

刑法视角下网络型诽谤罪的认定困境与司法完善研究

张祎凝

青岛科技大学法学院, 山东 青岛

收稿日期: 2026年4月23日; 录用日期: 2026年5月22日; 发布日期: 2026年6月8日

摘要

随着信息技术的深度发展, 网络空间已成为现代社会运行的基础性架构。虚拟空间中的诽谤行为因具备传播速率指数级增长、影响范围跨地域延伸、危害后果不可控等特征。虽然我国刑法及其相关司法解释, 在认定上初步形成了该罪的规制框架。但仍存在认定上困境, 即“诽谤事实”的界定不明确、“情节严重”的认定标准不统一、网络型诽谤罪与传统型诽谤罪区分标准模糊。笔者基于上述认定困境, 以刑法视角下, 探讨网络型诽谤罪的认定要件, 并根据刑法规范目的、法益保护属性, 厘清诽谤事实边界, 打破唯数量标准、综合认定“情节严重”标准, 明确两类诽谤罪的界限; 通过细化法律适用、规范证据认定、自诉公诉衔接、明确多元主体责任、加强司法指引等途径完善司法规则, 平衡言论自由与名誉权保护, 提高网络空间治理法治化水平。

关键词

刑法, 网络型诽谤罪, 诽谤事实, 情节严重

Research on the Identification Dilemmas and Judicial Perfection of Cyber Defamation Crime from the Perspective of Criminal Law

Yining Zhang

Law School of Qingdao University of Science and Technology, Qingdao Shandong

Received: April 23, 2026; accepted: May 22, 2026; published: June 8, 2026

Abstract

With the in-depth development of information technology, cyberspace has become the basic

framework for the operation of modern society. Defamatory acts in the virtual space are characterized by an exponential growth in communication speed, cross-regional extension of influence scope, and uncontrollable harmful consequences. Although China's Criminal Law and relevant judicial interpretations have initially established a regulatory framework for the conviction of this crime, prominent identification dilemmas still remain. Specifically, the definition of "defamatory facts" is ambiguous, the criteria for determining "serious circumstances" are inconsistent, and the distinguishing standards between cyber defamation crime and traditional defamation crime are vague. From the perspective of criminal law, this paper analyzes the constitutive elements of cyber defamation crime. In accordance with the normative purpose of criminal law and the attributes of legal interest protection, it clarifies the boundary of defamatory facts, abandons the sole quantitative standard, and adopts a comprehensive evaluation system for "serious circumstances", so as to distinguish the boundaries between the two types of defamation crimes. By refining the application of law, standardizing evidence determination, optimizing the connection between private prosecution and public prosecution, clarifying the liabilities of multiple subjects and strengthening judicial guidance, this paper aims to improve judicial rules, strike a balance between freedom of speech and the protection of reputation rights, and promote the rule of law in cyberspace governance.

Keywords

Criminal Law, Cyber Defamation Crime, Defamatory Facts, Serious Circumstances

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题的提出

随着移动互联网、自媒体、算法推荐等基础技术的推广，网络成为了信息传播和社会交往的主要工具。网络型诽谤不再局限于物理空间口头传播或者书面传播，而是以文字、图片、音频、视频、AI合成内容等形式，通过社交网络、短视频、论坛、微信等快速扩散，发挥跨地区、持续性的侵害效果，严重贬损他人名誉、造成网络暴力、舆情失控等危害，甚至危害社会秩序[1]。我国《刑法》对于诽谤行为的规定主要集中于第二百四十六条诽谤罪，其具体罪状表述为“捏造事实诽谤他人，情节严重的”[2]。2013年《关于办理利用信息网络实施诽谤等刑事案件适用法律若干问题的解释》(以下简称《网络诽谤解释》)以司法解释的形式规定“利用信息网络捏造事实诽谤他人，情节严重的”构成刑法第二百四十六条第一款规定的诽谤罪，更进一步明确了网络型诽谤罪中“情节严重”的认定标准与网络型诽谤罪的犯罪构成要件，初步建立了网络型诽谤罪的认定规则。

