·人智交互与信息行为·

生成式人工智能服务用户隐私保护三方随机演化博弈与仿真研究

王 艺[1] 李小龙[2] 杨 鹭[1] 朱梦蝶[1] 杨海平[1*]
（1. 南京大学信息管理学院，江苏 南京 210023；2. 厦门大学经济学院，福建 厦门 361005）

摘 要：［目的/意义］多主体协同参与是构建安全可信的生成式人工智能服务生态的关键。本文通过探究GAI服务用户隐私保护情景中相关主体策略行为与影响因素，拓宽了随机演化博弈理论的研究情景，加深了对GAI服务用户隐私泄露动态机制的理解，丰富了GAI隐私保护领域的理论与方法，并为推动GAI服务用户隐私保护实践提供了启示。［方法/过程］基于演化博弈理论及随机过程，构造时变用户隐私泄露风险函数，构建具有GAI服务场景特殊性的“用户—GAI服务提供者—政府”三方随机演化博弈模型，借鉴Itô随机微分方程理论和数值仿真分析三方主体行为策略的稳定性和演化规律。［结果/结论］研究发现：①三方初始意愿影响后续系统演化方向和演化速度，且仅在用户和政府均持有较高初始意愿时，GAI服务提供者才可能选择积极保护策略并使系统演化至理想状态。②随机扰动强度越大，三方博弈主体演化至稳定策略的波动性和不确定性越高，且GAI服务提供者对不确定因素更为敏感。③GAI服务市场优胜者和市场参与者均能在三方高初始意愿条件下收敛至理想状态的稳定策略，而市场生存者难以演化至理想状态；相较于通用型GAI服务提供者，垂直型GAI服务提供者策略具有更快的正向收敛速度。④政府增加对GAI服务提供者的罚金并提供适中奖励，会提高其积极保护隐私概率，进而降低其泄露隐私概率，而政府降低对用户不实举报惩罚力度，虽可提升用户披露隐私概率，但若对服务提供者的罚金过低，则无法促使其采取积极保护策略；政府增加用户维权赔偿或减少维权成本，均能推动其积极披露隐私。

关键词：隐私保护；生成式人工智能；随机演化博弈；政府监管；隐私泄露
DOI：10.3969/j.issn.1008-0821.2026.03.007
〔**中图分类号**〕TP183；G252.0 〔**文献标识码**〕A 〔**文章编号**〕1008-0821（2026）03-0081-18

Tripartite Stochastic Evolutionary Game and Simulation Research on User Privacy Protection of Generative Artificial Intelligence Services

Wang Yi[1] Li Xiaolong[2] Yang Lu[1] Zhu Mengdie[1] Yang Haiping[1*]
（1. School of Information Management，Nanjing University，Nanjing 210023，China；
2. School of Economics，Xiamen University，Xiamen 361005，China）

Abstract：［Purpose/Significance］Multi-subject collaborative participation is the key to building a secure and trustworthy Generative Artificial Intelligence（GAI）service ecosystem. This paper explores the strategic behaviors and influencing factors of relevant subjects in the context of GAI service user privacy protection，broadens the research scenarios of stochastic evolutionary game theory，deepens the understanding of the dynamic mechanism of privacy leakage for GAI service users，enriches the theories and methods in the field of GAI privacy protection，and provides enlightenment for promoting the practice of privacy protection for GAI service users.［Method/Process］This paper，based on evolutionary game theory and stochastic

收稿日期：2025-08-03
作者简介：**王艺**（2001-），男，博士研究生，研究方向：用户信息行为。**李小龙**（2000-），男，硕士研究生，研究方向：博弈论与实验经济学。**杨鹭**（2001-），女，硕士研究生，研究方向：媒介传播。**朱梦蝶**（1998-），女，博士研究生，研究方向：信息治理。
通信作者：**杨海平**（1967-），男，教授，博士生导师，研究方向：知识服务、数据安全。

processes，constructed a time-varying user privacy leakage risk function，and built a “user-service provider-government” three-party stochastic evolutionary game model with the particularity of the GAI service scenario. By applying Itô stochastic differential equation theory and numerical simulation，this paper analyzed the stability and evolution of the behavior strategies of these three parties.［Result/Conclusion］The findings indicate that：①The initial willingness of the government，service providers，and users significantly influences the direction and speed of subsequent system evolution；only when both users and the government exhibit high initial willingness could GAI service providers adopt proactive protection strategies，leading the system toward an optimal state. ②The greater the intensity of random disturbances，the slower the convergence speed of the three parties to a stable strategy，and GAI service providers are more sensitive to uncertain factors. ③Under conditions of high initial willingness from all three parties，market winners and participants converge to the ideal state，whereas market survivors find it challenging to reach this state；vertical GAI service providers demonstrate faster forward convergence compared to general GAI service providers. ④When the government increases the fines for GAI service providers and provides moderate rewards，it would increase the probability of their active privacy protection and thereby reduce the probability of privacy leakage. However，if the government reduces the punishment for users’ false reports，although it increases the probability of users disclosing privacy，if the fines for service providers are too low，it would not prompt them to adopt an active protection strategy. Increasing user rights protection compensation or reducing the cost of rights protection by the government could both encourage users to actively disclose privacy.

Keywords：privacy protection；generative artificial intelligence；stochastic evolutionary game；government regulation；privacy leakage

生成式人工智能(Generative Artificial Intelligence，GAI)作为新兴技术，近年来取得突破性进展并服务于公众。GAI服务意指利用GAI技术提供生成文本、图片、音视频等服务[1]。中国在AI领域具有数据丰富、资金充足、政策支持、用户规模等优势。截至2024年底，已备案GAI服务达302款，用户规模达2.49亿人[2]。但GAI服务依赖于海量数据训练，这使其不可避免地伴随着更高的隐私泄露风险，尤其在协同创作服务场景，用户个人信息被纳入训练语料库，进一步加剧这种风险。严重的隐私泄露既妨害互联网信息秩序，还阻碍GAI行业长效发展。因此，需逐步建立健全法律和规章制度，保障GAI服务中的用户隐私和数据安全。

本文中GAI服务背景下用户隐私保护核心是保证用户隐私权在使用GAI服务时不受侵犯，受侵害时可寻求法律保护。此过程中，政府、服务提供者和用户需依自身利益选择策略[3-5]，平衡潜在收益与风险，持续学习调整，最终平衡用户隐私有效保护与GAI服务发展。博弈理论与数值仿真方法[6]已广泛应用于用户隐私保护研究，但GAI服务具有数据处理复杂、交互模式多变、服务链条多元等特性[7]，其隐私保护面临数据迭代、算法安全漏洞、交互信息泄露、跨平台数据流转等新型风险[8]。传统静态用户隐私保护演化博弈模型难以分析其多主体持续动态博弈、隐私策略随数据与技术迭代快速演变的复杂局面。故亟须针对GAI服务用户隐私保护风险特征，构建具有场景特殊性的多主体用户隐私保护演化博弈模型，以精准揭示隐私保护策略的动态演化规律。

因此，本文基于演化博弈理论与随机过程，构建政府、GAI服务提供者、用户三方演化博弈模型，旨在：①探讨用户隐私保护的三方演化博弈过程及其影响因素，分析用户隐私泄露动态机制，理解政府、GAI服务提供者及用户在隐私保护决策中的互动关系、策略选择及其对用户隐私保护的影响。②分析成本、收益视角下各方行为选择的影响因素，引入随机干扰模拟外部冲击，进行稳定性分析。③探究参与主体初始意愿、随机扰动、GAI服务提供者类型、政府规制因素等对系统演化的不同影响。

本文在理论上：①构建具有GAI服务场景特殊性的三方随机演化博弈模型，借鉴Itô随机微分方程理论和数值仿真分析策略稳定性和演化规律，为GAI服务用户隐私保护提供新颖的研究视角与方法思路，拓宽了随机演化博弈理论的研究情景。②分析GAI服务应用场景特殊性，构造时变隐私泄露风

险函数，优化了既有的用户隐私演化博弈模型，深化了对于GAI服务用户隐私保护变化特征与演化规律的理解。在实践上，有助于引导政策制定、平台管理和用户行为，助力各方协同构建安全可信的GAI服务生态。

1 相关研究

1.1 GAI服务用户隐私保护研究

GAI技术发展伴生隐私泄露，易引起隐私危机并影响自身发展。当前，欧盟[9]、美国[10]、中国[1]等围绕GAI服务用户隐私保护建立法律体系，增设隐私保护机构，提供法律与机构保障。但既有研究表明，现有法律体系和治理框架应对GAI隐私风险仍有短板[11]（如公法约束效力有限、弱势群体保护缺失、损害赔偿制度不明晰、跨境法律标准不一等[12]）。GAI服务提供者常通过技术开发、规章制度、机构设置实现隐私保护与伦理自律。目前主要通过数据匿名化等技术降低泄露风险[13]，借AI打造监测平台辅助用户识别泄露隐患；中国主要数字科技企业设立隐私保护或技术伦理委员会，制定规章制度规范业务中的隐私保护，同时发布白皮书、倡议书推动行业隐私自律体系的完善[14]。用户因感知隐私风险、信息敏感程度、认知因素等个体差异，使用GAI服务的隐私边界管理主导规则亦不同[15]。研究表明，用户对GAI应用的隐私政策维度认知、风险感知对其持续使用意愿、隐私保护行为影响显著[16-17]。但当前用户隐私保护立法进程滞后于GAI发展[18]，难以准确定位规范数据安全与隐私问题，侵权责任划分、隐私保护与技术创新平衡等尚待探索[19]，体系上存在内容不完整、举措不兼容等问题[20]。而用户隐私保护带来的成本上升与数据“孤岛现象”可能阻碍GAI技术发展。同时，服务提供者常使用的告知用户同意办法流于形式，亦面临有效性困境[21]。上述问题共同对用户自身使用意愿、隐私保护态度与行为构成挑战。故GAI用户隐私保护需紧跟技术发展，构建隐私友好型技术架构和数据共享机制，平衡用户隐私保护与数据共享利用[22]，同时需挖掘国际及社会各方的合作潜力，共同确立并完善隐私保护的伦理原则与市场机制。

上述研究多聚焦政府、服务提供者、用户等单一主体行为，探讨GAI服务用户隐私保护风险、影响因素及应对措施等核心议题，但鲜有研究剖析三方利益关联及互动策略，为本文从多方利益博弈视角切入GAI服务用户隐私保护研究提供契机。

1.2 用户隐私保护相关利益方演化博弈研究

不少学者从静态利益博弈视角研究隐私保护相关利益方的意愿与行为策略[23-24]。但此类经典博弈论研究与现实情景有限理性和信息非对称的现象相悖。而演化博弈理论基于主体有限理性和信息不完全对称，结合博弈论和生物动态进化分析，为GAI服务用户隐私保护研究提供更贴合现实的动态框架[25]。

目前，已有学者基于演化博弈理论研究用户隐私保护和信息安全问题。社交媒体领域，多数研究将政府、平台、用户三方作为利益相关方。雷丽彩等[3]依据仿真结果提出政府监管优化建议。曲薪池等[26]指出政府应降低用户维权成本，加大对泄露隐私惩罚力度。丰米宁等[27]指出隐私边界可判别隐私保护过度或不足。亦有研究认为此领域主要是社交网络平台和用户的双方博弈[28]。医疗服务领域，朱光等[29]证实移动医疗服务商与用户的隐私策略关联问责成本、监管力度和隐私损失等。邱均平等[4]在利益相关方加入政府方后发现，用户披露隐私是最优演化路径的驱动力。其他在线服务业[5, 30]研究虽主体划分不同，但均认为政府作用关键，促进隐私保护需要提升用户、服务提供者保护收益，降低保护成本。

