AI-generated Fake Information and the Crime of Internet Fraud: Current Legal Challenges and Paths to Reform

Yanyan Liu[1,a,*]

[1]School of Law, Zhejiang Gongshang University, Hangzhou, Zhejiang, 310018, China
a. Liuyanyan0528@gmail.com
**corresponding author*

Abstract: Artificial Intelligence(AI)-generated fake information is increasingly becoming a new type of online fraud, and its legal characterization and application of constituent elements face many challenges. This study mainly analyzes and compares the similarities and differences between AI-generated fake information fraud and traditional means of network fraud, while discussing the legal definition of AI-generated fake information and its application to the existing crime of network fraud. The study found that the existing legal framework has problems such as unclear definitions of fake information and vague definitions of responsibility when dealing with AI-generated fraud, making it difficult to effectively combat and prevent such crimes. For this reason, it is recommended that laws should be amended to clarify the responsibility of each subject, technical identification should be strengthened, and the platform supervision mechanism should be improved, so that the legal risks and social threats posed by AI-generated false information can be better dealt with. The conclusion of the study shows that only through all-round and multi-level legal reform and refinement can the legal order and social security be effectively maintained.

Keywords: AI Generated Fake Information, Fraud, Liability Determination, Legal Challenges.

1. Introduction

With today's trend of rapid development of digital technology, artificial intelligence (AI) gradually penetrates all aspects of social life, especially the development of generative artificial intelligence technology, which makes it easier and more convenient for people to generate fake information. AI nowadays can deeply forge realistic images, audio, video, and even complex textual content, which provides a new tool and means for the criminal activities of cyber fraud.

The application of this new tool in online fraud is now more and more diversified as well. AI can generate fake personal identification information, it can also collect and store real personal information, and generate false avatars, mailboxes, and fake profiles to create fake social media accounts, which can be used to carry out fraud. What's more, the existence of AI face-swapping function makes it possible to generate smooth fake videos, and this kind of deepfake is often applied in phishing attacks, requesting money transfers by pretending to be friends and relatives, and other scenarios in online fraud.

Existing laws on cybercrime, such as Article 266 of the Criminal Law of the People's Republic of China and Articles 47 and 67 of the Network Security Law of the People's Republic of China, seem to be unable to be fully adapted to this type of crime. Therefore, the core of the research in this paper is to analyze the legal identification of AI-generated fake information in fraud, explore the applicability of the existing legal framework in responding to this type of crime, and put forward the necessity of legal reform.

2. Comparison of AI-Generated Fake Information and Traditional Cyber Fraud Patterns

Observing China's past jurisprudence, cybercrime is often dependent on computers, and unlike computers that operate in a framework and programmed manner, generative AI's in-depth use of data demonstrates human-like thinking, and it can process and understand natural language in specific situations [1].

In terms of elements of the crime, AI-generated fake information may be difficult to be directly applied to the traditional elements of the crime of fraud. For example, in some cases, AI-generated fake information does not directly lead to property damage, but rather causes non-property damage like psychological and reputation. According to traditional theories, the determination of culpability for this type of behavior may be more in line with some tort liability violations.

There are also differences between AI-generated fake information and traditional cyber fraud in terms of the scope of application of the law. Traditional network fraud usually involves the application of national criminal law and network security law, and the scope of application of the law is relatively clear. Article 266 of the Criminal Law of the People's Republic of China says, those defrauding relatively large amounts of public or private money and property are to be sentenced to three years or fewer in prison or put under limited incarceration or surveillance, in addition to fines; or are to be fined. Those defrauding large amounts of money and property or having involvement in other serious cases are to be sentenced to three to 10 years in prison, in addition to fines. Those defrauding extraordinarily large amounts of money and property or involved in especially serious cases are to be sentenced to 10 years or more in prison or given life sentences, in addition to fines or confiscation of property. If cases are governed by other stipulations of this law, those stipulations shall apply [2]. There are detailed provisions on the constituent elements of the crime of fraud, legal consequences, specific manifestations of behavior and the amount of property damage. In contrast, the scope of legal application of AI-generated fake information is vaguer and more complex. This kind of crime involves a wider range of areas and may touch on a number of laws such as criminal law, civil law, cybersecurity law, and personal information protection law at the same time.

3. The Legal Positioning of AI-Generated Fake Information in Internet Fraud

3.1. Legal Definition of AI-Generated Fake Information

The legal definition of AI-generated fake information is currently a controversial issue. The current definition of disinformation usually focuses on the act of spreading disinformation artificially, while for fake information generated by AI, the generation process involves the complexity of algorithms and models, and the ambiguity of the definition increases the difficulty of regulation [3].