近年来，网络型诽谤案件数量众多，多存在裁判尺度不一、认定逻辑不明确。特别是将“点击、浏览、转发”诽谤信息达到一定数量的情形认定为“情节严重”，引发了学界对其合宪性问题的争议[3]。李岚林(2025)分析了我国现行刑事实体规则适用存在法律关系主体的认定模糊，单位主体规制缺位，公众人物界定及其容忍义务规范缺失等问题[4]。加之网络诽谤案件有证据易隐匿、主体难隐匿、传播链条长等特点，自诉人举证困难、公诉困难，司法认定困难、追诉困难等困境[5]。因而，笔者将刑法规范目的和法益保护属性结合网络型诽谤罪认定困境梳理，并找寻其合适的司法完善路径，提高网络空间治理法治化水平。

2. 网络型诽谤罪的认定困境

(一) 网络型诽谤罪“诽谤事实”的界定不明确

诽谤事实的认定是判断网络型诽谤罪是否成立的重要标准，也是区分罪与非罪、此罪与彼罪的重要

依据。吴波(2021)指出,应注意将“捏造事实”的“事实”限定为足以败坏他人名誉、具有具体内容且令人可信的事实[6]。此观点为司法认定指明了方向,但网络语境仍不能完全涵盖所有情形。网络言论的表达碎片化、情绪化、隐喻化、无法言之有物,有很多内容并非是言之有物,而是根据真实事件夸大、裁剪、拼接、解释、或者是以讽刺、猜测的方式表述出来的,对于这类内容是否构成“诽谤事实”并无明确的标准。网络言论中的事实陈述属于客观事实,价值判断属于主观观点,前者可能会导致诽谤,而后者受到言论自由的保护。而网络表达往往将这两者混淆。对他人行为进行负面评价时,包含部分真实细节、主观臆断等,司法机关难以清晰分离事实要素和观点要素,容易把合理批评视为诽谤事实,扩张公民的言论空间。

网络型诽谤是明确指向具体事实,但通过语境暗示指向特定主体以损害其名誉,是否可以纳入“诽谤事实”,法律体系并没有作出回应。同时,网络信息实时更新与二次创作特点使原始事实和衍生虚假信息不断叠加,诽谤事实源头、内容、传播形式难以固定。此外,从场域理论的角度来看,网络虚拟场域与现实场域的叠加导致网络诽谤事实的侵害对象、影响范围与社会评价有所不同。原则上传统诽谤事实是指现实世界中的个人,网络诽谤事实则指向公众、公共机构、群体以至逝者等,不同主体的名誉权保护尺度不同,而现有“诽谤事实”制定却没有考虑场域的不同,所以对于公共事件监督、公众人物批评等内容予以认定过严,违背了网络场域言论的表达规律。

(二) 网络型诽谤罪“情节严重”的认定标准不统一

《网络诽谤解释》第二条规定了四项认定情形,其中以点击浏览五千次以上、转发五百次以上的数量标准成为司法实践最主要的适用依据,这也是“情节严重”的数量化认定标准。但这个标准并没有区分公开平台与私密平台、大众传播与小众传播。在微博、抖音等公开平台,信息可以分裂性传播,点击量、转发量往往会超过标准,但是并不会造成真正的名誉损失。微信朋友圈、微信群等半私密平台,信息只针对特定人群传播,无法达到量化标准,但是可能造成被害人的社会评价急剧下降、精神受到严重摧残。这种“唯数据论”的认定,低估了网络诽谤的实质危害,易出现“数据达标就入罪、数据不达标就出罪”的司法机制,违背了罪责刑相适应原则。更有学者以司法解释形式对“转发超过500次”界定为诽谤罪的“情节严重”情形已经构成对基本权利的限制,违反了法律保留原则,这是对法律的补充,违反了宪法所规定的权力框架[7]。

(三) 网络型诽谤罪与传统型诽谤罪区分标准模糊

传统诽谤罪主体只有自然人,多发生于熟人之间,主观上具有直接故意。网络型诽谤罪主体具有多元化、匿名化、组织化特点,除了自然人,网络公关公司、媒体平台、水军团队等单位主体较为多见,通过技术手段批量制造、传播虚假信息,主体间无现实关系,主观上具有营利、恶意炒作等动机,主体身份认定难度大。现行刑法没有规定单位作为诽谤罪主体,因此对组织化网络诽谤行为规制不畅,与传统诽谤罪主体规制的逻辑相悖。此外,传统诽谤以口头、书面等线下方式传播,传播速度慢、范围小,具有“点对点”“点对面”的传播特征。而网络诽谤基于互联网平台,即时、广泛、裂变式传播,信息可以在短时间内传播到全国甚至全球,传播路径不可控、不可逆,算法推荐会更为放大这种传播效果,形成“舆论海啸”[8]。传播特征决定了网络诽谤的危害后果远高于传统诽谤,已有认定规则未依据传播特征调整认定标准,简单套用传统诽谤罪的规制逻辑,不能彰显特殊的网络犯罪规制逻辑。

