International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(REVIEW ARTICLE)

Check for updates

Anatomy of a vishing call: A theoretically grounded review of mobile money fraud and AI voice cloning in Ghana

Afua Asantewaa Asante [1,*], Kwaku Gyamfi Boamah [1] and Ranjith Kumar Patil [2]

[1] *College of Computing, Cybersecurity Grand Valley State University, USA.*
[2] *College of Computing, Applied Computer Science Grand Valley State University, USA.*

International Journal of Science and Research Archive, 2026, 20(02), 030–041

Publication history: Received on 22 June 2026; revised on 29 July 2026; accepted on 31 July 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.20.2.1456

Abstract

Mobile money has become the financial backbone of Ghana, with transaction values reaching several trillion cedis annually and active accounts numbering in the tens of millions. This scale has attracted a parallel surge in fraud, much of it perpetrated through voice phishing (vishing): unsolicited telephone calls in which fraudsters impersonate telecommunications providers to manipulate subscribers into surrendering credentials or authorizing transfers. This paper reviews the state of vishing-enabled mobile money fraud in Ghana, synthesizing regulator data, industry reporting, documented case typologies, and the academic literature on social engineering and mobile money security. It organizes the dominant attack scripts - security-verification, loyalty-reward, misdirected-payment, and SIM-swap pretexts - into a common four-stage pattern of impersonation, pretext, manipulation, and extraction, and interprets this pattern through routine activity theory and persuasion theory. The paper then examines an escalating threat: the maturation of AI voice cloning, which can now replicate a target voice from seconds of audio, and which controlled studies show human listeners cannot reliably detect. Because Ghanaian vishing already succeeds using ordinary human voices, the paper argues that voice cloning constitutes a force multiplier that will erode voice familiarity, the last perceptual defense available to targets, and expand the pretext space from institutional to personal impersonation. The review contributes a consolidated threat taxonomy, a theoretically grounded attack model, and a research and policy agenda for providers, regulators, and awareness programs in Ghana and comparable mobile-money-dependent economies.

Keywords: Vishing; Voice Phishing; Social Engineering; Mobile Money Fraud; AI Voice Cloning; Deepfake Audio; Routine Activity Theory; Ghana; Financial Inclusion; Cybersecurity

1. Introduction

Mobile money has transformed financial access in sub-Saharan Africa, and nowhere more visibly than in Ghana. Introduced by MTN in 2009, mobile money (locally, "MoMo") now underpins a predominantly informal economy in which a large majority of economic activity occurs outside the formal banking sector [6], [7]. According to the Bank of Ghana's payment-systems oversight reporting, the total value of mobile money transactions reached roughly 3,010 billion cedis in 2024, up almost 57% from the prior year, against a base of tens of millions of active accounts [8]. This concentration of accessible, liquid value on ordinary handsets has, however, produced a parallel and rapidly growing target surface for fraud.

Among the many fraud modalities affecting mobile money, voice phishing - vishing - is distinctive and consequential. In a vishing attack, a fraudster telephones a subscriber, claims to represent a trusted institution (very often the telecommunications provider itself), and through a scripted interaction manipulates the victim into revealing a PIN or

* Corresponding author: Afua Asantewaa Asante.

one-time password, or into following prompts that move funds out of the wallet. Ghanaian regulators and industry bodies have documented sharp increases in such fraud: the Cyber Security Authority reported cyber-fraud losses arising from GH¢2.4 million in the first quarter of 2024 to GH¢14.94 million in the first half of 2025 [26], and the Ghana Chamber of Telecommunications reported that cyber fraud cost citizens GH¢4.4 million in the first quarter of 2025 alone, double the figure of the previous year [27]. These figures almost certainly understate the true burden, because a substantial share of victims never report [20], [26].

Simultaneously, technological development threatens to make vishing markedly more dangerous. Advances in generative artificial intelligence now permit the cloning of a human voice from only seconds of reference audio [3], and controlled human-subjects experiments consistently find that listeners cannot reliably distinguish high-quality synthetic speech from genuine speech, even after training [4], [5]. If the credibility of a vishing call has rested partly on the assumption that a familiar voice is authentic, that assumption is now unsafe. The convergence of a mature mobile money fraud ecosystem with accessible voice-cloning capability defines the problem this paper examines.

This paper is a structured review and threat analysis rather than an empirical study. Its objectives are fourfold: (1) to synthesize the current evidence on vishing-enabled mobile money fraud in Ghana from regulator, industry, and academic sources; (2) to organize the dominant attack scripts into a coherent, theoretically grounded model; (3) to analyze how AI voice cloning is likely to reshape this threat; and (4) to derive an agenda for research, practice, and policy. In doing so, the paper emphasizes throughout that where telecommunications providers such as MTN are named, they appear as impersonated brands and, in some documented instances, as institutions grappling with insider risk - not as wrongdoers; the abuse of a provider's brand is itself a central mechanism of fraud.