综上，已有研究大多从政府、服务提供者等单一主体行为出发，探讨GAI服务用户隐私保护的风险、应对措施等议题；博弈视角研究则多聚焦互联网平台用户隐私保护路径，鲜有涉足GAI服务用户隐私保护领域。故上述研究虽为本文提供借鉴，但仍存在可以丰富的地方：①大多基于政府、服务提供者、用户等单一视角研究GAI服务用户隐私保护路径，缺乏对三方交互策略的综合探讨；②大多在确定性条件下构建演化博弈模型，而忽视演化过程不可避免地受到情绪、道德风险、社会利益、技术效应等复杂因素影响，亦未考虑用户隐私风险的时变特性；③虽有研究已从演化博弈视角关注特定视角GAI服务特征[31]，但尚未针对用户隐私风

险变化特征、市场份额与应用范围异质性等情景特殊性进行深入分析。

因此，本文立足用户、服务提供者、政府的动态交互关系，基于随机演化博弈视角分析GAI服务中用户隐私保护议题，构建“用户—GAI服务提供者—政府”三方随机演化博弈模型。本文尝试突破传统互联网服务平台用户隐私保护博弈研究的静态隐私风险假设，引入时变用户隐私泄露风险函数，从成本、收益角度分析各方采取演化稳定策略条件及影响因素，综合考虑初始意愿、随机扰动、GAI服务应用场景和政府规制因素对三方决策影响，明确各方策略演变内在机理与博弈规律，以期为GAI服务用户隐私保护提供新颖研究视角与方法思路，深化对GAI服务用户隐私保护变化特征与演化规律的理解，推动政府政策、服务提供者内部管理与用户行为优化调整，助力各方协同构建安全可信的GAI服务生态。

2 理论基础与研究假设

2.1 理论基础

相较于传统AI，GAI隐私保护面临技术复杂性、数据驱动性、风险多样性等特性带来的挑战[13]。①技术层面，GAI技术算法复杂性使其存在技术黑箱，用户获取有效知情同意受阻[32]，且易面临数据记忆和模型反转攻击风险[33]。②数据方面，GAI需要大量数据以提升模型性能，部分功能需获知隐私信息方可完整使用，用户隐私数据保护与模型功能使用间存在矛盾[34]。③风险多样性方面，成员推理攻击、模型中毒和对抗性攻击[35]及微调造成的数据泄漏等亦是GAI亟待解决的隐私风险。而由于GAI应用场景广泛，其隐私风险亦呈现多样性、隐蔽性。不同类型GAI服务提供者采取不同程度隐私保护策略，易形成“数据孤岛”，阻碍行业发展。加之GAI发展迅猛，现有隐私保护法律体系尚未完全适应[18]。以上特性与挑战使得GAI服务用户隐私保护三方策略选择及其相互关系具有特殊性。

不同于传统互联网平台与基于机器学习算法应用，GAI因模型迭代、强交互性、数据规模膨胀等特征使其服务衍生出动态时变的用户隐私泄露风险：①传统机器学习算法应用训练过程与隐私风险模式相对固定，而GAI迭代优化模型时，可能因数据输入规模扩大、处理复杂度提升，导致隐私泄露概率随迭代次数呈非线性增长[8]。②传统应用多为单向一次性数据采集，GAI服务的实时对话特性催生过程性隐私损耗，用户多轮交互中暴露的情绪倾向、知识水平等碎片与情境化信息，可使GAI自动推断出更多个人信息[36]，形成传统情景不具备的渐进式隐私泄露。③传统平台数据规模增长较缓，隐私风险较为可控，而GAI依赖海量多模态用户数据，数据规模呈指数级膨胀，导致数据存储、传输、处理环节攻击面扩大，且数据关联性增强，均使隐私泄露风险在数据流动中不断演变[37]。这些显著区别于传统互联网用户隐私保护的风险特征与动态机制亦使得GAI用户隐私保护研究具有特殊性与必要性，并为GAI用户隐私保护博弈系统构建与研究假设提出奠定理论基础。

2.2 基本假设

假设1：GAI用户隐私保护博弈系统由GAI服务提供者、用户与政府组成，三者均为有限理性，其目标旨在实现自身利益最大化，行为上仅依据个人收益与成本的直接比较来做出决策，且高度灵活，能实时根据收益状况波动不断优化和调整自身策略，最终策略选择趋于最优状态[4]。

假设2：GAI服务用户博弈策略为针对个人隐私的披露或不披露。其会面临隐私悖论问题，意指用户担心隐私泄露，但又为获取GAI服务而不得不披露个人隐私，因而产生博弈冲突。相较其他技术与服务，用户使用GAI服务时需在知情同意基础上被动乃至主动披露更多信息以享受更全面精准的服务。而根据现有研究结果，用户通过GAI应用大多能在内容创作[38]、生产力辅助[39]、技能学习[40]等多维度获得便利，并取得收益$V_1(V_1>0)$。但GAI项目生命周期不同阶段均面临隐私泄露风险[11]，加之技术存在隐蔽性、复杂性与算法黑箱，用户受认知限制难以觉察隐私泄露，且知情同意机制模糊，故采取披露策略时，其隐私面临泄露风险，并随之产生直接损失$D_1(D_1>0)$。若用户察觉隐私泄露，会以概率$a(0\leqslant a\leqslant 1)$展开维权行为，维权成本为$D_2(D_2>0)$，可获得提供者赔偿收益$V_2$[26]$(V_2>0)$。因GAI服务存在多种隐私泄露风险且涉及多企业协作服务[11]，维权对象可能非实际泄露

者，给未泄露用户隐私而遭遇用户维权的提供者带来声誉损害，使其产生间接损失$P_1(P_1>0)$，此时其会支付保誉成本$B_1(B_1>0)$用以恢复。若用户不披露隐私，则无隐私泄露风险，无收益或赔偿，无需承受损失或成本。

假设3：GAI服务提供者博弈策略为针对用户个人隐私的积极保护或消极保护。当GAI服务提供者采取积极保护策略时，会投入更多资源用于维护用户隐私安全，保护成本增加至$C_1(C_1>0)$，并使用户隐私泄露初始风险显著降低至b_0，其会随着模型迭代次数、用户使用频率及多模态数据量累积而增加；而积极保护下提供者会在基础隐私保护技术上持续优化算法，提升保护水平，进而减缓隐私泄露风险的增加，甚至降低泄露风险至b_t（若政府消极监管则记为b_t'，$0\leqslant b_t\leqslant 1$）。同时，积极保护隐私可为提供者带来声誉收益$M_1(M_1>0)$。另外，若提供者选择消极保护，其投入成本会减少至$C_2(C_1>C_2>0)$，并增加隐私泄露初始风险至$c_0$，其仅在基础保护技术上随时间演化至$c_t$（若政府消极监管则记为$c_t'$，$0\leqslant c_t\leqslant 1$）。一旦发生用户隐私泄露事件，提供者虽会获得直接收益$M_2(M_2>0)$，但其声誉难免下降，导致流量受损、用户信任降低、品牌形象恶化及长期盈利能力削弱等负面影响，造成损失$C_3(C_3>0)$。同时，由于用户针对提供者展开的维权行为，泄露隐私的提供者产生赔偿损失V_2，未泄露隐私的提供者产生间接损失P_1和保誉成本B_1[26]。但此时，未泄露隐私的提供者通过维护声誉可扩大自身市场影响力，间接获得保誉收益$r_1(r_1>0)$。

本文认为，积极保护策略下，GAI动态隐私泄露风险的演化趋势由两类核心效应博弈决定：一是基础保护技术与技术优化措施形成的风险抑制效应，二是数据量增长、模型迭代及用户交互提升带来的风险驱动效应。若抑制效应显著强于驱动效应，风险会从初始水平b_0逐步降低（不低于0）；若抑制效应未能覆盖驱动效应，说明保护技术优化不足，风险将逐步上升（不超过1）。隐私泄露风险的动态演化过程呈典型非线性特征，而以保护技术水平为分段依据的指数函数型分段函数，能够表示这一非线性过程，从而为动态隐私泄露风险的量化与预测提供合理的数学表达范式。据此，本文尝试突破传统互联网服务平台用户隐私保护博弈研究的静态隐私风险假设，构造动态隐私泄露风险函数b_t，其随时间变化如式（1）所示。其中，μ_0、μ、δ_0、δ、λ、β分别表示基础保护技术、积极保护下技术随时间的优化、基准数据量、多模态数据量随时间的增加、模型随时间的迭代、披露用户的使用频率对泄露风险的影响系数，在演化过程中共同作用于隐私泄露风险。时变用户隐私泄露风险函数表达如式（1）所示：

$$b_t=\begin{cases}b_0e^{-(\mu_0+\mu\cdot t-\delta_0-\delta\cdot t-\lambda-\beta x)t}, & \text{if } \mu_0+\mu\cdot t\geqslant\delta_0+\delta\cdot t+\lambda+\beta x\\ 1+(b_0-1)e^{-(\delta_0+\delta\cdot t+\lambda+\beta x-\mu_0-\mu\cdot t)t}, & \text{else}\end{cases}$$
$$c_t=\begin{cases}c_0e^{-(\mu_0-\delta_0-\delta\cdot t-\lambda-\beta x)t}, & \text{if } \mu_0\geqslant\delta_0+\delta\cdot t+\lambda+\beta x\\ 1+(c_0-1)e^{-(\delta_0+\delta\cdot t+\lambda+\beta x-\mu_0)t}, & \text{else}\end{cases}\quad(1)$$

假设4：政府博弈策略为针对GAI服务用户隐私保护的积极监管或低投入监管。GAI能推动数字经济发展，提升政府财政收入[41]，可为政府带来的初始收益为$L_1(L_1>0)$。中国对GAI服务实行包容审慎和分类分级监管[1]，但目前尚无成熟法律与制度评估风险，亟须找到积极监管与鼓励发展间平衡点。参考《生成式人工智能服务管理暂行办法》[1]、“清朗·整治AI技术滥用”专项行动[42]等制度实践，若政府选择积极监管，需通过优化监管模式、增加执法人员等[43]加大力度，投入更高成本$N_1(N_1>0)$，但亦能提升政府公信力、公众满意度、提供者隐私保护水平和用户信任水平，促进用户积极披露个人隐私，使政府获得声誉收益L_2（$L_2>0$）和用户信任收益T[31]（$T>0$）。若用户不披露个人隐私而无法使用GAI服务，则政府不会获得声誉收益，且GAI行业发展亦会因用户流失受阻，政府财政收入下降至$L_3(L_1>L_3>0)$。若政府积极监管，会给予泄露用户隐私的提供者惩罚[1]（监管批评、经济罚款等，参考2024年底OpenAI被意大利数据保护局处以1 500万欧元罚款[44]），记为罚金$F(F>0)$。但用户维权存在不实举报[1, 45]，这既会增加政府工作负担，又损害提供者声誉，