3. 刑法视角下网络型诽谤罪的认定分析

(一) 网络型诽谤罪“诽谤事实”界定之厘清

根据《刑法》第二百四十六条和“两高”《网络诽谤解释》第一条可知,诽谤事实并非包含所有的虚假信息,而是特指能贬损他人人格、破坏他人名誉的具体虚假事实,并且必须与他人有所关联,具有可

被社会公众所知,降低被害人评价的性质。高峰(2025)从语义学和刑法解释学看,“诽谤”包括虚构事实和公开传播两方面内容,“捏造事实诽谤他人”并不是说,捏造和散布必须由同一人所完成,但捏造不散布不构成诽谤,只散布他人捏造的虚假事实,只要主观上“明知”且情节恶劣,同样可认定为“捏造事实诽谤他人”[9]。网络场域下的诽谤事实不单包括无中生有、凭空捏造,也包括对原始信息恶意篡改、断章取义、拼接剪辑形成的误导事实,且诽谤罪的主观方面包含间接故意,且本罪并非目的犯[10]。司法认定中必须要区分事实陈述和意见表达,对公共事物的合理批评、监督,即使部分内容失实,只要不是故意捏造损害他人名誉的事实,一律不属于诽谤事实范畴,坚持言论自由与名誉权保护的边界,不盲目扩大刑事处罚范围。

(二) 网络型诽谤罪“情节严重”认定标准之明确

网络诽谤犯罪的“情节严重”应当综合判断,否则容易造成机械司法和不当出入罪现象,进而导致刑法对言论自由的不当介入与过度限制[11]。针对网络诽谤的入罪判定规则,单次诽谤内容累计浏览、点击量超五千人或分享转发量满五百条,是结合线上信息扩散规律划定的量化评判依据。不过定罪不能单一拘泥于上述数值门槛,即便相关数据未达标,一旦致使受害人及其直系亲属出现精神失常、自伤乃至轻生等恶性结果,或是行为人此前两年已有诽谤类行政处罚前科、再度实施诽谤行为的,同样可归入情节严重范畴,这也是从法益实际受损层面开展实质性司法评判的具体体现。网络平台差异,公开社交场景与私人社群场景,朋友圈等封闭空间的传播应重点考量对被害人的实际侵害程度,而非数据。言论对象差异,对国家机关和工作人员的批评监督,应提高“情节严重”认定标准,保障公民监督权。司法认定要坚持主客观相统一,综合考量行为动机、传播范围、受众反应、被害人受损状况等,避免机械使用数据标准,实现罪责刑相适应。

(三) 网络诽谤犯罪与传统型诽谤罪之区分

无论是传统型诽谤罪还是网络型诽谤罪,都是通过散布人为捏造的虚假事实诽谤他人,继而导致了他人名誉权受损,情节严重应受刑罚处罚的行为。可以说,传统型诽谤罪与网络型诽谤罪具有基本相同的危害行为与极大范围重合的危害结果,在犯罪构成要件的主体、客体、主观方面都具有高度一致性。

网络型诽谤罪和传统型诽谤罪,只是在行为方法和危害结果上存在差异,二者虽然都是虚假散布虚假诽谤信息,但是网络型诽谤突破了传统诽谤的传统空间限制,利用网上线上网络空间向外扩散更为广泛,信息形式突破文字限制,AI换脸、虚假录音、P图等传统形式虚假信息的误导性更大,更容易让人相信。传统诽谤仅在传单、口头传播等传统传播途径中才能扩散,发散速度慢,范围小。而网络型诽谤通过社交网络、短视频等数字化平台,利用算法推荐、“水军”等手段就能够将信息快速扩散到全网。传统诽谤仅在熟人圈传播对被害人影响较大,传播在陌生地方传播对被害人影响极小,司法不在于传播数量的多少,而网络诽谤即使被害人是无名的,高点击、浏览、转发数量的诽谤信息也会对其心理与名誉造成损害,因此《网络诽谤解释》按“数量标准”认定“情节严重”是结合网络信息传播特点认定诽谤行为危害性的必然选择。

