What if we Defeated Cybercriminals with an AI-Generated Voodoo '*Curse Back*'? Considering a Socially Engineered Ethical Alternative to *Hacking Back*

Tim Pappa
Capitol Technology University, Laurel, Maryland, USA

tpappa@capitoltechu.edu

Abstract: Cybercriminals are influenced by beliefs in the supernatural and paranormal as much as we are. This practitioner's position paper will explore how artificially influencing a cybercriminal's belief that someone is cursing them models an ethical, socially engineered alternative for behaviorally disrupting cybercriminals. Rather than 'hacking back' cybercriminals, this paper asks how Nigerian cybercriminals using Voodoo to threaten victims and competitors would respond behaviorally if they believed they had been Voodoo cursed. Even Nigerian cybercriminals who do not believe in or practice Voodoo have demonstrated they are culturally and socially vulnerable to the suggestion they have been cursed, and they often threaten others with curses. This practitioner's paper visualizes an integrated framework of distrust and "sinister attribution error", explaining how people generally respond to imagined and IRL events based on their beliefs, but particularly how Nigerian cybercriminals who use real and imagined Voodoo cursing are vulnerable to these cognitive errors when trying to make sense of their unfavorable or troubling life circumstances and experiences. Visualizing this integrated framework would involve ingesting Nigerian cybercriminal communications, where there is use of Voodoo and other curses to fine-tune an LLM to generate artificial intelligence content mimicking what appear to be curses directed at Nigerian cybercriminals. However, this practitioner's paper will only attempt to position the behavioral foundations of this socially engineered ethical alternative to 'hacking back', for transgressive investigators with authorities and in some cases researchers or activists who believe there is a need to take a more aggressive approach to degrading Nigerian cybercriminal enterprises.

Keywords: Cybercrime, Voodoo, Sinister attribution error, Distrust, Transgressive ethics

1. Introduction

Nigerian cybercriminals have established a singular reputation around the world for enterprising socially engineered cyber campaigns, such as Business Email Compromise (BEC), for example. Media reporting and FBI cybercrime statistics have found that despite a marked increase in global law enforcement efforts to locally disrupt Nigerian cybercrime, last year Nigerian cybercrime activity rose more than 150 percent in just six months. In the past ten years, Nigerian cybercriminal rings have dominated fraud or scam markets, surpassing more than $50 billion in losses suffered by organizations and victims worldwide.

Nigerian cybercriminal research has increasingly focused on the evolution of these sub-culture cybercriminal communities and how responsive Nigerian cybercrime or "yahooboyism" has appeared to be to cultural and social environments in southern Nigeria (Adeniran, 2008; Lazarus, 2018; Nwobodo, 2024).

Nigerian cybercriminals, as another example of their reputation among other cybercriminals, have hosted training on Telegram and in Nigeria for overseas cybercriminals to study their social engineering methods. Nigerian "yahoo boys" more recently have been behind a growing number of sextortion attempts, especially targeting American teenagers, who increasingly have been taking their own life as a result. Many "yahoo boys" do not appear to fear the threat of law enforcement action (Uchenna, 2022).

Much of the cybercrime targeting victims inside and outside Nigeria is focused on people, not their networks. The victims of this kind of cybercrime are generally experiencing financial and psychological damage (Ojedokun and Eraye, 2012).

Melvin and Ayotunde (2011) and Aransiola and Asindemade (2011) both referred to the phenomenon of mostly Yoruba and Igbo Nigerian male and female "yahoo boys" or "yahoo yahoo" using real and imagined magic and spiritual powers to victimize people. This has been referred to as Nigerian "*yahoo plus*" or "yahoo boys" *plus* or even "Yahoo plus plus". This paper primarily focuses on this category of Nigerian cybercriminals.

Adejoh et al. (2019) interviewed law enforcement officials and "yahoo boys" about *yahoo plus*, finding that there was awareness of this approach to cybercrime with rituals including Voodoo to victimize people. Investigations have resulted in whoever is providing some of that ritual or spiritual support being charged as an accessory to the crime, in some cases. An example of their support is a cybercriminal taking necklaces or fabrics to an herbalist who prepares them with charms and the names of the intended victims. In another

example, a cybercriminal "yahoo boy" claimed he had been using it on victims but eventually stopped because of his own fear of repercussion. There are shared stories of yahoo plus cybercriminals being killed in accidents that many believed was a result of their use of Voodoo or charms to threaten other people, based on the author's experience supporting these investigations.