The paper makes four contributions. First, it consolidates fragmented regulator, industry, and academic evidence into a single account of the Ghanaian vishing threat, quantified where possible. Second, it advances a taxonomy of vishing pretexts (Fig. 4, Table I) and shows that these surface-level scripts reduce to one invariant four-stage structure (Fig. 5), a reframing with direct consequences for how awareness campaigns should be designed. Third, it interprets the phenomenon through routine activity and persuasion theory (Fig. 6), locating the vulnerability in a social protocol rather than a technical flaw and thereby explaining its resistance to technical-only remedies. Fourth, it develops a cue-erosion analysis of AI voice cloning (Fig. 8) and a layered, cloning-robust defense model (Fig. 9), and translates these into a concrete research and policy agenda. The remainder of the paper proceeds as follows: Section 2 describes the review method; Section 3 provides background on mobile money, social engineering, and the global vishing landscape; Section 4 characterizes the Ghanaian threat; Section 5 develops the theoretical attack model; Section 6 analyzes the AI voice cloning escalation; Sections 7 and 8 present implications and a research agenda; and Sections 9 and 10 discuss limitations and conclude.

2. Method of review

This review followed a narrative synthesis approach appropriate to an emerging, cross-disciplinary problem that spans criminology, information security, and development studies. Sources were assembled from four streams: (a) peer-reviewed academic literature on social engineering, mobile money fraud, and audio-deepfake perception; (b) official and regulatory reporting from Ghanaian institutions including the Bank of Ghana, the Cyber Security Authority, the National Communications Authority, and the Ghana Chamber of Telecommunications; (c) industry threat intelligence on vishing and voice cloning from established security vendors; and (d) documented, publicly reported fraud cases and provider advisories illustrating attack typologies. Sources were included where they bore directly on vishing, mobile money fraud, or voice-cloning capability and were dated primarily within the 2019–2026 window to reflect the current threat landscape, with foundational theoretical works included regardless of date. Evidence was then thematically organized into the threat taxonomy of Section 4 and the analytic model of Section 5.

3. Background

3.1. Mobile money and financial inclusion in Ghana

Mobile money allows subscribers to store and transfer value using a SIM-linked wallet without a conventional bank account, and it has become the dominant retail payment rail in Ghana [6], [7]. Its reach is a direct consequence of the country's economic structure: with most of the activity in the informal sector, mobile money supplies payment, savings, and remittance functionality to populations that formal banking has not served. The same features that drive adoption - simplicity, immediacy, and the collapse of financial action into a few handset prompts - also shape its risk profile.

Because a wallet can be emptied in the time it takes to complete a USSD sequence, the window between deception and irreversible loss is extremely short.

Three structural features of the Ghanaian mobile money system are particularly relevant to vishing risk. First, the SIM is the root of identity: because the wallet is bound to the SIM, control of the SIM tends to imply control of the wallet, which is why SIM-swap attacks are so damaging. Second, the system depends on a vast, distributed network of human agents for cash-in, cash-out, and registration; this network is essential to reach but introduces variability in verification quality and a documented insider-risk surface [7], [19]. Third, a national identity-linking exercise tied SIMs to a central identity credential, which improved know-your-customer coverage but also concentrated identity data whose leakage can supply the very details that make impersonation calls persuasive [18]. Each of these features is a legitimate design choice serving inclusion [31]; each also creates an exposure that vishing exploits.

3.2. Social engineering and the psychology of compliance

Social engineering is the manipulation of people into performing actions or divulging information that compromises security [11], [13]. Its effectiveness rests on well-documented principles of influence - authority, urgency, scarcity, reciprocity, liking, and commitment - that induce heuristic, low-scrutiny compliance rather than deliberate verification [10]. Voice is an especially potent channel for these levers: a call is synchronous, demands immediate response, denies the target time to reflect or verify, and conveys paralysis of confidence and authority [12]. Empirical studies of pretexting confirm that interactive, human-to-human deception is among the most successful attack forms [12], [13], and the harms of the resulting victimization extend well beyond money to shame, self-blame, and eroded trust [15], [28].

3.3. Vishing in the global threat landscape

Industry telemetry documents a pronounced global shift toward the voice channel. CrowdStrike reported a 442% increase in vishing activity between the first and second halves of 2024 [1], and Mandiant ranked voice phishing among the leading initial access vectors in its 2025 incident-response data [2]. Regulators have responded: the U.S. Federal Bureau of Investigation issued a public warning about AI-generated voice messages impersonating senior officials [14]. Hybrid “telephone-oriented attack delivery” schemes, which pair a message-based lure with a follow-up call, further illustrate how the voice channel is being industrialized. These global patterns provide the backdrop against which Ghana's mobile-money-specific vishing must be understood.

It is worth distinguishing the Ghanaian setting from the enterprise-focused vishing that dominates global reporting. In the enterprise case, the target is typically an employee whose compromise yields access to corporate systems, and the ultimate payoff for the attacker is mediated by further intrusion. In the mobile money case, by contrast, the target is an ordinary consumer and the payoff is immediate: the same call that achieves compliance also completes the theft, because the victim's wallet is the asset. This collapses the attack timeline and removes the post-compromise window in which enterprise defenders often detect and contain intrusions. It also changes the victim profile, shifting risk onto individuals - including those with limited digital literacy or recourse - rather than institutions with security teams. These differences make consumer mobile money vishing a distinct problem deserving dedicated study rather than a simple transplant of enterprise findings.

3.4. Positioning of this review

Against this literature, the present review occupies a specific gap. Existing Ghanaian scholarship on mobile money fraud is valuable but tends to adopt a stakeholder or institutional lens, examining controls, agent behavior, and regulatory gaps [7], [8], while regional surveys aggregate across countries and attack types [30]. The audio-deepfake literature, meanwhile, is technically deep but largely disconnected from the mobile money context in which cloned voices are most likely to cause consumer harm in Africa [4], [5], [32]. Few works bring these strands together to ask how the specific mechanics of Ghanaian mobile money vishing will interact with maturing voice-cloning capability. This review contributes that synthesis, and in doing so reframes the problem: not as two separate literatures - one on African mobile money fraud, one on deepfake audio - but as a single, converging threat that demands a combined analytic and policy response.

