量分析（语料库详情见下文数据部分）。此前基于语料库对《老友记》与自然发生的对话所做的比较表明，其剧本化对话具有许多自然对话的特征，同时也呈现出系统性的语言差异 [18]。

2.4 研究空白

此前对本体计量学的应用表明，可以跨叙事时间追踪语义得分 [15], ，从而将语义位置的变化表示为轨迹。现有的叙事应用主要考察文本或叙事片段层面的语义变化，而同一叙事群体中反复出现的角色之间的平行语义轨迹仍相对缺乏研究。这对于群像叙事尤为重要：多个反复出现的角色共同参与同一叙事语境，同时保持各自独特的对话与互动模式。因此，角色层面的本体计量分析有望将现有基于轨迹的方法从整体叙事变化拓展到群体中反复出现成员之间的语义变化。关系分析还进一步区分了边际协变与条件协变。成对相关则将现有的角色轨迹比较拓展到本体计量维度。

3 数据与方法

3.1 数据收集与预处理

该数据集包含《老友记》十季的完整剧本，来源于公开可用的资料 [19]。原始剧本在236集剧集中约包含一百万个词元；经过下文所述的预处理步骤后，分析语料库包含747,805个词元，平均每集约3,200个词元。每份剧本均按行提供对话，并标注说话人，同时嵌入舞台指示。在某些季中，最后两集被合并到同一个文件中，以保持情节连贯性。根据 IMDb [20]，我们将其全部重新划分为正确的剧集分集。

预处理分为几个步骤。首先对原始文本进行分词并转换为小写。对于下文所述的主要 Ousiometric 分析，将删除停用词。

说话人别名和缩写形式，例如表示瑞秋的“RACH”、表示钱德勒的“CHAN”以及表示莫妮卡的“MNCA”，均统一为完整的角色名称。包括笑声标记在内的非词汇标记也全部删除。例如，在 S02E10 中，我们将“MNCA”替换为“Monica”。

> MNCA：你作为国王真是太出色了。我真的印象深刻，确实如此。

为数据集进行基于词典的建模准备时，一个关键挑战是原始剧本中的场景标注缺乏一致性和标准化。原始剧本以逐行对话数据的形式提供，但并未始终包含结构化的场景级标注。由于场景边界对于分析互动模式和语义变化十分重要，我们构建了一个预处理流程，将剧集脚本分割为独立场景。我们根据剧本中包含的舞台指示、地点描述等上下文线索来识别场景边界。

虽然这种自动分段方法适用于大多数剧集，但我们发现第二季的几集中存在不一致之处：源数据并未明确标记所有场景转换。为确保基于场景的分析具有信度，我们对第二季的十集剧集进行了场景边界的人工标注。人工标注过程包括仔细审视每集的叙事流程，并在适当的转换点插入标准化的场景标记。更正后的标注在之后的所有分析中均保持一致。例如，在 S02E22《两个派对的那一集》中，

> [回到钱德勒和乔伊的派对。大家都在跳舞，玩得很开心。]

更正为

> [场景：回到钱德勒和乔伊的派对。大家都在跳舞，玩得很开心。]

预处理后，我们统计了表1.中六位主要角色在所有剧集中的词元分布。

语料库子集	唯一词元	总词元数
瑞秋	6,390	106,735
莫妮卡	6,022	90,411
菲比	6,492	90,400
Joey	6,645	99,609
钱德勒	6,896	96,234
Ross	7,182	107,063
其他人	9,291	157,353
全剧	**20,049**	**747,805**

表1： 《老友记》对话语料库的描述性统计。唯一词元表示分词后不同词形的数量，而总词元数表示词语出现次数的总数。

清理后的语料库涵盖《老友记》十季的内容，共包含747,805个词元和20,049个唯一词元。我们的角色层面分析聚焦于六位主要角色：瑞秋、莫妮卡、菲比、乔伊、钱德勒和罗斯。他们的对话合计约占整个语料库的78.96%。

3.2 本体计量学框架

在 RQ1 和 RQ2 中，我们主要关注权力和危险，因为它们是本体计量空间的两个主要维度，并且能够对角色组合之间的社会互动模式提供最具可解释性的表示。因此，词典中的每个词包含两个连续值：权力分值和危险得分。权力—危险—结构词典的描述性统计数据见表2。

维度	词典	均值		SD	最小值	最大值
权力	20,006	0.000	0.218	-0.612	0.758	
危险	20,006	0.000	0.218	-0.591	0.681	
结构	20,006	0.000	0.097	-0.382	0.432	

表2： 本研究所使用的权力—危险—结构词典的描述性统计。

基于该词典，《老友记》语料库的覆盖率如表3所示。

唯一词元	词典中的唯一词元	唯一覆盖率
20,049	7,756	38.69%
总词元数	词典内总数	总覆盖率
747,805	222,786	29.79%

表3： 基于权力—危险—结构词典的《老友记》语料库词典覆盖率。

3.3 RQ1：主要角色的语义轨迹

为了回答我们的第一个研究问题：“单个角色的语义画像如何随时间在权力-危险空间中演变？”，我们将所有剧本数据汇总为每个角色的时间序列。每个角色在每一集中的所有词元，都会汇总为平均权力得分和平均危险度得分。

$$P_{c,t} = \frac{1}{N_{c,t}} \sum_{i=1}^{N_{c,t}} p_i, \tag{1}$$

$$D_{c,t} = \frac{1}{N_{c,t}} \sum_{i=1}^{N_{c,t}} d_i, \tag{2}$$

其中$N_{c,\ t}$ 是角色 c在第 t集所说的对话词元数量，而 p_i 和 d_i 分别表示分配给词元 i的权力得分和危险得分。

为了更清晰地展示这些弧线，我们采用 LOESS（局部估计散点图平滑） [21] 进行平滑处理，跨度参数设为 10 集。平滑处理分别独立应用于每个角色的权力和危险序列。令

$$\hat{P}_{c,t} = \mathrm{LOESS}(P_{c,t};\ s), \tag{3}$$

$$\hat{D}_{c,t} = \mathrm{LOESS}(D_{c,t};\ s), \tag{4}$$

其中，$s = 10$ 表示平滑跨度，对应于包含十集的局部邻域。由此得到的平滑轨迹直观呈现了整部剧集中每个角色的语义轨迹。

平滑弧线展示了渐进式模式，但无法在统计意义上指示角色的情感何时发生结构性转变。因此，我们尝试检测变化点。连续剧电视作品中的叙事发展通常由事件驱动 [22]。具体的情节事件，例如关系的开始或结束、职业里程碑或重大人生决定，都可能导致该角色说话和互动方式发生离散变化。变点检测可以定位这些变化发生的剧集，从而使我们能够考察检测到的变点在描述性层面上是否与可识别的叙事发展相吻合。

为定位时间序列中序列统计特性发生变化的点，我们将 PELT（剪枝精确线性时间）算法 [23] 应用于每个角色的剧集层面的力量和危险时间序列。PELT 通过其惩罚似然框架提供了这样的判据：只有当拟合效果的改善超过增加额外分段的成本时，才保留断点 [23]。这可以防止将细微波动过度解读为具有叙事意义的变化 [24]。惩罚值依据敏感性标准确定。我们针对 0.5、1.0、1.5 和 2.0 的惩罚值进行了敏感性分析。我们比较了这些设置下检测到的变点数量及其位置。通过考察相邻惩罚值之间变点数量和位置的稳定性，选定惩罚值（0.5）。

词语偏移分析 作为补充性的词汇分析，我们使用加权平均词语偏移，考察沿权力和危险维度，哪些词语能够将每个角色与其余五位角色区分开来。对于每个角色，参考分布由该角色的对话构成，而比较分布则包含其余五位主要角色的对话。加权平均词语偏移根据相对词频的差异以及每个词语的权力或危险分值，将两种词汇分布之间的差异分解为词语层面的贡献。因此，绝对贡献最大的词语能够指出在相应语义维度上最能区分这两种分布的词汇项。此处词语偏移分析的单位不同于主要关系计量分析所采用的剧集层级单位。

3.4 RQ2：角色组合中的情感模式

我们对权力—危险语义空间进行组合层级分析。六位主要角色共有63个非空子集，每个子集都代表一种独特的叙事结构。图 1展示了该剧十季中所有63种组合的场景总数。图 2 按场景数量降序展示了所有组合的场景数量和词数。

对于每个组合，我们汇总所有恰好由该组角色共同出现的场景中的权力和危险得分。

数据筛选 在整个系列中，角色组合的出现频率差异极大。一些二元组（由两个角色组成的群体）和全体角色阵容会出现数百次，而某些群体出现的场景少于 10 个。我们采用两项筛选标准，以确保权力和危险分值估计的统计可靠性。只有同时满足以下两个条件的组合才会被保留用于分析：（1）至少出现在 10 个场景中；（2）角色在这些场景中说出的词语总数至少为 2,000。该标准确保有足够的观测数据，从而从语义空间中获得稳定且具有意义的结果。正如图1 和图2所示，这些阈值是根据全部 63 个组合的场景数量和词数分布确定的。筛选过程最终保留了 56 个组合用于后续分析，在全面覆盖与统计可靠性之间取得了平衡。在附录 A1中，表 A2列出了为进一步分析而移除的七个组合，表 A1 则显示全部 56 个组合均保留用于分析。