Many cybercriminal "yahoo boys" are Christian, but many Nigerians generally believe in supernatural power for success in whatever they are doing. This paper acknowledges that the use of ritual and charms and "juju" has similarities and differences to Voodoo but will refer to Voodoo hereafter to generally account for the range of uses of spiritual effects for criminal purposes online in Nigerian culture (Madukasi, 2023; Awodiran, Anwanna, and Idem, 2023; Akanle and Shadare, 2019; Whitty and Ng, 2017; Tade, 2013).

The author has also observed in past investigations cybercriminals paying for curses from Voodoo priests who then deliver or through an intermediary deliver or notify the target they have been cursed. There is a documented economy to this practice which may explain some of this behavior, but this paper will concentrate on exploring the behavioral vulnerabilities of Nigerian cybercriminal yahoo plus enterprises, especially if it is suggested or communicated that they have been cursed when presenting GenAI content that looks and feels like the communications and language they are familiar with.

Wang, Sun, and Zhu (2020) attempted to clarify some of the expansive and vague conceptual boundaries that characterize social engineering, especially in cyber contexts. Social engineering has historically been generally defined as trying to manipulate someone to obtain information with or without the use of technology (Breda, Barbosa, and Morais, 2017; Wang, Zhu, and Sun, 2021). However, Wang et al. considered more concentrated concepts such as social engineering that uses deception or social engineering that emphasize psychological exploitation. Their working definition in a cyber context described social engineering as a "type of attack wherein the attacker(s) exploit human vulnerabilities by means of social interaction…" Wang et al. continued that human vulnerabilities could present in cognitive and emotional environments and direct or indirect communication between another person or more people. Wang et al. referred generally to human vulnerabilities in emotion and feelings and in "human nature", which is foundational to this paper. This paper is exploring a richer behavioral context in some Nigerian communities where there is a heritage of Voodoo and cursing.

Even when considering a field where there is less demonization or stigmatization of Voodoo for example, such as astrology, we can still find some reflection of the general personality of most people and these cybercriminals. Bauer and Durant (1997) tested hypotheses that suggested people who are drawn to astrology generally have a religious orientation but no integration into organized religion, and that these people have personality traits that could be described as authoritarian. Bauer and Durant found support for that first hypothesis, suggesting that astrological interpretation attracts people looking for some meaning in their "metaphysical unrest". Generally, there appears to be a relationship between personal life crises and belief in astrology (Lillqvist and Lindeman, 1998). Andersson, Persson, and Kajonius (2022) found that the higher the narcissism in someone, the higher the belief in astrology. The suggestion is that there is a positive association because someone described or measured as narcissistic may be more likely to have a self-centered or egocentric worldview. These characterizations suggestively reflect some of the behaviors demonstrated by Nigerian cybercriminals referenced in this paper.

This practitioner's position paper will introduce some of the related work on "transgressive ethics", exploring how this approach to behaviorally exploiting Nigerian cybercriminals distrust can be normalized in practice. This paper will also discuss some of the development of how we understand distrust in groups and organizations and how "sinister attribution error" degrades decision making and grows distrust, especially among vigilant personalities we might approximate to be Nigerian cybercriminals. The paper will then visualize the integration of this framework as a model for transgressive practitioners.

2. Normalizing this GenAI Approach to Nigerian Cybercriminals in the Framework of "Transgressive Ethics" and Transgressive Investigators and Researchers

This practitioner's position paper recognizes that taking any kind of approach even to cybercriminals to behaviorally exploit them in some manner may be ethically questionable for most people.

As such, this paper seeks to normalize some movement in an ethical reference point in augmenting investigations and traditional investigative techniques against these cybercriminals. The growing literature on transgressions in ethical contexts provides some of that reference.

