

联系电话: 18804716708 (必留)  
联系邮箱: yanjun1023@163.com (必留)  
投稿邮箱: mtrhycb\_xuebao@cuc.edu.cn

(文章篇幅: 8000-10000 字)

## 偏好的博弈: 辟谣失灵现象及其治理

闫俊<sup>1</sup>, 宋彪<sup>2\*</sup>

(1. 内蒙古自治区信息化与数字经济推进中心, 呼和浩特 010070; 2. 内蒙古财经大学, 呼和浩特 010051)

**摘要:** 辟谣失灵是谣言治理过程中的一个重要研究话题, 目前大多数研究发现人们往往对负面信息赋予更多的关注, 导致了辟谣无效的现象, 其内在机理尚未清晰。与此同时, 偶有个例显示出相关主体在遭受负面谣言时, 不仅可以全身而退, 甚至借势成功, 这类成功辟谣的行为往往消弥危害于事前, 样本数量稀少, 导致了人们对谣言逆转相关研究的不足。本文从网民的风险偏好和信任偏好出发, 构建博弈模型进行分析, 旨在探索人们对负面信息关注和传播偏好的形成机理以及辟谣失灵的内在规律。结果表明, 人们关注正面消息还是负面消息, 取决于事件的模糊状态以及人们的风险偏好。人们对谣言信息的处理行为影响了受众的信任偏好形成, 受众对信息的信任程度相对较高时, 辟谣行为往往不能奏效, 降低受众对谣言的信任, 实现辟谣激励机制, 可以改善辟谣无效的困境。该结论尝试解释辟谣失灵的内在规律并给出相应的措施, 可以为人们治理谣言提供一定的借鉴。

**关键词:** 辟谣失灵; 博弈论; 偏好; 谣言治理

**中图分类号:** G206.3 **文献标识码:** A

## Game of Preference: The Failure of Rumor Refutation and Its Governance

YAN Jun<sup>1</sup>, SONG Biao<sup>2\*</sup>

(1. Inner Mongolia Informatization and Digital Economy Promotion Center, Hohhot, 010070, China; 2. Inner Mongolia University of Finance and Economics, Hohhot, 010051, China)

**Abstract:** The failure of rumor refutation is a significant research topic in the process of rumor governance. Most studies have found that people tend to pay more attention to negative information, leading to the ineffectiveness of refutation efforts. The underlying mechanisms of this phenomenon remain unclear. Meanwhile, there are occasional cases where the involved parties not only manage to escape the negative impact of rumors but also leverage the situation to their advantage. These successful refutation behaviors often mitigate harm in advance, but the scarcity of such cases has led to insufficient research on rumor reversal. This paper constructs a game theory model based on peoples' risk preferences and trust preferences to analyze the formation mechanisms of attention and dissemination preferences towards negative information, as well as the intrinsic rules of refutation failure. The results indicate that whether people focus on positive or negative messages depends on the ambiguity of the event and their risk preferences. People's handling of rumor information influences the formation of the audience's trust preferences. When the audience's trust in the information is relatively high, refutation efforts often fail. Reducing the audience's trust in rumors and implementing a refutation incentive mechanism can improve the dilemma of ineffective refutation. This conclusion attempts to explain the intrinsic rules of refutation failure and provides corresponding measures, offering some reference for people in rumor governance.

**Keywords:** Rumor Rejection Failure, Game Theory, preferences, rumor management

## 1 引言

谣言飞奔在前, 真理跛行其后, 已有学者研究表明即使信息完备的情况下, 人们仍然表现为偏信谣言<sup>[1]</sup>。切断谣言的源头以及打击造谣和传谣者, 是政府通常认为可以有效治理谣言的政策选项, 但是在现实中还可以经常看到, 很多人在面对事实或真相时, 仍然坚持错误信念而拒绝改变<sup>[2]</sup>。为什么会有相当数量的人在面对辟谣或澄清事实的信息时仍然相信谣言, 甚至在事实或真相被披露以后仍然拒绝承认和接受? 除了宏观的政治体制和信息传播结构的影响外, 人们对某个信息采信与否, 在信息供给与个体的内在特性之间互动机制上, 是否还存在着差异?

谣言作为一种特殊信息在中国历朝历代从来都不曾缺位, 而在其它国家, 各类谣言以及不实信息也是大量存在, 在政治领域尤为如此<sup>[2]</sup>。研究表明, 在互联网技术飞速发展的今天, 要阻止谣言的传播几乎是不可能的<sup>[3]</sup>。在融媒体和 5G 时代, 社交网络传播方式导致信息传统把关人所起作用受到削弱, 信息传播由“自上而下”向“自下而上”过渡, “后真相”时代信息发布由“前把关”向“后把关”模式的变革加重了信息不完备的情况。现实生活中, 信息的不完备性几乎是难以改变的事实<sup>[4]</sup>。更进一步, 我们发现即使信息完备透明也未必会使辟谣奏效, 信息对称的内涵已经发生嬗变, 如果信息供给机制设置不适当, 即便信息完备, 也无法激发网民有效的信任, 于是谣言难以制止。如 2017 年 5 月青岛发生的“棉花肉松”事件, 此前该谣言已在广西贺州、安徽肥东、浙江永康等地大量传播并经当地官方辟谣, 但经过一些网民的加工与传播, 再次引发青岛涉事企业的消费纠纷<sup>[5]</sup>。现实中显示偏见比无知离真理更远, 这启发我们解决“辟谣失灵”的问题, 亟待考虑辟谣无效的内在规律, 最终实现有效辟谣的目标。

## 2 网谣传播负面偏好的形成博弈分析

看的负面新闻报道越多, 越有可能对未来感到担忧, 可能使我们难以自拔<sup>[6]</sup>, 人们会通过从负面新闻中获取的信息来评估其资源所面临的实际或潜在威胁, 从而做出相应的防御行为<sup>[7][8]</sup>。此外, 负面新闻的故事主题、材料选择与内容叙事, 通常具有相对更大的情绪唤起能力。研究显示, 新闻所能激发的情绪越强烈, 越有可能获得受众的青睐并再次传播, 如在社交媒体中转发<sup>[9]</sup>。

我们以一个简单模型对谣言传播过程中人们负面偏好的形成进行描述, 假设 A、B 两人因为对同一事件传播两种相斥的信息, A 传播 (相信) 真消息, B 传播 (相信) 谣言, 在法制社会环境下, 设 A 为风险规避者, B 为风险偏好者。抽象假设双方提出利益诉求并由信息受众作为第三方裁决。为简化分析, 我们设全体受众认可信息带来的利益总价值标准化为 1, 并根据两方利益的互斥性, 假设其中如果 A 传播真信息并获得了一定比例受众的认可, 将有权得到总价值中的 $\theta_A$ 比例, 如果 B 传播谣言获得了一定比例受众的认可, 有权得总价值中的 $\theta_B$ 比例。其中 $\theta_A, \theta_B \in (0,1]$ , 并且 $\theta_A + \theta_B > 1$ 。由于双方的期待之和大于 1, 因此形成争端需要信息受众裁决。双方的目标是让自己获得的利益尽可能大, 第三方将会根据对是否谣言的判断做出分配裁决, 把第三方分配给 A、B 的比例记为向量 $(\theta_A^T, \theta_B^T)$ , 设 A 可以支持己方的有效因素为向量 $\vec{a} = (\alpha_1, \alpha_2, \dots, \alpha_m)$ , 设 B 可以支持己方的有效因素为向量 $\vec{b} = (b_1, b_2, \dots, b_n)$ , 在进行裁决时, 第三方会同时考虑 A 和 B 提出的理由, 则第三方的决策函数可以用向量 $(\alpha_1, \alpha_2, \dots, \alpha_m; b_1, b_2, \dots, b_n)$ 的函数 $f(\alpha_1, \alpha_2, \dots, \alpha_m; b_1, b_2, \dots, b_n) = (\theta_A^T, \theta_B^T)$ 表示。