4. The Ghanaian vishing threat landscape

Synthesizing regulatory data, industry reporting, and documented cases yields a consistent picture of how vishing operates against Ghanaian mobile money users. This section first characterizes the scale of the problem, then organizes the observed attack scripts into a taxonomy.

4.1. Scale and trajectory

Multiple independent Ghanaian sources indicate rapid growth. The Cyber Security Authority recorded cyber-fraud losses rising more than sixfold between early 2024 and the first half of 2025, alongside a near-doubling of reported cyber incidents [26]. The Ghana Chamber of Telecommunications independently reported a doubling of citizen losses to cyber fraud year-on-year in early 2025 [27]. Bank of Ghana financial-stability reporting has attributed a substantial share of financial-sector fraud cases to mobile money platforms [29]. A regional academic survey estimated social-engineering fraud losses across African mobile money markets in the billions of dollars, situating Ghana within a continental pattern [30]. Two cross-cutting features recur in this report: pervasive underreporting, which suppresses official totals [20], [26], and a significant insider-threat dimension, with observers noting that privileged access to subscriber data can enable unusually well-targeted approaches [19], [29]. The insider dimension matters for vishing specifically because it can supply the personal details - names, recent transactions, registration data - that make an impersonation call convincing [19], [21].

This growth is not incidental but structural. As Fig. 1 shows, the value flowing through mobile money has expanded steeply, and because fraud payouts scale with the value held in wallets, a larger mobile money economy is mechanically a larger target. The loss trajectory in Fig. 2 tracks this expansion. Compounding the problem is a weak enforcement position: of the thousands of SIM-swap-related fraud cases reported to Ghanaian cybercrime authorities in a recent year, only a small fraction resulted in arrests and, on available reporting, essentially none in convictions (Fig. 3). This enforcement gap has two effects relevant to vishing. First, it lowers the expected cost to offenders, sustaining offender motivation in routine-activity terms. Second, combined with slow or unsympathetic complaint handling, it discourages reporting, which in turn starves regulators and researchers of the data needed to size and target interventions — a self-reinforcing cycle of invisibility.

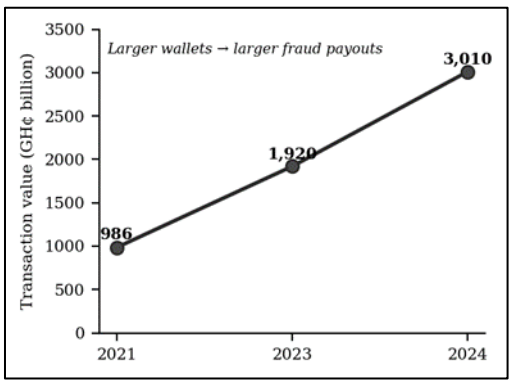

Figure 1 Growth in mobile money transaction value in Ghana, illustrating the expanding pool of liquid value available to fraudsters (figures in GH¢ billion; sources: Bank of Ghana and sector reporting).

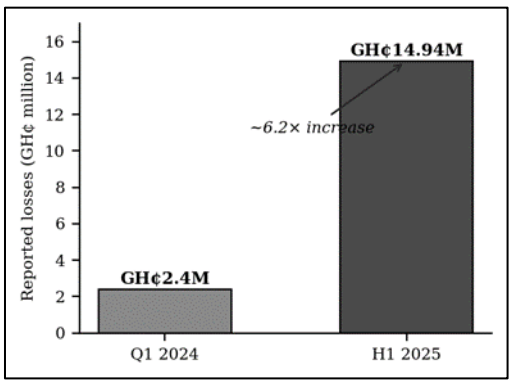

Figure 2 Reported cyber-fraud losses in Ghana rose roughly sixfold between Q1 2024 and the first half of 2025 (source: Cyber Security Authority, as reported in Ghanaian press).

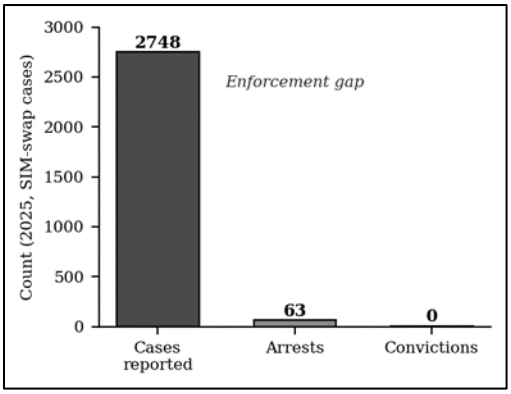

Figure 3 Illustrative enforcement funnel for SIM-swap-related mobile money fraud in Ghana: many reported cases, few arrests, and negligible convictions (source: Ghanaian cybercrime reporting).