Evers (2017) developed a typology of norm transgressions that provide some groundwork for how the question of ethics in using generative artificial intelligence might be conceptualized when considering who or what that generated content is applied against and in what context. Evers categorized this typology generally with states in mind, considering how states interact or respond to each other internationally when there is a transgression of what was believed to be a consensual norm. Evers defined *adaptive* transgressions as contexts or situations where states try to challenge shared expectations of how those consenting to those norms or "insiders" should act. The transgression occurs when one of the "insiders" tries to change whatever might be considered the norm among other insiders in that context or that group of relationships. Any kind of transgression of a norm in a group needs to be resolved, even if other insiders in that group or culture even were not explicitly aware of that norm or that the transgression was a transgression until it occurred.

Evers recognized that norm transgressions involve a process of shared understanding that form that resocialization of a new norm, if there is going to be any kind of resocialization because of that transgression. The author interprets this understanding of resocialization of norms following a transgression as there is opportunity to resocialize a new norm in a group when there has first been a demonstrated transgression by an insider. An insider in this paper's context may be a law enforcement investigator.

Macintyre and Chaves (2017) explored the dynamic of the "transgressive researcher" in environmental education, recognizing that it is increasingly clear that raising awareness or institutionalizing nature rights is not enough to drive major changes in peoples' behaviors toward the environment. They referred to Kulundu who suggested rethinking the activist and researcher as a "reflexive justice practitioner" who can situate "moral action and ethics as an ongoing dynamic relationship" between the researcher and those being researched or experiencing some application of research against them. Macintyre and Chaves suggested that there is space for this kind of demonstration and reflection, but they found that there is an "inherent tension" in their research between what they characterized as the "aggressive" nature of transgressive learning and the need for recognizing other worldviews and realities. They are talking about the consequences of confronting the impact of inequality and environmental degradation in the world.

In this paper, the author is suggesting that perhaps there is a need for transgressive researchers who have the tools and understanding of human behaviors to apply those tools and behavioral approaches more aggressively, against the kind of cybercriminals who are willing to drive teenagers and adults to kill themselves.

Strydom (2013) looked at transgressions a little more broadly, considering normativity or general social norms in a culture in politics or the economy and relationships as well. There can be a positive dimension to transgressions, such as technical transgressions in innovation and positive social justice change that began with transgressions of social norms, such as racially based laws.

As another example, Hoeyer and Jensen (2012) examined how the professional work ethics in intensive care units provided greater understanding of the change of treatment norms regarding resuscitation of patients considered to be brain-dead during a heart attack of some sort. While the patient medically could not be brought back to 'life' from that state of being brain dead, trying to resuscitate a patient during a cardiac event was sometimes referred to as "aggressive organ harvesting". Medical practitioners generally understood the need for trying to preserve organs that could fail quickly without the heart still preserving those organs, but because the attempts at resuscitation increasingly bothered some medical practitioners, many workers in intensive care units asked for more explicit guidelines and policies for informing decisions to resuscitate and that process. Hoeyer and Jensen found that the Danish medical practitioners they observed and interviewed during an ethnographic study of this practice in Danish intensive care units appeared to welcome a policy that suggested an even more 'aggressive' approach to organ procurement even if they disliked the practice personally. Practitioners were experiencing a localized moral dilemma, that until understood within their situations and contexts was more difficult to normalize the policy and public understanding of that practice. They characterized this as a "practical work ethic".

The practitioners in these studies or the "insiders" would likely have to be who would resocialize that next reference point in ethical approaches to countering Nigerian cybercriminals.

3. Emphasizing the Vulnerability of Nigerian Cybercriminals' Distrust and "Sinister Attribution Error"

Kramer (1994) found that people who feel that they are being evaluated or scrutinized tend to overestimate the degree to which they believe they are the target of someone's attention. There has been some

exploration of the differences between rational trust and irrational trust, where as an example someone may have rationally experienced repeated breaks in trust in a relationship and as another example someone may have an exaggerated belief there is reason to distrust others even if there are no experiences to base that on.

Kramer noted that much of the research at that point had been focused on individuals rather than groups or organizations. Kramer wanted to systematically study how and why people "overattribute personalistic and malevolent motives and intentions" to others, especially in defined social contexts.

