8th International Conference on Computational Social Science IC²S², July 19-22, 2022
The University of Chicago Booth School of Business, Chicago, IL, USA

A Computational Framework for Understanding Risk Factors in Cybercrime

Keywords: Online Victimization; Adolescents; Cyberbullying; Causality; Bayesian Statistics

Extended Abstract

The Internet has become an integral part of young people's lives. Minors (under 18 years old) already accounted for about one third of Internet users worldwide back in 2017 [1], and the COVID-19 pandemic has increased their online exposure. This has positive and negative consequences. On the one hand, digital technology and hyper-connectivity come with significant educational and economic opportunities and better access to information. It can be a game-changer, helping people unlock all their potential. On the other hand, the Internet also opens the door to new threats and opens new pathways for the ones already known. Given this, digital connectivity has made young people more accessible to bullies, sex offenders, and those who harm children, thus vulnerable minors being at greater risk. Around 10% of European children are already victims of cyberbullying (CB) on a monthly basis [2], and 49% have experienced a CB-related situation at least once [3]. These issues have gotten worse during the COVID-19 pandemic. Among the European children who had already reported being a victim of CB, 44% reported an increase in CB during the 2020 lockdown [3].

Law enforcement and cybersecurity professionals are making significant efforts to analyze technical tools and uncover strategies used to exploit digital vulnerabilities. However, social and psychological aspects of online threats are often neglected: What social or psychological features make someone more vulnerable? Are specific sectors of the population more vulnerable? Most importantly, what is the most effective way to identify potential victims to help them protect themselves? Understanding behavioral and developmental aspects around cybercrime is becoming increasingly important, underlying the need to shift from a reactive and punitive approach to prevention, deterrence, and early intervention.

In order to develop effective prevention and intervention strategies, it is essential to understand the role of the social, psychological, and technological factors involved in cybercrime, as well as their interplay. Traditionally, social scientists had focused on the valuable task of determine the human factors that are essential for identifying victims of cybercrime. However, such approaches can only gain insight into how these factors affect the cybercrime under study, but are not able to understand how these factors interact with each other, the relative weight of these factors in the presence of others, or the potential effects of confounders not measured in the study. This limits our understanding of the "whole picture" and the complex human process through which cybercrime has developed. On the other hand, the black box approach commonly used in computer science and Artificial Intelligence for prediction does not offer an effective and generalizable solution to these problems, especially when the data is sparse, which is usually the case.

This work proposes an iterative approach grounded in theory at a conceptual level and as a guiding process. Instead of fitting models to available data or devising experiments to identify factors, we start by looking at literature and analyzing data to hypothesize a prototype (causal) theory that drives the design of an appropriate data gathering tool (survey, interviews, etc.). Repeated iterations in the model-prediction-data gathering-model cycle are required to produce a significant understanding (see Figure 1).

Using this (iterative) framework, we start by conducting a literature review to study the human and technological factors that play the most relevant role in this phenomenon. When classifying these influencing factors, we identify three main groups: **personal**,

environmental, and **technological**. Personal factors may include age [4], gender [4], sexual orientation [5], low self-esteem, depression [6], or social anxiety [7] (although the last ones can also be a consequence of CB). Environmental factors include parental supervision and communication [4], minority ethnicity, or immigrant background [8]. Technological factors may include the number of hours spent online [4], and how the technology is used and known [9].

Then, we designed and analyzed a survey targeting those features that are relevant from our starting model viewpoint. The survey was answered by 840 children studying in (public and private) Spanish schools, aged between 13 and 17 (Mean=14.56, SD=0.9), with 48.8% identified as male and 44% as female. Figure 2 compares risk factors identified in the literature review with the percentage of children who have experienced CB-related risk situations. Although some factors display a straightforward correlation with risk, being a CB victim is a complex phenomenon with many interacting *forces*. For example, one could argue that more supervision of online activity at home is associated with less likelihood of being a victim, but only to a certain degree, which is higher. Of course, the direction of this relationship is not clear from the data. Participants who have experienced CB may be therefore supervised a lot at home. This is why we need to be clear about the causal model we are working with.