故假定在政府积极监管时，政府将对用户不实举报行为进行惩罚，记$f(f>0)$为政府对用户不实举报的罚金[26]。为鼓励GAI健康发展，政府会给予采取积极保护措施的提供者奖励扶持（参照欧盟《人工智能法案（2024/1689）》[46]），形成奖励成本$N_2(N_2>0)$。而当政府消极监管时，政府投入成本为$N_3(N_1>N_3>0)$。若用户披露隐私，因监管不力，隐私泄露初始概率大幅上升（上升幅度为Δp $(0\leqslant\Delta p\leqslant 1)$），造成GAI服务市场失序而面临行业危机、公信力下降及公共满意度下跌，致使政府声誉下降至$N_4(N_4>0)$，财政收入下降至$L_4(L_1>L_4>0)$。若用户不披露个人隐私而无法使用GAI服务，则政府无声誉损失，但GAI行业发展受阻，财政收入仍下降至L_3。

假设5：用户披露隐私、GAI服务提供者积极保护隐私和政府实施积极监管的初始意愿分别为$x_0,y_0,z_0\in[0,1]$，三方博弈策略随时间变化而动态调整。

2.3 收益矩阵

根据上述假设，不同策略下用户、GAI服务提供者及政府收益矩阵如表1所示，其中Ex_1至Ex_8、Ey_1至Ey_8、Ez_1至Ez_8分别表示三者选择相应策略时各自的收益。

表1 用户、GAI服务提供者及政府的收益矩阵

Tab. 1 Revenue Matrix of Users, GAI Service Providers and Government

用户	政府积极监管(z)		政府消极监管($1-z$)	
	GAI服务提供者积极保护(y)	GAI服务提供者消极保护($1-y$)	GAI服务提供者积极保护(y)	GAI服务提供者消极保护($1-y$)
披露(x)	$Ex_1=V_1-b_tD_1+ab_t(V_2-D_2)+T-a(1-b_t)f$	$Ex_2=V_1-c_tD_1+ac_t(V_2-D_2)+T-a(1-c_t)f$	$Ex_3=V_1-b_t'D_1+ab_t'(V_2-D_2)$	$Ex_4=V_1-c_t'D_1+ac_t'(V_2-D_2)$
	$Ey_1=-C_1+M_1+b_t(M_2-M_3)-a(b_tV_2+(1-b_t)(B_1+P_1-r_1))-b_tF+N_2$	$Ey_2=-C_2+c_t(M_2-C_3)-a(c_tV_2+(1-c_t)(B_1+P_1-r_1))-c_tF$	$Ey_3=-C_1+M_1+b_t'(M_2-C_3)-a(b_t'V_2+(1-b_t')(B_1+P_1+r_1)$	$Ey_4=-C_2+c_t'(M_2-C_3)c_t'-a(c_t'V_2+(1-c_t')(B_1+P_1-r_1)$
	$Ez_1=L_1-N_1+L_2-N_2$	$Ez_2=L_1-N_1+L_2$	$Ez_3=L_4-N_3-N_4$	$Ez_4=L_4-N_3-N_4$
不披露($1-x$)	$Ex_5=0$	$Ex_6=0$	$Ex_7=0$	$Ex_8=0$
	$Ey_5=-C_1+M_1+N_2$	$Ey_6=-C_2$	$Ey_7=-C_1+M_1$	$Ey_8=-C_2$
	$Ez_5=L_3-N_1-N_2$	$Ez_6=L_3-N_1$	$Ez_7=L_3-N_3$	$Ez_8=L_3-N_3$

3 演化稳定策略与系统均衡点

3.1 三方主体演化策略分析

根据用户、GAI服务提供者及政府的收益矩阵，用户披露隐私(x)、不披露隐私($1-x$)的期望收益及平均期望收益可表示为式（2）~（4），GAI服务提供者积极保护(y)、消极保护($1-y$)的期望收益以及平均期望收益可表示为式（5）~（7），政府积极监管(z)、消极监管($1-z$)的期望收益以及平均期望收益可表示为式（8）~（10）：

$$E_x=yzEx_1+(1-y)zEx_2+y(1-z)Ex_3+(1-y)(1-z)Ex_4 \tag{2}$$

$$E_{1-x}=0 \tag{3}$$

$$\overline{E_x}=xE_x+(1-x)E_{1-x} \tag{4}$$

$$E_y=xzEy_1+(1-x)zEy_5+x(1-z)Ey_3+(1-x)(1-z)Ey_7 \tag{5}$$

$$E_{1-y}=xzEy_2+(1-x)zEy_6+x(1-z)Ey_4+(1-x)(1-z)Ey_8 \tag{6}$$

$$\overline{E_y}=yE_y+(1-y)E_{1-y} \tag{7}$$

$$E_z=xyEz_1+x(1-y)Ez_2+(1-x)yEz_5+(1-x)(1-y)Ez_6 \tag{8}$$

$$E_{1-z}=xyEz_3+x(1-y)Ez_4+(1-x)yEz_7+(1-x)(1-y)Ez_8 \tag{9}$$

$$\overline{E_z}=zE_z+(1-z)E_{1-z} \tag{10}$$

依据演化博弈理论，当某种策略期望收益大于平均预期收益时，在群体中选择该策略的个体趋多，表现为选择该策略概率增大，这一过程可用复制动态方程来表示[47]。根据三方各自两种策

略的收益以及平均预益，可得三方的复制动态方程系统，如式（11）所示：

$$\begin{cases} F_x = \dfrac{dx}{dt} = x(E_x - \overline{E_x}) = x(1-x)f_1(y,z) \\ F_y = \dfrac{dy}{dt} = y(E_y - \overline{E_y}) = y(1-y)f_2(x,z) \\ F_z = \dfrac{dz}{dt} = z(E_z - \overline{E_z}) = z(1-z)f_3(x,y) \end{cases} \quad (11)$$

其中F_x、F_y、F_z分别表示用户、GAI服务提供者及政府的复制动态方程，且$f_1(y,z) = E_x - E_{1-x}$，$f_2(x,z) = E_y - E_{1-y}$，$f_3(x,y) = E_z - E_{1-z}$。

用户的复制动态方程F_x的一阶导数$\dfrac{dF_x}{dx} = (1-2x)f_1(y,z)$，其稳定策略必须满足$F_x = 0$且$\dfrac{dF_x}{dx} \leqslant 0$。当$(y,z)$满足$f_1(y,z) = 0$时，$F_x = 0$和$\dfrac{dF_x}{dx} = 0$，则$x \in [0,1]$都具有稳定性，用户的稳定策略不确定；当$(y,z)$满足$f_1(y,z) < 0$时，$F_x(x=0) = 0$和$\dfrac{dF_x}{dx}(x=0) < 0$，则$x = 0$具有稳定性，用户的稳定策略是不披露隐私；当$(y,z)$满足$f_1(y,z) > 0$时，$F_x(x=1) = 0$和$\dfrac{dF_x}{dx}(x=1) < 0$，则$x = 1$具有稳定性，用户的稳定策略是披露隐私。GAI服务用户和政府的稳定策略分析与用户的分析一致。

3.2 随机演化博弈模型构建

GAI服务隐私保护博弈系统作为复杂系统，其动力机制必然具有不确定性。故鉴于现实系统复杂性及不确定性扰动，为刻画随机扰动对三方动态决策过程的影响，将高斯白噪声引入复制动态方程中，如式（12）所示：

$$\begin{cases} dx(t) = x(t)(1-x(t))f_1(y,z)dt + \sigma x(t)(1-x(t))d\omega(t) \\ dy(t) = y(t)(1-y(t))f_2(x,z)dt + \sigma y(t)(1-y(t))d\omega(t) \\ dz(t) = z(t)(1-z(t))f_3(x,y)dt + \sigma z(t)(1-z(t))d\omega(t) \end{cases} \quad (12)$$

其中，$\omega(t)$服从标准一维Brown运动，作为一种无规则的随机涨落现象，可有效反映随机扰动对博弈主体的影响，$d\omega(t)$表示高斯白噪声，当$t > 0$时，步长$h > 0$，其增量$\Delta\omega(t) = \omega(t+h) - \omega(t)$服从正态分布$N(0,h)$；$\sigma$表示随机扰动的强度系数，$0 \leqslant \sigma \leqslant 1$。

3.3 均衡解的存在及稳定性分析

根据式（12）所示，假设当$t = 0$时，即三方演化博弈的初始时刻，$x(0) = 0$，$y(0) = 0$，$z(0) = 0$，此时进一步表明方程至少存在零解，即未受到干扰时，零解是方程均衡解。但现实中环境多变，需考虑随机因素对博弈策略的影响。式（12）的稳定性可根据随机微分方程稳定性判别定理进行判别，给定一个随机微分方程，如式（13）所示：

$$dx(t) = f(t,x(t))dt + g(t,x(t))d\omega(t), x(t_0) = x_0 \quad (13)$$

设存在函数$V(t,x)$与正常数c_1、c_2使得$c_1|x|^p \leqslant V(t,x) \leqslant c_2|x|^p, t \geqslant 0$。

①当存在正常数γ，使得$LV(t,x) \leqslant -\gamma V(t,x), t \geqslant 0$，则方程（13）的零解$p$阶矩指数稳定，且成立$E\left|x(t,x_0)\right|^p < (c_2/c_1)\left|x_0\right|^p e^{-\gamma t}, t \geqslant 0$。

②当存在正常数γ，使得$LV(t,x) \geqslant \gamma V(t,x), t \geqslant 0$，则方程（13）的零解$p$阶矩指数不稳定，且成立$E\left|x(t,x_0)\right|^p \geqslant (c_2/c_1)\left|x_0\right|^p e^{-\gamma t}, t \geqslant 0$。

根据上述引理，得出针对式（12）的稳定性判定依据，取$V(t,x) = x$、$V(t,y) = y$、$V(t,z) = z$、$c_1 = c_2 = 1$、$p = 1$、$\gamma = 1$，由于$x,y,z \in [0,1]$，因此$1-x$、$1-y$、$1-z$均为非负数，其不影响演化方向，可省略，并根据随机微分方程稳定性判别定理，若零解矩指数稳定，得到条件如式（14）所示：

$$\begin{cases} LV(t,x) = f_1(y,z)x \leqslant -x \\ LV(t,y) = f_2(x,z)y \leqslant -y \\ LV(t,z) = f_3(x,y)z \leqslant -z \end{cases} \quad (14)$$

若条件（14）同时成立，则零解矩指数稳定，用户、GAI服务提供者和政府会稳定于（不披露，消极保护，消极监管）策略，若均不成立，则三方将会选取（披露，积极保护，积极监管）策略。由于式（12）为非线性Itô随机微分方程，不能直接获取其解析解，故本文采用Euler-Maruyama方法进行数值求解，如式（15）所示：

$$\begin{cases} x(t_{n+1}) = x(t_n) + f(x(t_n))h + g(x(t_n))\Delta\omega_n \\ y(t_{n+1}) = y(t_n) + f(y(t_n))h + g(y(t_n))\Delta\omega_n \\ z(t_{n+1}) = z(t_n) + f(z(t_n))h + g(z(t_n))\Delta\omega_n \end{cases} \quad (15)$$

社会所期望的稳定状态为：用户披露隐私以获取服务，这既能促进GAI服务产业发展，亦可为

政府提供可观财政收益；GAI服务提供者积极保护用户隐私以提升市场声誉、规避行业信任危机；政府采取积极监管措施，以保障GAI服务产业健康发展，经政府背书后亦会提升用户信任。故为达到社会所期望的稳定策略(披露，积极保护，积极监管)，即稳定于纯策略均衡点(1, 1, 1)，本文将此策略作为主要分析对象。