设第三方正确的裁决结果为 $(\theta_A^*, \theta_B^*)$ , 第三方能够对谣言给出确定性判断, 错判的概率为 $p$ , 则正确判定的概率为 $1 - p$ , 则第三方错判给 A 和 B 双方各自增加 $s$ 单位总价值比例收益的概率为 $p/2$ , 其中 $s + \theta_i^* \in [0,1]$ 。因此对 A 和 B 而言, 其预期收益为式 (1) 所示:

$$E_{\text{确定}}(U(\theta_i^T)) = (1 - p)u(\theta_i^*) + \frac{p}{2} \times u(\theta_i^* - s) + \frac{p}{2} \times u(\theta_i^* + s) \quad (1)$$

式(1)中, $E$ 表示期望函数,  $U$ 表示总效用函数,  $u$ 表示效用函数, 以下各公式中含义相同。

但是由于事件模糊性特征导致第三方对是否谣言判断具有不确定性, 则第三方出错的概率增加为 $r$ , 此时对 A 和 B 来说, 其预期收益为:

$$E_{\text{模糊}}(U(\theta_i^T)) = (1 - p - r)u(\theta_i^*) + \frac{p + r}{2} \times u(\theta_i^* - s) + \frac{p + r}{2} \times u(\theta_i^* + s) \quad (2)$$

从风险规避角度来说,  $ru(\theta_i^*) > \frac{r}{2} \times [u(\theta_i^* - s) + u(\theta_i^* + s)]$ , 由公式 (1) 和公式 (2) 可知,

$$E_{模糊}(U(\theta_i^T)) = E_{确定}(U(\theta_i^T)) - \left[ ru(\theta_i^*) - \frac{r}{2} \times [u(\theta_i^* - s) + u(\theta_i^* + s)] \right] \quad (3)$$

因此在第三方对谣言判断具有模糊性的情况下, 风险规避者 A 传播真消息预期收益会减少, 而风险偏好者 B 传播谣言的预期收益会增大。因此人们在事件真相模糊程度增大, 判断难度的增加的情况下, 排除事件的利益因素影响, 就信息传播来说, 风险偏好者收益更高, 因此人们逐渐形成了传播 (相信) 谣言的负面偏好。不难发现, 对事件消息是否为谣言的判断成本高于  $ru(\theta_i^*) - \frac{r}{2} \times [u(\theta_i^* - s) + u(\theta_i^* + s)]$  时, 应当容忍信息与事件真相的一定偏离。

### 3 网民及谣言主体辟谣行为信号博弈模型

即使在完全信息下, 仍然有人坚持错误的言论, 拒绝接受真相。这说明, 在完全信息的条件下, 人们的裁决判断仍然具有非确定性的一面, 在现实生活中, 人们可能会面临存在风险的不确定性决策, 这导致其在很大程度上依赖于直觉和心理状态做出评估<sup>[10]</sup>, 这个过程中由于人们不同的价值观以及偏好等对真相的判断都存在或多或少的影响。

主体采取的行为可以作为主体相关消息所述真假的信号, 受众不知道消息所述的真相, 只能根据消息相关主体的行为外在表现来进行决策。主体涉及的消息具有两个方向, 分为正面消息和负面消息, 消息主体可以对消息采取两种行为, 分别为默认和辟谣。这是一个不完全信息的动态博弈, 具有信号博弈的基本特征。假设自然根据消息的属性分为两类, 正面消息的概率为  $r$ , 负面消息的概率为  $1 - r$ 。从主体发出行为信号, 受众判断消息的真相, 从默认行为相信正面消息或负面消息为真消息的概率为  $p$ , 相信正面消息或负面消息为假消息的概率为  $1 - p$ 。从辟谣行为相信正面消息或负面消息为真消息的概率为  $q$ , 相信正面消息或负面消息为假消息的概率为  $1 - q$ 。 $p$  和  $q$  我们可以理解为受众的信任偏好, 消息主体与受众双方的博弈如下见图 1 和图 2:

(1) 真相带给主体和受众的收益为  $(x, y)$ , 相关主体发生相关消息后, 假设受众和主体在封闭系统中, 受众的决策与信息的真相相符时, 将获得收益为  $b > 0$ , 反之遭受的损失为  $-b$ 。受众对正面消息的相信行为, 将为消息主体带来合理的收益  $\alpha$ , 反之对负面消息的相信行为, 将为消息主体带来损失为  $-\alpha$ 。

(2) 若受众不相信正面的假消息和负面的假消息, 真相维持主体和受众的原始收益, 即  $(x, y)$ 。

(3) 主体采取辟谣行为时, 主体将付出成本  $c$ , 在常规情况下  $c > 0$ 。

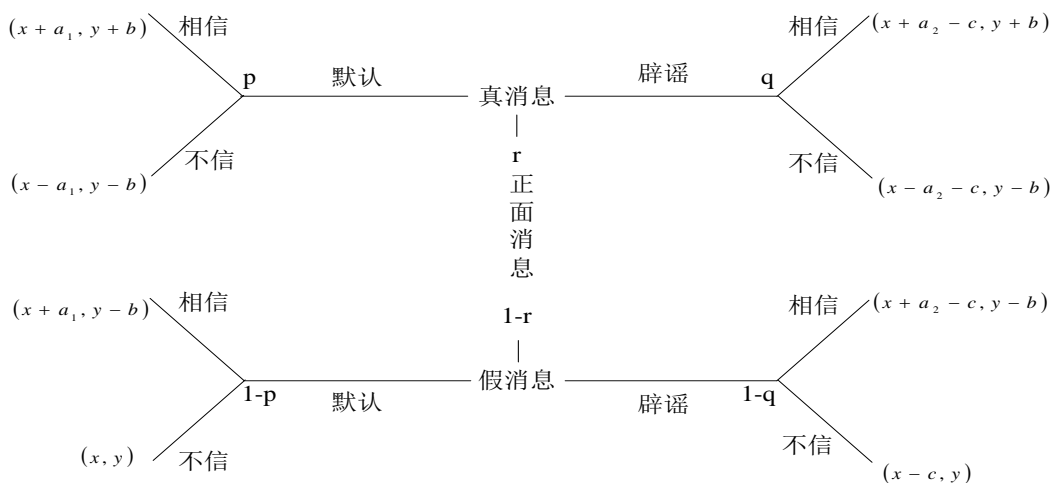


图 1 正面消息辟谣信号博弈

正面消息相关主体的策略:

(1) 自然赋予主体正面真消息, 消息相关主体选择默认策略; 自然赋予主体正面假消息, 消息相关主体选择默认策略。

(2) 自然赋予主体正面真消息, 消息相关主体选择默认策略; 自然赋予主体正面假消息, 消息相关主体选择辟谣策略。

(3) 自然赋予主体正面真消息, 消息相关主体选择辟谣策略; 自然赋予主体正面假消息, 消息相关主体选择默认策略。

(4) 自然赋予主体正面真消息, 消息相关主体选择辟谣策略; 自然赋予主体正面假消息, 消息相关主体选择辟谣策略。

正面消息受众的策略：

- (1)如果是正面真消息，受众选择相信策略；如果是正面假消息，受众选择不相信策略。
- (2)如果是正面真消息，受众选择相信策略；如果是正面假消息，受众选择相信策略。
- (3)如果是正面真消息，受众选择不相信策略；如果是正面假消息，受众选择不相信策略。
- (4)如果是正面真消息，受众选择不相信策略；如果是正面假消息，受众选择相信策略。