4.2. A taxonomy of vishing pretexts

Across the documented Ghanaian evidence, vishing calls cluster into a small number of recurring pretexts. Four dominate. The **security-verification pretext** frames the call as a routine account or SIM matter - for example, an urgent re-registration or a security check - and walks the subscriber through prompts that culminate in PIN or OTP disclosure; providers stress that legitimate staff never request the PIN [21], [25]. The **loyalty-reward pretext** frames the call as good news, informing the subscriber that they have won a promotion or qualified for a reward, and uses the resulting positive anticipation to secure compliance with prompts that instead extract funds [21]. The **misdirected-payment (reversal) pretext** tells the subscriber that money was mistakenly sent to their wallet and asks them to "return" it, weaponizing the target's honesty so that the reversal steps move the victim's own funds to the attacker [21]. Finally, the **SIM-swap pretext** either socially engineers the subscriber or a provider agent so that the attacker obtains a duplicate SIM, after which the mobile money PIN can be reset and the wallet emptied a modality that Ghanaian reporting identifies as among the fastest growing and that is frequently enabled by weak verification at some retail points and by leaked identity data [18], [20]. Adjacent modalities documented in the same sources include fake-delivery "customs fee" scams and cheap-data lures that harvest numbers for follow-on phishing [21].

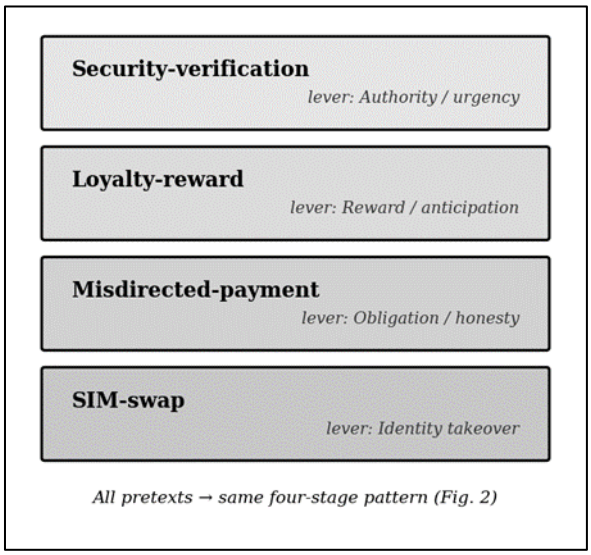

Figure 4 Taxonomy of documented vishing pretexts targeting Ghanaian mobile money users, grouped by the primary compliance lever each exploit.

Table I summarizes these pretexts along the dimensions most relevant to defense: the emotional or cognitive lever each exploit, the specific action the victim is induced to take, and the natural point of intervention. Reading the table by column rather than by row is instructive: the intervention points differ, but every row terminates in the same failure - an authorized-looking action performed by the victim which is why the downstream defense (out-of-band verification) is common to all of them.

Table 1 Vishing pretexts by lever, induced action, and intervention point.

Pretext	Lever	Induced action	Intervention point
Security-verification	Authority, urgency	Disclose PIN/OTP or run prompts	Rule: staff never ask for PIN
Loyalty-reward	Reward, anticipation	Run prompts to 'claim' reward	Skepticism toward unsolicited wins
Misdirected payment	Obligation, honesty	'Return' funds via prompts	Verify with provider before acting
SIM-swap	Identity takeover	Enable duplicate SIM issuance	Biometric lock on SIM replacement

5. A theoretically grounded attack model

5.1. The four-stage pattern

Although the pretexts differ at the surface, the documented cases share a common structure that can be represented as four stages. In the *impersonation stage*, an unsolicited caller claims affiliation with a trusted provider, borrowing its institutional credibility. In the *pretext stage*, the caller supplies a plausible, provider-consistent reason for the interaction a security check, a reward, a reversal. In the *manipulation stage*, the subscriber is guided through interactive prompts on their own handset, an experience that feels procedural precisely because legitimate providers use similar prompts. In the *extraction stage*, funds leave the wallet, typically discovered only afterward. The pretext is thus an interchangeable module within a fixed structure, which is why awareness messaging that targets a single script (e.g., “beware reward scams”) is easily evaded by substituting another.

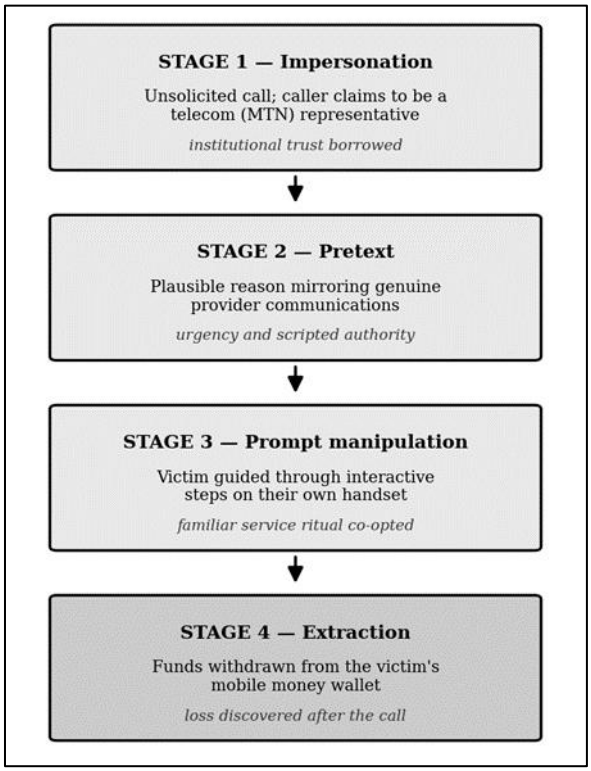

Figure 5 The four-stage vishing attack pattern, with the pretext stage as an interchangeable module.

