Signalisation cellulaire pour la detection des fraudes de contournement

Anne J. Kouam[1], Aline C. Viana[1], Philippe Martins[2], Cedric Adjih[1] et Alain Tchana[3]

[1]*INRIA Saclay, France* [2] *Telecom Paris, France* [3] *Grenoble INP, France*

La fraude de contournement, également connue sous le nom de fraude à la *SIMBox*, est l'une des plus sévères dans les réseaux cellulaires, générant des pertes annuelles de 3,11 milliards de dollars et des menaces pour la sécurité nationale. Un challenge majeur à sa détection est l'évolution constante de la fraude en vue de contourner les solutions publiées dans la littérature. Ce papier explore une nouvelle source de données résiliente à l'évolution de la fraude: la signalisation cellulaire. Au travers d'expérimentations avec des appareils *SIMBox* nous montrons son potentiel à distinguer les équipements *SIMBox* des téléphones ordinaires en temps-réel et avant que la fraude ne soit commise.

Mots-clefs : Fraude à la *SIMBox*, Latence de signalisation, Expérimentations

1 Introduction

Bien qu'initialement conçus pour étendre la couverture des communications vocales nécessaires aux activités de certaines entreprises (par exemple, les centres d'appel), les gateways VoIP GSM, encore appelés *SIMBox*, ont été détournés à diverses fins frauduleuses. Dans ce contexte, la fraude à la *SIMBox* [KVT21] est l'une des plus répandues dans les réseaux cellulaires, étant dans le top 3 des types de fraude à l'origine de pertes financières significatives pour les opérateurs mobiles. Comme illustré à la Figure 1, la fraude à la *SIMBox* consiste à intercepter des appels mobiles internationaux et à les faire aboutir, via un réseau VoIP, en tant qu'appels locaux vers le destinataire à l'aide de d'appareils *SIMBox*. Les fraudeurs contournent ainsi l'opérateur d'interconnexion légitime et les frais de terminaison internationaux pour payer les frais de terminaison locaux, moins élevés. Au-delà d'une perte de revenus directe pour les opérateurs, estimée à plus de 3,11 milliards de dollars par an [CFC21], la fraude à la *SIMBox* impacte négativement la disponibilité et la qualité du réseau pour les abonnés légitimes par surchage de cellules sous-approvisionnées. Pire que cela, les appareils *SIMBox* permettent d'écouter les conversations téléphoniques internationales, portant atteinte à la vie privée des utilisateurs et offrant la possibilité de faire de l'espionnage international. Enfin, la fraude à la *SIMBox* constitue une brèche de sécurité, exploitable par des terroristes pour dissimuler leurs activités dans un pays en tant qu'abonnés nationaux. Ce dernier aspect concerne tous les pays du globe, pouvant malheureusement induire des risques humanitaires critiques, ce qui atteste de la sévérité de cette fraude.

La détection de la fraude s'effectue au niveau de l'opérateur de destination. Elle implique l'analyse des *traces de l'activité cellulaire des abonnés* pour distinguer les cartes SIM utilisées dans la *SIMBox* des cartes SIM légitimes. Ainsi, la majorité des travaux de détection de la littérature (12 sur 14) analysent des traces *CDRs* (Call Detail Records), qui sont des évènements réseaux horodatés et géoréférencés (i.e., appels, SMS ou données) générés par des appareils mobiles dans leurs interactions avec un réseau d'opérateur. Ces méthodologies ne sont malheureusement pas robustes face à l'évolution constante de la fraude. En effet, les fraudeurs *SIMBox* affinent constamment leurs communications pour imiter les comportements humains ce qui les rend difficilement détectables par une analyse basée sur les CDRs. En outre, le temps nécessaire pour obtenir des informations pour la détection à partir des traces CDRs conduit à une fréquence de détection réduite (ex. une fois par semaine/jour) suffisante aux fraudeurs pour faire du profit. D'autre part, quelques contributions exploitent plutôt les *traces audio des appels* pour la détection. Bien que certaines de ces

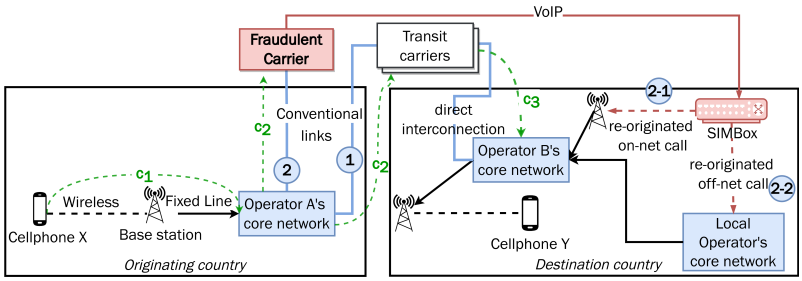

Fig. 1: Routage d'appel international: (Flow 1) schéma légitime, (Flow 2) schéma frauduleux.

méthodes semblent efficaces pour la détection en temps réel [RSB+15], elles nécessitent que les opérateurs analysent la qualité audio de tous les appels sur le réseau pour découvrir les éventuelles dégradations dues à la fraude. Ceci est difficilement réalisable en raison des réglementations limitant l'accès des opérateurs à l'audio des appels de leurs abonnés et des ressources importantes que cela demanderait.