为分析多种参数情形下对各主体策略选择的影响，本文利用MATLAB R2024a软件进行数值仿真，直观展示特定参数下三方演化路径和规律。本文参考已有研究[26, 29-30]，依据GAI行业相关报告的真实数据[48]与政策标准的原则性要求[1, 49]，并结合研究假设对部分难以参照真实数据赋值的参数进行自行设计，确保数据量纲一致性。本文还通过德尔菲法咨询学界、业内专家的意见进行数据校准，最终选取初始取值，如表2所示。同时，设定x, y, z初始值为$\Omega(0.2, 0.5, 0.8)$，取隐私泄露风险函数相关系数$\mu_0 = 1$、$\mu = 0.1$、$\delta_0 = 0.5$、$\delta = 0.05$、$\lambda = 0.05$、$\beta = 0.5$，进行多次仿真，取同等条件下10次重复实验平均值。经测试，本文将演化时间设置为3，时间步长为0.01，以便更清晰地展示各主体随时间的演化过程。

表2 参数初始值选择

Tab. 2 Selection of Initial Values for Parameters

参数	取值	参数	取值	参数	取值
V_1	10	C_1	20	L_4	10
V_2	10	C_2	10	N_1	15
T	5	C_3	5	N_2	10
D_1	15	P_1	20	N_3	5
D_2	25	B_1	20	N_4	10
f	10	F	15	a	0.8
M_1	10	L_1	20	b_0	0.3
M_2	15	L_2	10	c_0	0.7
r_1	5	L_3	5	Δp	0.1

4 演化博弈数值仿真分析

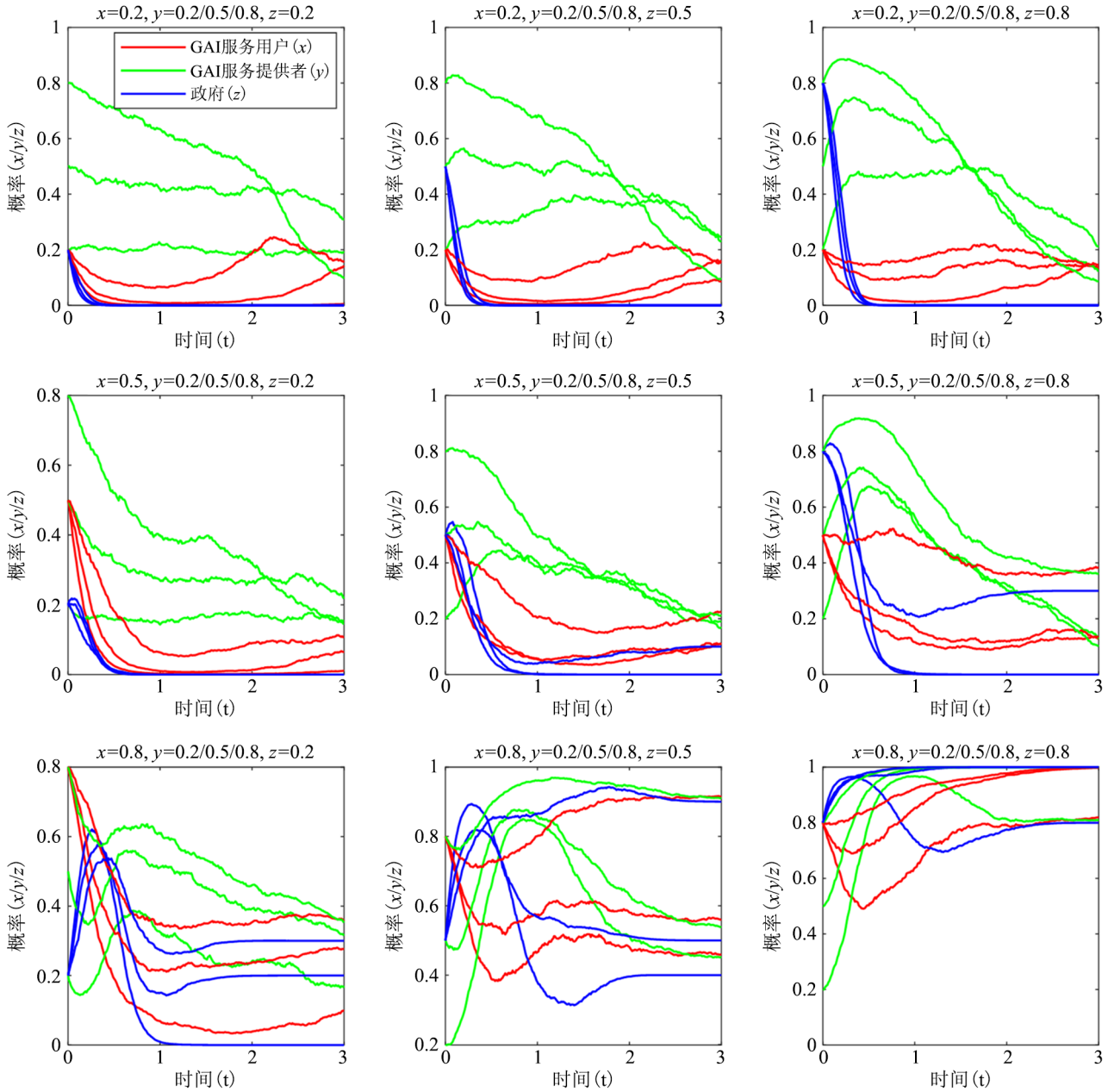
4.1 参与主体初始意愿对系统演化的影响

本文将用户披露、GAI服务用户提供者积极保护、政府积极监管的初始意愿划分为低、中、高三级，记为$x, y, z = \Omega(0.2, 0.5, 0.8)$，控制其他参数取值如表3所示，并设定初始随机扰动强度σ=0.5，以研究不同初始意愿对系统演化趋势的影响。仿真结果如图1所示。

结果表明，只有当初始概率x, y, z = (0.8, 0.2, 0.8)、(0.8, 0.5, 0.8)、(0.8, 0.8, 0.8)和(0.8, 0.8, 0.5)时，三方才均选择(或接近)社会所期望的策略。这表明当用户披露意愿和政府积极监管意愿均较高时，前者对GAI服务提供者提供利益驱动，后者对其泄露隐私行为形成严格约束，无论提供者初始积极保护意愿如何，都会因这双重作用向积极保护策略演化。故为保证各主体收敛于社会期望策略，实现用户隐私保护和行业健康发展，政府需先采取举措确保用户具有较高披露信息初始意愿，再积极承担社会责任、提升自身积极监管意愿，以促使更多提供者选择积极保护策略。

4.2 随机扰动对系统演化的影响

本文将随机扰动强度设置为σ=0.2、0.5、0.8，控制其他参数不变，以研究随机扰动强度对演化过程的影响。为简化分析，使演化结果更直观清晰，后文仅讨论三方初始意愿均相等时的演化情形，即三方初始意愿均同时设置为低、中、高三级。仿真结果如图2所示，不同程度随机扰动下，三方博弈主体演化策略行为呈一定幅度波动。三方初始意愿较高时，最终选择(披露，积极保护，积极监管)策略，但演化速度不同，这是因为博弈主体间的信息不对称，GAI服务提供者可及时整合市场信息，较快适应外部环境变化，政府可据此制定政策指引其采取应对策略，而用户处于信息弱势，策略选择速度受限。当三方初始意愿不高时，随时间增加，GAI服务提供者对随机干扰强度变化最敏感，可能在于静态均衡下其行为选择并非纯策略(仅部分积极保护)，外部冲击下不同提供者会随时调整策略以应对环境冲击；而用户和政府在静态均衡下最终选择纯策略(不披露，消极监管)，说明该纯策略对二者均占优，即使存在不同程度随机干扰，仅会影响演化过程波动，不会改变其最终策略选择。



注：各子图横坐标是演化博弈时间，纵坐标是主体选择策略的概率，红色实线、绿色实线和蓝色实线分别表示用户选择披露的概率(x)、GAI服务提供者选择积极保护的概率(y)和政府选择积极监管的概率(z)随时间的变化。下同。

图 1　参与主体初始意愿对系统演化的影响

Fig. 1　The Impact of Initial Willingness of Participating Subjects on System Evolution

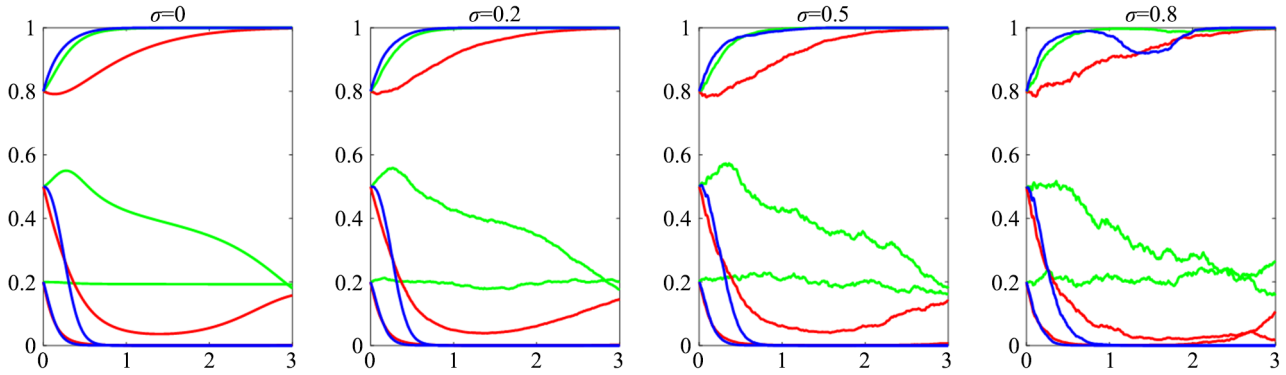


图 2　随机干扰强度对系统演化的影响

Fig. 2　Impact of Random Interference Intensity on System Evolution

4.3 GAI服务提供者类型对系统演化的影响

4.3.1 不同市场份额对系统演化的影响

根据BCG三四规则矩阵[50]，本文以市场份额为标准，将GAI服务提供者划分为三类：优胜者、参与者和生存者，推测其在系统演化中的情况差异：

①市场优胜者：在GAI市场中占据最大市场份额，拥有强大品牌影响力、技术创新和市场控制能力，持续推出创新服务，满足客户需求，肩负社会责任，看重声誉累积，为行业领导者或主导企业(如OpenAI等)。因其技术力和影响力更大，用户使用其服务获得收益或面临损失均较大，且政府对其实施监管和支持力度最高。②市场参与者：在GAI市场中拥有一定市场份额，具有较为成熟的服务体系和一定的市场竞争力，其大多专注于特定领域，较为重视企业声誉，但并非行业领导者，而可能处于市场份额中游或下游(如Perplexity等)。用户通过其服务能获得中等收益且可能面临一定损失，政府将对其实施中等程度监管和支持。③市场生存者：在GAI市场中市场份额微小，面临较大竞争压力，其服务可能缺乏差异性或创新能力。为维持生存，市场生存者需不断降低成本、提高效率或寻找新的市场机会，由于资源有限和市场地位较弱，通常处于行业边缘或尾部，市场声誉巩固需求较小(如You. com等)。其技术力和影响力有限，并非政府监管重点，故政府对其实施监管规制和奖励均较低，而用户使用其服务获得收益或面临损失均较小。

