Influence Assessment of Pornography-Related Cryptograms Based on ChatGPT-4o

Guancheng Chen *, #, Yize Wang #, Yijun Yang #, Sujia Yin #, Shiqi Li #, Hao Xu #

The School of Criminal Investigation, People's Public Security University of China, Beijing, China, 100038

* Corresponding author: frankchenguancheng@outlook.com

#These authors contributed equally.

Abstract. In order to optimise the network monitoring strategy, objectively and accurately analyse the propagation trend and characteristics of different types of pornographic cryptic expressions, and identify the cryptic expressions with greater influence, this article propose a method to assess the influence of pornographic cryptic expressions based on ChatGPT-4o. With the help of knowledge graph to make the content of cryptic text structured, and the frequency of cryptic appearances, the transmission path, and the number of people involved as the factors for assessing the influence of cryptic phrases, the pornography-related cryptic phrases are quantified. Firstly, this article use ChatGPT-4o to preprocess data and extract knowledge from the text corpus of pornography-related cryptic phrases through multiple rounds of Q&A, and construct a knowledge graph in the field of pornography-related cryptic phrases in a low-resource, fast and timely manner, and statistically analyse the frequency of cryptic phrases, propagation paths, the number of people involved and other types of factors based on the knowledge graph, and abstract the impact of different cryptic phrases into impact factors, and use the impact factors to depict the propagation trend and propagation characteristics for comprehensive analysis. The impact of different types of anaphors is abstracted into influence factors, and the influence factors are used to portray the propagation trend and characteristics for comprehensive analysis and judgement. By analyzing the existing cryptic data and calculating the impact factor of cryptic types, we obtained the changes of the impact factor of different cryptic types, verified the scientificity and validity of the impact factor calculation method, and provided a new method and idea for the evaluation of pornographic cryptic language. Combining the advantages of ChatGPT-4o and Knowledge Graph helps to grasp the change trend of cryptic communication in a timely manner, provides powerful support and guidance for network regulation, and is of great significance for protecting social morality and youth's health.

Keywords: Network Monitoring Strategy, Pornographic Cryptic Expressions Propagation, Cryptic Expressions Analysis, Influence Assessment, ChatGPT-4o.

1. Introduction

With the development of society and advancements in technology, the spread of pornography-related cryptic phrases has become a prominent issue in internet regulation, posing a threat to social morality and the health of young people. Current regulatory measures are mostly based on empirical analysis, which often leads to subjectivity [1], partiality, and a lack of adaptability. Given the variability of cryptic phrases, it is necessary to comprehensively assess the threat level and influence of different types of cryptic phrases and to capture the trends of emerging cryptic expressions promptly [2]. This allows regulatory resources to be focused on the most threatening areas, improving the efficiency of resource utilization, driving decision-making through quantitative data, and enhancing the scientific and effective nature of internet regulation [3].

To address the challenge of constructing knowledge graphs for pornography-related cryptic phrases due to small sample sizes, this paper proposes using ChatGPT-4o to build the cryptic phrase knowledge graph. Leveraging the powerful language understanding capabilities of ChatGPT-4o [4], knowledge extraction and graph construction can be completed at low cost and high quality without the need for data annotation or model training, by simply restricting content using question templates [5].

Additionally, a method for calculating influence factors based on the frequency of cryptic appearances, transmission paths, and the number of people involved is proposed to assess the influence of cryptic phrases, thereby formulating scientific and effective regulatory measures [6]. It is expected that this approach will improve the scientific and comprehensive monitoring of cryptic phrases, enabling precise suppression and effective prevention of cryptic phrase dissemination.

2. Influence Assessment Method for Pornography-Related Cryptic Phrases Case Types

Using large language models for data preprocessing eliminates the need for complex cryptic phrase lexicon construction and model training, directly leveraging the semantic understanding capabilities of the model to handle cryptic information, simplifying the process while ensuring accuracy. Compared to deep learning models, large language models can use existing capabilities to perform extraction tasks directly, offering high flexibility, supporting template-constrained tasks, and naturally achieving joint modeling of entities, events, and relationships [7]. Multi-turn Q&A illustrates the extraction logic, providing prior information and solving problems such as long-distance entity and relationship overlap.

2.1. Acquisition of Pornography-Related Cryptic Corpus

The foundation for constructing a knowledge graph in the domain of pornography-related cryptic phrases is the acquisition of relevant corpus. This paper builds the Yellow Hint Corpus (YHCs), which includes Yellow Hint Cases (YHC) and Anti-Yellow Hint Measures (AYHM) [8].