5.2. Routine activity and persuasion perspectives

Routine activity theory explains crime as the convergence of a motivated offender, a suitable target, and the absence of a capable guardian [23], and has been extended productively to cybercrime [24]. Applied here, the offender reaches the

target directly over the voice channel; target suitability arises less from wealth than from routine the ordinariness of transacting by handset and of receiving provider calls; and guardianship is weak at the decisive moment because the subscriber is alone with the caller, inside a familiar interaction ritual, with no institutional intermediary present. The wallet's technical safeguards remain intact but are circumvented through the victim's own hands. Persuasion theory supplies the complementary mechanism: authority and urgency [10] operate not as exotic manipulations but as amplifications of an interaction script the subscriber already trusts. Together the two lenses locate vulnerability not in technology but in a social protocol, which has direct implications for defense (Section 7).

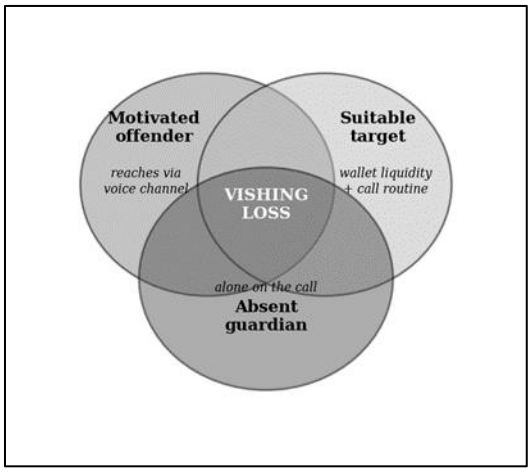

Figure 6 The routine activity convergence for mobile money vishing: a motivated offender, a routine-suitable target, and an absent guardian meet on the call itself.

This reading also clarifies why the vishing problem is resistant to purely technical fixes. Each of the three elements in Fig. 6 is sustained by structural conditions identified in Section 4: offender motivation is reinforced by weak enforcement and low conviction rates; target suitability is a byproduct of the very financial-inclusion success that makes mobile money valuable; and guardianship is absent precisely because the interaction is designed to be private and self-service. A defense strategy must therefore introduce a guardian into the moment of the call not a person, but a rule the subscriber carries with them. This is the logic behind the procedural defenses developed in Section 7, and it is also why the arrival of voice cloning, examined next, is so consequential: cloning attacks the subscriber's last internal guardian, the ability to recognize who is really speaking.

6. The escalating threat of ai voice cloning

6.1. Capability and accessibility

Neural speech synthesis can now clone a target voice from seconds of reference audio, and cloning tools are widely available commercially and as open source, with several found to lack meaningful safeguards against misuse [3], [16]. Surveys of deepfake creation and detection document rapid capability growth [17], [18], and industry measurement corroborates the trajectory in the fraud domain specifically: Pindrop recorded a more-than-tenfold rise in deepfake fraud attempts during 2024 [3], and security researchers have demonstrated real-time voice cloning within live vishing engagements [19]. The reference audio required is increasingly trivial to obtain, given the ubiquity of voice notes, social-media video, and recorded calls.

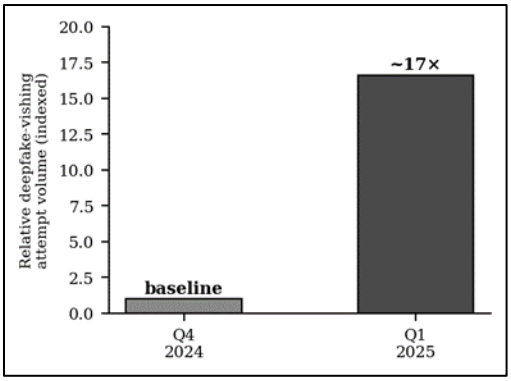

Figure 7 Indexed growth in deepfake-enabled voice-fraud attempts reported by industry sources, Q4 2024 to Q1 2025 (illustrative; source: vendor threat reporting).

Two properties of this capability are especially significant for the Ghanaian context. The first is cost asymmetry: producing a convincing clone is now cheap and fast, while detecting one reliably is not automated detectors remain imperfect and, critically, human listeners perform close to chance [4], [5]. The second is language and accent coverage: as synthesis models broaden their coverage of languages and local accents, the protective friction that an unnatural-sounding voice once provided diminishes, including for the Ghanaian languages in which many provider interactions occur. Together these properties mean that the technical barrier to cloned-voice vishing is falling while the mobile money target surface (Fig. 1) is expanding, and while automated detection research advances, it has not yet produced deployable, robust defenses for the live-call setting [32].