组合得分计算： 对于每个组合 g，令 $\mathcal{S}_g$ 表示 g出现过的场景集合。对于每个场景 $s \in \mathcal{S}_g$，令 P_s 和 D_s 分别表示该场景中所有话语的平均权力分值和危险分值。组合 g 的汇总分值为这些场景级数值的平均值：

$$\bar{P}_g = \frac{1}{|\mathcal{S}_g|} \sum_{s \in \mathcal{S}_g} P_s, \qquad \bar{D}_g = \frac{1}{|\mathcal{S}_g|} \sum_{s \in \mathcal{S}_g} D_s$$

自助法置信椭圆： 为了评估权力-危险空间中各组合位置差异的统计显著性，我们采用了非参数自助法程序。对于包含 $n_g = |\mathcal{S}_g|$ 个观测场景的每个组合 g，我们通过重新抽样生成 $B = 10,000$ 个自助法样本

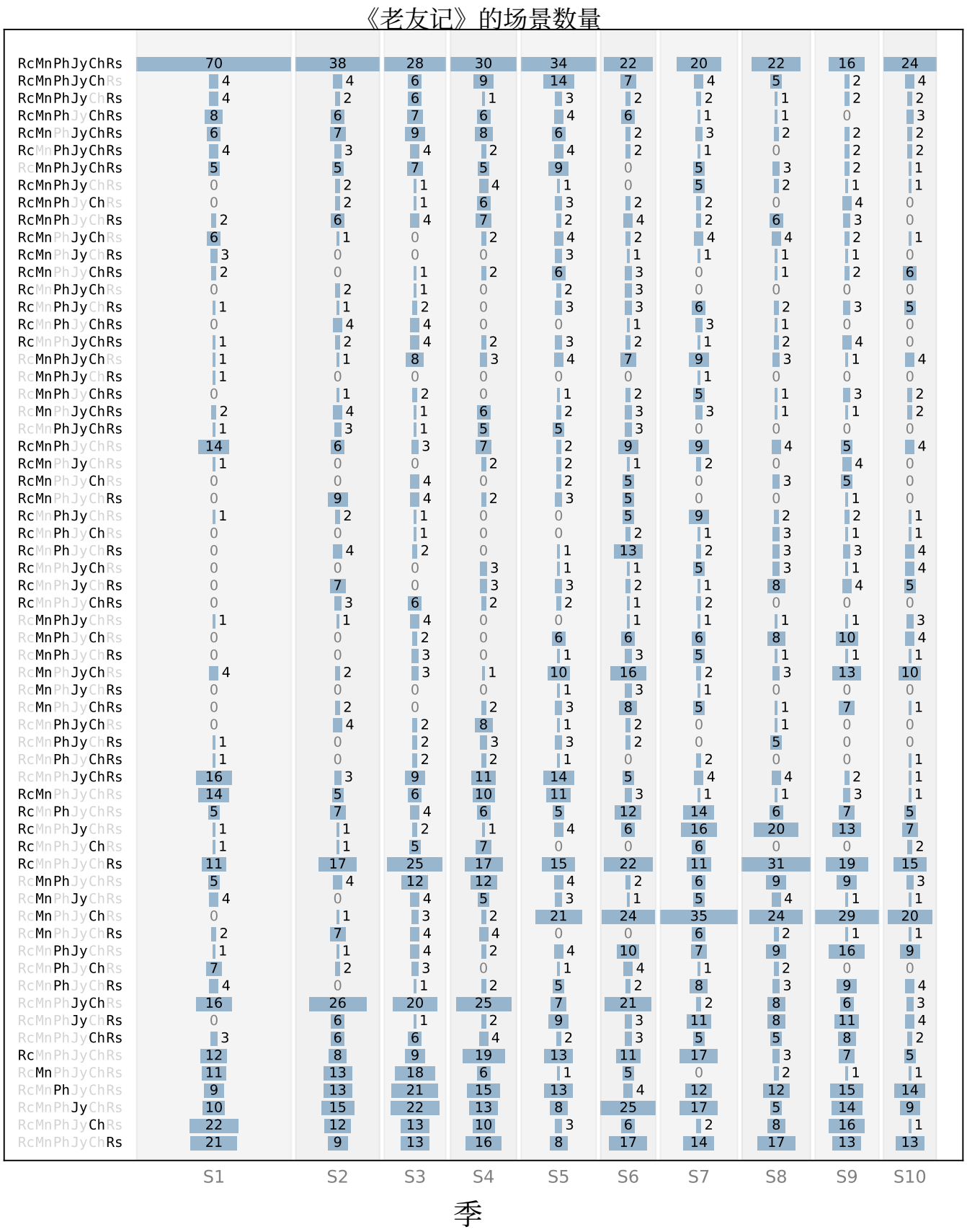


图1： 《老友记》各季的角色组合参与情况。该水平条形图展示了全部63种非空角色组合在每一季中的场景数量。各季 S1–S10）沿横轴排列，条形长度根据每一季中包含相应角色组合的场景数量进行缩放。

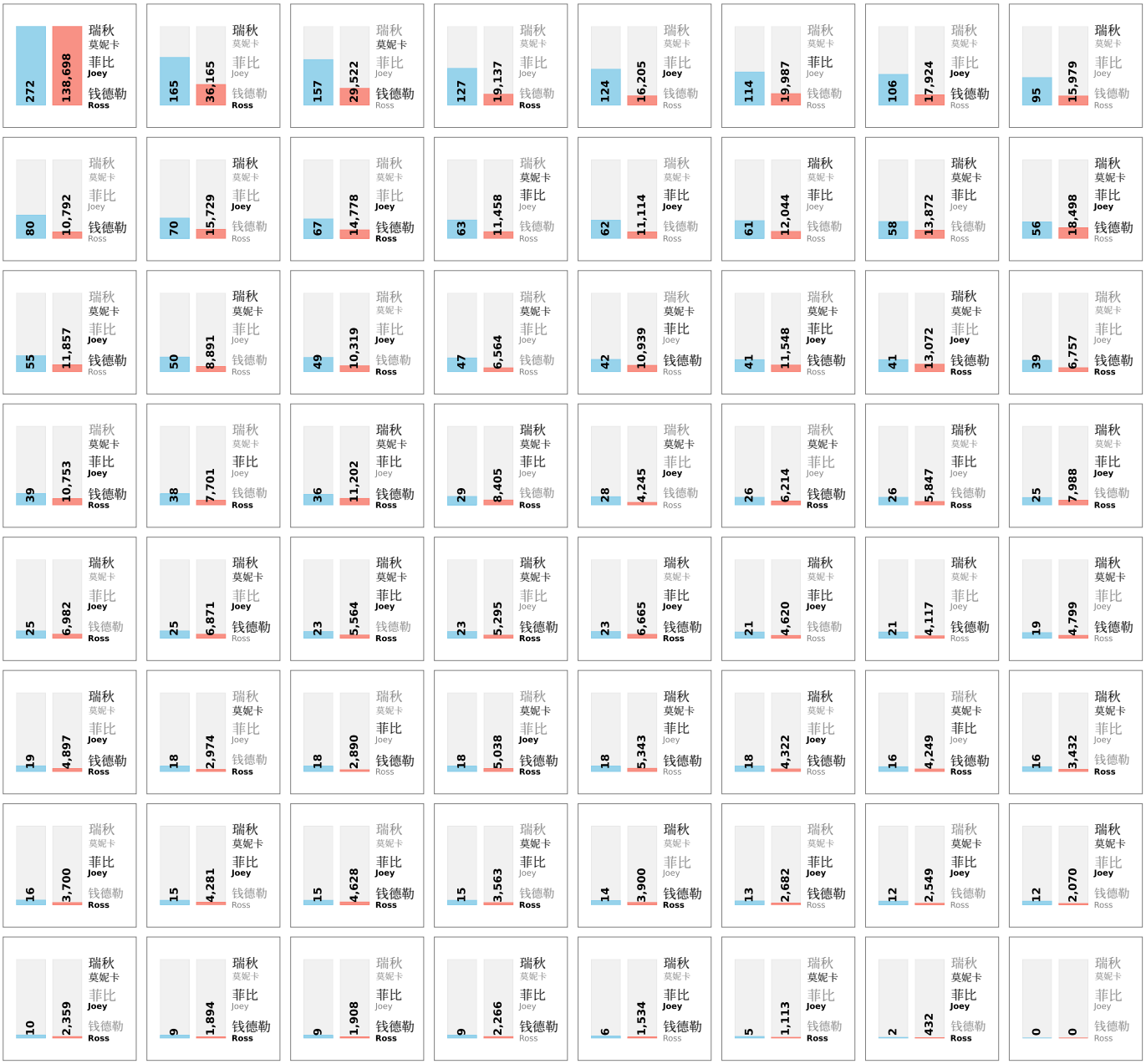


图2： 每个角色组合的场景数量和词数。蓝色柱表示场景数量，红色柱表示词数。各面板先横向排列，再向下排列，顺序按场景数量递减。

从其场景层级力量–危险坐标集合中进行有放回抽样 $\{(P_s, D_s)\}_{s \in \mathcal{S}_g}$。每个自助抽样都会生成一个均值向量 $\bar{\mathbf{X}}_g^{*(b)} = (\bar{P}_g^{*(b)}, \bar{D}_g^{*(b)})$，从而形成该组合估算语义位置的自助分布。

随后，根据自助法均值的协方差矩阵构建组合 g 的95%置信椭圆：

$$\hat{\Sigma}_g = \frac{1}{B-1} \sum_{b=1}^{B} \left(\bar{\mathbf{X}}_g^{*(b)} - \bar{\mathbf{X}}_g \right) \left(\bar{\mathbf{X}}_g^{*(b)} - \bar{\mathbf{X}}_g \right)^{\top}$$