The corpus in YHC mainly comes from two sources: first, from pornography-related cryptic cases collected from online forums, social media, and other public platforms; second, from various cryptic cases collected in the work of internet regulatory authorities. The cryptic cases from regulatory authorities are primarily gathered from monitoring and reported information, which includes chat records, webpage content, social media posts, and other documents. All such documents have undergone desensitization to protect personal privacy, such as usernames, IP addresses, and contact information. The knowledge extraction for constructing the knowledge graph mainly uses brief case descriptions.

Although both the public platform and regulatory authority cases contain typical pornography-related cryptic information, their writing styles differ significantly. The texts from public platforms tend to be informal and conversational, whereas the cryptic cases from regulatory authorities are more formal and standardized. This makes the extraction task challenging if using ordinary deep learning models.

The corpus consists of 1680 typical pornography-related cryptic cases, with 840 entries from public platforms and 840 from regulatory authorities, covering incidents from January 2020 to February 2023. Based on the classification system for cryptic phrases issued by internet regulatory authorities, the cases in YHC are categorized into seven major types and 60 specific methods: impersonation, shopping, seduction, fake warning, daily language, phishing virus, and other new illegal types. The corresponding regulatory measures in AYHM are prevention methods and treatments for these 60 specific methods. Although using ChatGPT to extract knowledge from text corpus is cost-effective without needing labeled data for training, to objectively verify ChatGPT's extraction capabilities, 100 entries of public platform corpus and 100 regulatory case data entries were randomly selected from the YHC, with entities, relationships, and events manually annotated by regulatory experts and professionals in related fields, serving as the gold standard for subsequent experiments to compare the extraction effects of ChatGPT and deep learning models.

2.2. Overall Method Flow

The construction and application of the knowledge graph for pornography-related cryptic phrases involve data preprocessing, various knowledge extraction techniques, the selection of influence factors for different cryptic types, and targeted regulatory measures. First, before constructing the knowledge graph, the text corpus YHCs undergoes preprocessing, which mainly involves transforming implicit expressions in the text into a standard format to facilitate subsequent extraction.

After data preprocessing, entity types and relationship types are determined, and a portion of the text is randomly selected. Different Q&A templates are used with ChatGPT-4o to extract knowledge from the corpus as required. By comparing the extraction results of different templates, the Q&A template used for extraction is finalized. Using the final template, entity extraction, relationship extraction, event extraction, and time extraction are performed for the entire corpus. After knowledge extraction, a knowledge graph is constructed using the Neo4j graph database, integrating the extracted results.

Based on the knowledge graph constructed using ChatGPT-4o, influence factors of different cryptic types are assessed by considering the frequency of cryptic appearances, transmission paths, and the number of people involved, characterizing propagation trends and features, and proposing corresponding regulatory measures and preventive methods.

2.3. Application of ChatGPT in Cryptic Extraction Tasks

2.3.1. Data Preprocessing Using ChatGPT

The data for this article was obtained from the Public Safety and Risk Laboratory, using Python's web scraping techniques to crawl websites such as Telegram. Due to the informal and cryptic nature of the collected pornography-related text corpus, directly processing the text could yield poor results, affecting the quality of the constructed knowledge graph. Therefore, data preprocessing is necessary. This mainly involves preprocessing implicit expressions in the text to facilitate recognition of various cryptic phrases and transmission times. The text often contains vague time markers like "someday" or "a few days later," requiring standardization. For this purpose, a template using ChatGPT-4o for standardizing time expressions is employed: "In the given text… replace all temporal information with precise year and month details to make the time information more specific."

2.3.2. Knowledge Extraction Using ChatGPT

Although ChatGPT-4o has human-like capabilities in general fields, its direct knowledge extraction in the domain of pornography-related cryptic phrases is not ideal. Therefore, corpus restrictions are required to better leverage ChatGPT-4o's capabilities and more accurately extract entities, relationships, events, and time from the corpus.

To perform the extraction task using ChatGPT-4o, a fixed template is needed to extract information from the corpus. The design of the template affects the extraction quality, as different templates yield different results. Template content also serves as a part of ChatGPT-4o's knowledge, aiding in information extraction. Therefore, labels should be embedded within the template, and using a multi-turn Q&A format for various extraction tasks significantly improves extraction accuracy. This is because, although ChatGPT-4o excels in human-like interaction, using it as an extraction tool requires it to produce content in a fixed format. Thus, a fixed question template helps achieve an optimal balance between performance and output format accuracy.

2.3.3. Storage of the Pornography-Related Knowledge Graph

Once the text extraction tasks using ChatGPT-4o are complete, the extracted entities, relationships, events, and times are saved using the Neo4j database, forming a visual knowledge graph. Neo4j is a graph-based database that uses a graphical data model to store and process data, supporting the modification and query operations of the knowledge graph using Cypher statements. Figure 1 shows a part of the constructed knowledge graph.