6.2. Why cloning is a force multiplier in Ghana

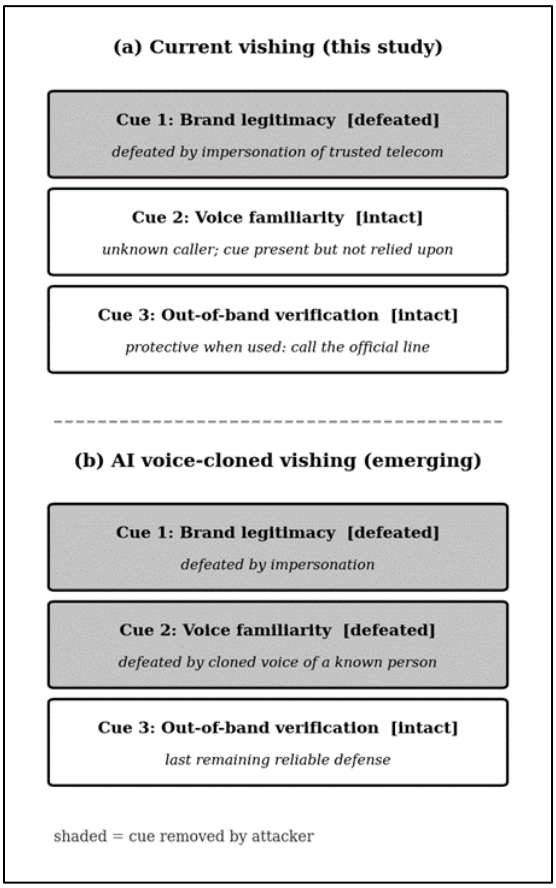

Figure 8 Detection cues available to targets in current vishing (a) versus AI voice-cloned vishing (b). Cloning removes voice familiarity, leaving out-of-band verification as the last reliable defense.

The central analysis of this review follows from combining two bodies of evidence. First, Ghanaian vishing already succeeds using ordinary, unfamiliar human voices: the documented attacks rely on institutional impersonation, not voice imitation [19], [21]. Second, controlled experiments show that human listeners cannot reliably detect synthetic speech, even after exposure to examples [4], [5]. It follows that voice cloning does not create the Ghanaian vishing problem it amplifies it along a specific axis. Where the current attack borrows the credibility of a provider brand, a cloned voice additionally borrows the credibility of a known individual, expanding the pretext space from institutional impersonation ("I am from MTN") to personal impersonation ("it's your brother; send the money"). This removes the one perceptual cue that a wary target might still rely on recognition of a familiar voice - as illustrated in Fig. 8.

6.3. A trajectory, not an event

It is useful to frame the threat as a trajectory rather than a single future event (Fig. 10). At present, Ghanaian vishing operates almost entirely with human voices and institutional impersonation, and it already succeeds. In the emerging phase, cloned voices begin to appear in fraud, adding personal impersonation to the attacker's repertoire and starting to erode the voice-familiarity cue. If the threat is left unaddressed, cloned-voice fraud could become routine, at which point perceptual defenses fail wholesale and only procedural defenses those that never relied on the ear remain effective. The practical value of this framing is that the defensive investments with the longest lead times, particularly public education in out-of-band verification and provider process changes, are most cheaply made now, before the emerging phase matures. Framing the problem as a trajectory also guards against two errors: complacency, which treats cloning as science fiction, and fatalism, which treats it as an unstoppable catastrophe. The evidence supports neither; it supports preparation.

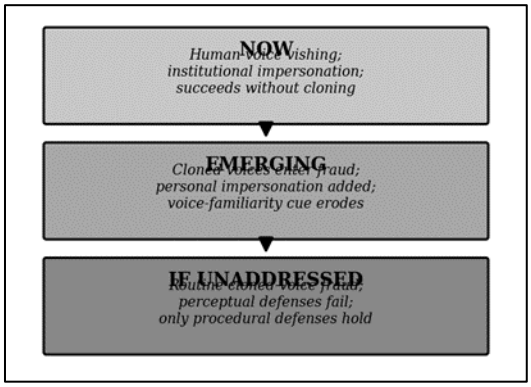

Figure 10 The escalation trajectory of mobile money vishing. Defensive lead times argue for acting during the current phase, before cloned-voice fraud becomes routine.

7. Implications and recommendations

The analysis implies a defense-in-depth posture in which responsibility is distributed across four layers of user, provider, regulator, and ecosystem rather than concentrated at any single point (Fig. 9). A useful design principle orders these layers by their robustness to voice cloning: measures that depend on a human correctly judging a voice are the most fragile, while measures that replace perceptual judgment with a verification procedure are the most durable. The recommendations below are organized accordingly.

Because vulnerability is a social protocol rather than a technical flaw, the most robust defenses are procedural and do not depend on the ear's ability to authenticate a voice - an ability the evidence shows to be unreliable [4], [5]. For **providers and mobile money operators**, the review supports prominent, repeated messaging of a single categorical rule - that legitimate representatives never ask customers to disclose a PIN or OTP or to act on wallet prompts during an unsolicited call - delivered through channels that reach rural and low-literacy populations, building on existing campaigns [21], [25]; in-app or on-prompt warnings at the moment of high-risk actions; strengthened biometric verification at every SIM-replacement point to close the SIM-swap vector [18], [20]; rapid transaction-freeze and low-friction reporting; and sustained attention to insider risk, given its documented role in enabling targeted approaches [19], [29]. For **regulators** (the Bank of Ghana, Cyber Security Authority, and National Communications Authority), the review supports coordinated cross-provider anti-fraud messaging, routine publication of standardized fraud statistics to counter underreporting and enable research, and oversight keeping pace with fintech growth [26], [29]. For **users and awareness programs**, the single most protective behavior is out-of-band verification: ending the call and dialing the provider's official published line, or confirming a personal request through a second channel, before acting. This last

recommendation is the one that survives the arrival of voice cloning, because it assumes the voice itself cannot be trusted as identity.