其中，$\bar{\mathbf{X}}_g = (\bar{P}_g, \bar{D}_g)$ 是观测均值。置信区域由满足以下条件的所有点 $\mathbf{x} = (P, D)$ 定义：

$$(\mathbf{x} - \bar{\mathbf{X}}_g)^{\top} \hat{\Sigma}_g^{-1} (\mathbf{x} - \bar{\mathbf{X}}_g) \le \chi^2_{2,0.95}$$

其中，$\chi^2_{2,0.95} \approx 5.99$ 是自由度为 2 的卡方分布的第 95 百分位数。该椭圆涵盖了组合的力量–危险位置的真实总体均值有 95% 的置信度落入的区域。

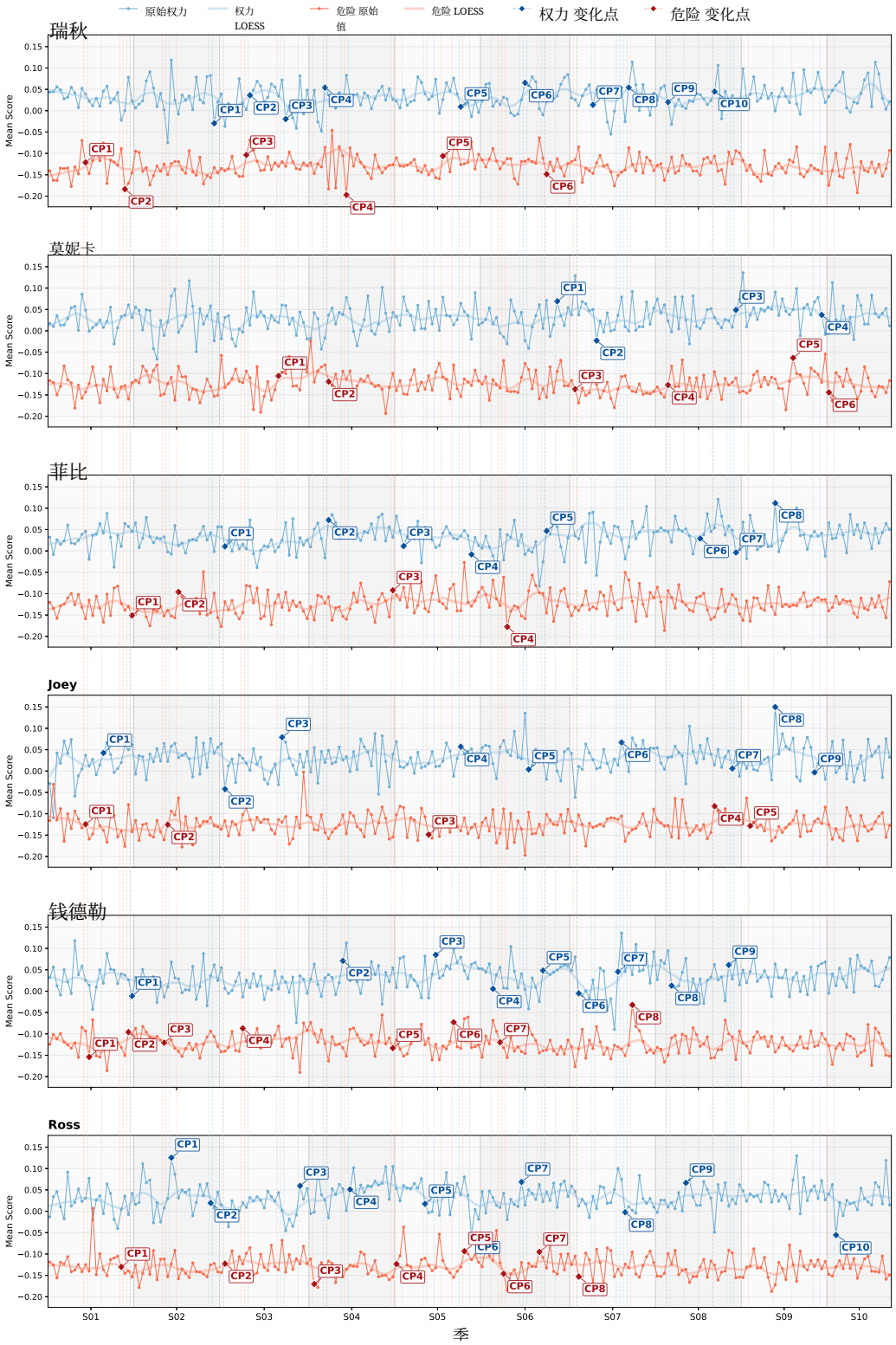


图3： 《老友记》中六位主要角色在所有剧集中的权力（蓝色）和危险（红色）轨迹。原始的每集平均分值以实心点和实线表示，而LOESS平滑曲线则概括了长期趋势。检测到的变化点表示相应权力或危险轨迹中的重大变化，这些变化由针对每个角色的变化点检测算法识别。垂直灰色带表示季边界。所有变化点的详细信息见附录 A2。

每个置信椭圆的面积，计算方式为$A_g = \pi \cdot \chi^2_{2,0.95} \cdot \sqrt{\det(\hat{\Sigma}_g)}$，可作为估计精度的指标。也就是说，组合出现在更多场景中会产生更小的椭圆，从而更精确地估计位置。

我们还需要定义一个全局质心，以确定哪些组合占据的位置在统计上与整个群体不同。让 $\bar{\mathbf{X}}_{\text{global}} = (\bar{P}_{\text{global}}, \bar{D}_{\text{global}})$表示全部56个保留组合均值的均值。该质心代表群体在权力-危险空间中的平均语义位置。因此，如果某个组合 g 的95%置信椭圆不包含$\mathbf{X}_{\text{global}}$，则该组合被归类为显著偏离全局质心。为刻画偏离方向，我们根据相对于全局质心的位置，将这些显著组合分为四个象限：高权力 + 低危险（$\bar{P} > P_{\text{global}},\ \bar{D} < D_{\text{global}}$）、高权力 + 高危险（$\bar{P} > P_{\text{global}},\ \bar{D} > D_{\text{global}}$）、低权力 + 低危险（$\bar{P} < P_{\text{global}},\ \bar{D} < D_{\text{global}}$）以及低权力 + 高危险（$\bar{P} < P_{\text{global}},\ \bar{D} > D_{\text{global}}$）。

组合级语义空间使用二维散点图进行可视化。每个保留的组合都显示为其观测均值坐标$(\bar{P}_g, \bar{D}_g)$处的一个点。每个点周围叠加95%自助法置信椭圆，并按椭圆面积缩放以表示估计精度。全局质心标记为参考点。椭圆不包含全局质心的组合将突出显示并标注。

角色组合的轨迹聚类分析：为捕捉组合轨迹中的动态模式，我们考察每个组合的力量–危险位置在十季叙事时间线中的变化。鉴于 RQ1 已证实单个角色会随剧集发展而变化，我们现在进一步判断组合是否呈现出某些变化模式，并能否将这些模式归纳为可解释的轨迹类型。

我们对角色组合的季节性力量—危险序列进行轨迹聚类分析。分析基于已识别出的 56 个角色组合展开。对于每个组合，我们按季汇总力量和危险分数，得到一组季均值数值$\bar{P}_{g,\,s}$ 并且$\bar{D}_{g,\,s}$ 对于每个组合 g 在第s（$s = 1$、2、.....、10）。为确保轨迹稳定性，我们进一步设定两项纳入标准。对于某个组合，如果该组合在某一季至少出现在三个场景中，则该季被视为有效。若某个组合在十季中至少有八季符合第一项标准，则将其纳入轨迹分析。在 RQ2 考察的 56 个组合中，有 15 个符合这两项标准，如附录 A5中的表格所示，附录 A3。

对于保留的组合，采用线性插值填补缺失的季节值，以确保所有轨迹均表示为完整的十季序列。对于无法插值的端点值，则以前向或后向填充的方式，采用最近一个观测季节的值。因此，每条轨迹都编码为一个10维序列 $T_g = \{\bar{P}_{g,s}, \bar{D}_{g,s}\}_{s=1}^{10}$，其中每个时间点都是由权力和危险分数组成的二维向量。在进行距离计算之前，使用z分数标准化对权力和危险分数进行标准化。

使用多变量动态时间规整（DTW） [25, 26]评估轨迹相似性。采用宽度为两季的Sakoe-Chiba带来约束对齐 [25]。

随后，基于DTW距离矩阵 [27]，使用完全连接法进行层次聚类。通过综合考量从2个到10个聚类的候选方案中的轮廓系数，确定聚类数量 [28]。

3.5 RQ3：角色的语义特征是否会在各剧集中系统性地协变?

为考察角色是否会随时间呈现协调的语义变化，我们使用原始的剧集层面的力量、危险和结构（PDS）时间序列。本分析关注主要角色的逐集波动是否存在协变。对于每个维度，我们计算六位主要角色中每一对角色时间序列之间的皮尔逊相关。六位角色在每个维度上可产生 $\binom{6}{2} = 15$ 个两两相关。正相关表示，相较于各自的序列均值，两个角色在各剧集中往往朝同一方向变化；负相关则表示相反的协变。统计显著性通过对 H_0 进行双侧检验来评估： $\rho = 0$，并对每个维度内的15项两两比较应用邦费罗尼校正。