In traditional online fraud cases, fake information is usually intentionally fabricated by the perpetrator, whereas in the case of AI, the content may be automatically generated by an algorithm, and the perpetrator is not directly involved in the process of fabricating the information. In such cases, whether the law should recognize that AI-generated content constitutes “fictitious facts” or “concealment of the truth” requires the law to reassess the nature and harmfulness of AI-generated fake information.

3.2. Application of AI-Generated False Information to the Elements of Existing Internet Fraud Offences

In the legal composition of network fraud, subjective intent, objective behaviour and damage results are the core elements. How AI-generated fake information applies to these elements is the key to studying its legal responsibility.

First, regarding the determination of subjective intent, in traditional network fraud cases, the intent of the perpetrator is embodied in the active creation of fake information and the implementation of fraud under the condition of awareness; however, when it comes to AI-generated fake information, the perpetrator may have just initiated the AI process, and the subsequent generation of the content is done automatically by the algorithm. In this case, the following three points can be included in the consideration of subjective intent: first, whether the perpetrator already possessed the intent to deceive others when he initiated the AI procedure; second, whether this intent can be shown on the specific content; and third, AI has acquired the ability to generate human-like natural language text through massive corpus. It may be able to generate content containing bias and injustice that is more ‘real’ than real to human beings, for the content that is extremely close to the truth and difficult to identify after it is generated, whether the perpetrator is aware that its fake is important, too [4]. The extent to which one or more of these three factors must be met to be considered as constituting subjective intent, and how different degrees should correspond to varying sentencing standards, should all be included as considerations for subjective intent.

Secondly, the determination of objective behaviour. The objective behaviour of the cyber fraud usually includes the manufacture and dissemination of fake information and the commission of fraud through it. In the scenario of AI-generated fake information, the act of generating content is done automatically by the algorithm. Although the perpetrator may not be directly involved in the specific manufacturing process of the content, generating fake information often requires instructions or specific requests from the perpetrator. This automated behaviour essentially covers the objective behaviour of the perpetrator, and generative AI is currently only the perpetrator's ‘criminal tool’ rather than the subject of the act [5]. Therefore, on the question of whether the objective behavioural elements of the crime of Internet fraud are met, the legal interpretation can be refined and clarified, so that the behaviour of AI in generating fake information is gradually adapted to the objective behavioural elements of the crime of fraud.

Finally, in the determination of the damage result, the establishment of network fraud offence usually requires the victim to have a wrong understanding due to fake information, so as to make property disposal, resulting in the occurrence of the damage result. In the case of AI-generated fake information, due to the high degree of simulation and deception of the content, the victim may be more easily misled, leading to the occurrence of the damage result. The law needs to assess whether AI materially aggravates the harm of the cyberfraud to determine whether it should affect the sentencing criteria.

4. AI-generated Fake Information Brings Challenges to Existing Legal Frameworks

4.1. Limitations of Forensic Technology and Resources

The challenge is not just a technical one, but also a challenge to the existing legal framework. The legitimacy and authenticity of the object usually rely on the ‘rules of evidence’ and ‘expert testimony’, however, these traditional techniques and procedures are being profoundly tested when dealing with AI-generated fake information.

According to the basic principles of evidence law, the legitimacy and authenticity of evidence need to be proved through ‘direct evidence’ and ‘circumstantial evidence’. In the case of generative AI, direct evidence (e.g., original documents or physical evidence) is often difficult to obtain, or is itself synthesised by AI technology and difficult to identify through traditional identification methods. This makes it necessary for the court, when determining whether a piece of evidence is legally valid, not only to judge the appearance and content of the evidence, but also to resort to complex technical analyses to determine the manner of its generation and authenticity.

4.2. The Issue of Full Coverage of the Chain of Responsibility

The issue of dividing legal responsibility for AI-generated fake information is a greater challenge. In the traditional legal system, the identification of criminal acts and the pursuit of responsibility mainly focuses on the direct perpetrator, while the dissemination and use of AI-generated fake information involves multiple subjects, including the technology provider, the platform company and the content creator, making it difficult to apply the existing principles of responsibility allocation.

First, from the perspective of causation, when multiple subjects are involved, the technology provider may claim that they only provide a neutral tool, and the platform side may claim that they only bear the responsibility for the dissemination of the information, rather than the creator of the information content. In such cases, the definition of liability involves the acts and faults of multiple subjects, and it is not sufficient to consider the traditional principle of ‘direct causation’ alone, but also the complex interactions between the various acts. This refinement of the causal relationship requires more in-depth study in legal theory.