We analyzed the survey using our theoretical framework in Figure 3, which shows a causal probabilistic graphical model (Bayesian network) prototype that helps us extract relevant information from the participants' responses. This causal graph, hypothesized from the existing literature, is questioned using the data to refine the survey design and the causal model itself. As we refine our theory through iterations, we will converge to understand what makes a minor vulnerable, predict their likelihood, and infer mechanisms to mitigate cyberbullying. This (Bayesian+causal) approach corrects two difficulties common to traditional statistical (or black-box Machine Learning) methods. Firstly, it foresees the right way to avoid confounders misinterpreting the data; secondly, (hierarchical) Bayesian methods regularize the effect of features on the predicted variable, thus smoothing out the effect of unbalanced data and extreme data outliers.

References

[1] The State of the World's Children 2017: Children in a Digital World. UNICEF (2017). https://www.unicef.org/media/48601/file

[2] Smahel, D., *et al*. EU Kids Online 2020: Survey Results from 19 Countries. London School of Economics and Political Science (2020). https://www.eukidsonline.ch/files/Eu-kids-online-2020-international-report.pdf

[3] Lobe, B., *et al*. How Children (10-18) Experienced Online Risks during the Covid-19 Lockdown: Spring 2020. Publications Office of the European Union (2021). https://doi.org/10.2760/066196

[4] Baldry, A., *et al.* Cyberbullying and cybervictimization versus parental supervision, monitoring and control of adolescents' online activities. *Children and Youth Services Review* 96, 302-307 (2019). https://doi.org/10.1016/j.childyouth.2018.11.058

[5] Llorent, V. J., *et al.* Bullying and cyberbullying in minorities: Are they more vulnerable than the majority group?. *Frontiers in psychology* 7 (2016). https://doi.org/10.3389/fpsyg.2016.01507

[6] Johnsson, L. S., *et al.* Online sexual abuse of adolescents by a perpetrator met online: A cross-sectional study. *Child and Adolescent Psychiatry and Mental Health* 13(32) (2019). https://doi.org/10.1186/s13034-019-0292-1.

[7] Pabian, S., *et al.* An Investigation of Short-Term Longitudinal Associations Between Social Anxiety and Victimization and Perpetration of Traditional Bullying and Cyberbullying. *Journal of Youth Adolescence* 45(2) (2016). https://doi.org/10.1007/s10964-015-0259-3

[8] Kim, S., *et al.* Cyberbullying and ICT use by immigrant youths: A serial multiple-mediator SEM analysis. *Children and Youth Services Review* 110 (2019). https://doi.org/10.1016/j.childyouth.2019.104621

[9] Kowalski, R., *et al.* A developmental approach to cyberbullying: Prevalence and protective factors. *Aggression and Violent Behavior* 45, 20-32 (2019). https://doi.org/10.1016/j.avb.2018.02.009