由于单集级话语计量分数是根据有限样本的对话词元估算得出的，我们还使用分半信度 [29, 30]评估 PDS时间序列的信度。对于每个剧集中的每个角色，词典匹配词元会被随机分成两个数量大致相等的子集。分别计算两半词元的权力、危险和结构分值，从而为每个角色和维度生成两条剧集层面时间序列。随后，使用 Pearson 相关系数 r计算两条半序列在各剧集间的相关性。由于每个相关系数都基于可用词元的一半，因此使用 Spearman–Brown校正 [29, 30],对所得系数进行调整。

$$r_{\text{SB}} = \frac{2r_{\text{half}}}{1 + r_{\text{half}}}, \tag{5}$$

其中， r_{half} 是两个独立评分部分之间的相关性。这会产生18个信度

估计值，分别对应三个PDS维度中的六个角色，并提供角色层面相关性分析所依据的测量稳定性估计。

最后，为了区分角色对特有的协变与群体中各角色共同呈现的变化，我们在控制其余四个角色后，计算每对角色之间的偏相关。偏相关根据原始单集级角色轨迹的协方差矩阵计算。对于每个 PDS 维度，令 Σ 表示六个角色轨迹的 6×6 协方差矩阵，令 $\Omega = \Sigma^{-1}$ 表示其精度矩阵。给定其余四个角色后，角色 i 与 j 之间的偏相关为

$$r_{ij \cdot \text{rest}} = -\frac{\Omega_{ij}}{\sqrt{\Omega_{ii}\Omega_{jj}}}. \tag{6}$$

这些偏相关量化了两条角色轨迹之间的残余线性关联，即在考虑它们与群体中其余角色的线性关系后所剩的关联。显著性通过双侧检验评估，并对每个维度内的 15 组角色配对比较进行邦费罗尼校正。

4 结果与讨论

4.1 RQ1：主要角色的情感弧线

4.1.1 整体权力—危险概况

我们考察将全剧汇总后，六位主要角色是否呈现出可区分的权力—危险度概况。在该分析中，我们将每集视为重复测量设计中的区组因素，并将角色作为被试内变量。分析共纳入236集，得到1,416个角色—剧集观测值。表4 呈现了每个角色在单集级别的平均权力和危险分值。

角色	权力 ($M \pm SD$)	危险 ($M \pm SD$)
瑞秋	0.0325 ± 0.0314	-0.1289 ± 0.0246
莫妮卡	0.0311 ± 0.0337	-0.1234 ± 0.0271
菲比	0.0330 ± 0.0308	-0.1222 ± 0.0271
Joey	0.0288 ± 0.0336	-0.1250 ± 0.0264
钱德勒	0.0316 ± 0.0338	-0.1215 ± 0.0248
Ross	0.0323 ± 0.0326	-0.1277 ± 0.0262

表4： 六位主要角色的权力和危险分值的描述性统计。

为检验观察到的描述性差异是否达到统计显著性，我们分别对权力和危险进行重复测量方差分析，将剧集视为重复测量单位，并将角色作为被试内因素。Mauchly检验表明，无论是权力还是危险，均没有证据显示违反球形性假设。

对于权力而言，角色的主效应不显著，$F(5, 1175) = 0.587$，$p = 0.710$，表明没有证据显示六位主要角色在剧集层面的力量分值上存在系统性差异。相比之下，角色对危险的影响具有统计显著性，$F(5, 1175) = 3.568$，$p = .003$，但相关效应量较小（$\eta_p^2 = .015$）。

我们分别针对权力和危险，对全部 15 对角色进行配对 t 检验，并在每个结果指标内应用邦费罗尼校正。校正后，权力方面没有任何角色对之间的差异达到显著水平。对于危险，只有瑞秋–钱德勒的比较在邦费罗尼校正后仍具有显著性，$t(235) = -3.421$，$p_{\text{adj}} = .011$。钱德勒的平均危险度得分（$M = -.1215$）略高于瑞秋（$M = -.1289$）。危险方面的其他角色对差异均未能通过校正。

总体而言，在剧集层面，六位角色大致处于权力–危险空间的同一区域。

值得注意的是，六位角色的平均权力得分均为正值，平均危险度得分均为负值，因此都位于该空间中共同的强势—安全区域。在本体计量术语中，这一共同基线可能反映了角色在整部剧集中反复参与的共同对话环境。作为一部群像情景喜剧，《老友记》将角色置于反复出现的社交场景和对话互动中，其中许多互动涉及对话幽默的形式，例如反讽及其他以幽默为导向的交流 [31]。这些共同的叙事和对话语境，可能促成了六位主要角色所处的大致相似的强势—安全区域。

4.1.2 词汇差异：词语迁移分析

尽管总体的权力—危险度概况大体相似，但根据对话内容，每个角色在词汇模式上仍存在差异。为描述性地考察这些差异，我们使用加权平均词语偏移，将每个角色与其余五个角色的合并对话进行比较。

图4 展示了钱德勒权力维度的一个示例。“词汇偏移”标识出对钱德勒与其他五个角色的对话之间差异贡献最为显著的词汇项。例如，对钱德勒而言，相对频繁地使用

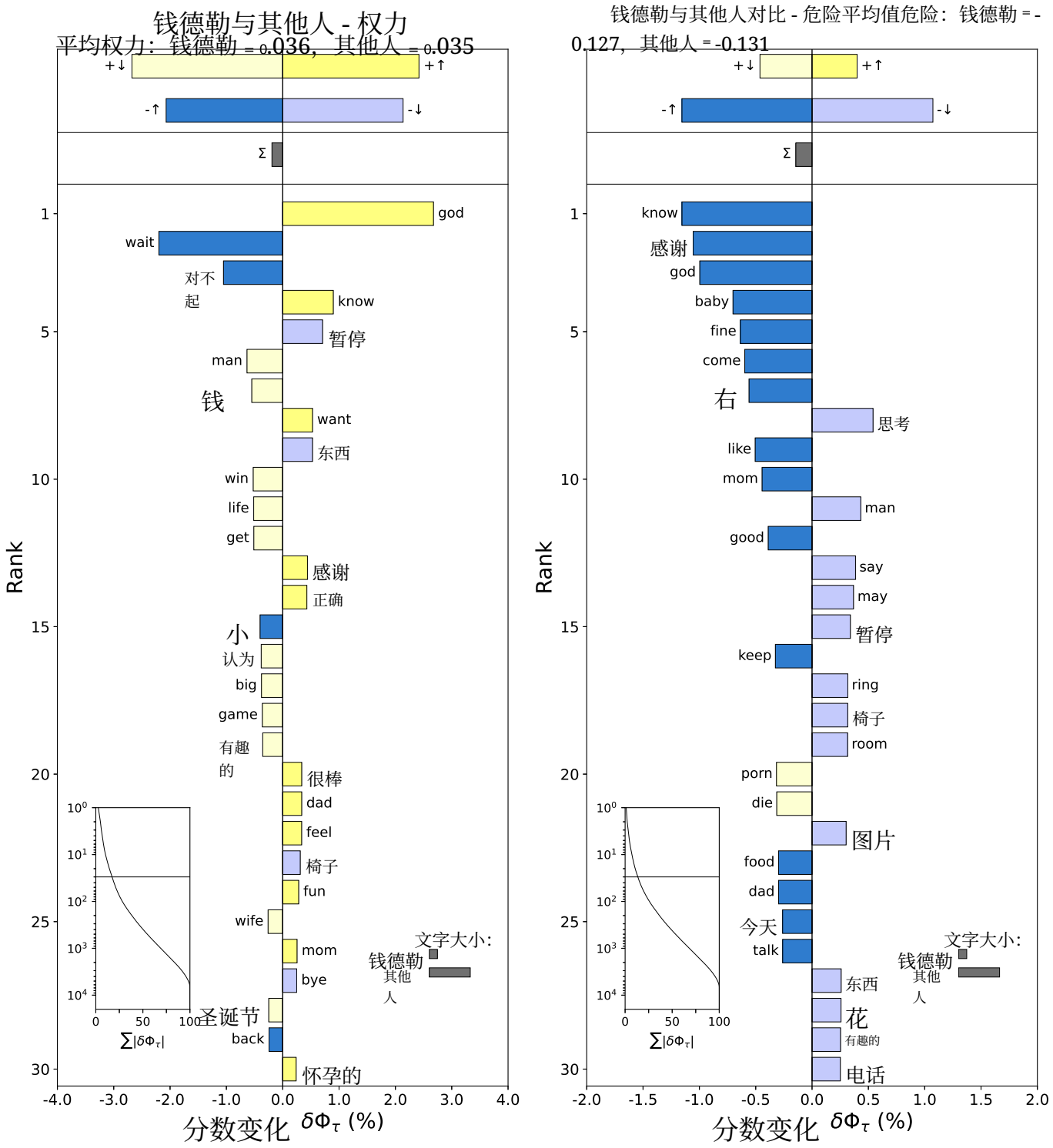


图4： 加权平均词语偏移，对比钱德勒的对话与其余五位主要角色合并后的对话在权力和危险度维度上的差异。黄色表示权力/危险得分相对较高的词语，蓝色表示权力/危险得分相对较低的词语。深浅不同的色调表示钱德勒使用这些词语的频率相对较高或较低。右侧的贡献会提高钱德勒的相对权力/危险度，左侧的贡献则会降低。图例还包括两侧的文字大小。