Secondly, in terms of criminal liability, the criminal act of AI-generated fake information faces the challenge of ‘identifying the perpetrator’. For example, when false information is disseminated and infringes on the rights of others, should the responsibility be attributed to the perpetrator who initially generated the information, or should the dissemination platform be held criminally liable, or in some cases should the joint and several liability of the algorithm developer be considered. These questions reveal the complexity of the chain of responsibility for AI-generated fake information, and also highlight the dilemma of the existing criminal law system in responding to new types of technological crimes.

In addition, from the perspective of compliance and regulation, AI-generated fake information has raised new compliance challenges for platform companies and technology providers. Under the existing legal framework, platforms usually argue that they should not be held liable as long as they do not actively intervene or participate in the generation of content, which is called an "information intermediary". However, with the widespread use of AI technology and the proactive role of platforms in recommending information, this neutrality and exemption from liability are being questioned. Whether platforms should assume stricter compliance obligations such as proactively monitoring and auditing content and establishing effective risk management mechanisms, is an area where the law could be improved.

5. Legal Responses to AI-Generated Fake Information Cyber Frauds

5.1. Upgrading of Forensic Technology and Resource Inputs

In response to the limitations of judicial appraisal technology and resources, multifaceted improvement measures must be taken. First, cooperation between judicial departments and technology enterprises, universities and research institutions should be promoted to develop more advanced identification technologies and tools, especially for the identification and analysis of AI-generated false information. Meanwhile, relevant technology departments can introduce AI-based forensic systems to improve the efficiency and accuracy of case adjudication. The government can also increase technical training and professional development in the judicial system to ensure that law enforcement officials and judges can understand and respond to new technologies. In terms of budget and resource allocation, the Government can prioritise the provision of sufficient financial support for these areas to ensure the continued development of technology and talent.

5.2. Clarifying the Legal Provisions of Responsibility

Regarding the complexity of the responsibility for AI-generated fake information, it is recommended that the legislature formulate new laws or amend existing laws in order to clarify the legal responsibilities of each subject. This includes refining the specific responsibilities of each party, such as algorithm developers, platform operators, and information disseminators. Such clear legal provisions will not only effectively define the responsibilities of each party, but also improve the enforcement of the law, while allowing for a more precise fight against the generation and dissemination of false information, thus improving the state of the law's response to new types of cybercrime.

5.2.1. Creating a ‘Technical Audit Obligation’ for AI Technology Providers

The legislature should develop specialised technical standards and compliance requirements that set out the basic principles for AI technology providers to follow when developing algorithms. For example, technical measures should be taken to prevent algorithms from being misused for illegal purposes, including deep forgery and fake information generation, which could also include ethical and security assessments of algorithms to ensure that they do not pose a significant risk to society if they are used inappropriately.

Also, mandatory algorithm examination and certification systems should be introduced, so that AI models that generate content are examined and approved by the government or an independent organisation. This move will ensure that AI algorithms meet national security and public interest requirements and reduce potential risks. At the same time, fines, licence revocations or other legal sanctions will be imposed on technology providers that violate the censorship requirements.

5.2.2. Strengthening the ‘Compliance Obligations’ of Platforms

Legislation should require platforms to establish effective mechanisms. There are already provisions requiring digital platforms to remove relevant content and products as soon as they receive information reported by users, and to inform digital service providers in a timely and adequate manner of the relevant measures to be taken. Failure to remove them in a timely manner can result in a fine of 6 percent of the platform's global turnover, so the platform can act immediately to remove or block the content when fake information is reported by users or third parties [6]. The law should also set specific timeframes (e.g. within 24 hours) and penalties in case of non-compliance to ensure that platform parties actively fulfil this obligation.

Platform parties should be required to introduce AI content recognition systems that automatically scan and detect whether uploaded content is AI-generated and flag up potential false information. Such systems can use machine learning algorithms to detect anomalous features in images, videos, audio and text to assist platform parties in identifying and dealing with fake content in a timely manner. The platform party should be held liable for compensation in the event that it fails to fulfil these obligations, resulting in the loss of users due to false information. The law may set out the upper limit of compensation and exemption conditions for the platform party, and clarify specific operational guidelines and procedures, such as how the victim can submit an application for compensation and how the platform party can respond and handle the application.

5.2.3. Legal Clarity on the Division of Responsibilities

Existing laws should clearly define the specific allocation of responsibilities among the various actors in the generation of fake information by AI. Platform parties should have initial responsibility for reviewing and monitoring the content uploaded on their platforms, while technology providers need to provide traceability mechanisms for the misuse of their technology. In addition, it can be stipulated that perpetrators should bear the main criminal and civil liability when using AI-generating technology to commit fraud, while technology providers and platform parties should bear joint and several liability accordingly.