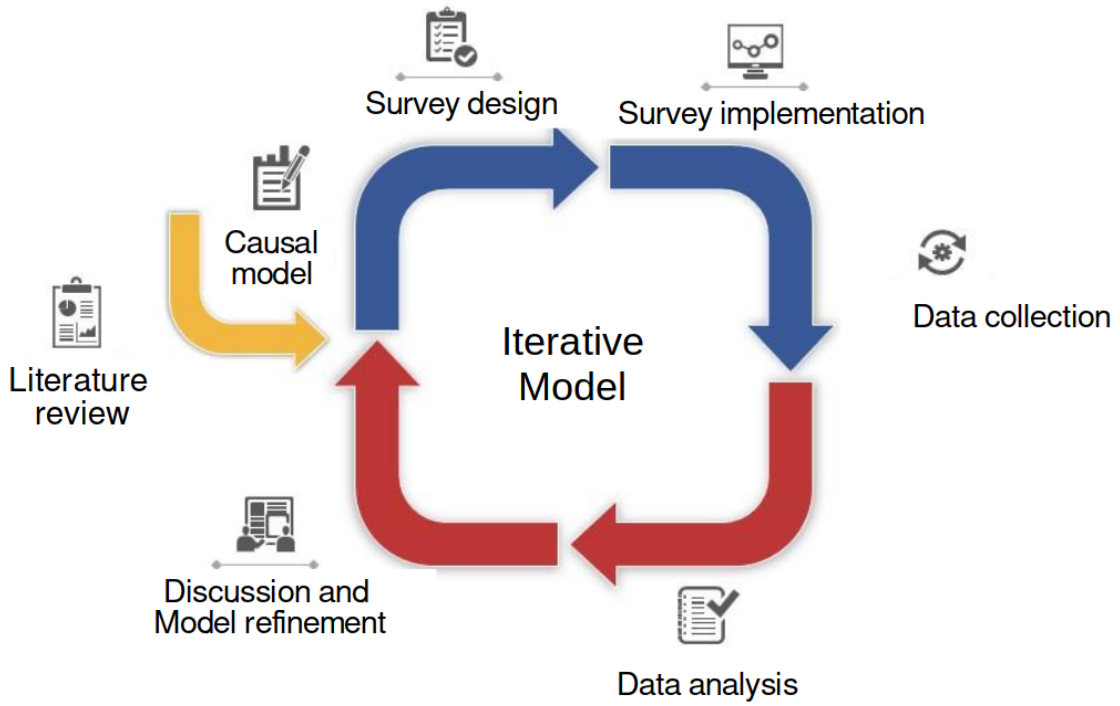

Figure 1. Diagram of the iterative process of acquisition of knowledge and understanding generation. Each iteration provides new information with which researchers can update their previous beliefs and design future experiments accordingly to obtain new data, and so on. We implemented the causal model using a Bayesian network. This has two benefits: (i) the conclusions are probabilistic (so we can define risk accounting for uncertainty); (ii) data is analyzed straightforwardly using the Bayesian net what improves the discussion and model refinement and avoids the use of hypothesis testing (using p-values) that promotes binary thinking about the factors.

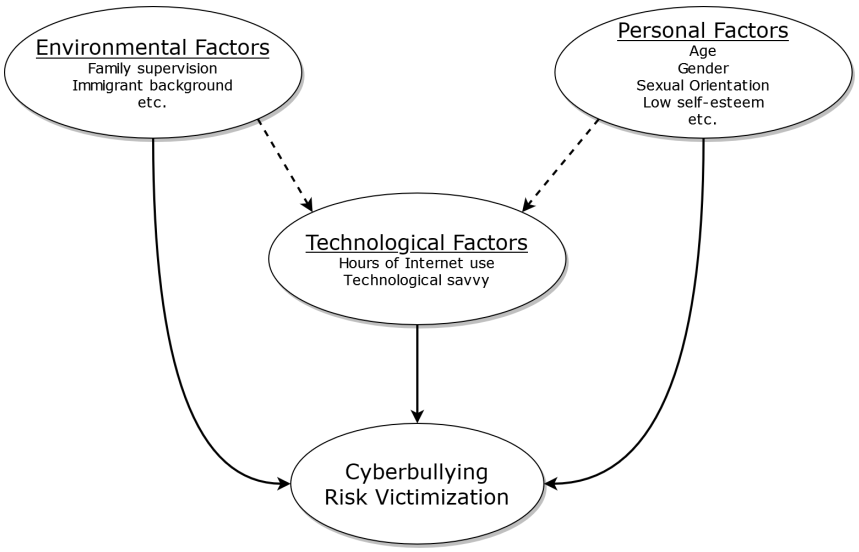

Figure 3. Graphical causal model of cyberbullying risk victimization from the potential personal, environmental, and technological risk factors analyzed in the literature review. The solid lines represent a causal relationship. The dashed lines represent a possible causal relationship (e.g., a hypothesis that we can test during the development of the research).