较高权力词“上帝”会提升其权力值，而相对频繁地使用“等待”和“对不起”等较低权力词则会使其权力值下降。此外，图中显示的均值是汇总得出的，

用于词移分解的按词元频率加权的词汇均值，因此不同于前述分析中报告的单集级均值；在前述分析中，每集构成一个分析单位。该

因此，这种词移揭示了被汇总后的力量分数所掩盖的词汇贡献。

4.1.3 主要角色的情感弧线

为了捕捉这些静态概况背后的叙事动态，我们考察每个角色的权力得分和危险得分如何随236集时间线演变。对于每一集，我们根据该集中的所有对话，计算每个角色的平均权力得分和危险得分。这些原始的单集级序列保留了数据的完整时间分辨率，但由于每集特定的主题、喜剧支线和偶然互动，呈现出显著的短期变异。为提取潜在的叙事信号，我们将LOESS（局部估计散点图平滑） [21] 应用于每个角色的权力和危险序列。平滑参数（跨度 = 10集）经过选择，以保留季节性趋势。图3 展示了全部六个角色的轨迹及平滑曲线。

趋势解读： 我们观察到，每个角色都呈现出某些形态特征。例如，瑞秋的危险轨迹在第一季的部分时期相对较高，随后逐渐下降。这一时间模式总体上与她早期叙事中逐步走向更加独立的转变相吻合 [7, 32]。对于莫妮卡而言，从第5季到第8季，权力得分呈上升趋势，并在第9季再次上升；这一时期与她和钱德勒关系的发展及结婚过程有所重叠 [7, 33]。对于菲比而言，第4季的权力得分出现低点，达到局部最低值。乔伊在两个维度上的变化最小，其得分在整个剧集中始终接近基线。钱德勒呈现出最明显的正向权力趋势，在第5季至第7季之间加速上升。根据这些观察，这些轨迹在时间变化的方向和幅度上存在描述性差异。

虽然 LOESS 平滑揭示了连续趋势，但无法识别情感轨迹发生结构性转变的离散时刻。为检测剧集中标志着权力或危险分值显著变化的此类变化点，我们对每个角色的时间序列应用了 PELT（剪枝精确线性时间）算法。表 A3 和附录A4 中的A2 表列出了所有检测到的变化点及其对应的剧集。

可能的变化点解读 检测到的变化点（变点）也与剧中的重大叙事事件呈现出描述性对应关系。如图 3 所示，钱德勒的权力轨迹是一个典型例子。他的大多数检测到的变点都出现在权力变化发生在他的人际关系和恋爱故事线的转折处。其中几个变点出现在该剧中间几季，尤其是第4季至第8季之间，此时钱德勒与莫妮卡的主要恋爱关系故事线逐渐发展。例如，CP2（S04E10）出现在他与凯西交往期间，而CP3（S05E12）则处于他与莫妮卡恋情的初期。之后，S06E04、S06E18、S07E03和S07E14处出现的变点，正值钱德勒与莫妮卡的关系逐步发展、迈向更深承诺的阶段，其中包括他们订婚前后的时期。因此，这些变点与重大叙事发展的接近表明，钱德勒词汇力量特征的某些变化与其长期恋爱关系弧线中的转折相吻合。其他角色的变点模式也为这些解读提供了补充性的描述依据。尤其是，钱德勒轨迹中识别出的若干变化，出现在莫妮卡的权力轨迹中独立检测到的变点附近，而这些变点对应着两人共同关系弧线的相关时期（莫妮卡权力轨迹中的CP1、CP2、CP3）。

4.2 RQ2：角色组合中的情感模式

4.2.1 角色组合的空间分布

接下来，我们考察不同角色组合在权力—危险语义空间中占据的不同区域。对于符合条件的 56 种组合，我们分别计算这些角色组合出现过的所有场景中的平均权力和危险得分。由此绘制出群体结构的空间分布图，其中每种组合都占据一个坐标，反映其典型的人际立场。

图 5 展示了所有保留的角色组合在力量–危险位置上的分布。图5中，每种组合均以一个六边形表示，分布于权力-危险空间中。每个六边形的中心表示该组合的平均值，颜色则表示其代表的组合。这一图示清晰直观地呈现了不同角色组合在权力-危险空间中占据的不同情感区域。

空间分布显示，首先，组合集中在危险度维度的下半部分——只有两个组合的危险得分高于 -0.14，且没有任何组合接近零或呈现正的危险得分。这种在所有群体中对低危险度的系统性偏向表明，无论哪些角色正在互动，该剧的对话都维持着一种与其日常喜剧群像基调相一致的人际安全感基线。不过，在没有进一步证据的情况下，不应将这一基线解读为该剧固有的

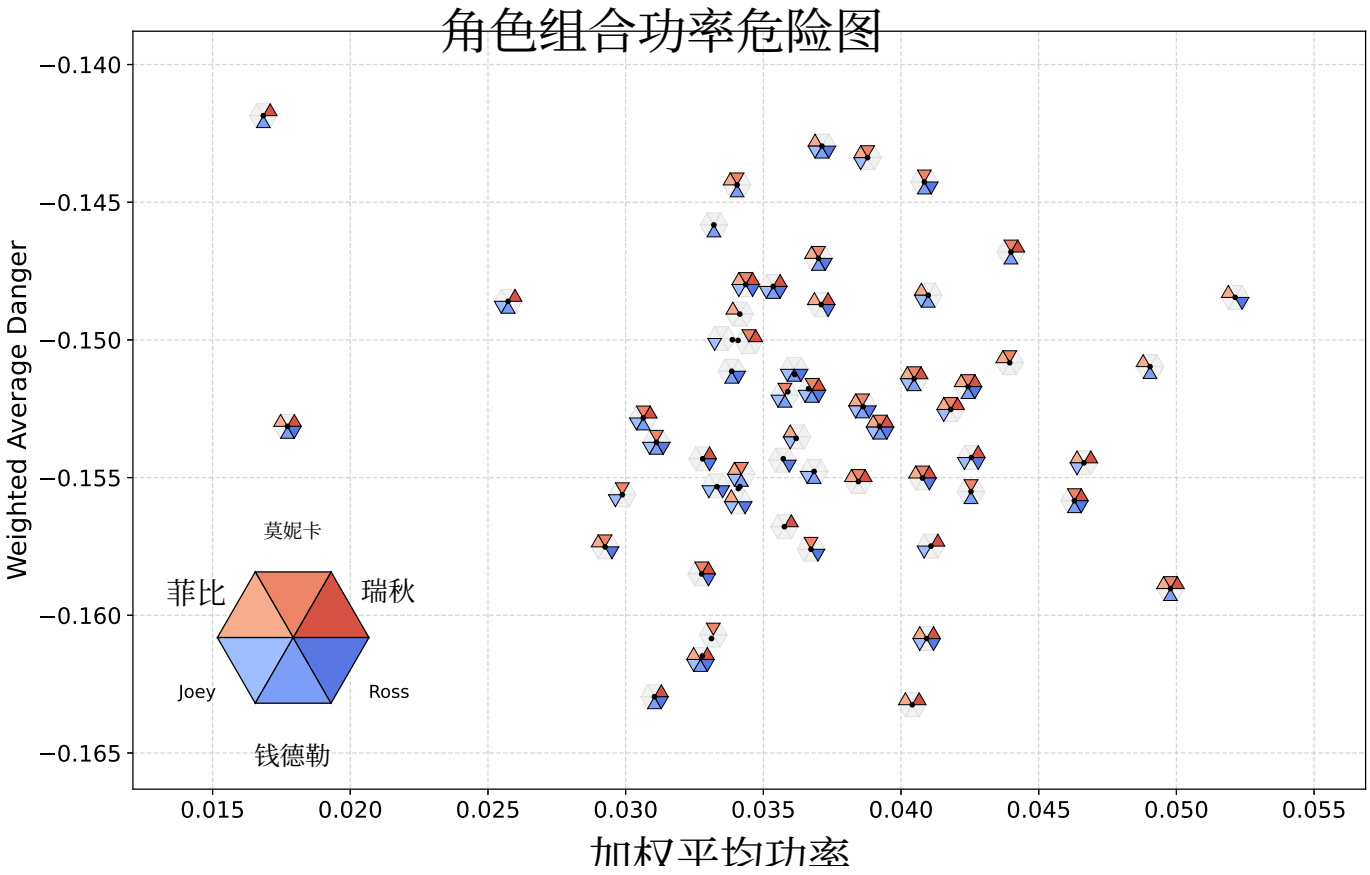


图5： 保留的56个角色组合的平均权力–危险位置。每个六边形代表一个角色组合，并使用不同颜色区分六个不同角色。每个六边形中心点的坐标表示平均权力得分和平均危险度得分。为便于可视化，仅对重叠标记应用了小幅位置偏移。

特性，因为词典中的零点是按词型计算的未加权平均值，而按普通频率加权的英语本身就位于其安全一侧。然而，要判断这一模式是否更普遍地反映了情景喜剧类型，还需要与其他叙事语料库进行比较。

其次，权力维度的离散度高于危险维度。权力分数范围从0.0137（瑞秋+菲比+钱德勒+罗斯）到0.0566（菲比+罗斯），约为危险度范围的四倍。从描述性统计来看，与危险度相比，权力在不同角色组合之间的变化范围更大，这表明在本语料库中，权力可能对群体构成更加敏感。

我们还考察了在每个维度上处于极端位置的组合，无论其置信椭圆是否排除了质心。这些极端组合确定了《老友记》中群体的语义边界。在权力维度上，考察权力最高的 5 个组合后，我们注意到一个模式：每个组合中都有菲比。这一模式表明，包含菲比的组合往往位于所观测权力范围的较高端。在危险度维度上，