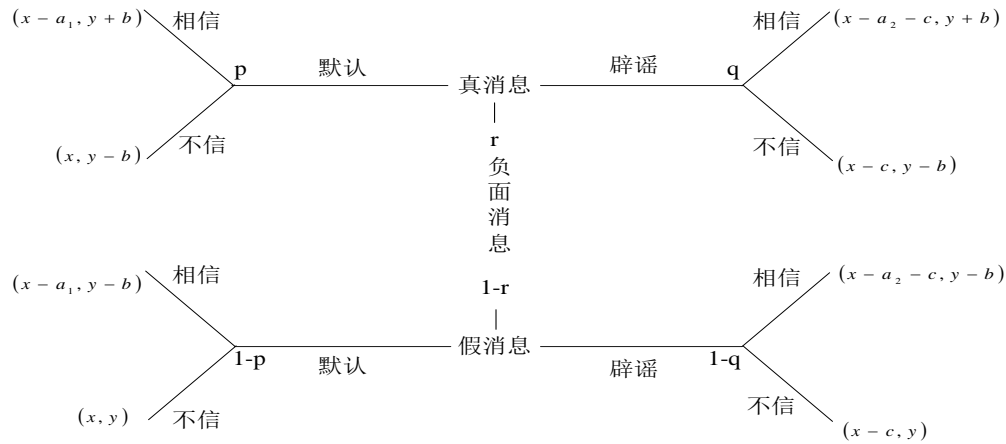


图2 负面消息辟谣信号博弈

负面消息相关主体的策略：

- (1)自然赋予主体负面真消息，消息相关主体选择默认策略；自然赋予主体负面假消息，消息相关主体选择默认策略。
- (2)自然赋予主体负面真消息，消息相关主体选择默认策略；自然赋予主体负面假消息，消息相关主体选择辟谣策略。
- (3)自然赋予主体负面真消息，消息相关主体选择辟谣策略；自然赋予主体负面假消息，消息相关主体选择默认策略。
- (4)自然赋予主体负面真消息，消息相关主体选择辟谣策略；自然赋予主体负面假消息，消息相关主体选择辟谣策略。

负面消息受众的策略：

- (1)如果是负面真消息，受众选择相信策略；如果是负面假消息，受众选择不相信策略。
- (2)如果是负面真消息，受众选择相信策略；如果是负面假消息，受众选择相信策略。
- (3)如果是负面真消息，受众选择不相信策略；如果是负面假消息，受众选择不相信策略。
- (4)如果是负面真消息，受众选择不相信策略；如果是负面假消息，受众选择相信策略。

#### 4 辟谣信号博弈均衡分析

大体上来说，这个辟谣信号博弈的精炼贝叶斯均衡解应该满足三个条件：第一是受众对主体行为出现的概率有一个推断（信念），这个推断要符合贝叶斯法则，是一个后验概率；第二是在后验概率条件下，信号接受者的行动要使信号接受者的效用最大化；第三是给定信号接受者的战略，信号发出者使自己的效用最大化。不完全信息动态博弈的均衡解有三类：分离均衡、混同均衡和准分离均衡，本文的研究只有两个信号、两个类型，因此精炼贝叶斯均衡可不考虑准分离均衡。

##### 4.1 针对正面消息的分离均衡

正面消息的分离均衡可能存在两个**纯战略分离均衡**，第一个分离均衡是真的正面消息主体选择默认策略，假的正面消息主体选择辟谣策略；第二个分离均衡是真的正面消息主体选择辟谣策略，假的正面消息主体选择默认策略。

第一个分离均衡的后验概率条件应该满足：

$$\mu(\text{真消息}|\text{默认}) = 1, \mu(\text{假消息}|\text{默认}) = 0 \text{ 和 } \mu(\text{真消息}|\text{辟谣}) = 0, \mu(\text{假消息}|\text{辟谣}) = 1$$

根据该后验概率，则由贝叶斯法则有  $\mu(\text{真消息}|\text{默认}) = \frac{rp}{rp+(1-r)(1-p)} = 1$ ，此时  $p=1$ 。

$$\mu(\text{假消息}|\text{辟谣}) = \frac{(1-r)(1-q)}{rq+(1-r)(1-q)} = 1, \text{ 此时 } q=0。$$

在该后验概率条件下，受众看到消息主体默认，选择相信正面消息，看到消息主体辟谣，选择不相信正面消息，能够使受众的利益最大化，要求条件是： $b > 0$ 。

同时，真消息选择默认，假消息选择辟谣，要能使信息主体价值最大化，这一条件是： $a_1 < a_2 + c$  和  $c < 0$ ， $c < 0$ 不满足经济理性假设，第一个分离均衡没有实际意义。

第二个分离均衡的后验概率条件应该满足：

$$\mu(\text{真消息}|\text{辟谣}) = 1, \mu(\text{假消息}|\text{辟谣}) = 0 \text{ 和 } \mu(\text{真消息}|\text{默认}) = 0, \mu(\text{假消息}|\text{默认}) = 1$$

根据该后验概率，则由贝叶斯法则有  $\mu(\text{真消息}|\text{辟谣}) = \frac{rq}{rq+(1-r)(1-q)} = 1$ ，此时  $q=1$ 。

$$\mu(\text{假消息}|\text{默认}) = \frac{(1-r)(1-p)}{rp+(1-r)(1-p)} = 1, \text{ 此时 } p=0。$$

在该后验概率条件下，受众看到消息主体辟谣，选择相信正面消息，看到消息主体默认，选择不相信正面消息，能够使受众的利益最大化，要求条件是： $b > 0$ 。

同时，真消息选择辟谣，假消息选择默认，要能使信息主体价值最大化，这一条件是： $c > 0$ ，该条件可以成立，因此第二个分离均衡可能存在。

#### 4.2 针对正面消息的混同均衡

此信号博弈的混同均衡有两个，一个是真正面消息和假正面消息都混同于沉默策略，一个是真正面消息和假正面消息都混同于辟谣策略。

第一种情况，根据混同均衡的含义，如果真正面消息和假正面消息都采取沉默状态，受众选择相信，则要满足以下条件：

$$\text{在混同均衡下，后验概率等于先验概率，} \mu(\text{真消息}|\text{默认}) = \frac{rp}{rp+(1-r)(1-p)} = r,$$

$$\mu(\text{假消息}|\text{默认}) = \frac{(1-r)(1-p)}{rp+(1-r)(1-p)} = 1-r, \text{ 可以算出 } p = 0.5。 \text{ 根据上述后验概率，受众的反应如}$$

果是相信，则需要满足相信的效用要大于不相信的效用： $E(\text{相信}) = p(y+b) + (1-p)(y-b)$ ，

$E(\text{不信}) = p(y-b) + (1-p)y$ 。可以求出， $p > 1/3$  时，且  $a_1 > a_2 - c$  时，该混同均衡存在。

类似的，第二种情况，真正面消息和假正面消息都采取辟谣策略， $q = 0.5$ 。受众的反应如果是相信，则需要满足相信的效用要大于不相信的效用，可以求出， $q > 1/3$  时，且  $0 < a_1 < a_2 - c$  时，该混同均衡存在。

#### 4.3 针对负面消息的分离均衡

负面消息的分离均衡可能存在两个纯战略分离均衡，第一个分离均衡是真的负面消息主体选择默认策略，假的正面消息主体选择辟谣策略；第二个分离均衡是真的负面消息主体选择辟谣策略，假的负面消息主体选择默认策略。

第一个分离均衡的后验概率条件应该满足：

$$\mu(\text{真消息}|\text{默认}) = 1, \mu(\text{假消息}|\text{默认}) = 0 \text{ 和 } \mu(\text{真消息}|\text{辟谣}) = 0, \mu(\text{假消息}|\text{辟谣}) = 1$$

根据该后验概率，则由贝叶斯法则有  $\mu(\text{真消息}|\text{默认}) = \frac{rp}{rp+(1-r)(1-p)} = 1$ ，此时  $p=1$ 。

$$\mu(\text{假消息}|\text{辟谣}) = \frac{(1-r)(1-q)}{rq+(1-r)(1-q)} = 1, \text{ 此时 } q=0。$$