2.3.4. Method for Assessing Case Type Influence Factors

Due to limitations in manpower and resources, priority should be given to monitoring the most harmful and widespread pornography-related cryptic phrases and raising public awareness. Thus, research on the propagation trends and features of different cryptic types is necessary. Through the knowledge graph, statistical analysis of cryptic factors (frequency, transmission path, number of people involved, etc.) is performed. First, correlation and factor analysis determine the relationship between cryptic factors and their influence on propagation characteristics. Each factor's impact is quantified as a weight, with experts scoring them, and the influence of the cryptic phrases is calculated. Finally, the weight values undergo fitting analysis to determine the influence factor calculation method for propagation trends and characteristics. Expert-scored weights are estimated using maximum likelihood and normalized for fitting, yielding the fitting function that shows weight variation trends as follows:

$$\omega_1 = 1 - e^{-\beta \times \sum_{i=1}^{n} N_i} \tag{1}$$

$$\omega_2 = 1 - e^{-\lambda \times \sum_{i=1}^{n} A_i} \tag{2}$$

$$\omega_3 = \frac{\sum_{i=1}^{n} \left(2 - e^{\ln 4 \times \Delta t_i}\right)}{n} \tag{3}$$

$$\omega = \omega_1 + \omega_2 + \omega_3 \tag{4}$$

In the formula, ω represents the comprehensive influence factor of a specific type of cryptic phrase, n denotes the total number of cases of this cryptic type, ω1 is the factor for the number of people involved, N represents the number of people involved (measured in "tens of people"), ω2 is the factor for occurrence frequency, A represents occurrence frequency (measured in "hundreds of times"), ω3 is the factor for transmission path, and Δt_i represents the time difference between the occurrence of a specific cryptic phrase and a specified time (measured in "years"), typically chosen as six months. β and λ are hyperparameters used for calculating the influence factors, generally set to 1.

Analyzing the fitting effect of the influence factor calculation function reveals a good performance in showing the basic trends: the more frequent the occurrence and the greater the number of people involved, the larger the influence factor, indicating that greater attention is needed. However, when the frequency and number of people reach a certain level, the growth rate of the influence factor slows down and remains below 1, which helps mitigate the impact of outlier data on the evaluation. As for the transmission path, the more concentrated the path, the larger the influence factor.

2.4. Evaluation Metrics

In the knowledge extraction process, commonly used evaluation metrics include Precision, Recall, and the F1-score. Table 1 shows the confusion matrix, and the formulas for Precision, Recall, and F1-score are given below [9]:

Table 1. Confusion matrix

Confusion matrix		Actual	
		1	0
Predicted	1	TP	FP
	0	FN	TN

$$P = \frac{TP}{TP+FP} \tag{5}$$

$$R = \frac{TP}{TP+FN} \tag{6}$$

$$F1 = 2 \times \frac{P \times R}{P+R} \tag{7}$$

3. Experimental Results and Analysis

The results extracted using ChatGPT-4o and those of the unsupervised knowledge extraction models with better current performance were compared and analyzed using Precision [10], Recall, and F1-score metrics. Table 2 presents the statistics of entities and relationships in the labeled dataset.

Table 2. Statistics of dataset labelling results

Dataset	Number of Entities	Number of Relationship Types
Judicial Cases	3144	14
Field Cases	2100	7

3.1. Evaluation of ChatGPT Extraction Tasks and Template Selection

Three different question templates were used for knowledge extraction with ChatGPT-4o: Template 1 for rough Q&A, Template 2 for single-turn detailed Q&A, and Template 3 for multi-turn detailed Q&A. Although the differences in the results of these three templates can be observed intuitively, quantitative comparisons were also conducted for rigor. The detailed results are shown in Table 3.

Table 3. Statistics of different template extraction results

Template	Entity Extraction			Relationship Extraction		
	P	R	F1	P	R	F1
Template 1	52.81	51.09	51.23	16.46	13.72	14.64
Template 2	66.76	63.69	65.25	65.18	63.48	67.62
Template 3	68.25	67.26	68.25	68.25	66.56	68.15

Comparing the Precision, Recall, and F1-score across different templates, it is evident that Templates 2 and 3 significantly outperformed Template 1, with an increase of over 14 percentage points in entity extraction F1 and more than 48 percentage points in relationship extraction F1. The improvement is due to adding extraction-related information in the templates, providing an extraction paradigm for various tasks while constraining extraction results to avoid difficulties in synonym normalization. The improvement in relationship extraction was primarily because 14 types of relationships were defined, and the responses were constrained in Templates 2 and 3. Comparing Templates 2 and 3 shows that the multi-turn Q&A approach has certain advantages, as the answer to the previous question serves as a prompt for the next, enhancing the logical flow. Therefore, Template 3 was used in subsequent comparisons for evaluating the accuracy of knowledge extraction.

