DOI: http://dx.doi.org/10.34069/AI/2020.29.05.53

Criminological Analysis and Filtering of Sites with Aggressive Content

Криминологический Анализ и Фильтрация Сайтов с Агрессивным Контентом

Received: April 2, 2020 Accepted: May 8, 2020

Written by:
Nguyen Huy Binh[210]
https://orcid.org/0000-0002-5185-4239

Abstract

The work carried out a study of the current problem of minimizing risks and damage from aggressive impacts on the Internet, in social networks. The methodology of system and forensic analysis has been used. Using methods of mathematical and information-logical analysis, comparative and cognitive analysis of counteraction to aggressive impact on the visitor or owner of the site (moderator of the forum) is carried out. Methods of aggressive behavior (disinhibition, trolling, cyber-bullying and astroturfing), their local and global goals - up to the capture and management of communicative resources for the purpose of affecting the user - have been analyzed. Aggressiveness on the site has become a public problem, legally not only in Vietnam, but also in many countries, including Russia, France, South Korea, etc. Hosting distances itself from an important problem for it as well. The aggressiveness of a site is classified by the site's aggressiveness scale. Criminological identification of aggression and aggressor is a complex and multidimensional problem. Was carried out systemic and practical, criminological analysis of aggressive behavior on sites. Was possible to formulate and propose measures, approaches to filtering and combating aggressive behavior on the site, as well as to formalize it mathematically, for subsequent research. The analysis will help to counter the aggressor and support the victim.

Keywords: aggressive influence, aggressor, website, criminological, systemic, analysis, identification.

Аннотация

В работе проведен системный и криминологический анализ актуальной проблемы минимизации рисков и ущерба от агрессивных воздействий в сети Интернет, в социальных сетях. Методами системного анализа, криминологического, психологического и математического анализа проводится анализ целей, задач и решений по противодействию агрессивному воздействию на посетителя или владельца сайта (модератора форума). Проанализированы категории агрессивного поведения (растормаживание, троллинг, кибербуллинг и астротурфинг), их локальные и глобальные цели-вплоть до захвата и управления коммуникативными ресурсами с целью воздействия на пользователя. Агрессивность на сайте стала общественной проблемой, легально не только во Вьетнаме, но и во многих странах, включая Россию, Францию, Южную Корею и др. Хостинг также дистанцируется от важной для него проблемы. Агрессивность сайта классифицируется по шкале агрессивности сайта. Криминологическая идентификация агрессии и агрессора - это сложная и многомерная проблема. Был проведен системный и практический, криминологический анализ агрессивного поведения на объектах. Удалось сформулировать и предложить меры, подходы к фильтрации и борьбе с агрессивным поведением на сайте, а также формализовать его математически, для последующего исследования. Анализ поможет противостоять агрессору и поддержать жертву.

Ключевые слова: агрессивное воздействие, агрессор, сайт, криминологический, системный, анализ, идентификация.

[210] PhD in Law, Deputy head of the criminal police Department, People's Police Academy, Hanoi, Vietnam.

Introduction

On sites it's often possible to find not only elements of subculture, but also clearly aggressive content, especially, placed and activated in "Anonymous" mode. A user with difficulties in real communication may be tempted to "free of all rules and control" interpersonal communication. It can try to self-establish or compensate for the lack of attention of others. Anonymity allows not only to misinform the interviewer, but also to try to influence him in an aggressive way, to use him for his evil purposes. There is even a new term - "cyber-aggressiveness" or "web-aggressiveness", introduced in 2007 by D. Shabro as a category of deviant behavior, a phenomenon of social disinhibition. It can manifest itself not only in the form of trolling, cyberbullying and astroturfing, etc. (Bulatova, 2017; Vorontsova, 2016; Yaseen Al-Shwani and Shaker Mahmood, 2019).

Most analysts see trolling as a speech form of provocation to escalate conflict in communications. This form has also met before Internet networks, but in web communication the consequences and forms are tightened by anonymity and equal access to communicative space. Goals are also changing: not only the acquisition (interception) of data, negative impact on the image of the interviewer, etc., but also global - the capture and management of communicative resources to influence the owner or user. Aggressiveness on the site is a significant factor for society of aggressive social and psychological impact not only on the user, but also on the society (Bochaver and Hills, 2014). This is, so far, a virtually right-wing mechanism for manipulating communications. Hosting can't handle it, and doesn't want to do it.