在该后验概率条件下，受众看到消息主体默认，选择相信负面消息，看到消息主体辟谣，选择不相信负面消息，能够使受众的利益最大化，要求条件是： $b > 0$ 。

同时，真消息选择默认，假消息选择辟谣，要能使信息主体价值最大化，这一条件是： $0 < a_1 < a_2 + c$  和  $c < 0$ ， $c < 0$  不满足假设，因此第一个分离均衡不存在。

第二个分离均衡的后验概率条件应该满足：

$$\mu(\text{真消息}|\text{辟谣}) = 1, \mu(\text{假消息}|\text{辟谣}) = 0 \text{ 和 } \mu(\text{真消息}|\text{默认}) = 0, \mu(\text{假消息}|\text{默认}) = 1$$

根据该后验概率，则由贝叶斯法则有  $\mu(\text{真消息}|\text{辟谣}) = \frac{rq}{rq+(1-r)(1-q)} = 1$ ，此时  $q=1$ 。

$$\mu(\text{假消息}|\text{默认}) = \frac{(1-r)(1-p)}{rp+(1-r)(1-p)} = 1, \text{ 此时 } p=0。$$

在该后验概率条件下，受众看到消息主体辟谣，选择相信负面消息，看到消息主体默认，选择不相信负面消息，能够使受众的利益最大化，要求条件是： $b > 0$ 。

同时，真消息选择辟谣，假消息选择默认，要能使信息主体价值最大化，这一条件是： $a_1 > a_2 + c$ ，该条件可以成立，因此第二个分离均衡可能存在。

这个分离均衡的意义是，有些时候，辟谣反而使受众相信负面消息，而沉默策略有时受众

反而不相信负面消息。

#### 4.4 针对负面消息的混同均衡

此信号博弈的混同均衡有两个，一个是真负面消息和假负面消息都混同于沉默策略，一个是真负面消息和假负面消息都混同于辟谣策略。

第一种情况，根据混同均衡的含义，如果真正负面消息和假负面消息都采取沉默状态，受众选择相信，则要满足以下条件：

$$\text{在混同均衡下，后验概率等于先验概率，} \mu(\text{真消息}|\text{默认}) = \frac{rp}{rp+(1-r)(1-p)} = r, \\ \mu(\text{假消息}|\text{默认}) = \frac{(1-r)(1-p)}{rp+(1-r)(1-p)} = 1-r, \text{ 可以算出 } p = 0.5. \text{ 根据上述后验概率，受众的反应如}$$

果是相信，则需要满足相信的效用要大于不相信的效用期望： $E(\text{相信}) = p(y+b) + (1-p)(y-b)$ ， $E(\text{不信}) = p(y-b) + (1-p)y$ 。可以求出， $p > 1/3$  时，且  $0 < a_1 < a_2 + c$  时，该混同均衡存在。

第二种情况，后验概率等于先验概率， $\mu(\text{真消息}|\text{辟谣}) = \frac{rq}{rq+(1-r)(1-q)} = r$ ， $\mu(\text{假消息}|\text{辟谣}) = \frac{(1-r)(1-q)}{rq+(1-r)(1-q)} = 1-r$ ，可以算出  $q = 0.5$ 。根据上述后验概率，受众的反应如果是相信，则需要满足相信的效用要大于不相信的效用：可以求出， $q > 1/3$  时，且  $a_1 > a_2 + c$  时，该混同均衡存在。

从该模型的基本结论可以看出，正面消息和负面消息可以成立的第二个分离均衡表明，受众对相关人默认的信息持不相信态度，以及对相关人辟谣的信息持相信态度在一定条件下可以成立。因此，一些相关人对真正正面消息进行辟谣已达到保密的目的以及对虚假正面消息采取默认的行为、对负面虚假信息辟谣同时对负面真实消息也进行辟谣的行为等，在某种程度上形成了受众对相关人信号的反向选择偏好。

在对正面信息和负面信息达到一定信任程度的情况下，无论信息相关者对正面消息和负面消息采取默认还是辟谣策略，甚至不论消息的真假，某种条件下受众都可以达到一个混同均衡，维持相信相关信息的稳定状态，这种条件与信息相关者采取默认还是辟谣行为所获得收益的差异有直接关系，简单来说就是当信息相关者采取的行为方向与其利益趋同时，其辟谣往往是不奏效的。当然我们不难发现，针对负面假消息来说，如果能够采取恰当的激励机制使辟谣成本小于 0，可以达到分离均衡并能够实现有效辟谣。

## 5 群体参与网络谣言治理的激励与合作博弈分析

事实上传谣网民往往能够获得足够的利益，从而导致了网络谣言频频发，更为关键的是传谣网民获得该利益的成本很低，谣言传播的每个始作俑者甚至可以因为在整个事件中的作用微乎其微，最后导致无法追责。在这种情况下，由于高频率谣言的发生，人们强化了对负面谣言相信和传播的偏好，网民群体的认知偏好结构影响了对谣言的判断<sup>[11]</sup>，而政府监管机构无法事无巨细投入过量精力，那么激励普通网民参与谣言治理合作将是一个有效并且必要的途径。

谣言的传播网络平台具有不可推卸的责任，为能够在具体可执行层面实现监管，我们针对网络平台进行博弈模型构建，这样可以将监管部门对普通网民参与治理合作这项工作聚焦到企业一级即网络平台，基于此措施来推动各种网络平台加强自律及事前监管。

我们设定在对网络谣言传播进行监管时，政府监管机构为激励普通网民参与辟谣，需要对网络平台收取一定金额的保证金，网络平台作为利益既得方，对谣言传播网民进行必要的企业级治理。总体流程即若遇传谣网民在网络平台传播网络谣言，由普通网民进行发现并且参与治理，政府收到普通网民反馈后进行稽查，并根据稽查情况则从保证金划拨给参与治理的普通网民进行奖励，同时对网络平台采取没收保证金的处罚。

针对网民群体的合作行为进行博弈模型构建，做出如下假设：普通网民作为个体都是“有限理性”的，在进化博弈论中，认为博弈方可以通过不断的学习和模仿而做出决策选择。同样网民之中的网络谣言治理参与者之间也存在着学习和模仿的过程，当某个网络谣言治理参与者因为某一策略的选择而获得较好收益时，将会导致其他普通网民复制该策略的倾向，增强了参与网络谣言治理的积极性等，因此，普通网民的上述性质符合进化博弈论的基本假设。

### 5.1 群体参与谣言治理行为博弈模型

我们设定三个有限理性的参与者，分别为甲、乙、丙。甲、乙、丙在谣言治理合作过程中，有两种策略可以选择，即参与 G 和不参与 N 两种策略，则策略空间为 **{Y 参与, N 不参与}**。博

弈方甲、乙、丙采取参与的概率分别为  $p_{甲}$ ,  $p_{乙}$ ,  $p_{丙}$ , 博弈方采取不参与策略的概率分别为  $1 - p_{甲}$ ,  $1 - p_{乙}$ ,  $1 - p_{丙}$ , 其中  $0 \leq p_{甲} \leq 1$ ,  $0 \leq p_{乙} \leq 1$ ,  $0 \leq p_{丙} \leq 1$ . 明确参与谣言治理的普通网民都是具有独立性的个体, 参与谣言治理和不参与谣言治理主要取决于在参与过程中获得的收益和遭受的损失。

合理设置以下限制条件并构建谣言治理参与博弈矩阵, 如表 1 所示:

(1) 甲、乙、丙三方在谣言治理过程中都采取不参与策略时, 收益为 0。

(2) 甲、乙、丙三方在网络谣言治理过程中都采取参与策略时, 监管机构稽查成功付给普通网民参与直接总收益为期望  $E$ , 网民参与越多, 节点直接收益越小, 分配系数为  $\theta_{甲}$ 、 $\theta_{乙}$ 、 $\theta_{丙}$ ,

其中  $0 \leq \theta_{甲} \leq 1$ ,  $0 \leq \theta_{乙} \leq 1$ ,  $0 \leq \theta_{丙} \leq 1$ , 且  $\theta_{甲} + \theta_{乙} + \theta_{丙} = 1$ , 则甲、乙、丙三方各自直

接获得监管机构奖励收益为  $\theta_{甲}E$ 、 $\theta_{乙}E$ 、 $\theta_{丙}E$ , 由于普通网民群体参与了网络谣言治理, 这个过程中参与的网民越多, 那么提供的证据链就越全, 则监管机构进行谣言稽查成功的可能性越大。为宏观社会环境带来稳定的舆论环境而转化为自身增值方面的收益, 兼之监管机构对普通网民参与治理这种模式信任的增加从而带来一定的收益为  $S_{甲} > 0$ 、 $S_{乙} > 0$ 、 $S_{丙} > 0$ , 每个参与治理的普通网民的非直接收益随着参与谣言治理的其它网民的数量增加而增加。

(3) 普通网民选择参与网络谣言治理策略时, 投入的成本为  $\gamma_i I_i$ , ( $i=甲、乙、丙$ ), 其中  $1 < \gamma_i < \infty$  为投入标准转化系数,  $I_i$  为直接投入, 我们设定每个普通网民投入的成本随参与谣言治理的其它网民的数量而增加, 并且随着时间的增加而增加, 因为当参与谣言治理的网民越多, 网络谣言将越不容易发生而导致发现困难。

(4) 普通网民选择参与网络谣言治理时, 对网民参与治理整个激励体系带来的溢出效应  $Y_i$ , ( $i=甲、乙、丙$ ), 这个溢出效应为其它参与普通网民获得。

(5) 当已经处在治理合作中普通网民选择不参与网络谣言治理时, 该网民的退出无法获得监管机构收益, 可视为惩罚为  $\theta_i E$ , 这个惩罚被另两个参与的普通网民平均获得, 如果两个普通网民退出参与治理时, 各自退出的惩罚为  $\theta_i E$ , 这个惩罚被另一个普通网民全部获得。

收益矩阵如表 1 所示:

表 1 收益矩阵

|   |                 | 乙   |   | 丙   |   |
|---|-----------------|---|---|---|---|
|   |                 | 参与 $p_{乙}$  | 不参与 $1 - p_{乙}$   | 参与 $p_{丙}$  | 不参与 $1 - p_{丙}$   |
| 甲 | 参与 $p_{甲}$      | $Y_{乙} + Y_{丙} + \theta_{甲}E + S_{甲} - \gamma_{甲}I_{甲}$<br>$Y_{甲} + Y_{丙} + \theta_{乙}E + S_{丙} - \gamma_{丙}I_{丙}$<br>$Y_{甲} + Y_{乙} + \theta_{丙}E + S_{丙} - \gamma_{丙}I_{丙}$ | $Y_{丙} + (\theta_{甲} + 0.5\theta_{乙})E + S_{甲} - \gamma_{甲}I_{甲}$<br>$Y_{甲} + Y_{丙}$<br>$Y_{甲} + (\theta_{丙} + 0.5\theta_{乙})E + S_{丙} - \gamma_{丙}I_{丙}$ | $Y_{乙} + (\theta_{甲} + 0.5\theta_{丙})E + S_{甲} - \gamma_{甲}I_{甲}$<br>$Y_{甲} + (\theta_{乙} + 0.5\theta_{丙})E + S_{乙} - \gamma_{乙}I_{乙}$<br>$Y_{甲} + Y_{乙}$ | $(\theta_{甲} + \theta_{乙} + \theta_{丙})E + S_{甲} - \gamma_{甲}I_{甲}$<br>$Y_{甲}$<br>$Y_{甲}$ |
|   | 不参与 $1 - p_{甲}$ | $Y_{乙} + Y_{丙}$<br>$Y_{丙} + (\theta_{乙} + 0.5\theta_{甲})E + S_{乙} - \gamma_{乙}I_{乙}$<br>$Y_{乙} + (\theta_{丙} + 0.5\theta_{甲})E + S_{丙} - \gamma_{丙}I_{丙}$                     | $Y_{丙}$<br>$Y_{丙}$<br>$(\theta_{甲} + \theta_{乙} + \theta_{丙})E + S_{丙} - \gamma_{丙}I_{丙}$   | $Y_{乙}$<br>$(\theta_{甲} + \theta_{乙} + \theta_{丙})E + S_{乙} - \gamma_{乙}I_{乙}$<br>$Y_{乙}$   | $0$<br>$0$<br>$0$   |
|   |                 | 参与 $p_{丙}$  |   | 不参与 $1 - p_{丙}$   |   |

根据上述条件构建各参与节点收益方程:

网民甲在谣言治理过程中采取参与治理策略的期望收益为:

$$\begin{aligned}
 U_{甲}(G) = & (Y_{乙} + Y_{丙} + \theta_{甲}E + S_{甲} - \gamma_{甲}I_{甲})p_{乙}p_{丙} + (Y_{丙} + (\theta_{甲} + 0.5\theta_{乙})E + S_{甲} - \gamma_{甲}I_{甲})(1 - \\
 & p_{乙})p_{丙} + (Y_{乙} + (\theta_{甲} + 0.5\theta_{丙})E + S_{甲} - \gamma_{甲}I_{甲})p_{乙}(1 - p_{丙}) + ((\theta_{甲} + \theta_{乙} + \theta_{丙})E + S_{甲} - \\
 & \gamma_{甲}I_{甲})(1 - p_{乙})(1 - p_{丙})
 \end{aligned} \tag{4}$$

网民甲在谣言治理过程中采取不参与策略的期望收益为:

$$U_{\text{甲}}(N) = (\gamma_{\text{乙}} + \gamma_{\text{丙}})p_{\text{乙}}p_{\text{丙}} + \gamma_{\text{丙}}(1 - p_{\text{乙}})p_{\text{丙}} + \gamma_{\text{乙}}p_{\text{乙}}(1 - p_{\text{丙}}) \quad (5)$$

网民甲的平均期望收益为:

$$\bar{U}_{\text{甲}} = p_{\text{甲}}U_{\text{甲}}(G) + (1 - p_{\text{甲}})U_{\text{甲}}(N) \quad (6)$$

网民甲采取参与的复制动态方程为:

$$\begin{aligned} \frac{dp_{\text{甲}}}{dt} &= p_{\text{甲}}[U_{\text{甲}}(G) - \bar{U}_{\text{甲}}] = p_{\text{甲}}(1 - p_{\text{甲}})(U_{\text{甲}}(G) - U_{\text{甲}}(N)) \\ &= p_{\text{甲}}(1 - p_{\text{甲}})(\partial_{\text{甲}}Ep_{\text{丙}} - \partial_{\text{乙}}Ep_{\text{乙}} + s_{\text{甲}} - \gamma_{\text{甲}}I_{\text{甲}} + 0.5\partial_{\text{乙}}Ep_{\text{丙}}(1 + p_{\text{乙}}) - 0.5\partial_{\text{丙}}Ep_{\text{乙}}(1 + p_{\text{丙}})) \end{aligned} \quad (7)$$