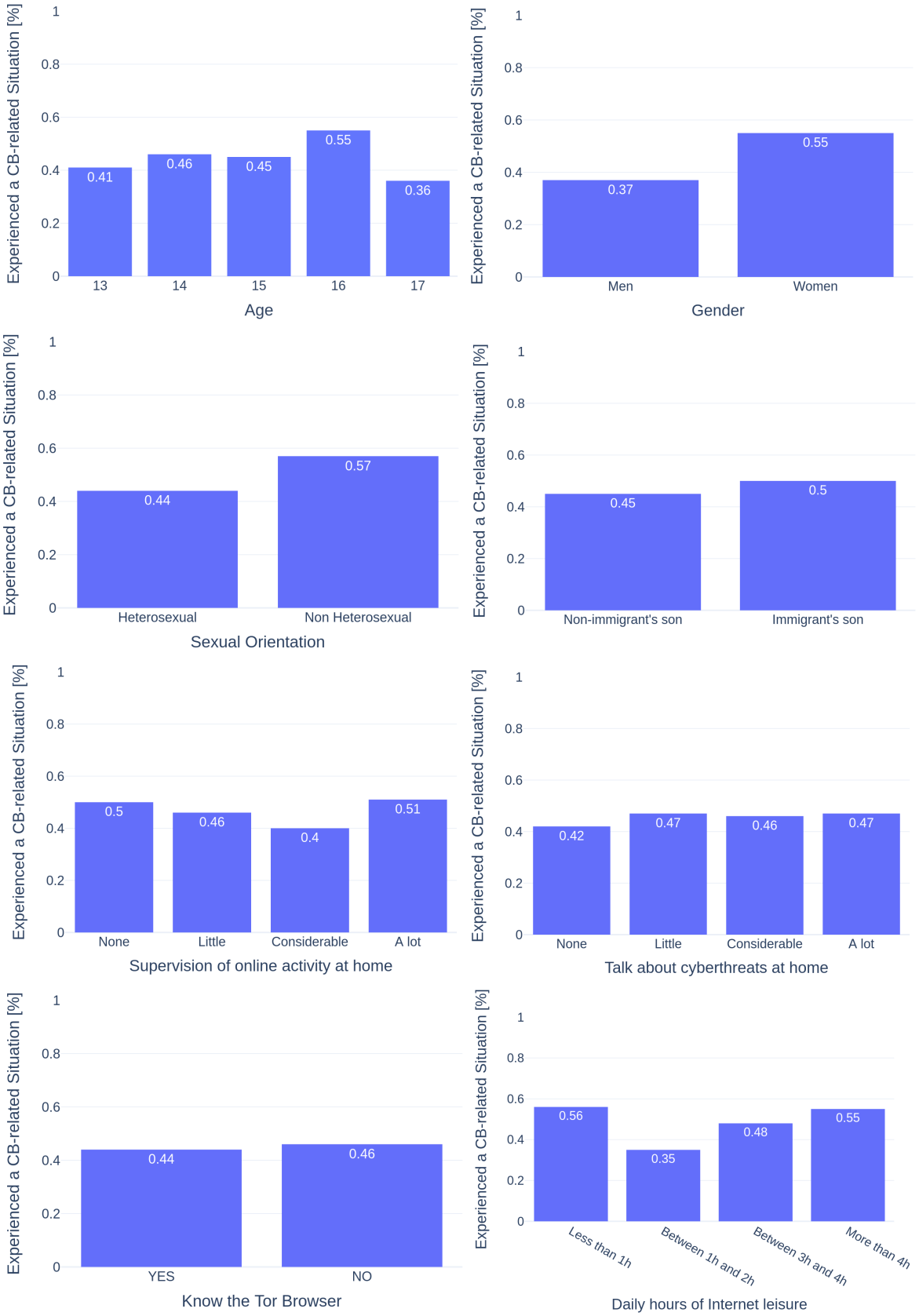

Figure 2. Statistics obtained from the responses to the survey conducted with minors in Spain. In all the sub-figures, the Y-axis represents the percentage of the population that has experienced some situation related to Cyberbullying (CB). The X axis of each sub-figure represents a potential personal, environmental, or technological risk factor.