3.2. Entity Extraction Results

Given the high specialization of the pornography-related cryptic domain, limited labeled data, and the fact that no labeled data were used for training the model, the comparative models mainly included unsupervised and distant supervised models for extracting knowledge from text. However, for some common tasks, such as cryptic phrase identification and transmission path identification, more mature entity extraction models were used as baselines to further demonstrate the advanced capabilities of ChatGPT-4o.

3.2.1. Baseline Models for Entity Extraction

In the entity extraction subtask, ChatGPT-4o was used to extract named entities, which mainly included cryptic users, transmission paths, cryptic types, number of people involved, and transmission media. Since cryptic users and transmission paths may involve names and places, mature Chinese name and place recognition models were used as comparative models. Although the extracted results by ChatGPT-4o include entity types (e.g., "cryptic user: Zhang"), when calculating recognition accuracy, no distinction is made between user and path types, but the specific role of each name or place can be determined by the accuracy of relationship extraction.

In the comparison, well-performing models in named entity recognition were selected, and some labeled YHC pornography-related cryptic cases were used to fine-tune the pre-trained models to achieve good results in the current dataset. The traditional deep learning models involved included:

(1) Bi-LSTM-CRF Model: This model combines LSTMs (Long Short-Term Memory networks) with CRFs (Conditional Random Fields), capturing both the mapping relationship between samples and annotations and the relationships among annotations. It was trained on an open-source dataset to predict labels and extract entities, achieving good performance across various datasets.

(2) BERT-CRF Model: Similar in principle to Bi-LSTM-CRF, but using BERT as the emission matrix for CRF training, showing strong performance in named entity recognition tasks.

(3) FGN (Fusion Glyph Network) Model: This model integrates glyph information to assist Chinese named entity recognition, using internal Chinese character information, and has shown great results across several datasets.

(4) LEMON (Lexicon Memory) Model: This model uses lexicon-based memory to recognize named entities, combining the characteristics of words and characters, achieving good results on public datasets.

MECT (Multi-Metadata Embedding Based Cross-Transformer) Model: This model uses multiple metadata to improve Chinese named entity recognition by capturing semantic information from the structure of Chinese characters.

In addition to traditional deep learning models, several large language models similar to ChatGPT have shown good performance on various tasks. Therefore, comparisons were also made with such models, including:

(1) Huawei Pangu NLP Model: A Chinese pre-trained model with over 100 billion parameters, optimized for the Chinese language, performing well across many tasks.

(2) Alibaba Tongyi Qianwen Model: A language model with strong understanding capabilities, integrating multimodal knowledge to provide efficient generation capabilities.

All Q&A for large models used Template 3 to minimize differences in extraction effects caused by varying templates and guide the models to generate the correct answers.

3.2.2. Result Display

Table 4 displays the performance of the comparative models and ChatGPT-4o on entity extraction.

Table 4. Display of results of various entity extraction methods

Model	Judicial Cases Dataset			Field Cases Dataset		
	P	R	F1	P	R	F1
Bi-LSTM-CRF	62.15	60.26	60.74	57.11	58.21	58.72
BERT-CRF	63.84	61.27	61.27	58.16	59.12	59.33
FGN	64.87	63.79	63.11	60.19	61.12	60.12
LEMON	65.12	65.21	65.14	61.21	63.17	58.12
MECT	67.15	67.14	68.14	65.18	66.16	65.21
Huawei Pangu	69.16	68.87	69.67	66.97	65.12	68.01
Alibaba Tongyi	70.26	69.12	70.11	67.14	67.25	66.24
ChatGPT-4o	71.83	69.25	71.25	68.18	68.89	68.17

Comparing the entity extraction results of ChatGPT-4o with those of mature deep learning models, the effect is similar. ChatGPT-4o's entity extraction slightly outperformed that of deep learning models in network regulatory scenarios involving pornography-related cryptic phrases, with the F1-score being 1.67 percentage points higher. This might be due to the implicit nature of pornography-related cryptic cases, which presents challenges for deep learning models trained on fixed-format data and formal language, whereas ChatGPT-4o’s training data is broader and not limited by expression formats.

Other large language models such as Huawei Pangu and Alibaba Tongyi Qianwen, though performing well in extraction tasks without training, still fall short compared to ChatGPT-4o: in the public platform cases, Huawei Pangu's performance was 2.24 percentage points lower than ChatGPT-4o, and Alibaba Tongyi Qianwen's performance was 1.85 percentage points lower. In regulatory field cases, Huawei Pangu lagged by 3.59 percentage points, and Tongyi Qianwen by 1.58 percentage points. Therefore, for entity extraction tasks, ChatGPT-4o performed better.

3.3. Relationship Extraction Results

Table 5 presents the results of relationship extraction by comparison models and using ChatGPT-4o as a tool.