同理, 网民乙采取参与的复制动态方程为:

$$\begin{aligned} \frac{dp_{\text{乙}}}{dt} &= p_{\text{乙}}[U_{\text{乙}}(G) - \bar{U}_{\text{乙}}] = p_{\text{乙}}(1 - p_{\text{乙}})(U_{\text{乙}}(G) - U_{\text{乙}}(N)) \\ &= p_{\text{乙}}(1 - p_{\text{乙}})(\partial_{\text{乙}}Ep_{\text{丙}} - \partial_{\text{甲}}Ep_{\text{甲}} + s_{\text{乙}} - \gamma_{\text{乙}}I_{\text{乙}} + 0.5\partial_{\text{甲}}Ep_{\text{丙}}(1 + p_{\text{甲}}) - 0.5\partial_{\text{丙}}Ep_{\text{甲}}(1 + p_{\text{丙}})) \end{aligned} \quad (8)$$

网民丙在谣言控制中采取参与策略的复制动态方程为:

$$\begin{aligned} \frac{dp_{\text{丙}}}{dt} &= p_{\text{丙}}[U_{\text{丙}}(G) - \bar{U}_{\text{丙}}] = p_{\text{丙}}(1 - p_{\text{丙}})(U_{\text{丙}}(G) - U_{\text{丙}}(N)) = p_{\text{丙}}(1 - p_{\text{丙}})(\partial_{\text{丙}}Ep_{\text{甲}} - \\ &\partial_{\text{乙}}Ep_{\text{乙}} + s_{\text{丙}} - \gamma_{\text{丙}}I_{\text{丙}} + 0.5\partial_{\text{乙}}Ep_{\text{甲}}(1 + p_{\text{乙}}) - 0.5\partial_{\text{甲}}Ep_{\text{乙}}(1 + p_{\text{甲}})) \end{aligned} \quad (9)$$

令  $\frac{dp_{\text{甲}}}{dt} = 0$ ,  $\frac{dp_{\text{乙}}}{dt} = 0$ ,  $\frac{dp_{\text{丙}}}{dt} = 0$  得出 14 个均衡点, 分别为  $D_1(0,0,0)$ ,  $D_2(1,0,0)$ ,

$D_3(0,0,1)$ ,  $D_4(0,1,0)$ ,  $D_5(1,0,1)$ ,  $D_6(1,1,0)$ ,  $D_7(0,1,1)$ ,  $D_8(1,1,1)$ ,

$$D_9 \left( 1, \frac{\partial_{\text{甲}}E - s_{\text{丙}} + \gamma_{\text{丙}}I_{\text{丙}} + 0.5\partial_{\text{乙}}E}{(\partial_{\text{甲}} + \partial_{\text{丙}} - 0.5\partial_{\text{乙}})E}, \frac{\partial_{\text{甲}}E - s_{\text{乙}} + \gamma_{\text{乙}}I_{\text{乙}} + 0.5\partial_{\text{丙}}E}{(\partial_{\text{甲}} + \partial_{\text{乙}} - 0.5\partial_{\text{丙}})E} \right),$$

$$D_{10} \left( \frac{\partial_{\text{乙}}E - s_{\text{丙}} + \gamma_{\text{丙}}I_{\text{丙}} + 0.5\partial_{\text{甲}}E}{(\partial_{\text{乙}} + \partial_{\text{丙}} - 0.5\partial_{\text{甲}})E}, 1, \frac{\partial_{\text{甲}}E - s_{\text{乙}} + \gamma_{\text{乙}}I_{\text{乙}} + 0.5\partial_{\text{丙}}E}{(\partial_{\text{甲}} + \partial_{\text{乙}} - 0.5\partial_{\text{丙}})E} \right),$$

$$D_{11} \left( \frac{\partial_{\text{乙}}E - s_{\text{丙}} + \gamma_{\text{丙}}I_{\text{丙}} + 0.5\partial_{\text{甲}}E}{(\partial_{\text{乙}} + \partial_{\text{丙}} - 0.5\partial_{\text{甲}})E}, \frac{\partial_{\text{甲}}E - s_{\text{丙}} + \gamma_{\text{丙}}I_{\text{丙}} + 0.5\partial_{\text{乙}}E}{(\partial_{\text{甲}} + \partial_{\text{丙}} - 0.5\partial_{\text{乙}})E}, 1 \right),$$

$$D_{12} \left( 0, \frac{\gamma_{\text{丙}}I_{\text{丙}} - s_{\text{丙}}}{\partial_{\text{丙}}E + 0.5\partial_{\text{甲}}E}, \frac{\gamma_{\text{乙}}I_{\text{乙}} - s_{\text{乙}}}{\partial_{\text{乙}}E + 0.5\partial_{\text{甲}}E} \right),$$

$$D_{13} \left( \frac{\gamma_{\text{丙}}I_{\text{丙}} - s_{\text{丙}}}{\partial_{\text{丙}}E + 0.5\partial_{\text{乙}}E}, 0, \frac{\gamma_{\text{甲}}I_{\text{甲}} - s_{\text{甲}}}{\partial_{\text{甲}}E + 0.5\partial_{\text{乙}}E} \right),$$

$$D_{14} \left( \frac{\gamma_{\text{乙}}I_{\text{乙}} - s_{\text{乙}}}{\partial_{\text{乙}}E + 0.5\partial_{\text{丙}}E}, \frac{\gamma_{\text{甲}}I_{\text{甲}} - s_{\text{甲}}}{\partial_{\text{甲}}E + 0.5\partial_{\text{丙}}E}, 0 \right)$$

由微分方程得出雅克比矩阵:

$$\begin{vmatrix} J_{11} & J_{21} & J_{31} \\ J_{12} & J_{22} & J_{32} \\ J_{13} & J_{23} & J_{33} \end{vmatrix} \quad (10)$$

其中, 矩阵元素如公式所示:

$$\begin{aligned} J_{11} &= (1 - 2p_{\text{甲}})(\partial_{\text{甲}}Ep_{\text{丙}} - \partial_{\text{乙}}Ep_{\text{乙}} + s_{\text{甲}} - \gamma_{\text{甲}}I_{\text{甲}} + 0.5\partial_{\text{乙}}Ep_{\text{丙}}(1 + p_{\text{乙}}) \\ &\quad - 0.5\partial_{\text{丙}}Ep_{\text{乙}}(1 + p_{\text{丙}})) \end{aligned}$$

$$J_{12} = p_{\text{乙}}(1 - p_{\text{乙}})(0.5\partial_{\text{甲}}Ep_{\text{丙}} - \partial_{\text{甲}}E - 0.5(1 + p_{\text{丙}})\partial_{\text{丙}}E)$$



$$J_{13} = p_{\text{丙}}(1 - p_{\text{丙}})(\partial_{\text{丙}}E + 0.5\partial_{\text{乙}}E(1 + p_{\text{乙}}) - 0.5\partial_{\text{甲}}Ep_{\text{乙}})$$

$$J_{21} = p_{\text{甲}}(1 - p_{\text{甲}})(0.5\partial_{\text{乙}}Ep_{\text{丙}} - 0.5\partial_{\text{丙}}Ep_{\text{丙}} - \partial_{\text{乙}}E)$$

$$J_{22} = (1 - 2p_{\text{乙}})(\partial_{\text{乙}}Ep_{\text{丙}} - \partial_{\text{甲}}Ep_{\text{甲}} + S_{\text{乙}} - \gamma_{\text{乙}}I_{\text{乙}} + 0.5\partial_{\text{甲}}Ep_{\text{丙}}(1 + p_{\text{甲}}) - 0.5\partial_{\text{丙}}Ep_{\text{甲}}(1 + p_{\text{丙}}))$$

$$J_{23} = p_{\text{丙}}(1 - p_{\text{丙}})(0.5\partial_{\text{乙}}Ep_{\text{甲}} - \partial_{\text{乙}}E - 0.5\partial_{\text{甲}}E(1 + p_{\text{甲}}))$$

$$J_{31} = p_{\text{甲}}(1 - p_{\text{甲}})(\partial_{\text{甲}}E + 0.5\partial_{\text{乙}}Ep_{\text{乙}} - 0.5\partial_{\text{丙}}Ep_{\text{乙}})$$

$$J_{32} = p_{\text{乙}}(1 - p_{\text{乙}})(\partial_{\text{乙}}E + 0.5\partial_{\text{甲}}E(1 + p_{\text{甲}}) - 0.5\partial_{\text{丙}}Ep_{\text{甲}})$$

$$J_{33} = (1 - 2p_{\text{丙}})(\partial_{\text{丙}}Ep_{\text{甲}} - \partial_{\text{乙}}Ep_{\text{乙}} + S_{\text{丙}} - \gamma_{\text{丙}}I_{\text{丙}} + 0.5\partial_{\text{乙}}Ep_{\text{甲}}(1 + p_{\text{乙}}) - 0.5\partial_{\text{甲}}Ep_{\text{乙}}(1 + p_{\text{甲}}))$$