由于三者间技术、服务、资金等竞争优势和声誉巩固需要的差异，三者作为GAI服务提供者参与演化系统过程时，便天然存在初始参数差异，可能对系统演化方向和速率产生不同影响。表3展示三者间初始参数差异情况，并设定δ_0(0.7、0.5、0.3)，δ(0.07、0.05、0.03)，λ(0.07、0.05、0.03)，β(0.7、0.5、0.3)，μ_0(1.5、1、0.5)，μ(0.15、0.1、0.05)分别代表市场优胜者、参与者、生存者情形下的隐私泄露风险模型参数。仿真结果如图3所示，在初始意愿均较高时，市场优胜者和市场参与者情形下，三方会选择(披露，积极保护，积极监管)策略，且前者收敛于稳定策略的速度要高于后者，在中等初始意愿下，在市场优胜者情形三方也会收敛至社会期望策略，这是因为市场优胜者影响力较大，政策规制力度更大，用户更愿意相信市场龙头并积极披露隐私以获取服务。而市场生存者缺乏政策支持和用户信任，导致政府对其监管力度较弱，用户不愿披露隐私，其将始终倾向于消极保护，故为鼓励其发展，政府可肩负市场调控责任，制定针对性支持政策，呼吁用户使用其服务。

表3 市场优胜者/参与者/生存者参数初始值选择

Tab. 3 Initial Value Selection for Market Winners/ Participants/Survivors Parameters

参数	取值	参数	取值	参数	取值
V_1	13/10/7	C_1	25/20/15	L_4	10
V_2	13/10/7	C_2	13/10/7	N_1	18/15/12
T	5	C_3	7/5/3	N_2	13/10/7
D_1	15	P_1	25/20/15	N_3	5
D_2	25	B_1	25/20/15	N_4	10
f	10	F	18/15/12	a	0.8
M_1	13/10/7	L_1	25/20/15	b_0	0.2/0.3/0.4
M_2	15	L_2	10	c_0	0.7
r_1	7/5/3	L_3	5	Δp	0.1

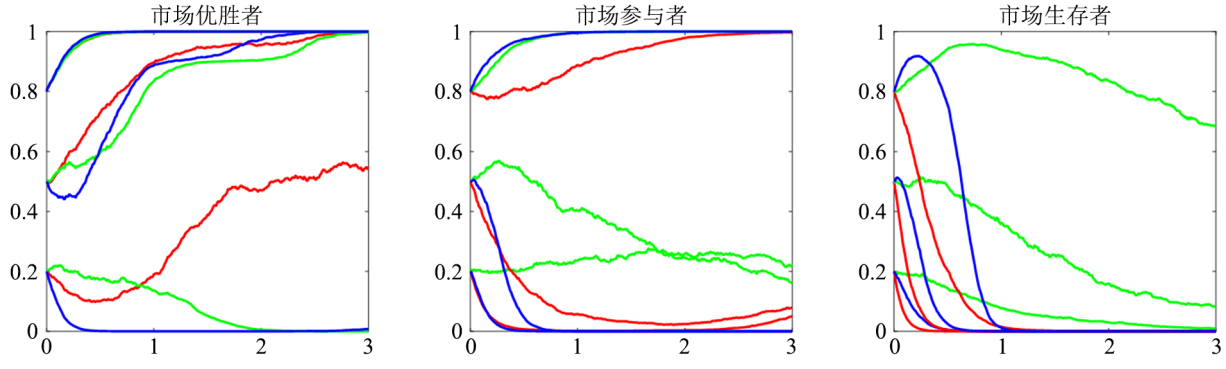


图3 不同市场份额对系统演化的影响

Fig. 3 Impact of Different Market Shares on System Evolution

4.3.2 不同应用场景对系统演化的影响

依照所提供服务应用场景差异，可将GAI服务提供者划分为通用型和垂直型。通用型GAI涉及大规模、周期性数据收集、更新、清洗、预处理和模型训练，故需开发和实施更为先进复杂的隐私保护技术，以确保用户数据安全性和隐私性，隐私保护成本更高。而垂直型GAI服务提供者数据规模相对较小且来源集中，可实施针对性隐私保护技术，其保护成本相对更低。在政府层面，假设通用型和垂直型GAI服务作为GAI服务两类分支均备受政府重视，政府奖惩程度相等。但垂直型GAI服务提供者处理数据往往更敏感，一旦发生数据泄露或被恶意利用，可能引发严重社会问题和安全风险。因此，政府可能针对不同行业制定专门监管政策和法规(如对于医疗、金融等垂直领域，付诸更多监管成本)。可见，通用型和垂直型GAI服务提供者参与到演化系统过程时，亦存在初始参数上的差异，并可能产生系统演化趋势差异，两者间初始参数设置情况如表4所示，并设定δ_0(0.5、0.2)，δ(0.05、0.02)，λ(0.05、0.02)，β(0.5、0.3)，μ_0(1、1.3)，μ(0.1、0.2)分别代表(通用型、垂直型)GAI情形下的隐私泄露风险模型参数。仿真结果如图4所示，通用型GAI服务符合标准博弈系统演化规律，而垂直型GAI服务提供者因其积极保护对用户必要性和更高成本效益，无论初始意愿如何，即使政府监管缺失，其均会实施积极保护策略，用户亦会披露隐私以获取服务。当积极保护初始意愿较低时，用户起初不愿披露隐私，但会随积极保护力度加强转变策略，不断选择披露隐私以获取相关服务。若垂直型GAI服务提供者因随机干扰而降低积极保护意愿，用户披露意愿会小幅下降，此时政府需引导提供者策略以避免社会风险与安全风险。

表4 通用型/垂直型参数初始值选择

Tab. 4 Selection of Initial Values for General/Vertical Parameters

参数	取值	参数	取值	参数	取值
V_1	10	C_1	20/15	L_4	10/7
V_2	10	C_2	10	N_1	15/18
T	5	C_3	5/7	N_2	10
D_1	15/20	P_1	20	N_3	5
D_2	25	B_1	20	N_4	10
f	10	F	15	a	0.8
M_1	10	L_1	20	b_0	0.3/0.2
M_2	15	L_2	10	c_0	0.7
r_1	5	L_3	5	Δp	0.1

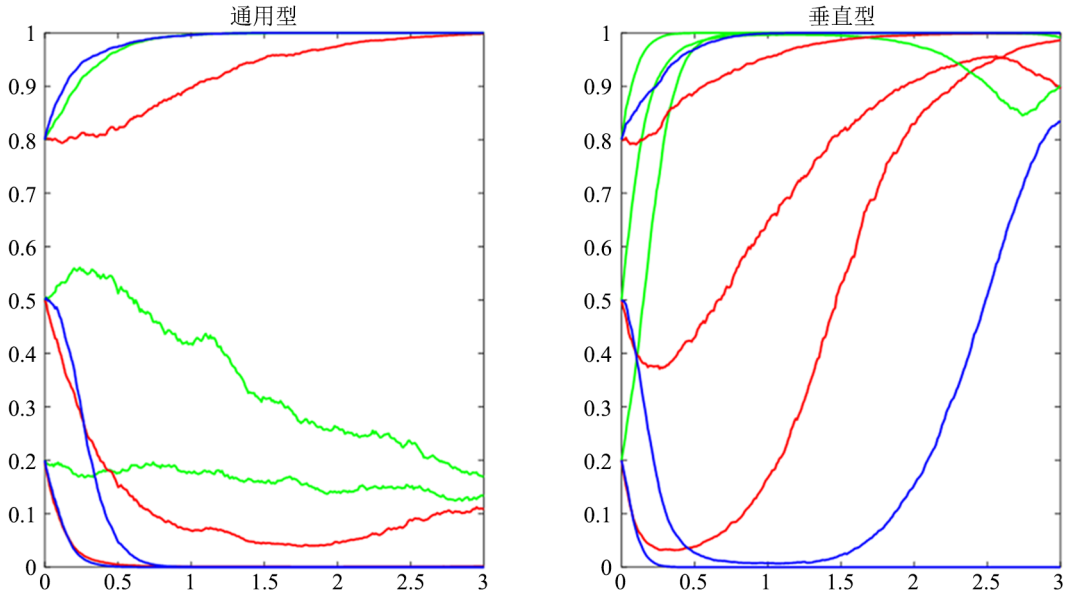


图4 不同应用场景对系统演化的影响

Fig. 4 Impact of Different Application Scenarios on System Evolution

4.4 政府规制因素对系统演化的影响

4.4.1 政府奖励/惩罚对系统演化的影响

控制其他参数不变，分别在$(f,F)=(5,5)$、$(5,15)$、$(15,5)$、$(15,15)$的情况下(对应于用户与提供者均低惩、用户低惩且提供者高惩、用户高惩且提供者低惩、用户与提供者均高惩)进行数值模拟。仿真结果如图5所示，只有$(f,F)=(5,15)$时，三方参与主体稳定策略才有可能达到社会期望策略，否则用户和服务提供者均不会稳定于披露策略和积极保护策略。故政府采取积极监管策略时，对用户处罚力度应相对较低，对GAI服务提供者处罚力度应适当升高。

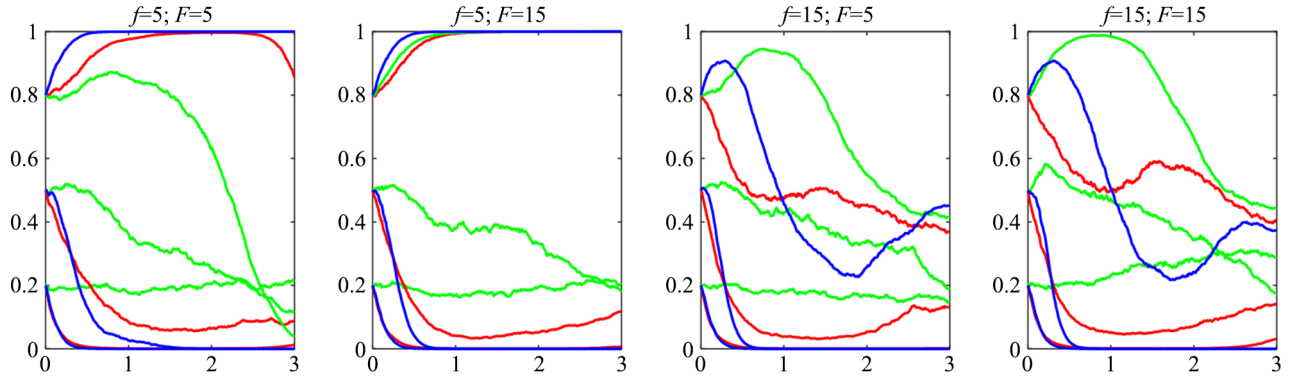


图5 政府惩罚对系统演化的影响

Fig. 5 Impact of Government Penalties on System Evolution

控制其他参数不变，分别在N_2(0、5、10、20)的情况下(对应于政府对GAI服务提供者积极保护不奖励、低奖励、适中奖励、高奖励)进行数值模拟。仿真结果如图6所示，结果表明，只有N_2=10时，三方才可能达到社会期望策略，否则即使政府严格积极监管，用户和提供者也不会达到社会期望，而奖励过高时，虽然一开始会促使提供者积极保护，但政府会因财政压力而出现监管缺失，提供者随后也会不断演化至消极保护。故政府采取积极监管策略时，对提供者的积极保护行为应给予适中奖励，不宜过高或过低。