Table 5. Display of results of various relationship extraction methods

Model	Judicial Cases Dataset			Field Cases Dataset		
	P	R	F1	P	R	F1
GraphRel	42.53	40.14	41.30	41.37	39.76	40.54
CopyRL	43.27	40.09	41.62	43.11	40.61	41.82
CASREL	59.63	52.94	56.09	56.38	51.65	53.91
Huawei Pangu	62.78	60.13	61.43	62.19	59.94	61.04
Alibaba Tongyi	63.19	61.08	62.12	63.16	61.21	62.14
ChatGPT-4o	68.98	66.51	67.72	67.16	67.72	68.12

From the data in Table 5, it is evident that large language models outperform deep learning models in relationship extraction tasks. This is largely due to the fact that the template used for large language models constrained the possible types of relationships, and also due to the limitations of small sample sizes—even with partial labeled data for fine-tuning, deep learning models still fail to fully perform in pornography-related cryptic case text corpora. Among the same models, extraction from public platform texts outperformed extraction from regulatory field texts, primarily because public platform texts are more standardized, with fewer implicit relationships. Among the large language models, ChatGPT-4o outperformed the other two models by over 5 percentage points in F1-score.

Combining the results of entity extraction and relationship extraction across various deep learning models, large language models, and ChatGPT-4o, it can be concluded that the use of ChatGPT-4o for constructing knowledge graphs offers higher accuracy and is significantly more cost-effective, highlighting its advantages.

3.4. Influence Factor Assessment

After determining the method for calculating the influence factors of different cryptic types, it was necessary to verify its feasibility and accuracy. In this study, the influence factors for seven types of pornography-related cryptic phrases were calculated based on the collected data, and the changes over time were plotted as line graphs. Data from 2020 to 2023 were analyzed on a semi-annual basis, starting from January 2020. The parameters β and λ for frequency influence factors and number of people involved were both set to 1. The x-axis represents the time point, with "2020.1" indicating January 1, 2020, and "2020.6" representing June 30, 2020, while the y-axis represents the influence factor.

By analyzing the trends in influence factors, it is possible to understand the changes and developments of each type of pornography-related cryptic phrase. Table 6 provides the keywords and scenario corpus for pornography-related cryptic phrases.

Table 6. A corpus of pornography-related cryptic keywords and scenarios

Indicator	Keyword	Scenario Corpus
Flight	LinkedIn Brush	Brush platform / LinkedIn brush
	Brush Platform	Brush platform / LinkedIn brush
	Telegram Brush	Telegram account / Telegram brush
	Flight Number	Flight number
	Member	member
	User	Open flight user
	GAY	Ma Dou 91 / GAY dating
	Studio	Media studio / 91 Media
	Domestic Video	91 welfare / 91 domestic / 91 video
Brothel	User	51 brothel user exchange group
	Brothel	Chengdu beautiful brothel verification list
	Tea Party	Chengdu Little Moon Phoenix tea gathering
	Door-to-Door	Baoding door-to-door / brothel / tea
	Resources	Handan high-definition resources
	High-end	Xi'an billion-dollar boy dream lover

	Full Set	Changsha brothel ~ KTV ~ full set
	Massage	Lanzhou brothel / Lanzhou resources / massage
	Group	Repair group / cracked group
	Come in and Pick	Chengdu brothel resources report room
Door-to-Door	User	High-end door-to-door service user resources
	Hangzhou	Hangzhou escort / brothel / resources
	Guangzhou	Guangzhou escort / repair / door-to-door
	Shenzhen	Shenzhen Jinbian clubhouse, available for delivery
	Beijing	Beijing Concubine Selection Channel
	Chengdu	Chengdu door-to-door / brothel / resources
	Door-to-Door	National fly-in door-to-door services
	Xiamen	Xiamen young models, full set services
	Shanghai	Shanghai courtesan
	Wuxi	Wuxi virgin door-to-door
Escort	User	High-end / door-to-door service user resources
	Escort	High-quality escort / nationwide escort
	Shenzhen	High-quality escort
	Beijing	Beijing escort / McLaren
	Shanghai	Shanghai escort / students
	Guangzhou	Guangzhou Fengming Academy
	Hangzhou	Hangzhou escort / Ferrari
	Guangzhou Escort	Escort / repair / door-to-door
	Guangzhou Repair	Escort / repair / door-to-door
	Guangzhou Student	Student / influencer hidden recordin
Repair	User	Flower viewing / soliciting / escort / repair
	Beijing	Beijing repair clubhouse
	Clubhouse	Shenzhen selection repair
	Phoenix	Chongqing Phoenix eating chicken group
	Pile Driver	Midnight repairman / pile driver
	Team	Philippines group rental / largest Chinese circle
	Paradise	Tianjin Paradise
	Happy Grove	Malaysia Happy Grove
	Full Set	Changchun repair full set spa/ktv
	Guangzhou	Guangzhou tea house, celebrity mansion
Student	Lolita	Pure student / jk / lolita / welfare sister
	Welfare Sister	Welfare club / student stockings Douyin
	Contrast	High school lolita white tiger
	Escort	Beijing escort / student / model
	Deflower	Middle school virgin deflowering
	Oral Sex	Eating melon / curiosity / museum
	Female Dog	Rare student / female dog exposure
	Young Foot	Student young foot / footjob
	Student	Student resources / 91 hidden recording
	Hidden Recording	Contrast hidden recording student
One Hand	User	User resources contrast self-recording
	Beauty	Welfare sister beauty self-stimulation
	Pin-hole	High school lolita white tiger pinhole recording
	Cheap Price	Stockings student / cheap price Douyin
	Teacher	Young teacher / student dating / teacher wife
	Pure Male	Hidden recording / student / homemade pure male
	Fat Butt	91 fat butt student Chinese
	Seven Wolves	Seven Wolves grey production group connection
	Junior Sister	Contrast hidden recording junior sister