危险度最低的组合都包含瑞秋，并且都不包含莫妮卡。这些模式在描述上与她们支持性友谊动态相一致。

4.2.2 自助法置信椭圆

接下来，我们为每个组合应用自助法95%置信椭圆。对于每个角色组合，我们以有放回抽样的方式对其场景进行10,000次重抽样。然后，我们重新计算每个自助抽样的平均权力和危险坐标，并使用自助分布的协方差矩阵定义置信区域。结果显示，56个组合中有5个组合的结果具有显著性，这意味着它们的置信椭圆不包含全局质心。

然而，经过多重比较的邦费罗尼校正后，这些结果均不显著。没有任何组合通过邦费罗尼校正。

4.2.3 角色组合的轨迹聚类

在 RQ2 中分析的 56 种角色组合里，有 15 种符合轨迹纳入标准：至少出现在八季中，且每个有效季至少有三个场景。这 15 种组合包括剧集中最常见的群体搭配，共占群体对话场景的绝大多数（详情见附录表A5，附录A3）。

对这15个组合的季节性权力—危险轨迹计算得到的多变量DTW距离矩阵应用采用“complete”连接的层次聚类 [26]。对从2到10的聚类数量评估轮廓系数，并在四个聚类时达到最大值0.2668 [28]。最大轮廓系数仅为0.2668，表明聚类分离度较弱，几乎没有证据支持存在明确定义的聚类结构。

4.3 RQ3：角色的语义画像是否在各剧集中系统性地共同变化?

4.3.1 角色级皮尔逊相关性

我们首先考察六位主要角色在权力、危险和结构方面是否表现出协调的逐剧集变化。图6 展示了从原始剧集层面时间序列计算得出的两两皮尔逊相关，并对每个维度内15组角色配对比较进行了邦费罗尼校正。

对于权力，仅有三对呈现显著的正相关性：莫妮卡–瑞秋（$r = 0.22$）、莫妮卡–菲比（$r = 0.24$）和莫妮卡–钱德勒（$r = 0.26$）。我们注意到，所有显著的权力关联都涉及莫妮卡，尽管这些相关性的幅度较小。对于危险，仅有两对仍然显著：莫妮卡–钱德勒（$r = 0.20$）和乔伊–钱德勒（$r = 0.28$）。结构呈现出更广泛的显著协变模式，其中六对在邦费罗尼校正后仍然显著：瑞秋–乔伊（$r = 0.22$）、莫妮卡–乔伊（$r = 0.24$）、菲比–钱德勒（$r = 0.22$）、菲比–罗斯（$r = 0.21$）、乔伊–钱德勒（$r = 0.21$）以及钱德勒–罗斯（$r = 0.20$）。总体较弱的相关性表明，这种共同变化并不意味着强烈的成对同步。然而，在邦费罗尼校正后仍然显著的关联揭示了更为局部的时间协变模式。

在三个维度中都可以明显看出一种一致的模式：全部45个成对相关系数均为正。这种系统性的正向趋势表明，角色的单集级轨迹可能包含一个共同成分，因此角色往往会朝相同的方向变化。

因此，我们进一步评估单集级轨迹的测量信度，并使用偏相关来确定在控制与其他角色共享的变异后，任意成对关联是否仍然存在。

4.3.2 分半信度

在进一步解读这些成对相关之前，我们使用分半信度评估单集级PDS轨迹的测量信度。在18个角色–维度组合中，随机划分的两半词元之间的所有相关性均为正且具有统计显著性。应用Spearman–Brown校正后，信度估计值范围为0.330 至0.619，表明单集级测量具有中等稳定性。结果见附录A6中的表A4

信度在不同维度和角色之间均存在差异。Power的平均Spearman–Brown信度最高（$\bar{r}_{SB} = 0.524$），高于Structure（$\bar{r}_{SB} = 0.445$）和Danger（$\bar{r}_{SB} = 0.437$），这表明，与Danger或Structure的变化相比，该维度上逐集变化的测量更为一致。最高的单项信度出现在罗斯的Power轨迹（$r_{SB} = 0.619$），其次是瑞秋的Power（$r_{SB} = 0.572$）、乔伊的Structure（$r_{SB} = 0.567$）和乔伊的Power（$r_{SB} = 0.565$）。最低的估计值出现在钱德勒的Danger（$r_{SB} = 0.330$）、瑞秋的Structure（$r_{SB} = 0.344$）和菲比的Power（$r_{SB} = 0.377$）。这些结果表明，单集级PDS轨迹在独立对话中包含可重复的信号。因此，可以结合这种中等程度的测量信度来解读较弱的成对相关。

4.3.3 偏相关

为确定皮尔逊相关中观察到的成对关联是否反映了特定角色对之间的关系，我们在控制其余四个角色的情况下进一步计算了偏相关。如图 7所示，在将群体中的其他角色纳入考量后，显著关联的数量大幅减少。对于权力而言，只有莫妮卡–钱德勒的正向关联在邦费罗尼校正后仍然显著（$r_{\text{partial}} = 0.21$）。相比之下，标准皮尔逊分析发现了三项显著的权力关联：莫妮卡–瑞秋、莫妮卡–菲比

，以及莫妮卡–钱德勒 . Thus, 尽管

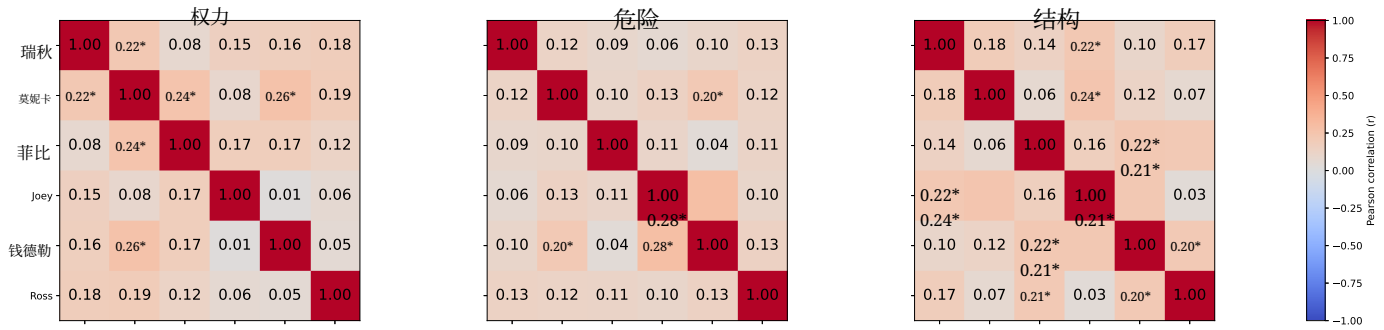


图6： 根据原始剧集层面时间序列计算得出的六位主要角色在剧集层面的力量、危险和结构轨迹之间的两两皮尔逊相关。红色表示正相关，蓝色表示负相关。正值表示两个角色在不同剧集中往往朝相同方向变化。星号表示在每个维度内进行15次两两比较并应用邦费罗尼校正后仍具有统计显著性的相关性（$p_{\text{adj}} < 0.05$）。

莫妮卡与多个角色呈现正向协变，但在控制其他四个角色的轨迹后，只有她与钱德勒的关联仍具有统计显著性。

危险维度也出现了类似的减少。已识别出的两项显著关联中，只有乔伊–钱德勒在偏相关分析中仍具有显著性（$r_{\text{partial}} = 0.26$）。对于结构维度，15项偏相关均未通过邦费罗尼校正。因此，偏相关结果大幅削弱了标准相关分析中观察到的两两结构。

尽管如此，在控制其余四位角色的轨迹后，力量维度中的莫妮卡–钱德勒关系和危险维度中的乔伊–钱德勒关系仍可被检测到，且均为正向，这为有限的角色对特定条件关联提供了证据。这两项保留下来的关联实际上对应着叙事中广为人知的关系。莫妮卡和钱德勒从朋友发展为长期恋爱伴侣，而乔伊和钱德勒在整部剧集中也维持着亲密友谊。他们反复共同出现在场景和叙事情境中。

单集级轨迹显示出显著的时间变化，多个角色都出现了多个变点。其中一些转变发生在可识别的叙事发展附近，这表明 ousiometric 轨迹的变化可能与叙事语境的变化相对应。

在组合层面，力量的变化范围比危险更广，并且似乎对群体构成的变化更加敏感。然而，不同角色组合之间的差异并不足以证明存在明显不同的组合类型，而轨迹聚类同样显示出很少具有稳健性或清晰分离的结构。

在两两层面，角色轨迹通常会在不同剧集中朝同一方向变化，但这些关联程度有限；在控制与群体中其他角色共享的变异后，这些关联大多会减弱。只有力量维度中的莫妮卡—钱德勒组合和危险维度中的乔伊—钱德勒组合仍保留可检测的配对特异性条件关联，而结构则未显示出类似的配对特异性关系。这表明，主要角色的时间协变主要由群体中共享的变异所驱动。

5 结论

本研究将本体计量学的力量—危险框架应用于《老友记》的完整十季对话语料库，并从三个层面考察语义动态：单个角色、角色组合和两两轨迹；此外，在两两分析中还纳入了结构。

在个体层面，将全剧中的对话汇总后，六位主要角色彼此相似，整体上都处于语义空间中相近的力量—安全区域。

这些发现将本体计量学的应用从描述语言中的整体模式，拓展到考察长篇群体叙事中的语义变化。通过沿着可解释的语义维度呈现对话，该框架使我们能够在共同的语义空间内比较语言如何随角色、对话分组和时间而变化。在《老友记》中，这一视角表明，权力和危险作为不断变化的叙事与互动语境的维度，更具解释力。因此，本体计量学为计算叙事分析提供了一种表征方式，将语言中可测量的变化与不断演变的语境联系起来。