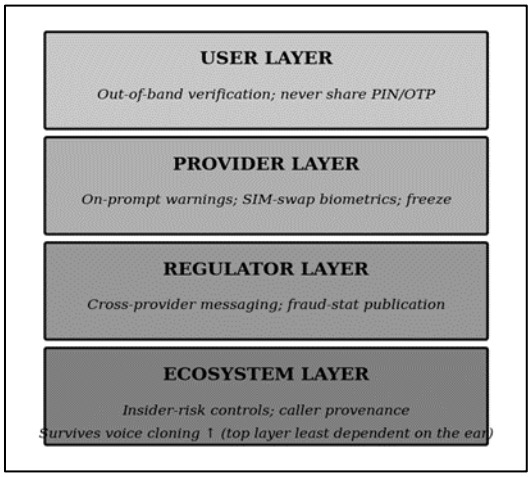

Figure 9 A layered defense model. Higher layers depend less on perceptual voice judgment and therefore remain effective as AI voice cloning matures.

A further implication concerns the design of awareness messaging itself. Because the pretext is an interchangeable module (Fig. 5), campaigns built around specific scam stories "beware the reward scam," "beware the wrong-transfer scam" are structurally fragile: attackers simply rotate to an unlisted pretext. The taxonomy in Fig. 4 suggests a more durable approach that teaches the invariant structure rather than the surface stories, training subscribers to recognize the shared shape of the attack an unsolicited call, a request to act on prompts, a demand for a credential regardless of the narrative wrapper. This structural framing is also more compatible with the personal-impersonation pretexts that voice cloning will enable, since it does not rely on the caller claiming to be from a provider at all.

8. Research agenda

The synthesis points to five priorities. **(1) Victim-centered empirical work.** Qualitative and survey studies of Ghanaian vishing victims are needed to establish prevalence, loss distributions, and the determinants of underreporting, complementing the stakeholder-level focus of existing studies [7], [8]. **(2) Controlled susceptibility experiments.** Listening studies with Ghanaian participants should measure detection accuracy and confidence for AI-cloned versus genuine voices in mobile money scenarios, using consenting voice donors and full ethics approval, to convert the cue-erosion model of Fig. 8 from proposition to measurement. **(3) Intervention evaluation.** Field experiments should test whether procedural-guardianship messaging - especially the out-of-band verification rule - measurably reduces compliance with simulated, ethically administered vishing prompts. **(4) Technical countermeasures in context.** Research with providers should assess caller-provenance signals, post-call wallet-prompt cooldowns, and anomaly detection on prompt-proximate transfers under Ghanaian operating constraints. **(5) Cross-national comparison.** Comparative study across mobile-money-intensive markets such as Kenya, Nigeria, Tanzania, and Uganda [30] would test the transferability of the four-stage model and of the defenses proposed here.

9. Limitations of this review

As with any review, the scope of this study should be understood explicitly, though each boundary is one the contribution can accommodate. First, this is a narrative rather than a systematic review; it does not follow a pre-registered search protocol, and a subsequent systematic review with formal screening criteria would usefully complement it. This choice is appropriate to the aim here, which is to synthesize a fragmented, fast-moving, cross-disciplinary evidence base into a coherent analytic framework rather than to exhaustively enumerate a settled literature.

Second, a substantial portion of the quantitative evidence on Ghanaian mobile money fraud originates in regulator statements and industry or press reporting rather than peer-reviewed research. These figures are treated here as indicative of scale and direction rather than as precise measurements, and the persistent underreporting of victimization means that all reported totals should be read as lower bounds. This does not weaken the paper's central

argument, which rests on the qualitative structure of the attacks and on well-established experimental findings about human perception of synthetic speech, not on the exact magnitude of any single statistic.

Third, direct evidence of AI voice cloning operating in Ghanaian mobile money fraud specifically remains limited. The force-multiplier argument advanced here is therefore a theoretically and empirically grounded projection rather than a report of an already-widespread phenomenon. This is by design: the paper's purpose is to anticipate a convergence before it fully arrives, precisely so that defensive investment can precede rather than follow mass victimization. The research agenda in the preceding section is structured to test this projection.

Finally, the four-stage attack model and the pretext taxonomy are deliberate abstractions. They are intended as organizing frameworks that make the threat legible and defensible, not as exhaustive catalogues of every attack variant; individual cases will contain idiosyncrasies the model does not capture. The value of the abstraction lies in what reveals the shared structure beneath diverse surface scripts and in the defensive strategy that structure implies.

10. Conclusion

Vishing-enabled mobile money fraud is a large and growing threat in Ghana, driven not by technical compromise but by the manipulation of the trust relationship between subscribers and their providers. This review synthesized regulator, industry, and academic evidence to show that a small set of interchangeable pretexts - security verification, loyalty rewards, misdirected payments, and SIM swaps - all resolve into a common four-stage attack pattern best understood through routine activity and persuasion theory. Against this backdrop, the maturation of AI voice cloning represents a genuine escalation: because Ghanaian vishing already succeeds with ordinary voices, and because humans cannot reliably detect synthetic speech, cloning will function as a force multiplier that erodes voice familiarity and extends impersonation from institutions to individuals. The appropriate response is to shift defenses from perceptual judgment to procedure - above all, out-of-band verification - and to pursue the empirical, experimental, and policy agenda set out here before cloned-voice fraud becomes routine.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

Funding

This research received no external funding.

Note on named providers

References to MTN and other providers describe brand impersonation by third parties and documented sector-wide challenges; they do not allege wrongdoing by the providers named.