Analyzing the influence factors and their synthesis, the "Flight" category of cases has an influence factor related to the time of occurrence concentrated around 0.56, indicating that such cases persist almost continuously across large samples. The influence factors for both the amount involved and the number of people involved in the "Flight" category remain consistently high, approaching 1.

The influence factors for "Welfare" cases show an arc-like shape, with higher values in early 2020 and late 2022. This trend mainly results from a surge in pornography-related "Welfare" cases involving fraudulent sales of masks, protective suits, and other pandemic-related materials during those periods, whereas such cases were less common at other times.

For "Impersonation" cases, both the amount involved and the number of people involved were relatively high until mid-2021. However, after mid-2021, public awareness campaigns by law enforcement improved people's understanding of these tactics, leading to increased prevention and a decline in both metrics. Analyzing the influence factor for the time of occurrence reveals a cyclical pattern, which can be linked to the two major shopping periods, "6.18" and "11.11." These are high-risk times for impersonators posing as customer service or delivery personnel. The influence factor for cases occurring in the first half of the year (ending June 30) was higher than for cases occurring in the latter half, but both were above the average value of 0.56.

The "Brothel," "Door-to-Door," and "Escort" categories exhibit high variability in the time-related influence factors due to smaller case sample sizes. From the metrics of the amount involved and the number of people affected, more people fell victim to the "Brothel" type, mainly targeting the elderly through scams exploiting their concern for their children. The "Escort" type involved fewer victims, but primarily targeted corporate executives and business owners, resulting in larger monetary amounts.

Analyzing the influence factors for "Repair" cases shows a marked increase in both the amount involved and the number of people affected in early 2021, with the time-related influence factor stabilizing around 0.56. This was mainly due to the surge in "One Hand" cases beginning in 2021, which posed a significant threat to public financial safety.

Based on the comprehensive analysis of influence factors, targeted regulatory measures can be developed. For frequently occurring and economically damaging cases like the "Flight" category, awareness campaigns and appropriate regulatory actions should be increased. For "Brothel" scams that frequently target the elderly, targeted public awareness and preventive measures should be implemented. For high-profile individuals commonly targeted by "Escort" scams, campaigns explaining the common transmission methods of malware could help prevent the root cause of pornography-related scams. When shortages of highly sought-after resources arise, targeted awareness campaigns about the "Welfare" category should be conducted, advising people to verify the credentials of sellers and avoid falling for scams due to low prices.

Overall, "Flight" and "Repair" cases, which are more frequent and cause significant damage, should be subject to continuous awareness campaigns and regulatory efforts, while focusing on monitoring financial and social activities of key populations. Other pornography-related cases should be closely monitored for their cyclical nature, with timely identification of criminal tactics and effective preventive measures proposed. Figures 1 through 4 illustrate the time trends of the individual and comprehensive influence factors for various case types, with the x-axis representing the time point (e.g., "2020.1" for January 1, 2020) and the y-axis representing the influence factor magnitude.