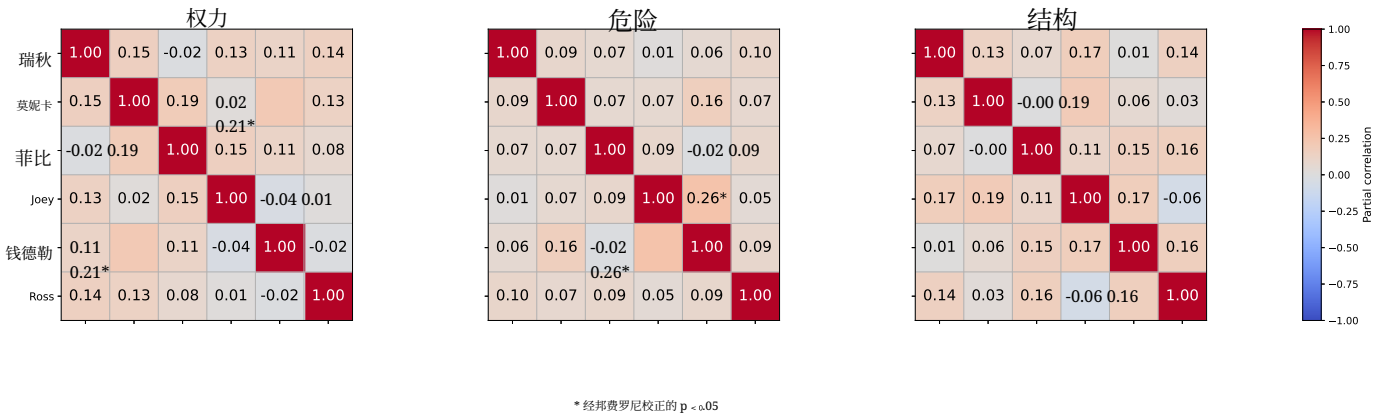


图7： 六位主要角色在剧集层面的力量、危险和结构轨迹之间的偏相关。对于每一对角色，偏相关均根据原始的剧集层面时间序列计算，同时控制其他四位角色的轨迹。红色表示正相关，蓝色表示负相关。正值表示，在考虑群体中其余角色相关的变化后，两位角色往往朝相同方向变化。星号表示在对每个维度内的15组两两比较进行邦费罗尼校正后，仍具有统计显著性的相关性（$p_{\text{adj}} < 0.05$）。

6 局限性与未来工作

本研究存在若干局限性，同时也为未来研究指明了方向。首先，分析基于单部电视剧。因此，若干发现可能反映的是《老友记》或情景喜剧类型特有的属性，而非叙事互动的一般属性。要确定这些模式能否推广至《老友记》之外，有必要对多部电视剧及多种类型进行比较分析。例如，这类比较可以检验权力—安全偏差是否也存在于其他情景喜剧中，以及以冲突或威胁为核心的类型（如犯罪剧）是否会在权力–危险空间中产生系统性不同的分布。

其次，互动分析仅限于六位主要角色：瑞秋、莫妮卡、菲比、乔伊、钱德勒和罗斯。因此，63种可能的配置及后续的组合层面分析，均仅以这一核心群体为范围。《老友记》的场景中经常出现反复登场的配角，例如恋爱伴侣、家庭成员和同事，他们的出现可能会改变对话中的人际动态。因此，归为同一种主要角色组合的两个场景，其完整的社会配置可能并不相同。未来的研究可以将配角纳入组合，并考察第三方角色的出现是否会改变互动的权力—危险特征。

第三个局限性源于本体计量学本身基于词典的特性。每个词汇项都被赋予一个相对固定的分值，而对话或场景的语义位置则通过汇总这些与语境无关的词汇值来估算。这种表征无法充分捕捉对话中意义的语境变异性。同一个词根据语用语境的不同，可能传达出不同程度的权力或人际危险。这一点对于情景喜剧尤其如此，因为其中常见讽刺和反讽，而且话语的互动意义可能与其字面内容存在显著差异。本研究分析的权力—危险分数应仅被解读为词典衍生的语义信号。未来工作可以在会话层面发展语境化权力—危险测量。此类扩展将有助于确定哪些互动模式能够经受语境变异。

最后，电视叙事会通过文字以外的多个渠道传达意义，包括面部表情、手势和身体动作。这些信号可能强化或反驳对话的语义内容，尤其是在反讽和肢体喜剧都十分重要的情景喜剧中。将这一框架扩展到多媒体层面，将有助于更全面地说明视听叙事中的互动意义是如何构建的。

致谢

作者感谢美国国家科学基金会奖项 #2242829（在线语料、知识与故事科学）、MassMutual 提供的基础支持，以及一笔匿名慈善捐赠。

参考文献

[1] M. Taboada。情感分析：语言学视角综述。2016年。[2] S. M. Mohammad。9——情感分析：从文本中检测效价、情绪及其他情感状态。载于 H. L. Meiselman 编，《style id='3'>情绪测量》，第 201–237 页。Woodhead Publishing，2016年。[3] P. Stone、J. Kirsch 和 C. C. Associates。《style id='6'>通用探究者：内容分析的计算机方法。《通用探究者：内容分析的计算机方法》。麻省理工学院出版社，1966年。[4] A. M. van der Veen 和 E. Bleich。机器学习时代基于词典的情感分析的优势。《style id='9'>PLOS ONE》，20，2025年。[5] A. Esuli 和 F. Sebastiani。Sentiwordnet：一种公开可用的观点挖掘词汇资源。载于国际语言资源与评测会议，2006年。[6] P. S. Dodds、T. Alshaabi、M. I. Fudolig、J. W. Zimmerman、J. Lovato、S. Beaulieu、J. R. Minot、M. V. Arnold、A. J. Reagan 和 C. M. Danforth。本体计量学：意义的本质与权力—危险—结构框架相契合，而非效价-唤醒-支配框架。《style id='15'>Science Advances》，12(19):eadr4039，2026年。[7] D. Crane 和 M. Kauffman。《老友记》。电视剧，1994年–2004年。Bright/Kauffman/Crane Productions 和华纳兄弟电视公司；最初由 NBC 播出。[8] J. Eder、F. Jannidis 和 R. Schneider。《虚构世界中的角色：理解文学、电影及其他媒体中的想象性人物》。2010年。[9] U. Margolin。行动者与行动：作为叙事人物塑造基础的行动。《style id='20'>Poetics Today》，7:205，1986年。[10] D. Bamman、B. T. O’Connor 和 N. A. Smith。学习电影角色的潜在角色人格。载于计算语言学协会年会，2013年。[11] D. K. Elson、N. Dames 和 K. McKeown。从文学虚构作品中提取社交网络。载于计算语言学协会年会，2010年。[12] F. Brahman、M. Huang、O. Tafjord、C. Zhao、M. Sachan 和 S. Chaturvedi。“让你的角色讲述自己的故事”：一个以角色为中心的叙事理解数据集。载于 M.-F. Moens、X. Huang、L. Specia 和 S. W.-t. Yih 编，《style id='29'>计算语言学协会会议论文集：EMNLP 2021》，第 1734–1752 页，多米尼加共和国蓬塔卡纳，2021年11月。计算语言学协会。[13] J. Boucher 和 C. E. Osgood。《波莉安娜假说》。《言语学习与言语行为杂志》，8(1):1–8，1969年。[14] P. S. Dodds、E. M. Clark、S. Desu、M. R. Frank、A. J. Reagan、J. R. Williams、L. Mitchell、K. D. Harris、I. M. Kloumann、J. P. Bagrow、K. Megerdoomian、M. T. McMahon、B. F. Tivnan 和 C. M. Danforth。人类语言揭示了普遍积极偏向。《美国国家科学院院刊》，112(8):2389–2394，2015年。[15] M. I. Fudolig、T. Alshaabi、K. Cramer、C. M. Danforth 和 P. S. Dodds。通过累积词时中的情感计量波动分解书籍结构。《人文与社会科学传播》，10(1):187，2023年。[16] X. Bost、V. Labatut、S. Gueye 和 G. Linarès。叙事平滑：用于分析电视剧剧情的动态对话网络。2016年IEEE/ACM社会网络分析与挖掘国际会议（ASONAM），第1111–1118页，2016年。[17] B. Mills。《情景喜剧》。2009年。[18] P. Quaglio。电视对白：《老友记》与自然对话。2009年。[19] Corbari, Ederson de Moura。《老友记》剧本：按季整理的《老友记》剧本抄本。https://edersoncorbari.github.io/friends/，2025年。访问日期：2026-02-11。[20] IMDb。《老友记》（1994年–2004年）剧集列表，2025年。访问日期：2025-06-03。[21] W. S. Cleveland。稳健的局部加权回归与散点图平滑。《美国统计学会杂志》，74:829–836，1979年。[22] M. J. Porter、D. L. Larson、A. Harthcock 和 K. B. Nellis。重新（定）义叙事事件：考察电视叙事结构。《大众电影与电视杂志》，30:23 – 30，2002年。[23] R. Killick、P. Fearnhead 和 I. A. Eckley。以线性计算成本实现变点的最优检测。《美国统计学会杂志》，107(500):1590–1598，2012年。[24] C. Truong、L. Oudre 和 N. Vayatis。离线变点检测方法的选择性综述。《信号处理》，167，2020年。[25] H. Sakoe 和 S. Chiba。用于口语词识别的动态规划算法优化。《IEEE声学、语音与信号处理汇刊》，26(1):43–49，1978年。[26] D. J. Berndt 和 J. Clifford。使用动态时间规整在时间序列中寻找模式。载于：KDD研讨会，1994年。[27] F. Murtagh 和 P. Contreras。层次聚类算法：概述。《威利跨学科评论：数据挖掘》