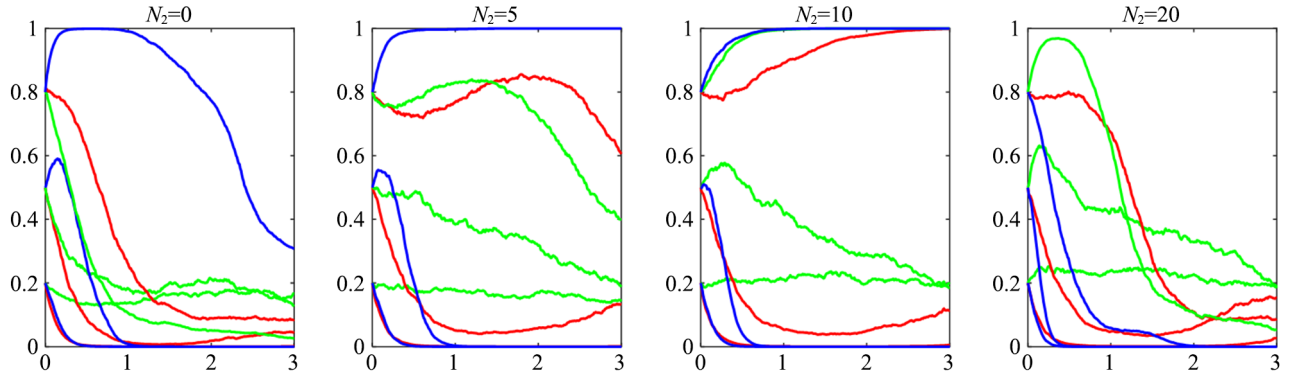


图6 政府奖励对系统演化的影响

Fig. 6 Impact of Government Rewards on System Evolution

4.4.2 维权赔偿金与成本对系统演化的影响

控制其他参数不变，本文分别在V_2(0、5、10、20)的情况下进行数值模拟，其分别代表(无赔偿金、低赔偿金、适中赔偿金、高额赔偿金)。仿真结果如图7所示，结果表明，只有V_2 = 10和20时，三方才有可能达到社会期望策略，且维权赔偿金越高，系统演化速度越快，且可进一步促使中等初始意愿的三方达到社会期望策略。因此，政府在设定维权赔偿机制时，应当根据实际情况，尽可能高地对出现隐私泄露的GAI服务提供者收取维权赔偿金，并返还给维权用户，以保障用户权益。

控制其他参数不变，分别在D_2(0、5、15、25)

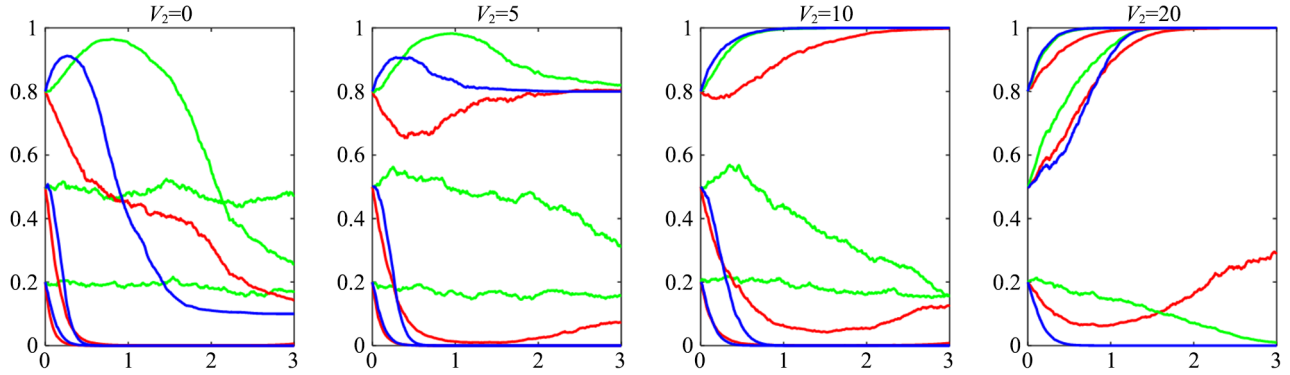


图 7 维权赔偿金对系统演化的影响

Fig. 7 Impact of Compensation for Rights Protection on System Evolution

的情况下进行数值模拟，其分别代表(无维权成本、低维权成本、适中维权成本)、高维权成本。仿真结果如图 8 所示，上述不同成本下，三方均有可能达到社会所期望的策略，维权成本越低，其可能性越大。当维权成本较低或无维权成本时，不同初始意愿的三方均能达到社会的期望策略。故政府在设定维权成本机制时，为降低维权难度以促使用户披露信息，防止 GAI 服务提供者泄露隐私，应采取相应措施让维权行为更加高效便捷，尽可能降低用户维权成本。

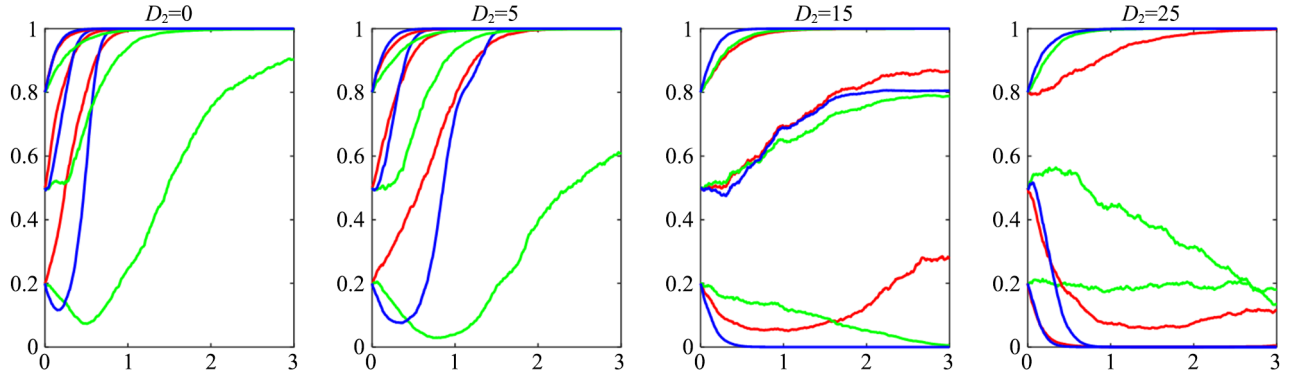


图 8 维权成本对系统演化的影响

Fig. 8 Impact of Rights Protection Costs on System Evolution

4.5 参数的敏感性分析

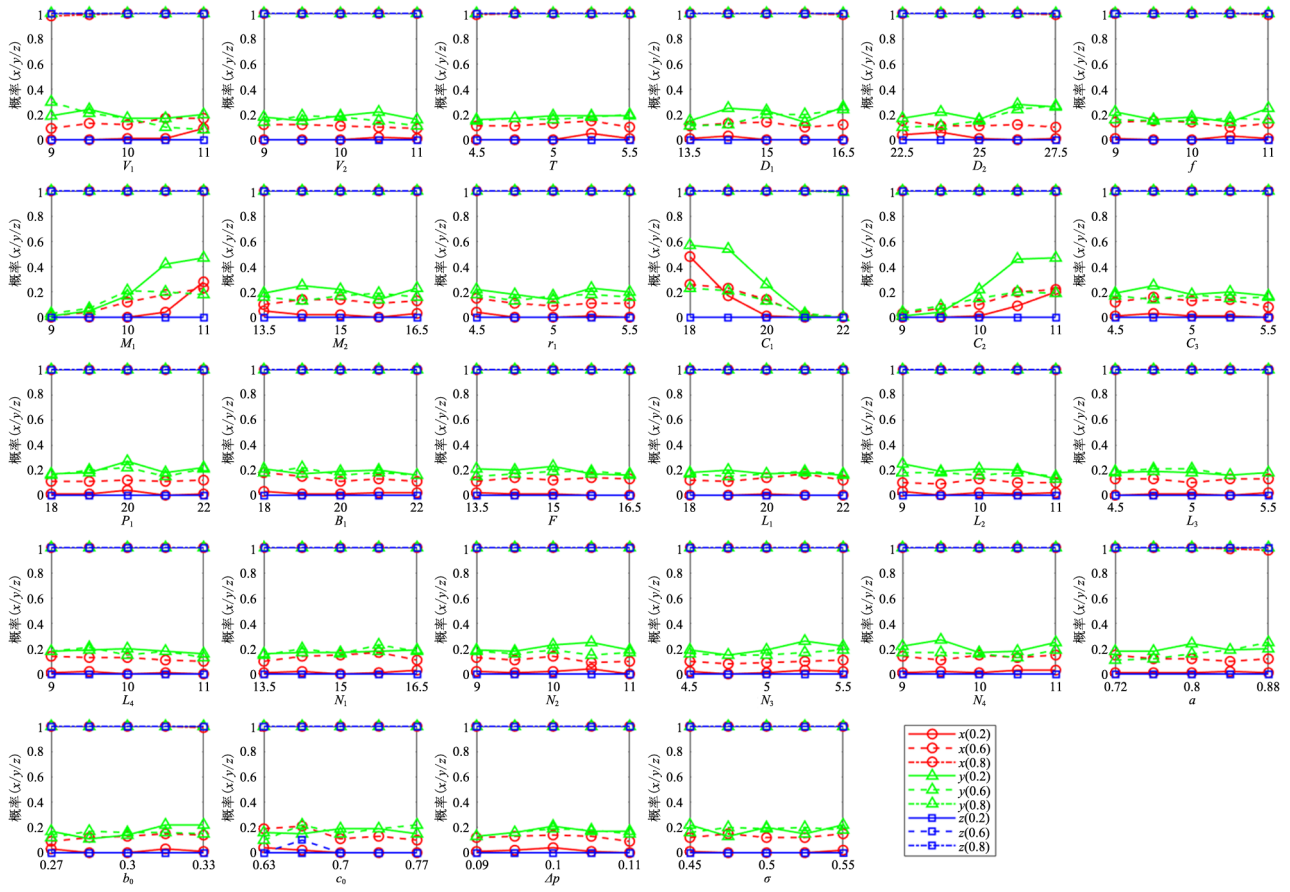
前述分析均基于表 3 初始参数设定以及 σ=0.5 前提下，对其他参数影响进行仿真探讨。但 GAI 服务行业复杂多变，行业内部和外部均存在影响三方相关参数波动的因素，此处分析表 3 初始参数及随机干扰强度系数 σ 共 28 个参数分别变动−10%、−5%、+5% 和+10% 时，博弈系统的演化结果如何变化，详见图 9。由于政府的演化策略较为稳定，同等条件下，单一参数的波动对其演化结果的影响不大，且初始意愿均较高时，三方演化结果受参数波动的影响有限，此处不展示。结果表明，积极保护声誉收益的增加、积极保护成本的减少及消极保护成本的增加均促使 GAI 服务提供者和用户在系统演化末期(t = 3)分别采取积极保护和披露信息策略，且两者对积极保护成本参数的变动均最为敏感，故在现有初始参数情境下，首要任务是政府和 GAI 服务提供者实现政企合作，推动技术研发，促使用户隐私积极保护成本下降，进而推动行业规范发展。