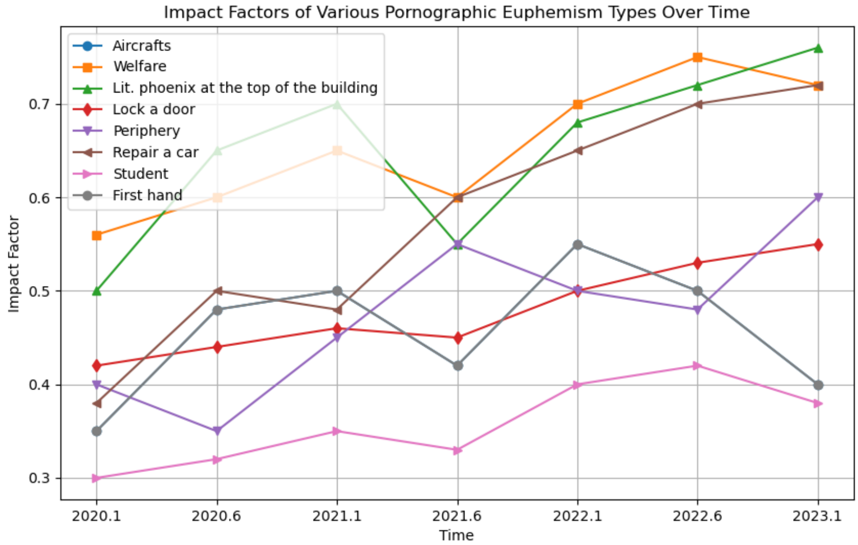

Figure 1. Trends over time in the time-to-case impact factor for each type of case

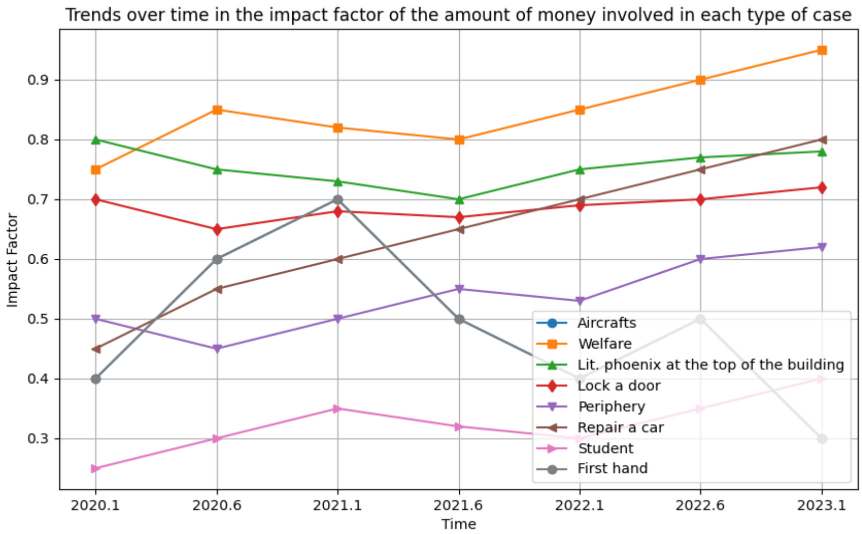

Figure 2. Trends over time in the impact factor of the amount of money involved in each type of case

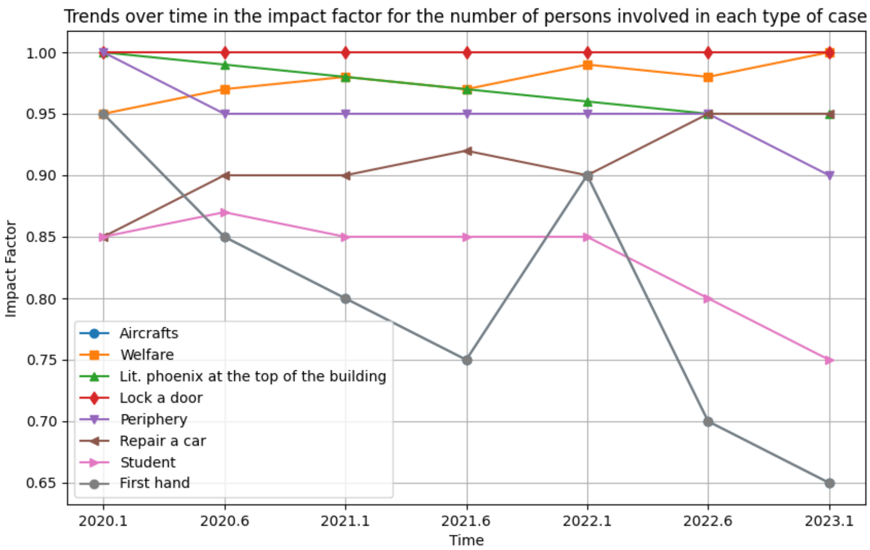

Figure 3. Trends over time in the impact factor for the number of persons involved in each type of case