4. 刑法视角下网络型诽谤罪的司法完善路径

(一) 细化法律适用标准,明确认定边界

网络诽谤存在的司法适用问题主要是由于传统诽谤罪规范对网络场域的适用性差导致的,需要以罪刑法定与法益保护为视角,对构成要件进行细化。从行为方式来看,应根据言论内容区分事实陈述和意见表达,事实陈述应以合理查证后的基本真实为限度,意见表达应以不损害他人人格尊严为尺度[12]。依据语义学和刑法解释原则将“捏造事实诽谤他人”解释为“捏造虚假事实诽谤他人”,否定仅有恶意散布也是诽谤,将司法解释规定“明知散布”的情形与刑法相契合,排除类推效力的争议。对“诽谤事实”

进行界定,严格区分事实与意见、恶意捏造与合理批评。此外,可将《网络诽谤解释》第5条第2款所规制的行为理解为“网络造谣+起哄闹事”这一复合行为的方法可以在不违反罪刑法定的前提下兼顾法益保护的需求[13]。

(二) 规范证据收集与认定,破解举证难题

网络型诽谤罪的诉讼证明应当注意合理分配举证责任,即由控方承担构成要件和可罚性条件的举证责任,辩方承担存在违法阻却事由和责任阻却事由的举证责任[14]。证据收集方面,立法应当强化网络服务提供者存、提、供责任,明确平台存储诽谤信息的期限、调取步骤、应急处置时限,拒不配合取证的平台实施行政处罚和民事追责。二是在证据认定方面,司法机关应当建立盖然性与关联性相统一的证据认定标准,对电子数据真实,重点审查来源、传输路径、修改痕迹。对传播范围,主要通过平台后台原始数据确定,不能简单地通过公开显示数据认定。对行为人明知,主要结合行为人身份、信息内容、传播方式、是否被辟谣或投诉综合推定。法院应当主动、及时依申请或者依职权要求公安机关协助取证,包括IP地址、账号、传播路径、获利情况等。

(三) 优化诉讼程序,完善自诉与公诉衔接机制

我国诽谤案件刑事诉讼机制长期以自诉为原则,公诉为例外,但随着网络诽谤事件频发,网络诽谤事件受害人自行诉讼的负担也更大[15]。以“杭州快递诽谤案”为例,折射出我国网络诽谤案件自诉转公诉的程序衔接缺乏立法规定,“被害人撤诉讼”由于缺乏法律支撑而无法彻底解决该问题[16]。笔者认为应当明确自诉转公诉的标准、情形,将被害人受到网络暴力、造成重大身心伤害、造成社会秩序混乱等作为公诉的考量因素,由检察机关依职权进行审查,改变被害人以舆情启动的偶然性局面。此外,可理顺程序冲突,公诉程序吸收自诉程序的重要价值,在相关指导意见的基础上进一步扩大公诉范围。自诉程序中应当加强司法救济和公权力配合,针对被害人诉讼能力弱问题,建立法律援助、释明引导、调查令等制度支持被害人正确行使诉权。同时,明确公安机关的受理、初查、取证义务,报案后,公安机关应及时固定证据、排查行为人、评估危害程度,符合立案条件的依法侦查,不符合刑事标准的依法处罚,通过程序优化既避免公权力强行介入私人纠纷,又不致维权无门。

(四) 明确多元主体责任,强化协同规制

网络型诽谤是捏造、传播、转发、网络平台构成,完全靠一个自然人完成不了。因而要实现分层责任、协同规制。在自然人责任方面,要区分出恶意发起者、组织者、恶意推波助澜者与一般转发者,从重追究主观恶性大、传播作用重的主体责任,对于明知是诽谤信息仍进行教育、从宽处理的主体,实行宽严相济。在单位主体责任方面,将网络公关公司、营销号机构、新闻媒体等单位纳入刑法规制范畴,对于以营利为目的组织造谣传谣、破坏网络秩序的单位实行双罚制,追究直接责任人和单位罚金责任。压实网络平台的刑事责任边界,对明知是诽谤信息不删除、不屏蔽、不制止的平台和责任人,以诽谤罪共犯、拒不履行信息网络安全管理义务罪追究责任,实现责任与刑事风险匹配。