En raison des limitations susmentionnées des solutions de détection existantes, la fraude aux *SIMBox* continue de sévir [CFC21]. Face à cela, *cet article décrit le premier travail de la littérature introduisant l'analyse de la signalisation cellulaire pour la détection de la fraude à la SIMBox.*

Motivation: La signalisation cellulaire, contrairement à l'audio des appels, a l'avantage d'être facilement accessible sous forme de logs collectés à plusieurs nœuds dans un réseau d'opérateur. Elle est encadrée par les normes 3GPP et véhicule des informations sur les appareils mobiles qui, bien que non privées, peuvent servir à identifier la fraude en temps-réel et avant tout dommage. De plus, les fraudeurs n'ont aucun accès ni contrôle sur la signalisation cellulaire et peuvent à peine l'interférer car elle implique principalement les composants des opérateurs (c'est-à-dire les cartes SIM à l'intérieur des équipements *SIMBox* et les stations de base du réseau d'opérateur). Par conséquent, contrairement aux CDRs, la signalisation cellulaire est résiliente aux techniques de fraude avancées qui visent à déguiser le comportement de communication des utilisateurs frauduleux.

Contributions et organisation: Dans ce papier, nous proposons l'exploitation de la latence de signalisation pour la détection de la fraude à la périphérie du réseau cellulaire (i.e., au niveau de la station de base). Premièrement (§3) nous mettons sur pied un testbed avec des appareils *SIMBox* du fabricant Hybertone (le plus répandu du marché international [GoA18]) déployés dans un réseau d'opérateur à l'intérieur d'un environnement isolé (i.e., une cage de Faraday). Ensuite (§4), nous analysons la latence de la signalisation cellulaire de la procédure d'attachement au réseau, dans différents scénarios, pour la *SIMBox* en comparaison aux téléphones. Les résultats montrent que la *SIMBox* présente une latence totale en moyenne 8X supérieure à celle des téléphones classiques qui permettrait de la distinguer. Enfin, la section §5 discute les résultats obtenus tout en détaillant les étapes futures de généralisation et exploitation de ces résultats.

2 Background: Architecture de la *SIMBox*

La *SIMBox* est un dispositif standard fonctionnant comme une passerelle VoIP GSM. Elle reçoit le trafic d'appel dévié en tant que client VoIP et le route en relançant des appels mobiles cellulaires à l'aide de nombreuses cartes SIM. La *SIMBox* crée en permanence des UEs (User Equipment) *"virtuels"* en *combinant des cartes SIM et des antennes cellulaires*. Dans cette association, l'antenne cellulaire assure la communication sans fil avec le réseau et la carte SIM identifie et authentifie l'UE formé. La *SIMBox* fonctionne par l'interaction de trois types de composants matériels :

- La *gateway* est un appareil avec un ensemble d'antennes cellulaires assurant la communication sans fil dans une plage de fréquences donnée. Elle reçoit le trafic VoIP entrant et le distribue aux antennes.
- La *SIMBank* est un dispositif doté de nombreux emplacements SIM. Elle gère les cartes SIM de la *SIMBox*, c'est-à-dire leur ajout, retrait et le transfert de leurs données et de leur état à d'autres composants.
- Le *serveur de contrôle* est un serveur Web fournissant les fonctions de contrôle de la *SIMBox*, c'est-à-dire la *combinaison des cartes SIM et antennes cellulaires* et la configuration de l'ensemble de l'architecture. Il peut être hébergé en ligne pour faciliter l'accès à distance à partir d'une interface web.

Paramètres		**Valeurs**
PC Hôte (BS, MME, SGW)		Intel(R) Core(TM) i9-10900K CPU@3.70GHz, 16GB RAM, Gigabit Ethernet controller
Cellule	Largeur de bande	5MHz FDD
	Configuration	SISO
	Bande de fréquence	Downlink center frequency: 1845 MHz, Band 3
Cartes SIM programmables		SysmoSIM-SJS1 de Sysmocom
Téléphones		Huawei Mate 8, Xiaomi Redmi Note 9
Appareils *SIMBox* Hybertone		SIMBank: SMB32; Gateway: 2x GoIPx8; Serveur de Controle (v. 2021/06/30)

Tab. 1: Spécifications des composants du testbed.

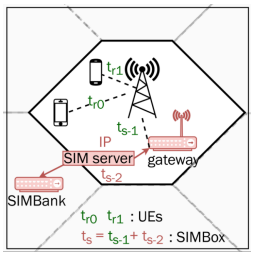

Fig. 2: Latence de signalisation.