Kramer included past work and his own prior research in developing what he characterized at the time as the "notion" of *sinister attribution error*.

Kramer highlighted early work by Colby (1981) who defined "paranoid cognitions" as something similar, in terms of "persecutory delusions" that someone who targeting you or wishes to harm you. Much of this work framed these delusions as more pathological, rather than the subsequent research that characterized this kind of behavior as much more normal and every day. Someone's general attitudes and beliefs about other people can also influence how they might respond to heightened belief that someone is targeting them in some way. One of Kramer's studies found that newer people in an organization may struggle with more uncertainty about the way others are behaving toward them or what others think about them, whereas people who have been a member of an organization for longer have more information about their standing in the organization and their relationships with others.

Another study Kramer conducted involved participants sitting in a dark room in a cubicle with walls so they could not see the other participants. Participants were informed that they needed to keep a dot on the screen in the middle of the quadrant with a joystick. If participants observed the dot moving, other participants were trying to move the dot into the center of their own quadrant to score higher or make more money. Participants were not told that the dot was stationary, however, the dot appeared to be moving at times for participants because of uncontrolled eye movements. Participants were given breaks during the study, during which they were interviewed and later given a questionnaire regarding their opinion about the motivations and behaviors of other participants. This study was designed to amplify rumination about someone's motives and then measure how that rumination even in an artificial environment can cause distrust.

Kramer also found that ruminating about someone's intentions and motives in this kind of social context can increase that person's confidence in their attribution and judgment. Ruminating following a negative event can particularly increase negative thinking.

Kramer (1998, 1999) years later continued to emphasize how important the history of someone's interaction with others is in shaping their perceptions or the origins of distrust and suspicion. Kramer referred to Taylor and Fiske (1975) who explained that social context is so important because relationships and knowledge of people influences their attention and the importance that person places on those interactions and beliefs about them.

Especially hypervigilant people who need to ruminate about someone's motives and intentions for their own safety are generally more likely than others to generate "raw data" that they came up with to explain why someone is behaving in some manner or why events appear to be occurring. Because the "raw data" came from whoever is ruminating and vigilant in their observation of others, they are more likely to be confident in their own interpretation of behaviors and events. This is relevant in this paper, considering the focus of any 'curse back' is the member or leader of a cybercriminal organization using Voodoo to threaten others.

Kramer characterized hypervigilance as a form of "heightened search and appraisal" of threatening social information, which suggests someone's sensemaking can become nonadaptive to all kinds of information. Kramer discussed a study in the 1980s where clinicians working with senior populations found some correlation between gradual hearing loss and the onset of perceptions that could be described as paranoid. People who tended to suffer hearing loss also tended to express perceptions believed by many to be paranoid. The natural parallel in this paper is a criminal boss who loses confidence in some of his information collection and the people who provide information to him, so he is less sure of what he is hearing and what he is reading.

With this foundation, we can visualize how this framework may look when integrated with our understanding of social engineering and these behaviors related to distrust and "sinister attribution error".

4. Visualizing an Integrated Framework of social Engineering and "Sinister Attribution Error" to Model a GenAI Approach to 'Cursing Back' These Cybercriminals

To visualize this integrated framework, this paper will characterize a common scenario from the referenced research literature and the author's experience supporting investigations focused on Nigerian cybercriminals.

In this theoretical scenario, a rising Nigerian "*yahoo plus plus*" cybercriminal has started going to Voodoo priests to buy curses on some of his competitors. He made money quickly compared to some his peers because of his skills at manipulating victims online with his command of the English language and his use of several female friends he provides some cut of the profit to when they help him convince victims they are real. His business has struggled lately, however, perhaps because of heightened law enforcement attention and because of the increasingly crowded Nigerian cybercriminal market.

While he has never really practiced Voodoo, he is familiar with some of the practices and has heard many stories about how Voodoo got somebody killed and how a curse on a competitor helped restore profits. Because of his challenges lately, he decided to go to a Voodoo priest to buy curses on some of his competitors. The priest recently communicated to these targets that a curse had been placed on them.