International legal action to combat web-based aggression faces many obstacles, mainly due to the lack of a relevant legislative framework. However, in a number of countries, draft laws are being drafted to combat aggression on websites and on the Internet. For example, in South Korea, France and Germany are already punished for cyberbullying, which led to legal misconduct (up to 10 years in prison). Such measures are beginning to be worked out in Vietnam.

There is also a network consulting "Anticybermobbing", where you can get real help. In Moscow - online agency of advice on safe use of the Internet "Children online". It's important to always bear in mind that the key state value is the social and legal state (Khodusov, 2020).

The work carries out a systematic and relatively practical analysis of problems of criminological analysis of aggressive behavior on sites, as a result of which measures, approaches to filtering and combating aggressive content and behavior on sites are proposed.

Theoretical framework

Criminological analysis of the content of sites requires that the criminologist look at the event as criminal behavior, from crime went to complex problems of society, prevention of a criminogenic event (Luneev, 2015, c.308). Criminology is in crisis in many countries, not only because of the pandemic, but because of the lack of effective means to combat crime.

Prevention of a criminal act is a difficult task, requiring not only financial and material investments, but also maintaining a theoretical and case law base. Research on crime problems in Viet Nam is carried out by employees of research and educational legal institutions, laboratories and centers, forming teams of specialists in law, IT, psychology, psychiatry and other areas, as well as the public.

Joint and multi-dimensional cooperation is important, not formal, but with the development and testing of pilot projects (Yuzhanin, 2020). The motivation of aggression, tools that criminalized civil use, with the extraction of self-interest, etc. are taken into account. In modern criminological studies of aggressive actions on the Internet, websites usually estimate by the rate of violent crime (relative value, for example, per 1000 visitors of the site), age and anti-social activity, etc.

Statistical analysis of aggressive behavior on the site (on the Internet) should be accompanied by identification:

1) dynamic characteristics (tempo, frequency, etc.);
2) levels of social group, network (meta, macro, mesa- or micro-environment);
3) spheres and groups of activity (economic, political, spiritual, etc.).

Here, without system analysis-synthesis, structural problems cannot be solved.

Mathematical and situational information-logical modeling is also needed (Kaziev, 2007).

Jurisprudence in Vietnam takes into account not only technological innovations of investigation. It's engaged in the prevention and prevention of aggressive network effects on the person, especially on the teenager. Technological advances develop network interactions (Artyushina, 2019), having great evolutionary potential, but their use can be directed to harm, to be latent and remote.

The results of exorbitant network aggressiveness can be:

1) attacks on the rights-freedoms of the individual, which led to a real suicide in networks (virtually);
2) inducement to suicide (assistance in networks in its commission);
3) network activity that encourages suicidal behavior;
4) involvement in the commission of network actions dangerous to human life;
5) illegal distribution of personal information in networks;
6) threats of death (causing serious harm) of a person, removal of his organs (tissues), and abduction;
7) coercion to sexual influence, virtual corruption of a minor, etc.

Aggression in the form of physical force often gives way to intellectual and mental influence. The affecting person often turns out to be an advanced IT user, emotionally resilient (Buz, 2017). Such a method is faster, psychologically more "comfortable" ("for one click"), effective (easily hidden) and intellectual (Ignatov, Višneviecki, & Kashkarov, 2018). The forms of aggression themselves are also changed and technologically updated, becoming latent. Careful system-legal analysis and organizational efforts are required to filter aggressive content on websites, experiments in the Internet (Olinder and Gamburov, 2017).

Let's take a look at the task of dynamically filtering websites with aggressive content. Most often, URL filtering, generation and application of "black" or "white" URL lists are used for filtering. To do this, you monitor web resources, user behavior on sites. For example, in Russia the Register of "the black websites" of Rostekhnadzor works No. FZ-139, 28.07.2012, FZ-149, Article 15.1 (the State registry, 2020). It blocks sites after repeated complaints and monitoring. There are many reasons for blocking:

1) due to child pornography, advertisements; narcotic substances;
2) methods of and calls for suicide;
3) prohibited by courts, etc.

Blocking is performed only by lists:

1) not allowed updating of lists (bases);
2) does not prevent an advanced user from bypassing such lists, for example, using IP addresses or using Anonymous public services;
3) updates the problem of restricting access to the entire site, not to its separate harmful page.

It's prohibited in Russia already more than 20,000 web pages. Several thousand decisions were made on web pages calling for suicide on social networks.

Internet censorship is also used for algorithms of filtering web resources by blocking lists of domain names of aggressive sites. Different approaches vary from country to country. Vietnam can be classified as a "middle" class - filtering isn’t as strict as in China and not as weak as in Canada, where a special instruction on web censorship circumvention was even prepared (Civisec, 2007), which was also suitable for systems of filtering aggressive web resources.