5 讨 论

5.1 结论

本文基于演化博弈理论及随机过程，构建包含政府、GAI 服务提供者和用户开展 GAI 服务用户隐私保护的随机演化博弈模型，研究发现：

1）在标准随机演化博弈系统中，三方初始意愿影响后续系统演化方向和速度，且仅在用户和



注：图中横坐标是相关参数的取值在−10%~10%的变动范围，纵坐标是三方在演化末期(t = 3)分别选择披露、积极保护和积极监管的概率，x(0.2)、x(0.5)和x(0.8)分别代表三方低、中、高初始意愿时用户的披露概率(x)，y(0.2)、y(0.5)和y(0.8)分别代表三方低、中、高初始意愿时提供者积极保护概率(y)，z(0.2)、z(0.5)和z(0.8)分别代表三方低、中、高初始意愿时政府积极监管概率(z)。

图 9　参数的全局敏感性分析

Fig. 9　Global Sensitivity Analysis of Parameters

政府均持有较高初始意愿时，GAI服务提供者才可能选择积极保护策略，并使系统达到理想均衡点。其核心逻辑在于GAI强数据驱动性需要海量多模态隐私数据支撑模型迭代与功能实现[37]，且强交互性催生过程性隐私损耗[36]，若用户初始意愿低、不愿披露隐私，提供者会因数据缺口过度投入数据获取而轻隐私保护，或因数据不足削弱商业价值、压缩保护成本空间；同时，GAI模型迭代易导致隐私泄露概率非线性增长[8]，应用场景广泛使市场保护标准容易混乱，且现有法律适配滞后[18]，若政府初始意愿低、监管与引导缺位，提供者会因缺乏外部约束节省保护成本，或因积极保护成本高昂难以竞争。故GAI服务中三方利益耦合性远强于传统场景，单一主体高意愿无法破解数据需求与隐私保护、监管与成本的矛盾，唯有用户积极披露隐私和政府实施积极保护策略的协同状况才是提供者实施积极保护策略并实现理想效果的前提。

在影响因素方面，用户披露个人隐私后获得的个性化推送与内容创作服务、政府信任收益、泄露损失、维权赔偿与成本等是影响用户策略选择的重要因素。提供者积极保护隐私的声誉收益、保护成本、政府罚金与奖励等是影响提供者策略选择的主要因素；政府监管获得的声誉收益、投入成本、奖励成本，以及监管不力导致的声誉损失等是影响政府策略选择的重要因素。GAI服务提供者模型迭代次数、隐私保护技术优化、用户交互强度、数据规模与多模态特征、政府监管力度等是影响用户隐私泄露概率的重要因素，其存在动态演化特征，由此使得GAI服务用户隐私保护的影响因素显著区别于先前对传统互联网平台服务静态风险的研究结论[26-30]。在积极保护策略下，提供者基础隐私保护技术与技术优化措施所产生的风险

抑制效应要高于模型迭代次数、用户交互强度、数据规模提升所带来的风险驱动效应，可有效实现GAI行业中隐私泄露风险的动态管控。

2）随机扰动强度越大，演化博弈系统演化过程的波动性和不确定性越强，且相较于用户和政府，GAI服务提供者对不确定因素更为敏感。这是由于提供者需处理大量的、多模态用户数据[37]，并通过复杂、多次迭代的算法模型[8]进行隐私保护和内容生成，实时响应用户需求，这会使得其在随机扰动下更易出现数据处理波动和隐私泄露风险，且其提供的GAI模型本身的技术复杂性[13]与黑箱特性[32]亦会放大扰动影响。此外，提供者需在数据收集、存储、处理、生成多环节落实隐私保护[11]，机制内生脆弱性使随机扰动进一步增加环节的复杂性与不确定性。而用户的博弈决策多聚焦数据安全、过程简单，且通常缺乏专业数据处理与隐私知识，故对扰动敏感性低；政府则通过政策法规与监管进行宏观调控，决策过程相对稳定，可依托长期政策调整应对扰动。

3）GAI服务市场优胜者和市场参与者均能在三方高初始意愿条件下收敛至理想状态的稳定策略，这符合标准系统演化规律。而任何情形下，市场生存者都难以演化至理想状态；相较于通用型GAI服务提供者，垂直型GAI服务提供者策略具有更快的正向收敛速度。不同于既往研究大多未区分服务提供商/平台类型开展分析，本文依据市场份额与应用场景对GAI服务提供商进行划分后，发现不同类型提供者的演化博弈过程及结果存在显著差异。其原因是：GAI属技术密集型领域，市场生存者的资金、技术力和服务水平有限，用户不会优先选择对其实施积极隐私披露行为。且由于市场生存者众多且分散，影响力有限，政府不会耗费大量监管资源规制其隐私保护行为，导致市场生存者成为监管盲区。对比通用型GAI服务提供者，垂直型GAI服务提供者深耕特定领域形成的用户群体同质性高、需求聚焦，能精准定制隐私保护机制，降低隐私保护成本；长期积累的专业数据处理经验与针对性算法优化技术，使其能高效平衡数据价值挖掘与隐私风险控制，缩短博弈策略演化周期；其依托专业壁垒构建可信技术标签，易使用户形成路径依赖，主动配合数据共享与策略迭代，形成良性互动闭环；垂直领域明确具体的行业监管框架与合规标准，能驱动其以更高安全阈值约束行为，减少隐私泄露的随机扰动，加速策略收敛至稳定均衡状态。

4）随着政府对GAI服务提供者奖励和罚金增加，提供者积极保护隐私概率均上升，其泄露用户隐私概率下降，但若奖励成本过高，政府可能因财政压力而不愿积极监管，从而无法规制提供者稳定于积极保护策略。结合GAI隐私保护的技术黑箱[32]、模型迭代中泄露概率非线性增长[8]、强交互催生过程性隐私损耗[36]、海量多模态数据扩大攻击面[37]及法律适配滞后[18]等挑战，政府对GAI服务提供者的奖惩设计需适配其特性：增加奖励可缓解其应对数据记忆、模型反转攻击[33]等的保护成本压力，提高罚金能放大其因微调数据泄漏等风险的损失，二者共同推动其提升积极保护概率，但奖励过高会因GAI监管复杂度高加剧政府财政压力，难使易形成“数据孤岛”的提供者稳定于积极保护。而政府增加GAI服务用户维权赔偿，可激励用户积极披露隐私，因GAI技术黑箱[32]加剧信息不对称，高不实举报罚金易致用户披露信息后维权被误判而抑制披露意愿，故需对用户实行低罚金，同时提升维权赔偿可倒逼提供者规避高额损失而积极保护；降低维权成本能解决GAI举证难问题，激励用户披露与维权，此结论与曲薪池等[26]一致，且政府需同步强化监管应对用户维权意愿减弱风险。奖惩方面，政府对提供者应实施“高额罚金+适中奖励”设计，对用户需实施“低罚金+高赔偿+低维权成本”设计，以推动三方演化至理想策略。

5）政府的演化策略在各参数波动10%范围内表现得均较为稳定，而积极保护声誉收益的边际上升、积极保护成本的边际降低以及消极保护成本的边际上升，均能促使GAI服务提供者与用户采纳积极保护与披露策略，且GAI服务提供者与用户均

对积极保护成本参数的变动表现出最高敏感性。这可能在于政府在GAI隐私治理中的核心角色是规则制定者与监管者，其目的在于维护行业整体合规与社会公共利益，故往往采用鲁棒性监管工具，这使得其策略对参数短期波动具备天然抗干扰性；GAI服务提供者因逐利性，将涉及技术研发[13]等刚性支出的成本作为隐私保护首要考量，受动态风险放大成本、声誉影响，积极保护声誉收益上升、成本下降及消极保护成本上升均推动其转向积极策略，且声誉收益亦能助力打破“数据孤岛”。用户则因GAI技术黑箱[32]和过程性损耗[36]等风险对隐私泄露较为敏感，其披露意愿随提供者保护策略动态调整。据此，当前应重点构建政企协同机制，推进隐私保护技术研发与优化，降低积极保护成本，为GAI隐私治理规范化与可持续发展提供支撑。

5.2 建议

本文结合研究结论，针对各主体提出以下建议：

政府需构建动态适配且注重分类规制的GAI隐私治理体系：考虑到GAI隐私保护成本高，且泄露后损失涉及用户权益与行业信任，同时过高奖励会加重政府财政压力、削弱监管动力，因此对提供者的罚金应覆盖全链条损失，应大幅提升罚金，强化罚金执行力，确保违法成本高于合规成本，避免“罚而不痛”，奖励采用分阶段模式，初期给予一定比例研发补贴来激励技术投入，后期转为声誉认证与政策倾斜；鉴于不同类型GAI提供者差异显著(如市场优胜者和参与者均实现隐私保护正向收敛、市场生存者因资金技术不足常处于监管盲区、垂直型依托专业优势可快速实现隐私保护正向收敛，通用型因场景泛化易出现多模态数据泄露漏洞)，需推动市场优胜者牵头制定行业隐私保护标准，以标杆作用打破“数据孤岛”；对市场参与者强化全流程动态监测，重点管控多模态数据交叉分析及强交互场景隐私损耗，推动其对标优胜者；为市场生存者搭建技术共享与补贴机制，开放低成本保护工具，消除监管盲区；为垂直型提供者提供定制化合规指导，支持联合协会制定领域规范，借收敛快的优势打造示范案例；为通用型提供者建立风险预警机制，监测模型迭代漏洞、实行跨场景数据共享审批备案。又因GAI技术黑箱导致用户维权时举证困难，且过高不实举报罚金会抑制用户维权意愿，故应适度减轻不实举报罚金，同时建立数据流向追溯工具、明确提供者举证倒置责任，实行“基础+浮动”赔偿与政府先行赔付，且实时调整监管策略以适配市场变化。

GAI服务提供者应打造贴合技术特性且精准匹配场景的隐私保护能力：结合自身业务特性，垂直型提供者可依托用户需求聚焦、数据专业性强优势，针对GAI强交互易引发过程性隐私损耗问题，采用多轮交互渐进式采集数据方式，减少碎片信息过度暴露；通用型提供者面对多模态数据规模大、模型迭代易累积漏洞现状，需在每轮迭代前开展隐私泄露概率预评估，通过分类脱敏与权限分级存储数据，降低攻击面；市场优胜者应依托技术资金优势构建全链路体系，数据收集遵循“最小必要原则”，迭代引入联邦学习等隐私计算技术实现“数据可用不可见”，并开放工具包带动行业协同；市场参与者需聚焦短板补齐，通过泄露概率预评估模型、数据分类脱敏与权限分级存储，解决模型迭代漏洞累积等问题，定期对标优胜者自查整改；市场生存者为解决数据不足与保护成本高等难题，可加入行业隐私保护联盟等类似机构，共享脱敏非敏感数据、使用联盟低成本保护工具，在控制成本的前提下满足基础合规要求。同时，针对GAI技术黑箱阻碍用户知情同意且用户对保护策略认知不足等问题，GAI服务提供者应优先研发可解释性隐私保护技术，以动态弹窗、短视频演示等直观方式告知用户数据用途与保护措施，建立隐私反馈闭环，实时推送交互中的风险提示并根据用户建议优化方案，提升用户信任度。