References

[1] CrowdStrike, *2025 Global Threat Report*, CrowdStrike, Inc., 2025. [Online]. Available: https://www.crowdstrike.com

[2] Mandiant, *M-Trends 2026*, Mandiant/Google Cloud, 2026. [Online]. Available: https://www.mandiant.com

[3] Pindrop, *2025 Voice Intelligence and Security Report*, Pindrop Security, Inc., 2025. [Online]. Available: https://www.pindrop.com

[4] N. M. Müller, K. Pizzi, and J. Williams, "Human perception of audio deepfakes," in *Proc. 1st Int. Workshop on Deepfake Detection for Audio Multimedia (DDAM)*, ACM, 2022, pp. 85–91. doi: 10.1145/3552466.3556531.

[5] K. T. Mai, S. Bray, T. Davies, and L. D. Griffin, "Warning: Humans cannot reliably detect speech deepfakes," *PLOS ONE*, vol. 18, no. 8, e0285333, 2023.

[6] B. Maurer, "Mobile money: Communication, consumption and change in the payments space," *Journal of Development Studies*, vol. 48, no. 5, pp. 589–604, 2012.

[7] I. Akomea-Frimpong, C. Andoh, A. Akomea-Frimpong, and Y. Dwomoh-Okudzeto, "Control of fraud on mobile money services in Ghana: An exploratory study," *Journal of Money Laundering Control*, vol. 22, no. 2, pp. 300–317, 2019. doi: 10.1108/JMLC-03-2018-0023.

[8] Bank of Ghana, *Annual Payment Systems Oversight Report 2024*, Bank of Ghana, 2025. [Online]. Available: https://www.bog.gov.gh

[9] J. Whisker and M. E. Lokanan, "Anti-money laundering and counter-terrorist financing threats posed by mobile money," *Journal of Money Laundering Control*, vol. 22, no. 1, pp. 158–172, 2019.

[10] R. B. Cialdini, *Influence: The Psychology of Persuasion*, Rev. ed. New York, NY, USA: Harper Business, 2006.

[11] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks," *Journal of Information Security and Applications*, vol. 22, pp. 113–122, 2015.

[12] M. Workman, "Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security," *Journal of the American Society for Information Science and Technology*, vol. 59, no. 4, pp. 662–674, 2008.

[13] F. Mouton, L. Leenen, and H. S. Venter, "Social engineering attack examples, templates and scenarios," *Computers & Security*, vol. 59, pp. 186–209, 2016.

[14] Federal Bureau of Investigation, "Malicious actors use AI-generated voice messages to impersonate senior officials," Public Service Announcement PSA250515, Internet Crime Complaint Center, May 2025. [Online]. Available: https://www.ic3.gov

[15] S. Kemp and N. Erades Pérez, "Consumer fraud against older adults in digital society: Examining victimization and its impact," *International Journal of Environmental Research and Public Health*, vol. 20, no. 7, 2023.

[16] Y. Mirsky and W. Lee, "The creation and detection of deepfakes: A survey," *ACM Computing Surveys*, vol. 54, no. 1, pp. 1–41, 2021.

[17] M. Westerlund, "The emergence of deepfake technology: A review," *Technology Innovation Management Review*, vol. 9, no. 11, pp. 39–52, 2019.

[18] Group-IB, "The voice of fraud: Deepfake vishing and the new age of social engineering," 2025. [Online]. Available: https://www.group-ib.com/resources/research-hub/voice-of-fraud/

[19] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[20] L. E. Cohen and M. Felson, "Social change and crime rate trends: A routine activity approach," *American Sociological Review*, vol. 44, no. 4, pp. 588–608, 1979.

[21] E. R. Leukfeldt and M. Yar, "Applying routine activity theory to cybercrime: A theoretical and empirical analysis," *Deviant Behavior*, vol. 37, no. 3, pp. 263–280, 2016.

[22] Cyber Security Authority (Ghana), "Mobile money fraud," public awareness resource, 2025. [Online]. Available: https://www.csa.gov.gh

[23] M. Button, C. Lewis, and J. Tapley, "Not a victimless crime: The impact of fraud on individual victims and their families," *Security Journal*, vol. 27, no. 1, pp. 36–54, 2014.

[24] Bank of Ghana, *Financial Stability Report*, Bank of Ghana, 2023–2024. [Online]. Available: https://www.bog.gov.gh

[25] M. Adongo, "Mobile money social engineering attacks in African countries: A survey," SSRN Working Paper, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5257020

[26] J. Aron, "Mobile money and the economy: A review of the evidence," *The World Bank Research Observer*, vol. 33, no. 2, pp. 135–188, 2018.

[27] Z. Almutairi and H. Elgibreen, "A review of modern audio deepfake detection methods: Challenges and future directions," *Algorithms*, vol. 15, no. 5, art. 155, 2022.

[28] B. Gyamfi Kwaku, A. Afua, T. Ashley and O. Fening Kwadwo: "Artificial intelligence integration in cyber incident response teams to enable faster containment, forensic accuracy, and resilient business continuity," International Journal of Science and Research Archive, 2025, 17(01), 1263–1280

[29] A. Afua, "Evaluating Confidential Computing Runtimes to Enforce Verifiable Privacy Guarantees and Automated GRC Controls In Multi-Tenant Cloud Data Processing Workflows, "International Journal of Computer Applications Technology and Research Volume 14 , ISSN: -2319–8656