(五) 加强司法指引与宣传,提升治理效能

司法裁判具有规则的制定和行为的引导功能,依法治理网络诽谤,要以典型案例和司法指引明确裁判尺度,传递行为边界。最高法、最高检应不断发布指导案例,围绕“情节严重”“主观明知”“自诉转公诉”“平台责任”等热点问题归纳裁判规则约束自由裁量权,实现同案同判,可以通过白皮书、典型案例通报、庭审公开等形式明晰网络言论底线,告知人们网络不是法外之地,做好行为预期。法治宣传是源头治理的关键,司法机关要协同网信、公安、平台开展普法教育,针对大学生、新媒体从业人员、普通网民等重点人群,以案释法讲解诽谤罪构成、后果及维权渠道,消除“匿名无责”“法不责众”“转发无罪”的错误思维,推动平台建立前置提醒、风险提示、一键维权等功能,把法律规则嵌入产品中,实现技术治理与司法治理的协同统一,通过司法指引和社会宣传双向促进公民网络法治意识的提升,推动形成

尊法、守法、理性表达的网络生态，从源头减少网络诽谤。

5. 结论

网络型诽谤罪本质上仍是诽谤罪，旨在规制利用网络等现代化信息传输手段，捏造、散布虚假事实，损害他人名誉的网络诽谤行为，属于互联网时代下诽谤罪的一种新形态。根据刑法规范目的与法益保护本质，厘清诽谤事实的规范属性，明确捏造事实的判断原则。摒弃唯数量论、实质危害优先的情节严重认定标准。以法益为本，厘清本罪与传统型诽谤罪的相关犯罪的关系。司法完善通过细化法律适用标准、规范证据认定、完善自诉公诉衔接、明确多元主体责任、加强司法指导和宣传，建立实体准确、程序通畅、协同高效的治理体系。只有恪守罪刑法定、罪责刑相适应、权利保障均衡的原则，才能有效地惩治网络诽谤，保护公民的名誉权、规范网络言论秩序，推进网络空间健康有序发展。

参考文献

- [1] 木巴拉克·马合木提. 自媒体运营视角下的网络诽谤行为规制[J]. 职工法律天地, 2025(14): 76-78.
- [2] 韩雪梅, 马一凡. 网络诽谤行为的刑法规制研究[J]. 绥化学院学报, 2025, 45(2): 38-40.
- [3] 郭昊东. “网络诽谤入罪”的合宪性问题研究——以“点击、浏览、转发”条款为中心[J]. 浙江万里学院学报, 2026, 39(1): 27-36.
- [4] 李岚林. 网络影射型诽谤的刑事规制: 教义学阐释与体系建构[J]. 西南民族大学学报(人文社会科学版), 2025, 46(9): 29-38.
- [5] 孙宋龙, 赵聆伶. 网络诽谤自诉案件事实认定与证明问题研究[J]. 证据科学, 2025, 33(4): 445-460.
- [6] 吴波. 法教义学视角下网络诽谤犯罪的司法认定[J]. 华东政法大学学报, 2021, 24(3): 185-192.
- [7] 李晓明. 诽谤行为是否构罪不应由他人的行为来决定——评“网络诽谤”司法解释[J]. 政法论坛, 2014, 32(1): 186-191.
- [8] 刘佳敏. 网络暴力犯罪的程序法治理痛点及其应对——以网络诽谤案件为视角[J]. 中国刑警学院学报, 2025(3): 111-119.
- [9] 高峰. 网络诽谤争议问题的法律辨析与治理路径[J]. 新闻知识, 2025(2): 55-59+94.
- [10] 杜少尉. 网络暴力治理视野下诽谤罪的法益重塑与主观方面辨析[J]. 河南大学学报(社会科学版), 2024, 64(6): 42-48+153.
- [11] 金鸿浩, 杨迎泽. 网络诽谤犯罪“情节严重”的综合判断[J]. 国家检察官学院学报, 2022, 30(3): 111-126.
- [12] 王武, 阮素静, 王若男. 网络诽谤犯罪的司法认定及程序适用[J]. 人民司法, 2025(10): 12-15.
- [13] 王志远, 姚沉辰. 法益保护与罪刑法定之间: 《网络诽谤解释》第5条第2款之“限缩适用”[J]. 警学研究, 2021(6): 57-68.
- [14] 孙远, 史达. 论网络型诽谤罪之诉讼证明[J]. 证据科学, 2024, 32(1): 17-27.
- [15] 盖胜男, 吴海玲. 网络诽谤案件的公诉化诉讼机制构建[J]. 河北青年管理干部学院学报, 2024, 36(2): 101-106.
- [16] 赵天水. 网络诽谤犯罪自诉转公诉的域外考察及借鉴[J]. 江苏警官学院学报, 2021, 36(4): 63-69.