The intelligence of the filter algorithm, analysis of the topics of web pages is important. A heuristic approach is commonly used (Kalafati, Moiseyev, Starkov and Shushkova, 2008). You must use all types of filtering to reduce heuristic errors. For example, page analysis omissions and error locking. Dynamic filtering error - within 5-10%, for surfing - large.

Virtual content analysis for aggressiveness is a popular phenomenon, as virtual judicial representation (the website of the court, the bar and another law firm) implies a legal environment that allows employees to conduct their business with the most efficiency by applying network capabilities, IT and digital signature.

Methodology

The alternative to the URL filtering method is dynamic filtering, where content is analyzed at the time of access or on the fly. The browser blocks the web page from loading when its

content is identified as unwanted. The work proposes a method that allows to ensure the maximum speed of decision-making according to the question: does the content of the analyzed page correspond to the analyzed topic? In the method used, when the answer is negative (i.e., the topic is undesirable), access is blocked.

As a rule, all experience and all algorithms cannot be immediately applied to the Vietnamese zone, as to any other, such as Runet. Filtration methods are several. The easiest way to do this is to implement content filtering as a client application. This does not exclude the application of the more complex method of the dedicated filter server of the company, the organization.

As the most relevant but also more complex method, an external filtering server is offered on the hosting provider site, or a cloud service (as filtering outsourcing). For example, the "parental control" option is sometimes sufficient. For aggressive impacts, attacks and large organizations, we offer the use of a dedicated filter server. It has more capabilities to configure and administer the filter, does not require highly qualified specialists. The third method proposed in the work - external filter-server is a compromise solution, considering that the https-filter is bypassed, blocked by IPFW.

A mathematical modeling method is used to restrict (lock) the work. The procedure of simulation of expert criminological assessment of site blocking factors is proposed. For these factors the method of average integral values is built.

A system is also used that will allow you to assess the levels of current aggressiveness of the site using the multidimensional scaling method.

Meta-tag auditing techniques (such as Refresh) and Cross Site Scripting, SQL injections, and PHP inclusions have also been used.

Results and discussion

The list of common aggressive actions on the site grows, though wrestling is going on. It's important to protect the site in a comprehensive manner, eliminate vulnerabilities, and resist sustainable negatives and regulations, without making it difficult to protect the work of specialists, without ignoring the declared rules and restrictions of work.

Models of possible aggression, methods of computer implementation - many, based on requirements classified by:

1) goals of realization, vulnerability of infrastructure;
2) sources of destruction (both external and internal);
3) information technology, as well as political or other nature, contributing to the realization of the goal of aggressive influence;
4) potential or realized damage taking into account the limit damage threshold (Ugolnicki and Usov, 2020) in each environment.

When constructing a legal model of aggressive impact, the multi-stage effects are taken into account:

1) analysis of potential object state to which aggressive action is directed;
2) initiation of cyberconflict and its strengthening;
3) multiple (maximum) damage;
4) consolidation of the outcome of the conflict.

Models (diagrams) in the above stages include not only IT-impact, but also psycho-legal; rely on classes of threats that violate confidentiality and integrity of data, leading to dependence of potential users. This leads to psycho-legal demoralization, a decrease in the level of readiness for resistance. Examples are unsustainable personalities.

The main targets of impacts may be:

1) discrediting, justifying the actions of the aggressor and allowing attracting other sympathizers;
2) reduction of psychological readiness to repel the main phase of aggression;
3) reinforcing and resonating effect of the striking action of the aggressor due to participation of third parties.

In addition to content filtering, code access and web logs of user activity analysis, personnel training, security audit, active license anti-virus packages, Firewall, secondary authentication, etc. We need a comprehensive approach to IT auditing. One of its components is the audit of security and vulnerability - systemic, with competent security policy (Kaziev, Kazieva & Kaziev, 2017).

The audit of the site should assess in a comprehensive manner all its capabilities and competitiveness, strategies of reaction to aggressive visitors. For this purpose, methods of continuous identification of users are effective (Cochegurov and Martynova, 2020).

The comprehensive audit of the site implements the assessment: security - risks, security, vulnerabilities (Višniewski, 2018). Recently, the biometric standard of access, its capabilities (Olejnik and Castelluccia, 2013) and recommendations, which need not only to be argued, but also ranked by importance, have also begun to be used.