In an investigation of this Nigerian "yahoo plus plus" cybercriminal, this would be a moment to apply this ethical socially engineered alternative approach to amplifying his distrust of others and his attribution to what or who is behind the next step of real and imagined occurrences in his life that he might believe is a result of competitors countering his curse or protecting themselves from his curse. A prototypical LLM that has been fine-tuned with a body of ingested Nigerian cybercriminal communications could be promoted to create artificially generated content in a dialect and manner he might expect that communicates he has also been cursed by named or unnamed competitors. In this scenario, an unknown individual using a misattributed cellular telephone could send him a text claiming that a victim's family recently placed an expensive curse on him.

These curses are notional, of course, but the influence of the curse is in this Nigerian "*yahoo plus plus*" misattributing any misfortune or harassment or negative behaviors from those around him to these curses. The GenAI content in this scenario is only textual but could include audio and video in particular situations. Rather than any demonstration or practice of preparing a charm or curse, this approach is only creating textual communications to suggest or notify this cybercriminal that a curse has been placed on them.

Some examples could include the following simple communications:

- **Ogun ma pa eh**
- Ogun (the Yoruba god of iron) will kill you

- **Iku ni mo pe ooooooooo**
- Death I call you ooooooooo

- **Dem swear for you? You go chop beat**
- Was a curse put on you? You are going to get beaten up

These are understandably aggressive and threatening textual communications, however, these are notional. There is no identifiable sender and there does not have to be any suggestion or attribution to a named competitor or named victim, because the recipient in this scenario would make his own attribution or guess. There does not need to be any continued engagement other than a single statement, but an LLM fine-tuned to generate these kinds of artificial communications and content could support continued communication.

5. Discussion

This practitioner's position paper proposed a socially engineered alternative to behaviorally disrupting cybercriminals, namely Nigerian "yahoo plus plus" cybercriminals familiar with the use of Voodoo to threaten. The author argued that this proposed integrated framework can be transgressively ethical, depending on the people like law enforcement investigators with the authorities to apply it who resocialize that reference point for how we investigative and counter Nigerian cybercrime.

This paper introduced some of the behavioral foundations of how everyone is vulnerable to distrust and paranoia, in fact, rather quickly. The research referenced in this paper has emphasized how people who are more likely to be hypervigilant and distrustful such as these Nigerian cybercriminals are arguably much more vulnerable than people may realize.

This approach with this proposed integrated framework using fine-tuned LLM artificially generated communications to prompt these statements resembling Voodoo curses and other charms could be effective, especially in an online environment where people are plausibly anonymous and tend to threaten each other. Whether the target of this 'cursing back' can prove he or she was cursed or not makes little difference. Kramer established in this paper that as the targets of what appears to be harassment or malice continue to ruminate on the possibilities or explanations of that targeting, those targets tend to come up with their own explanations or "raw data" for why something is happening and who might be responsible. This "sinister attribution error" can quickly drive someone who believes they are being targeted to increasingly distrust others and degrade or warp their sensemaking or interpretation of every interaction.

This proposed integrated framework models how practitioners can apply this approach using GenAI in a limited manner and limited to a predicated target of an investigation.

Ultimately, a goal of nearly every investigation of cybercriminals who are willing to hurt others and who do not fear law enforcement action is to degrade or disrupt their perceptions of trust and decision making, so that law enforcement can disrupt their criminal activity and mitigate that cybercriminal in a more ethical manner.

References

Adejoh, S.O., Alabi, A.A., Adisa, W.B. and Emezie, N.M., 2019. "Yahoo boys" phenomenon in Lagos metropolis: A qualitative investigation.

Adeniran, A.I., 2008. The Internet and Emergence of Yahooboys sub-Culture in Nigeria. *International Journal of Cyber Criminology*, *2*(2).

Akanle, O. and Shadare, B.R., 2019. Yahoo-plus in Ibadan: Meaning, Characterization and Strategies. *International Journal of Cyber Criminology*, *13*(2).

Andersson, I., Persson, J. and Kajonius, P., 2022. Even the stars think that I am superior: Personality, intelligence and belief in astrology. *Personality and Individual Differences*, *187*, p.111389.

Aransiola, J.O. and Asindemade, S.O., 2011. Understanding cybercrime perpetrators and the strategies they employ in Nigeria. *Cyberpsychology, Behavior, and Social Networking*, *14*(12), pp.759-763.