均衡点分析如表 2 所示：

表 2 均衡点的局部稳定性分析

| 均衡点   | 条件  | 结果                                   |
|---|---|--------------------------------------|
| $D_1 (0,0,0)$<br>$\prod_{i=\text{甲、乙、丙}}(S_i - \gamma_i I_i)$ | 少数网民有正向收益，其它网民为负向收益时<br>$\prod_{i=\text{甲、乙、丙}}(S_i - \gamma_i I_i) > 0, \sum_{i=\text{甲、乙、丙}}(S_i - \gamma_i I_i) < 0$ | 演化稳定点                                |
| $D_2 (1,0,0)$   | $\gamma_i I_i - S_i < 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_3 (0,0,1)$   | $\gamma_i I_i - S_i < 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_4 (0,1,0)$   | $\gamma_i I_i - S_i < 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_5 (1,0,1)$   | $\gamma_i I_i - S_i > 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_6 (1,1,0)$   | $\gamma_i I_i - S_i > 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_7 (0,1,1)$   | $\gamma_i I_i - S_i > 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_8 (1,1,1)$   | 多数网民正向收益，少数网民负向收益时<br>$\prod_{i=\text{a、b、c}}(\gamma_i I_i - S_i) > 0, \sum_{i=\text{a、b、c}}(\gamma_i I_i - S_i) < 0$   | 演化稳定点                                |
| $D_9$   | $\gamma_i I_i - S_i > 0$ 时行列式为正，迹值为正  | 不稳定点，（ $\gamma_i I_i - S_i < 0$ 时鞍点） |
| $D_{10}$  | $\gamma_i I_i - S_i > 0$ 时行列式为正，迹值为正  | 不稳定点，（ $\gamma_i I_i - S_i < 0$ 时鞍点） |
| $D_{11}$  | $\gamma_i I_i - S_i > 0$ 时行列式为正，迹值为正  | 不稳定点，（ $\gamma_i I_i - S_i < 0$ 时鞍点） |
| $D_{12}$  | $\gamma_i I_i - S_i < 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_{13}$  | $\gamma_i I_i - S_i < 0$ 时行列式为负，迹值不确定   | 鞍点                                   |
| $D_{14}$  | $\gamma_i I_i - S_i < 0$ 时行列式为负，迹值不确定   | 鞍点                                   |

## 5.2 进化路径分析

由模型设定，普通网民在监管机构的激励下积极参与是这个网络谣言治理合作体系得以存在的基础。普通网民参与谣言治理获得的收益大于其投入成本，即 $\partial_i E + S_i > \gamma_i I_i$ ， $i = a, b, c$ ，其中 $\partial_i E \approx 0$ 。同时，任何一个普通网民的初始策略点和进化策略点都应该在三维空间 $V = \{0 \leq p_{\text{甲}} \leq 1, 0 \leq p_{\text{乙}} \leq 1, 0 \leq p_{\text{丙}} \leq 1\}$ 之内，如图 3 所示。

当存在网民在网络谣言治理过程中采取不参与策略，则少数普通网民可以获得正向收益，其它普通网民获得负向收益时， $p_{\text{甲}} = 0$ 、 $p_{\text{乙}} = 0$ 、 $p_{\text{丙}} = 0$ 是稳定策略，这时候由于关于谣言的证据链条不足，因此监管机构不能核实是否谣言并将保证金以奖励的形式弥补大多数节点的投入成本，仅有直接证据网民有正向收益，此时 $D_2$ 、 $D_3$ 、 $D_4$ 、 $D_{12}$ 、 $D_{13}$ 、 $D_{14}$ 为鞍点，且只能进化到点 $D_1 (0,0,0)$ 。

当多数普通网民在参与谣言治理过程中有正向收益，少数普通网民在参与谣言治理过程中

为负向收益时,  $p_{甲} = 1$ 、 $p_{乙} = 1$ 、 $p_{丙} = 1$ 是稳定策略, 此时监管机构应该能够将保证金奖励转化可以弥补大多数普通网民参与谣言治理的投入成本, 这样网民参与网络谣言治理体系的长期博弈结果将为{参与, 参与, 参与}, 此时 $D_5$ 、 $D_6$ 、 $D_7$ 为鞍点, 且只能进化到 $D_8$  (1,1,1)。

当 $\gamma_i I_i - S_i > 0$ 时,  $D_9$ 、 $D_{10}$ 、 $D_{11}$ 为不稳定点, 并且各自向鞍点 $D_{12}$ 、 $D_{13}$ 、 $D_{14}$ 演化, 其演化的可能性和每个面的区域2面积直接相关, 即面积 $M$ 越大则向鞍点 $D_{12}$ 、 $D_{13}$ 、 $D_{14}$ 演化可能性将越大, 为:

$$M = \frac{\partial_{乙} E - S_{丙} + \gamma_{丙} I_{丙} + 0.5 \partial_{甲} E}{(\partial_{乙} + \partial_{乙} - 0.5 \partial_{甲}) E} + \frac{\partial_{甲} E - S_{丙} + \gamma_{丙} I_{丙} + 0.5 \partial_{乙} E}{(\partial_{甲} + \partial_{丙} - 0.5 \partial_{乙}) E} + \frac{\partial_{甲} E - S_{乙} + \gamma_{乙} I_{乙} + 0.5 \partial_{丙} E}{(\partial_{甲} + \partial_{乙} - 0.5 \partial_{丙}) E} \quad (11)$$

当 $\gamma_i I_i - S_i < 0$ 时,  $D_9$ 、 $D_{10}$ 、 $D_{11}$ 为鞍点, 各自向稳定点 $D_8$  (1,1,1)演化, 演化的可能性与每个面的区域1面积有关, 面积越大向稳定点 $D_8$  (1,1,1)演化可能性越大, 为:  $M = 3 -$

$$\left( \frac{\partial_{乙} E - S_{丙} + \gamma_{丙} I_{丙} + 0.5 \partial_{甲} E}{(\partial_{乙} + \partial_{丙} - 0.5 \partial_{甲}) E} + \frac{\partial_{甲} E - S_{丙} + \gamma_{丙} I_{丙} + 0.5 \partial_{乙} E}{(\partial_{甲} + \partial_{丙} - 0.5 \partial_{乙}) E} + \frac{\partial_{甲} E - S_{乙} + \gamma_{乙} I_{乙} + 0.5 \partial_{丙} E}{(\partial_{甲} + \partial_{乙} - 0.5 \partial_{丙}) E} \right) \quad (12)$$