用户需强化聚焦风险认知且注重理性维权的隐私保护意识：鉴于GAI存在过程性隐私损耗、模型迭代导致泄露概率非线性增长等特殊风险，且用户常因技术盲区难以识别风险，应主动学习GAI隐私保护专项知识，重点关注多轮交互中的信息暴露风险与提供者技术迭代带来的防护变化，借

助政府或行业协会推出的风险评估工具，在披露隐私前审慎评估泄露可能性；在政府出台维权政策基础上，用户可通过官方渠道了解多模态数据泄露举证要点，利用免费法律咨询服务降低维权难度，密切关注政策动态调整，在隐私泄露时，依据相关赔偿标准积极维权，既保障自身权益，亦为三方协同实现GAI隐私治理理想状态提供支撑。

6 结 语

本文构建具有GAI服务场景特殊性的三方随机演化博弈模型，拓宽了随机演化博弈理论的研究情景，为GAI服务用户隐私保护研究引入新颖研究视角，构造时变隐私泄露风险函数，优化既有的用户隐私演化博弈模型，加深对GAI服务用户隐私泄露动态机制的理解，丰富GAI隐私保护领域的理论与方法，为推动隐私保护实践提供理论依据和对策启示。未来研究仍有一些可改进之处：①本文主要讨论政府、GAI服务提供者与用户三方的用户隐私保护策略，后续可纳入更多主体进行分析(如GAI利益相关者之间的跨平台数据交换引发的隐私和敏感数据泄露等问题)。②本文研究假设与模型构建虽尽可能模拟还原现实情景，但面临理想化博弈情景的固有局限，未能考虑GAI服务用户边际效益的差异性、政府奖惩机制在现实监管情景中的有效性等现实问题，后续可围绕此类问题做进一步探索。③演化博弈理论推演和仿真实验尚缺少实证数据支撑，后续可尝试利用实证分析方法，通过问卷调查或行业数据等真实数据以对模型仿真结果进行佐证。

参 考 文 献

[1] 中华人民共和国中央人民政府. 生成式人工智能服务管理暂行办法 [EB/OL]. [2025-09-11]. https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm.

[2] 中国互联网络信息中心. 第55次《中国互联网络发展状况统计报告》[R/OL]. [2025-09-11]. https://www3.cnnic.cn/NMediaFile/2025/0428/MAIN17458061595875K4FP1NEUO.pdf.

[3] 雷丽彩，郭芷欣. 基于政府监管的社交媒体用户隐私保护演化博弈分析 [J]. 数字图书馆论坛，2024，20 (8)：39-50.

[4] 邱均平，徐中阳，陈锡慧. 基于三方演化博弈的在线健康社区用户隐私披露行为研究 [J]. 情报理论与实践，2023，46 (1)：24-36.

[5] 胡小飞，曾聪，郭宇雯. 快递物流业个人信息隐私保护的演化博弈分析 [J]. 现代情报，2019，39 (6)：142-148.

[6] Traulsen A，Glynatsi N E. The Future of Theoretical Evolutionary Game Theory [J]. Philosophical Transactions of the Royal Society of London Series B，Biological Sciences，2023，378 (1876)：20210508.

[7] 张欣. 生成式人工智能的数据风险与治理路径 [J]. 法律科学(西北政法大学学报)，2023，41 (5)：42-54.

[8] 张小燕. 生成式人工智能数据研究综述：风险、挑战与治理 [J]. 图书情报工作，2025，69 (9)：136-148.

[9] 杨强. AI与数据隐私保护：联邦学习的破解之道 [J]. 信息安全研究，2019，5 (11)：961-965.

[10] 苗运卫. 生成式AI赋能图书馆中的读者信息分类保护 [J]. 图书馆论坛，2024，44 (8)：34-43.

[11] Ye X B，Yan Y H，Li J，et al. Privacy and Personal Data Risk Governance for Generative Artificial Intelligence：A Chinese Perspective [J]. Telecommunications Policy，2024，48 (10)：102851.

[12] Golda A，Mekonen K，Pandey A，et al. Privacy and Security Concerns in Generative AI：A Comprehensive Survey [J]. IEEE Access，2024，12：48126-48144.

[13] Feretzakis G，Papaspyridis K，Gkoulalas-Divanis A，et al. Privacy-Preserving Techniques in Generative AI and Large Language Models：A Narrative Review [J]. Information，2024，15 (11)：697.

[14] 肖红军，张丽丽. 中国企业数字科技伦理发展：演变历程、最新进展与未来进路 [J]. 产业经济评论，2024 (2)：153-171.

[15] 严炜炜，王妍妍，宋佳慧. 用户数据采集对人工智能服务放弃使用意愿的影响研究——基于信息敏感度和隐私视角 [J]. 情报科学，2024，42 (12)：20-30，41.

[16] Dahabiyeh L，Taha N，Thneibat M，et al. Privacy Awareness in Generative AI：The Case of ChatGPT [J]. Interactive Technology and Smart Education，2026，23 (1)：25-48.

[17] Lee C，Kim J，Lim J S，et al. Generative AI Risks and Resilience：How Users Adapt to Hallucination and Privacy Challenges [J]. Telematics and Informatics Reports，2025，19：100221.

[18] Murdoch B. Privacy and Artificial Intelligence：Challenges for Protecting Health Information in a New Era [J]. BMC Medical Ethics，2021，22 (1)：1-5.

[19] 蒋徐鑫. 人工智能模型中数据泄露的法律风险防范 [C]//《新兴权利》集刊2023年第2卷——生成式人工智能法律问题研究文集. 西华大学，2024：69-84.

[20] 肖红军，阳镇. 数字科技伦理监管：美国进展与中国借鉴 [J]. 财经问题研究，2023 (6)：73-86.

[21] 张涛. 生成式人工智能训练数据集的法律风险与包容审慎规制 [J]. 比较法研究，2024 (4)：86-103.

[22] Quach S，Thaichon P，Martin K D，et al. Digital Technologies：Tensions in Privacy and Data [J]. Journal of the Academy of Marketing Science，2022，50 (6)：1299-1323.

[23] 杨少杰，张辉，彭英杰，等. 基于博弈论的个性化LBS用户位置隐私保护方案 [J]. 计算机应用与软件，2023，40 (5)：312-318，349.

[24] 冯景瑜，杨锦雯，张瑞通，等. 抗位置隐私泄露的物联网频谱共享激励机制 [J]. 计算机研究与发展，2020，57 (10)：2209-2220.

[25] 黄凯南. 演化博弈与演化经济学 [J]. 经济研究，2009，44 (2)：132-145.

[26] 曲薪池，侯贵生. 基于三方演化博弈的平台信息安全治理研究 [J]. 现代情报，2020，40 (7)：114-125.

[27] 丰米宁，朱光，杨嘉韵. 基于演化博弈的社交网络隐私保护研究 [J]. 情报杂志，2017，36 (9)：127-132，85.

[28] 朱光，曹雪莲，孙玥. 社交网络环境下隐私保护投入的博弈策略分析——基于演化博弈的视角 [J]. 情报科学，2017，35 (7)：25-30.

[29] 朱光，刘虎，陈婧，等. 移动医疗服务情境下的隐私问责研究——基于演化博弈的视角 [J]. 现代情报，2018，38 (12)：32-39.

[30] 徐璐，钟涛，潘禹辰，等. 大数据时代招聘平台用户数据隐私保护研究 [J]. 情报理论与实践，2022，45 (12)：68-75.

[31] 郭海玲，卫金金，刘仲山. 生成式人工智能虚假信息协同共治研究 [J]. 情报杂志，2024，43 (9)：121-129，165.

[32] 黄锴. 生成式AI对个人信息保护的挑战与风险规制 [J]. 现代法学，2024，46 (4)：101-115.

[33] Yang Y Q，Zhang B N，Guo D X，et al. Generative AI for Secure and Privacy-Preserving Mobile Crowdsensing [J]. IEEE Wireless Communications，2024，31 (6)：29-38.

[34] Rajaobelina L，Prom Tep S，Arcand M，et al. Creepiness：Its Antecedents and Impact on Loyalty When Interacting With a Chatbot [J]. Psychology & Marketing，2021，38 (12)：2339-2356.

[35] Cilloni T，Fleming C，Walter C. Privacy Threats in Stable Diffusion Models [EB/OL]. [2025-09-11]. https://doi. org/10.48550/arXiv. 2311. 09355.

[36] Staab R，Vero M，Balunović M，et al. Beyond Memorization：Violating Privacy via Inference With Large Language Models [EB/OL]. [2025-09-11]. https://doi. org/10. 48550/arXiv. 2310. 07298.

[37] 李亚玲，蔡京京，柏洁明. 生成式大模型引发的隐私风险及治理路径 [J]. 智能科学与技术学报，2024，6 (3)：394-401.

[38] Doshi A R，Hauser O P. Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content [J]. Science Advances，2024，10 (28)：eadn5290.

[39] Noy S，Zhang W. Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence [J]. Science，2023，381 (6654)：187-192.

[40] Fan L，Deng K Y，Liu F X. Educational Impacts of Generative Artificial Intelligence on Learning and Performance of Engineering Students in China [J]. Scientific Reports，2025，15：26521.

[41] 中国互联网络信息中心. 生成式人工智能应用发展报告(2024) [R/OL]. [2025-09-11]. https://www. cnnic. cn/n4/2024/1216/c88-11196. html.

[42] 中华人民共和国国家互联网信息办公室. 中央网信办深入开展“清朗·整治AI技术滥用”专项行动第一阶段工作 [EB/OL]. [2025-09-11]. https://www. cac. gov. cn/2025-06/20/c_1752129980667315. htm.

[43] 和军，杨慧. ChatGPT类生成式人工智能监管的国际比较与借鉴 [J]. 湖南科技大学学报 (社会科学版)，2023，26 (6)：119-128.

[44] ROME(AP). Italy’s Privacy Watchdog Fines OpenAI for ChatGPT’s Violations in Collecting Users Personal Data [EB/OL]. [2025-09-11]. https://apnews. com/article/italy-privacy-authority-openai-chatgpt-fine-6760575ae7a29a1dd22cc666f49e605f.

[45] 邱遥堃. 举报治网的理论反思与制度改进 [J]. 人大法律评论，2022 (2)：123-141.

[46] Official Journal of the European Union. European Union Artificial Intelligence Act [EB/OL]. [2025-09-11]. https://artificialintelligenceact. eu/the-act/.

[47] Cressman R. The Stability Concept of Evolutionary Game Theory：A Dynamic Approach [M]. Berlin：Springer Science & Business Media，2013：14-17.

[48] 南方都市报. 南都数字经济治理研究中心发布《生成式AI用户风险感知和信息披露透明度测评报告(2024)》[EB/OL]. [2025-09-11]. https://epaper. oeeee. com/epaper/A/html/2024-12/20/content_21496. htm.

[49] 全国网络安全标准化技术委员会. 生成式人工智能服务安全基本要求 [EB/OL]. [2025-09-11]. https://www. tc260. org. cn/upload/2024-03-01/1709282398070082466. pdf.

[50] Reeves M，Deimler M，Stalk G，et al. The Rule of Three and Four：A BCG Classic Revisited [M]. Hoboken：John Wiley & Sons，Inc. 2013.

（责任编辑：李汇森）