Bauer, M. and Durant, J., 1997. Belief in astrology: a social-psychological analysis. *Culture and Cosmos: A Journal of the History of Astrology and Cultural astronomy*, *1*(1), pp.55-71.

Breda, F., Barbosa, H. and Morais, T., 2017. Social engineering and cyber security. In *INTED2017 Proceedings* (pp. 4204-4211). IATED.

Chukwuka, O.U., 2022. INTERNET FRAUD: THE MENACE OF 'YAHOO BOYS'AND THE DECEITFULNESS OF RICHES. *Sapientia Global Journal of Arts, Humanities and Development Studies*, *5*(2).

Evers, M.M., 2017. On transgression. *International Studies Quarterly*, *61*(4), pp.786-794.

Hoeyer, K.L. and Jensen, A.M., 2013. Transgressive ethics: Professional work ethics as a perspective on 'aggressive organ harvesting'. *Social Studies of Science*, *43*(4), pp.598-618.

Kramer, R.M., 1994. The sinister attribution error: Paranoid cognition and collective distrust in organizations. *Motivation and emotion*, *18*, pp.199-230.

Kramer, R.M., 1998. Paranoid cognition in social systems: Thinking and acting in the shadow of doubt. *Personality and Social Psychology Review*, *2*(4), pp.251-275.

Kramer, R.M., 1999. Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual review of psychology*, *50*(1), pp.569-598.

Lazarus, S., 2018. Birds of a feather flock together: the Nigerian cyber fraudsters (Yahoo Boys) and hip hop artists. *Criminology, Crim. Just. L & Soc'y*, *19*, p.63.

Lillqvist, O. and Lindeman, M., 1998. Belief in astrology as a strategy for self-verification and coping with negative life-events. *European Psychologist*, *3*(3), pp.202-208.

Macintyre, T.K.J. and Chaves, M., 2017. Balancing the Warrior and Empathic Activist: The role of the'Transgressive'Researcher in Environmental Education. *Canadian Journal of Environmental Education (CJEE)*, *22*, pp.80-96.

Madukasi, F.C., 2023. Cybercrime and Ritualism: the Menace of â€ œYahoo Boysâ€ and the Demonization of Traditional Igbo Religion and Society. *GLOBAL ONLINE JOURNAL OF ACADEMIC RESEARCH (GOJAR)*, *2*(2).

Melvin, A.O. and Ayotunde, T., 2011. Spirituality in cybercrime (Yahoo Yahoo) activities among youths in South West Nigeria. In *Youth culture and net culture: Online social practices* (pp. 357-380). IGI Global.

Nwobodo, R.E., A Critical Examination of 'Yahoo-Yahoo'among Nigerian Youths as a Breach of African Culture and Values.

Ogundele, A.T., Awodiran, M.A., Idem, U.J. and Anwana, E.O., 2023, January. Cybercrime activities and the emergence of Yahoo Boys in Nigeria. In *2023 International Conference On Cyber Management And Engineering (CyMaEn)* (pp. 313-320). IEEE.

Ojedokun, U.A. and Eraye, M.C., 2012. Socioeconomic lifestyles of the yahoo-boys: A study of perceptions of university students in Nigeria. *International Journal of Cyber Criminology*, *6*(2), p.1001.

Strydom, P., 2013, April. Normativity and transgression: A brief conceptual analysis. In *paper presented*.

Tade, O., 2013. A spiritual dimension to cybercrime in Nigeria: The 'yahoo plus' phenomenon. *Human Affairs*, *23*(4), pp.689-705.

Wang, Z., Sun, L. and Zhu, H., 2020. Defining social engineering in cybersecurity. *IEEE Access*, *8*, pp.85094-85115.

Wang, Z., Zhu, H. and Sun, L., 2021. Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods. *Ieee Access*, *9*, pp.11895-11910.

Whitty, M.T. and Ng, M., 2017. Literature review for UNDERWARE: UNDERstanding West African culture to pRevent cybercrimEs. *Report for the National Cyber Security Centre as part of a group of studies funded in the Research Institute in Science of Cyber Security*.