To restrict the site and block it, we offer the following mathematical model (simulation procedure). The site liquidation rates are set by experts (criminologists and other experts of the expert group). We take them as average integral values of the form:

$$m_x = \int_x^{x+1} \mu(x) l(x) dx \Big/ \int_x^{x+1} l(x) dx \,,$$

where in the numerator - the number of prohibitions in the group of "equal-weight by aggressiveness" for *x* years, and in the denominator - the average number of sites for the period [*x; x+1*], $\mu(x)dx$ - the ban probability (function of distribution of probabilities), $\mu(x)$ - the rate of the ban on group, *l(x)* - the rate of replenishment of group (emergence of the new websites):

$$\mu(x) = \frac{k(l_x - l_{x+1})}{(l_x + l_{x+1})},$$

where *k* is the acceleration factor, it can be taken, for example, as *2 (k = 2).*

Indicators are considered absolute (not the most informative for dynamics), relative (for structure analysis, quality of absolute indicators, for example, indices of search engines), dynamic (trend) and qualitative (popularity). For the analysis of a condition of the website, its rank variables which can be quantitative (comparative), serial (ordinal) or rank are used (on group). When generating a rating, it's necessary to choose the type of loyalty variables or aggressiveness of the site to be taken into account.

The aggressiveness of the site can, in our opinion, be classified by the following scale:

1) the level of aggressiveness is maximum (aggressive actions are accompanied by the competence of the aggressor, are effective to a sufficient extent, and the countermeasures taken are ineffective);
2) level - high;
3) level - medium;
4) level - permissible;
5) the level is weak;
6) safety level.

To implement filtering and, if necessary, site blocking, you can consider the appropriate status indicators:

1) 3xx - status code, request redirection, for example, 305 - access code, which is possible only through the proxy server;
2) 4xx - error code, problems on the site, for example, 403 - prohibition code to view pages of the request;
3) 5xx is the error code detected by the server, e.g. 502 is the invalid response code of the server in the query chain;
4) robots.txt - command list file for automatic scanning of the site;
5) site map (.xml) - pages to be indexed in the folder behind the site start page;
6) closed from page indexing (via robots.txt, Noindex tag).

At a minimum, use secure HTTPS and the Refresh meta-tag audit. When placing media content that is uncontrolled in a rigid way (names, values, protocols), the attacker can get (very likely) full page control, for example, using Cross Site Scripting (XSS) - attacks through user scripts, SQL injections, PHP inclusions. And there and before attacks (XSRF or DDoS) - "near": on forums it's possible to covertly connect links.

A site security policy is required, providing security infrastructure, organizational measures and security identification procedures (electronic logs, administration, etc.). This policy is based not only on filtering, but also on the delineation of rights. For each user-process relationship pair, list the valid access types. The hierarchy only regulates the tolerance class, and administrators, the security subsystem, can modify rights. It's used more frequently by mandate access or by access category in the hierarchy. If the classification level of the user is sufficient, the categories include hierarchical categories at the process classification level. The development of an effective filtering procedure is based on analysis of visitor behavior, adaptation of policies and protection of the site to current or

predicted behavior (Zhong, et al 2014). For example, you can read and write information on an access group's mandate.

Conclusions

"Virtual communication" has become a common phenomenon. Especially for teenagers. But it's such communication that can easily implement introduction into personal space with aggressive intentions. The illusion of anonymity and the resulting illusion of impunity in the mind of the influencing person provoke some users to conflict and aggression. A direct (planned) goal is also possible - to humiliate a particular person. This is evident in networks in the form of an attack on the author of the blog, moderator of the forum, owner of the account of the social network, etc. It also affects the fact that the aggressor may not be satisfied with his life, thus getting rid of dissatisfaction with himself and increasing self-esteem by reducing the self-esteem of the interviewer.

The criminological approach to identifying the fact of aggression and the aggressor itself is a complex, relevant and multilateral problem. The system analysis of the problem and the proposed approaches and models will be useful in future studies on the problem. They will help counter the aggressor and support the victim. The new problem of modernity requires both criminological support and psychotherapeutic support. But IT support for identifying and "bana" an unwanted visitor is particularly needed. The work also analyzed errors and measures to audit the site.

Bibliographic references

Artyushina, O.V. (2019) Violent Crime and IT Technology // LEX RUSICA. №9 (154). pp. 78-83. DOI: 10.17803/1729-5920.2019.154.9.077-084.

Buz, S.I. (2017) Concept and main criminological characteristics of violent crimes in modern Russia // Journal of the Krasnodar University of the Ministry of Internal Affairs of the Russian Federation. №3(37). -pp.8–11.