知识发现 2 2012
，， 。[28] P. J. Rousseeuw。轮廓系数：帮助解释和验证聚类分析的图形工具。计算与应用数学杂志，20：53–65，1987年。[29] C. E. Spearman。根据错误数据计算的相关性。英国心理学杂志，3：271–295，1910年。[30] W. BROWN。心理能力相关性的一些实验结果。英国心理学杂志，*1904-1920*，3(3)：296–322，1910年。[31] M. Dynel。超越笑话：会话幽默的类型。语言学指南，3：1284–1299，2009年。[32] 维基百科。瑞秋·格林。https://en.wikipedia.org/wiki/Rachel_格林，2026年。访问日期：2026年6月7日。[33] 维基百科。莫妮卡·盖勒。https://en.wikipedia.org/wiki/Monica_盖勒，2026年。访问日期：2026年6月7日。

A 附录

A1 字符组合 保持不变

角色组合	场景数量	词数
瑞秋、莫妮卡、菲比、乔伊、钱德勒、罗斯	414	167,293
瑞秋 ，Ross	177	35,331
莫妮卡，钱德勒	161	26,885
Ross	135	12,548
Joey	133	11,268
乔伊，钱德勒	130	20,790
菲比	128	14,275
瑞秋	99	10,517
钱德勒	91	7,322
乔伊、钱德勒、罗斯	72	15,447
瑞秋、乔伊	71	14,815
瑞秋 ，菲比	69	12,463
菲比，乔伊	66	10,658
莫妮卡，菲比	65	10,364
瑞秋 ，莫妮卡，菲比	64	14,021
莫妮卡	63	4,763
莫妮卡，乔伊，钱德勒	59	12,654
乔伊，罗斯	54	10,193
瑞秋 ，莫妮卡	53	8,585
瑞秋、莫妮卡、菲比、乔伊、钱德勒	48	14,741
莫妮卡，菲比 ，钱德勒	43	10,280
钱德勒 ，Ross	42	6,609
莫妮卡、菲比、乔伊、钱德勒	38	10,241
菲比 ，Ross	38	6,597
莫妮卡、菲比、乔伊、钱德勒、罗斯	35	8,276
瑞秋 ，莫妮卡，菲比 ，Ross	33	8,974
瑞秋、莫妮卡、乔伊、钱德勒、罗斯	33	9,980
瑞秋 ，菲比 ，Ross	30	6,767
莫妮卡、乔伊	29	4,248
莫妮卡，Ross	27	3,487
莫妮卡，钱德勒 ，Ross	27	5,881
瑞秋 ，莫妮卡，钱德勒 ，Ross	26	6,062
瑞秋、菲比、乔伊、罗斯	25	7,622
瑞秋 ，莫妮卡，菲比 ，钱德勒 ，Ross	24	8,939
瑞秋 ，莫妮卡，Ross	24	5,003
瑞秋、莫妮卡、乔伊、钱德勒	24	6,128
瑞秋、乔伊、罗斯	24	6,751
瑞秋、乔伊、钱德勒、罗斯	23	5,865
瑞秋、菲比、乔伊	23	4,864
瑞秋、莫妮卡、菲比、乔伊、罗斯	23	4,658
瑞秋 ，钱德勒	22	4,080
莫妮卡、乔伊、钱德勒、罗斯	21	5,570
菲比 ，钱德勒	21	3,046
瑞秋 ，莫妮卡，菲比 ，钱德勒	19	4,905
瑞秋 ，莫妮卡，钱德勒	19	4,819
瑞秋、乔伊、钱德勒	18	4,292
菲比、乔伊、钱德勒、罗斯	18	4,674
菲比、乔伊、钱德勒	17	3,369
莫妮卡，菲比 ，钱德勒 ，Ross	17	4,587
瑞秋、菲比、乔伊、钱德勒、罗斯	17	4,716
菲比、乔伊、罗斯	16	3,360
瑞秋、莫妮卡、菲比、乔伊	15	4,444
莫妮卡，菲比 ，Ross	14	2,936
瑞秋 ，钱德勒 ，Ross	14	3,761
莫妮卡、菲比、乔伊	13	2,455
瑞秋 ，菲比 ，钱德勒 ，Ross	12	2,252

表A1： 组合仍用于组合层级分析。

角色组合	场景数量	词数
瑞秋、莫妮卡、乔伊	11	1962
瑞秋、莫妮卡、乔伊、罗斯	9	1963
瑞秋 ，菲比 ，钱德勒	8	2183
菲比 ，钱德勒 ，Ross	7	1435
瑞秋、菲比、乔伊、钱德勒	6	1130
莫妮卡、乔伊、罗斯	5	731
莫妮卡、菲比、乔伊、罗斯	2	427

表 A2： 已删除以供进一步分析的组合。

A2 每个角色的变化点

维度	角色	#	位置
权力	瑞秋	CP1	S02E23
		CP2	S03E09
		CP3	S03E19
		CP4	S04E05
		CP5	S05E19
		CP6	S06E13
		CP7	S07E07
		CP8	S07E17
		CP9	S08E04
		CP10	S08E17
权力	莫妮卡	CP1	S06E22
		CP2	S07E08
		CP3	S08E23
		CP4	S09E23
权力	菲比	CP1	S03E02
		CP2	S04E06
		CP3	S05E03
		CP4	S05E22
		CP5	第6季第19集
		CP6	S08E13
		CP7	S08E23
		CP8	S09E10
权力	Joey	CP1	S01E16
		CP2	S03E02
		CP3	S03E18
		CP4	S05E19
		CP5	S06E14
		CP6	S07E15
		变点7	S08E22
		CP8	S09E10
		CP9	S09E21
权力	钱德勒	CP1	S01E24
		CP2	S04E10
		CP3	S05E12
		CP4	S06E04
		CP5	S06E18
		CP6	S07E03
		CP7	S07E14
		CP8	S08E05
		CP9	S08E21
权力	Ross	CP1	S02E11
		CP2	S02E22
		CP3	S03E23
		CP4	S04E12
		CP5	S05E09
		CP6	S05E22
		CP7	S06E12
		CP8	S07E16
		CP9	S08E09
		CP10	S10E03

表 A3： PELT 检测出的《老友记》中六位主要角色权力维度的变点。

维度角色编号		位置
危险	瑞秋	CP1 S01E11 CP2 第1季第22集 CP3 第3季第8集 CP4 第4季第11集 CP5 第5季第14集 CP6 第6季第19集
危险	莫妮卡	CP1 S03E17 CP2 S04E06 CP3 S07E02 CP4 S08E04 CP5 S09E15 CP6 S10E01
危险	菲比	CP1 S01E24 CP2 S02E13 CP3 S04E24 CP4 S06E08
危险	Joey	CP1 S01E11 CP2 S02E10 CP3 S05E10 CP4 S08E17 CP5 S09E03
危险	钱德勒	CP1 S01E12 CP2 S01E23 CP3 S02E09 CP4 S03E07 CP5 S04E24 CP6 S05E17 CP7 S06E06 CP8 S07E18
危险	Ross	CP1 S01E21 CP2 S03E02 CP3 S04E02 CP4 S05E01 CP5 S05E20 CP6 S06E07 CP7 S06E17 CP8 S07E03

表 A4： PELT 检测到的《老友记》中六位主要角色危险维度的变化点。

A3 轨迹聚类保留的角色组合

g

组合	有效季数	场景总数
钱德勒	8	88
莫妮卡 + 钱德勒	8	159
瑞秋 + 莫妮卡 +菲比 + 乔伊 + 钱德勒	8	44
瑞秋 + 莫妮卡 +菲比 + 乔伊 + 钱德勒 + 罗斯	10	414
莫妮卡 + 菲比	9	63
菲比 + 乔伊	8	63
瑞秋	10	99
瑞秋 + 莫妮卡 +菲比	9	63
瑞秋 +菲比	10	69
乔伊 + 钱德勒	9	128
瑞秋 +罗斯	10	177
Ross	10	135
乔伊 + 钱德勒 +罗斯	9	71
菲比	10	128
Joey	10	133

表A5： 为轨迹聚类保留的角色组合

A4 分半信度

表A6： 六位主要角色的剧集层面力量、危险和结构分值的分半信度。对于每个角色和维度，Pearson's r 衡量的是基于236集中的对话词元随机分成两半后计算出的分值之间的相关性。最后一列报告斯皮尔曼–布朗校正信度系数。

角色	维度	r_{half}	p	r_{SB}
瑞秋	权力	0.400	< .001	0.572
	危险	0.257	< .001	0.409
	结构	0.207	0.001	0.344
莫妮卡	权力	0.341	< .001	0.509
	危险	0.307	< .001	0.470
	结构	0.245	< .001	0.394
菲比	权力	0.232	< .001	0.377
	危险	0.263	< .001	0.416
	结构	0.255	< .001	0.406
Joey	权力	0.394	< .001	0.565
	危险	0.369	< .001	0.539
	结构	0.395	< .001	0.567
钱德勒	权力	0.337	< .001	0.504
	危险	0.198	0.002	0.330
	结构	0.364	< .001	0.534
Ross	权力	0.449	< .001	0.619
	危险	0.298	< .001	0.459
	结构	0.272	< .001	0.428