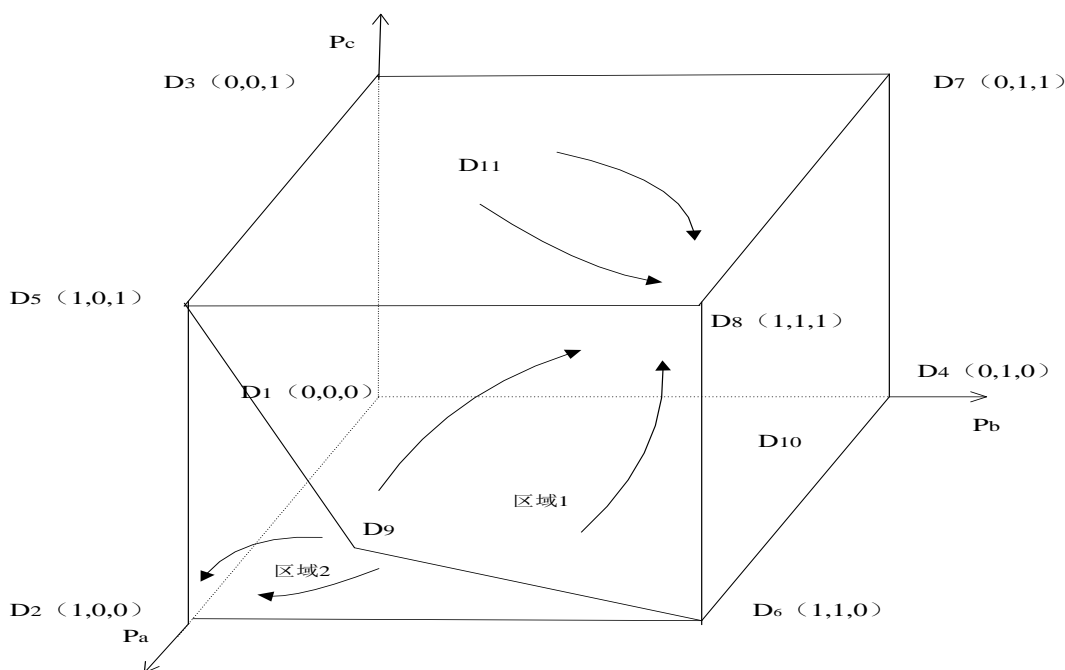


图3 各策略点三维空间图

### 5.3 进化博弈结果分析

对  $M$  求偏导, 分析整个谣言治理博弈系统, 考虑各参数对普通网民参与策略的影响:

$$\frac{\partial M}{\partial E} = \frac{-S_{丙} + \gamma_{丙} I_{丙}}{[(\partial_{乙} + \partial_{乙} - 0.5 \partial_{甲}) E]^2} + \frac{-S_{丙} + \gamma_{丙} I_{丙}}{[(\partial_{甲} + \partial_{丙} - 0.5 \partial_{乙}) E]^2} + \frac{-S_{乙} + \gamma_{乙} I_{乙}}{[(\partial_{甲} + \partial_{乙} - 0.5 \partial_{丙}) E]^2} \quad (13)$$

由上式可以看出, 当 $\gamma_i I_i - S_i < 0$ 时和 $\gamma_i I_i - S_i > 0$ 时情况不同, 得到结论 1。

结论 1: 其他条件不变, 参与谣言治理间接收益足够大时, 谣言监管机构的直接奖励越小, 进化到稳定点 (1, 1, 1) 的可能性越大。当参与谣言治理的间接收益弥补不了网民投入成本时, 谣言监管机构的直接奖励越大, 进化到稳定点 (1, 1, 1) 的可能性越大。

$$\frac{\partial M}{\partial S} = \frac{1}{(\partial_{乙} + \partial_{丙} - 0.5 \partial_{甲}) E} + \frac{1}{(\partial_{甲} + \partial_{丙} - 0.5 \partial_{乙}) E} + \frac{1}{(\partial_{甲} + \partial_{乙} - 0.5 \partial_{丙}) E} \quad (14)$$

由上式可以得到结论 2。

结论 2: 其他条件不变, 网络环境越优化, 监管机构越依赖群体治理合作, 普通网民间接收益越大, 进化到稳定点 (1, 1, 1) 的可能性越大。

为计算方便令 $\partial_{甲} = \partial_{乙} = \partial_{丙}$ , 经计算可知 $\frac{\partial M}{\partial \partial_{甲}} < 0$ , 得到结论 3。

结论 3: 其他条件不变, 参与谣言治理的网民越多, 对谣言核实的证据链条越完整, 参与的分配系数越小, 则进化到稳定点 (1, 1, 1) 的可能性越大。

同理简化表述, 当 $\frac{\partial M}{\partial \gamma_i I_i} < 0$ , 则得到结论 4。

结论 4: 其他条件不变, 普通网民参与网络谣言治理的成本越小, 进化到稳定点 (1, 1, 1) 的可能性越大。

## 6 结论

辟谣失灵的根源在于人们对负面信息赋予更高的信任偏好，而这种偏好的产生与人们的风险偏好具有密切的关系，当事件的真相越模糊，风险偏好者将越相信负面信息。谣言无论是正面信息还是负面信息，从信号博弈角度分析，一旦形成对信息的高度信任，受众对谣言信息的信念将很难被辟谣行为所打破，从这种层面看，高信任度信息传播平台失责对网络谣言泛滥的现象难辞其咎。政府一定要更加严格的监管具有一定信任度信息源的信息发布，如权威网站，网络大 V、社交群体这类高信任度信息源的信息传播行为。对信息传播源头来说，目前网络舆情场存在一种传谣成本较低，辟谣成本较高的态势，如果监管机构能够采取一定的激励措施，以实现对辟谣者的激励为路径加强谣言治理，将有效改变辟谣失灵的整体现状。

激励网民群体参与网络谣言治理，可以进化成群体参与网络谣言治理的稳定体系，能够改善网民已经形成的谣言传播偏好，同时相对一些虚假信息发布源与受众的关系，信息相关方应注意建立与受众更高的信任关系，以降低受众对谣言信息源头的信任度，将会实现较好的辟谣效果。同时这也要求信息相关方在日常的行为中，采取合适的信息相关行为，避免受众形成针对信息相关方信号的反向选择偏好。

#### 参考文献 (References):

- [1] Soroush Vosoughi, Deb Roy and Sinan Aral, "The Spread of True and False News Online", Science, Vol. 359, No. 6380, 2018. 38(05):127-133.
- [2] 马得勇. "匹配效应": 政治谣言的心理及意识形态根源[J]. 政治学研究, 2018, 142(5): 54-66.
- [3] Alessandro B, Mauro C, Alexandru D G, et al. Science vs Conspiracy: Collective Narratives in the Age of Misinformation[J]. Plos One, 2015, 10(2):e0118093.
- [4] 那艺. 贺京同: 行为经济学与实验经济学的学术分野[J]. 经济学动态, 2019(07):109-122.
- [5] 王超. 辟谣何以失灵?——一个信息传播效果视角的解释框架[J]. 情报杂志, 2019, 38(05):127-133.
- [6] Rupa Jose, E. Alison Holman, Roxane Cohen Silver, Community organizations and mental health after the 2013 Boston Marathon bombings, Social Science & Medicine, Volume 222, February 2019:367-376.
- [7] McCluskey J J, Swinnen J, Vandemoortele T. You Get What You Want: A Note on the Economics of Bad News [J]. Information Economics & Policy, 2015 (30) :1-5.
- [8] Svensson H M, Albaek E, Van Dalen A, et al. Good News in Bad News: How Negativity Enhances Economic Efficacy [J]. International Journal of Communication, 2017 (11) :165-187.
- [9] Dobele A, Lindgreen A, Beverland M, et al. Why Pass on Viral Messages? Because They Connect Emotionally [J]. Business Horizons, 2007 (4) :291-304.
- [10] 周晓阳, 王黎琴, 冯平平, 扈袁权, 温浩宇. WSR 方法论视角下基于信任关系、前景理论和犹豫模糊偏好的群决策研究[J]. 管理评论, 2020, 32(07):66-75.
- [11] 郎玫, 简楠. "公""私"之辩: 公民认知偏好结构如何影响公共议题的生成[J]. 管理评论, 2021, 33(02):298-310.