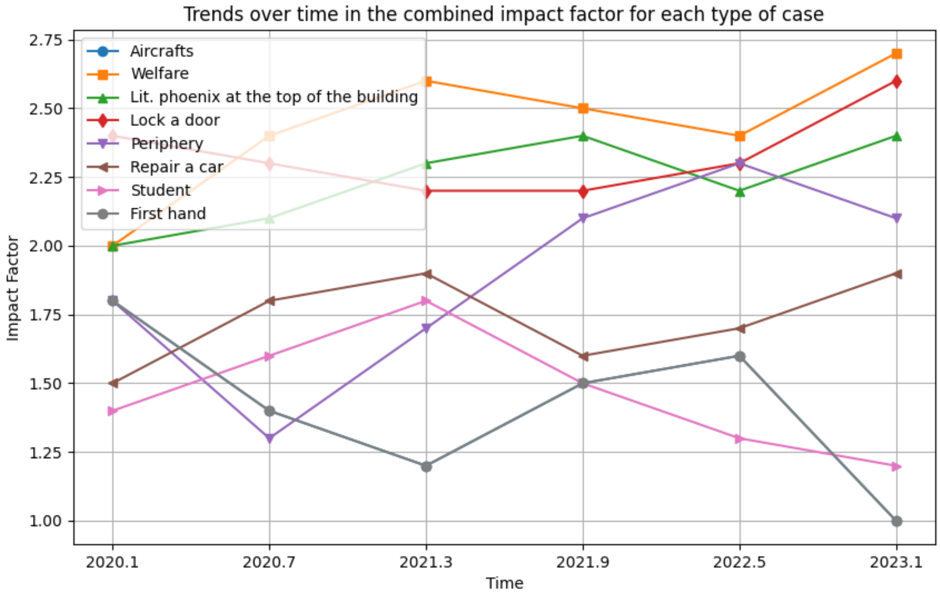

Figure 4. Trends over time in the combined impact factor for each type of case

4. Conclusion and Outlook

In this paper, we utilized the powerful language understanding capabilities of ChatGPT to process texts related to pornography-related cases, completing tasks such as data preprocessing and constructing a knowledge graph for the pornography-related domain. Based on the constructed knowledge graph, we conducted a comprehensive assessment of various factors such as incident time, amount involved, and the number of people involved in different cryptic types, to make targeted regulatory decisions.

The proposed method of using ChatGPT for data preprocessing and knowledge extraction to construct knowledge graphs, compared to using traditional deep learning models, does not require labeled training data, thereby mitigating the challenges of insufficient domain-specific corpus for training deep learning models. For the pornography-related domain, where the available corpus is limited, using a general-purpose language model like ChatGPT allows for faster deployment without the need for time-consuming model training typical of deep learning methods. This enables timely assessment of trends in pornography-related cases and the formulation of targeted regulatory measures, thereby providing a viable approach for safeguarding public welfare. Moreover, using ChatGPT to handle and understand various types of texts with a small number of samples provides a potential solution to meet the needs of special vertical sectors like public security.

Furthermore, in response to the increasing frequency of pornography-related cases, this paper proposes a regulatory strategy assessment method based on three factors: incident time, amount involved, and number of people involved. Influence factors were introduced to evaluate trends for different types of cases, which can guide targeted regulatory actions, optimize the allocation of resources, improve regulatory efficiency, and enhance public awareness and prevention of widespread pornography-related scams.

References

[1] Ye J, Chen X, Xu N, et al. A Comprehensive Capability Analysis of GPT-3 and GPT-3.5 Series Models [EB/OL]. 2023.

[2] Gilardi F, Alizadeh M, Kubli M. ChatGPT Outperforms Crowd-Workers for Text-Annotation Tasks [EB/OL]. 2023.

[3] Kuzman T, Mozetic I, Ljubesic N. ChatGPT: Beginning of an End of Manual Linguistic Data Annotation? Use Case of Automatic Genre Identification [EB/OL]. 2023.

[4] Huang F, Kwak H, AN J. Is ChatGPT Better than Human Annotators? Potential and Limitations of ChatGPT in Explaining Implicit Hate Speech [EB/OL]. 2023.

[5] Kutela B, Msechu K, Novat N, et al. Uncovering the Influence of ChatGPT's Prompts on Scientific Writings Using Machine Learning-Based Text Mining Approaches [EB/OL]. 2023.

[6] Gao T, Yao X, Chen D. Simcse: Simple Contrastive Learning of Sentence Embeddings [C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2021: 6894 - 6910.

[7] Barbieri F, Camacho-Collados J, Anke L E, et al. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification [C]//Findings of the Association for Computational Linguistics: EMNLP 2020, 2020: 1644 - 1650.

[8] Van Hee C, Lefever E, Hoste V. Irony Detection in English Tweets: SemEval-2018 Task 3 [C]//Proceedings of the 12th International Workshop on Semantic Evaluation, 2018: 39 - 50.

[9] Deepa M D. Bidirectional encoder representations from transformers (BERT) language model for sentiment analysis task [J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12 (7): 1708 - 1721.

[10] Wang Y, Yao Q, Kwok J T, et al. Generalizing from a few examples: A survey on few-shot learning [J]. ACM Computing Surveys (csur), 2020, 53 (3): 1 - 34.