Second, the law could provide clear conditions for each party to be exempted from liability, for example, the platform party could be exempted from further legal liability if it removes the offending content within a reasonable period of time. Similarly, technology providers can be exempted from liability under certain conditions if they develop and reasonably review the use of their algorithms within a compliant framework. Clear conditions for recovery and exemption from liability will help to reduce legal disputes and increase the incentives for compliance by all parties.

5.2.4. Strengthening Public and Industry Oversight Mechanisms

There are significant technological gaps between online digital platforms and traditional government regulators when it comes to combating the spread of fake information. Large information asymmetries make it practically difficult for regulators to deal with the ‘deep fake’ that appears online in a timely way, and to determine accurately the realistic scale of regulation [7].

In response to this type of problem, the government can set up a specialized regulatory agency or committee responsible for overseeing the dissemination of AI-generated fake information and the fulfillment of the responsibilities of each subject. At the same time, AI regulators should establish a public reporting platform to encourage the public to report deepfake, and provide appropriate legal protection and reward mechanisms for whistleblowers, and take public monitoring and reporting as an important criterion for evaluating deepfake, and penalize users of platforms that have been reported for repeated violations, and warn users who maliciously use deep fake technology to carry out illegal activities. Penalize platform users who have been repeatedly reported to have violated the law, and penalize users who maliciously use the deepfake technology to carry out illegal activities by issuing warnings, fines and blocking their numbers, so as to ensure that the supervision work can cover more practical scenarios [8].

Industry associations and enterprises should be encouraged to jointly formulate self-regulatory conventions or codes of conduct for AI-generated content to enhance overall compliance through self-management and mutual supervision in the industry. At the same time, an AI technology ethics committee can be set up to review potential abuses of AI-generated content and issue ethical guidelines to help companies better understand and comply with legal and ethical standards.

6. Conclusion

This study focuses on the application of AI-generated fake information in the application of the constituent elements of Internet fraud, and provides an in-depth analysis of the limitations of the existing legal framework. The study first compares the differences between AI-generated fake information and traditional online fraud in terms of specific modes and legal characterization, and clarifies the application problems and application suggestions that may arise from the constituent elements of traditional fraud crimes in the face of AI technology. At the same time, the study points out the complexity brought by AI-generated fake information to the determination of responsibility, especially the existence of legal blind spots in the allocation of responsibility among multi-party subjects such as algorithm developers, platform operators and information disseminators.

After sorting out the challenges of the existing legal framework, the study proposes three specific countermeasures, including strengthening technological identification means, revising laws and regulations, and reinforcing platform responsibility. These measures aim to fill legal loopholes, improve law enforcement efficiency, and provide institutional safeguards to deal with cybercrime caused by AI-generated fake information.

With the further development of AI technology, the means of generating fake information will become more complex and hidden, so legal research should also continue to pay attention to the challenges posed to legal rules by the advancement of AI technology. Future research can explore the application of AI-generated fake information in other types of crimes, as well as the role of international cooperation in dealing with cross-border cybercrime, in order to better respond to the global challenges, and to safeguard the rule of law order and security of society.

References

[1] *Zhang Chenghao, Zhang Yong. (2023) Source Governance of Generative AI: Risk Concerns and Criminal Regulation of Deep Data Utilization[J]. Hubei Social Sciences,(11):127-135.*

[2] *Standing Committee of NPC, (2023) Criminal Law of the People's Republic of China, China Legal Publishing House,(12),*

[3] *Wang meichen. (2024) Analysis of Legal Issues Regarding the Promotion of the Proliferation of False Information by Generative Artificial Intelligence: Taking "Deepfake" as an Example [C]// The Journal of Legal Studies, Volume 3, 2023 - A Collection of Research on the Rule of Law for Promoting the Integrated Clustering Development of Strategic Emerging Industries.School of Law, Shanghai Maritime University;8.*

[4] *He jing. (2023) Challenges and Strategies for the Development of Intelligent Communication Based on ChatGPT [J]. Yuejiang Academic Journal,(03):67-73+173-174.*

[5] *Liu shuiling, Dong Wenkai. (2024)The Impact and Response of Generative Artificial Intelligence such as Sora on the Judical Identification of Fraud Crimes[J]. Issues on Juvenile Crimes and Delinquency,(03):114-124.*

[6] *European Commission, Digital Services Act (DSA), 2020(12), 2024(10), https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en*

[7] *Zhang Yuanting. (2022) The Legal Regulation of "Deepfake" Abuse in the Age of Artificial Intelligence [J]. Theory Monthly. (09):118-130.*

[8] *Shang Haitao. (2023) New Paradigm and System of Legal Regulation of “Deep Fake” Technology [J]. Hebei Law Science,41(01):23-42.*