Bulatova, E.I. (2017) Networking communication strategies: Trolling // Journal of the St. Petersburg State University of Culture and Arts. №2(31), pp. 75-78. Retrieved from: https://cyberleninka.ru/article/n/setevyekommunikativnye-strategii-trolling (date of appeal: 20.02.2020).

Bochaver, A., Hills, K. (2014) Cyberbullying: Harassment in the space of modern technology // Psychology. Journal of the Higher School of Economics. Vol.11, N3. –pp. 177-191.

Civisec. (2007) Guide to Bypass Internet Censorship for All. The Citizen Lab, University of Toronto. Retrieved from: http://www.civisec.org/sites/all/themes/civisec/guides/everyone's-guide-rustian.pdf (date of appeal: 10.02.2020).

Cochegurov, E.A., Martynova Yu. A (2020) Aspects of continuous user identification based on free texts and hidden monitoring // Programming and Computer Software, Vol. 46, No.1, pp. 12-24.

Ignatov, A.N., Višneviecki, K.V., Kashkarov, A.A. (2018) Conceptual bases, directions and measures to counter criminal violence // Legal science and practice: Journal of the Nizhny Novgorod Academy of the Ministry of Internal Affairs of the Russian Federation. №2 (42). pp. 133-141.

Kaziev, V.M. (2007) Introduction to analysis, synthesis and modeling of systems. –M.: Binom. Laboratory of knowledge. Intuit. – 244 p. ISBN 5-94774-511-9. Retrieved from: https://www.intuit.ru/studies/courses/83/83/info (date of the address: 20.02.2020).

Kaziev, V.M., Kazieva, B.V., Kaziev, K.V. (2017) Foundations of legal informatics and informatization of legal systems, 2nd ed. -M.: INFRA-M, 336 pp. ISBN-online: 978-5-16-104376-9. Retrieved from: http://znanium.com/catalog/product/545154.

Kalafati, Y.D., Moiseyev, K.V., Starkov, S.O., Shushkova, S.A. (2008) Technology of Storage and Processing of Electronic Documents with Intellectual Search Properties // International J. Information Theories and Appl. (IJITA), Vol.15, 184.

Khodusov, A. (2020). Social legal state as a constitutional value // Amazonia Investiga, 9(25), pp.471-478. URL: https://amazoniainvestiga.info/index.php/amazonia/article/view/1096

Luneev, V.V. (2015) Course of world and Russian criminology. In 2 volumes, vol. I. General part / V.V. Luneev. -M.: Jurayt, -1003 p. ISBN 978-5-9916-1751-2.

Olejnik, L., Castelluccia, C. (2013) Towards Web-Based Biometric Systems Using Personal Browsing Interests. International Conference on Availability, Reliability and Security. – Regensburg. - pp.274-280. DOI: 10.1109/ARES.2013.362.

Olinder, N.V., Gamburov, E.A. (2017) On the Results of the Experiment "Search and Perception of Identity Information on the Internet" and its Use in Crime Investigation // Forensic Expert, No.4, pp.29-31.

State register of prohibited sites (2020). Retrieved from: https://eais.rkn.gov.ru/ (date of appeal 23.03.2020).

Ugolnicki G.A., Usov A.B. (2020) Dynamic models of coordination of private and public interests in the sphere of economic corruption. J. Comput. Syst. Sci. Int. 59, pp.39–48. https://doi.org/10.1134/S1064230720010128

Višniewski, A.S. (2018) Deceptive system for detection of hacking, based on analysis of website visitors behavior // Cybersecurity issues, №3(27), pp.54-58. DOI:10.21681/2311-3456-2018-3-54-62.

Vorontsova, T.A. (2016) Trolling and Fleiming: Speech Aggression in Internet Communications // Journal of Udmurt University. History and philology. Vol.26, N2 - pp.109-116.

Yaseen Al-Shwani, N.A.; Shaker Mahmood, I. (2019). The crime of exploiting children in prostitution via internet (legal study), *Revista de la Universidad del Zulia*, 10 (28), 315-330. https://produccioncientificaluz.org/index.php/rluz/article/view/30813/31853

Yuzhanin, M.A. (2020). The specifics of manipulative influence in contemporary social interactions and communications // Amazonia Investiga, vol.9 (26), pp.125-133. https://doi.org/10.34069/AI/2020.26.02.14.

Zhong, J., Yan, C., Yu, W., Zhao, P., Wang, M. (2014) A Kind of Identity Authentication Method Based on Browsing Behaviors. Seventh International Symposium on Computational Intelligence and Design. -Hangzhou. pp. 279-284. DOI: 10.1109/ISCID.2014.205.