Step	**Description**	**Direction**
1	EMM: Attach request	UE → eNB
2	Authentication request	UE ← eNB
3	*Authentication response*	*UE → eNB*
4	NAS Security mode command	UE ← eNB
5	*NAS Security mode complete*	*UE → eNB*
6	RRC Security mode command	UE ← eNB
7	*RRC Security mode complete*	*UE → eNB*
8	Attach accept	UE ← eNB
9	RRC connection reconfiguration	UE ← eNB
10	*RRC Connection reconfiguration complete*	*UE → eNB*
11	*EMM : Attach complete*	*UE → eNB*

Tab. 2: Procédure d'attachement.

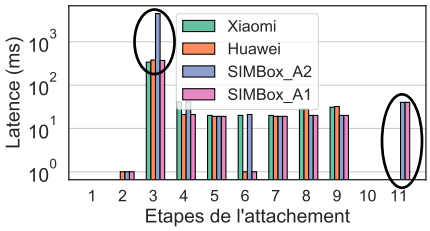

Fig. 3: Valeur de latence par étape en échelle logarithmique.

À partir de ces composants, nous distinguons deux types d'architectures *SIMBox*: *standalone* et *distribuée*. Dans l'architecture standalone, la *SIMBox* est constituée uniquement d'une gateway munie d'emplacements SIM, qui peut donc gérer toutes les fonctions des composants à la fois. L'architecture distribuée, quant à elle, fait intervenir plusieurs équipements : au moins une gateway, au moins une SIMBank, et le serveur de contrôle. Elle donne la possibilité aux fraudeurs de simuler des déplacements virtuels à leurs cartes SIM pour contourner les détecteurs analysant le comportement de mobilité des utilisateurs [KVT21].

3 Environnement expérimental

Le tesbded utilisé dans cette étude repose sur une suite Amarisoft 4G et les spécifications liées aux composants utilisés sont reportés dans la Table 1. L'EnodeB, le MME et le SGW sont colocalisés sur le même PC, où est effectué le traitement de la bande de base. Le traitement SDR est réalisé par une tête radio USRP B210 connectée au PC via une interface USB3. Le traitement en bande de base communique avec la tête radio via une API libuhd 4.3.0 TRX. Le testbed déploie une seule cellule 4G dont les paramètres radio sont décrits dans la Table. En outre, nous utilisons pour notre étude deux téléphones, et trois architectures *SIMBox* formés par une SIMBank, deux gateways et un serveur de contrôle Hybertone.

4 Latence de signalisation cellulaire

L'intuition de détection que nous developpons se base sur le fait que la signalisation de la *SIMBox* a une latence relativement élevée par rapport à des téléphones ordinaires, en raison de l'architecture de la *SIMBox*.

En effet, dans les réseaux cellulaires, les UE se composent de deux éléments distincts, l'équipement mobile (ME) et la carte SIM. Le ME est un dispositif matériel qui contient un processeur, une mémoire, un émetteur-récepteur, une batterie, etc. Une carte SIM physique est insérée dans l'équipement mobile ou une carte SIM électronique est activée par un abonnement. La combinaison du ME et de la SIM avec un abonnement actif permet à l'UE de communiquer avec le réseau cellulaire.

Pour les UEs de la *SIMBox*, la combinaison du ME (gateway) et des cartes SIM est soit *manuelle* pour l'architecture standalone, soit *logique* au niveau du serveur de contrôle pour l'architecture distribuée. Pourtant, la plupart des déploiements *SIMBox* sont distribués en raison des limitations induites par l'architecture *SIMBox* standalone décrites en §5. Ainsi, comme le montre la Fig. 2, la latence de communication des UEs *SIMBox* est $t_s = t_{s-1} + t_{s-2}$, où $t_{s-1} \approx t_{r0} \approx t_{r1}$ est la latence régulière due au traitement des opérations

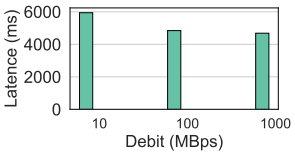

Fig. 4: *SIMBox* (A2) en fonction du débit.

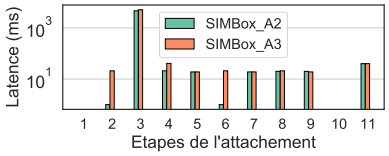

Fig. 5: Comparaison de latences (A2) et (A3).

de signalisation et à la propagation du signal sans fil dans la cellule. D'autre part, t_{s-2} est la latence supplémentaire subie par la *SIMBox* en raison de l'interaction de ses composants par IP.

Résultats: Nous expérimentons cette latence en réalisant trois architectures *SIMBox*: (A1) *architecture standalone*, (A2) *architecture distribuée locale* où les composants de la *SIMBox* sont dans le même réseau local et (A3) *architecture distribuée remote* où le serveur de contrôle est sur Internet et la SIMBank s'y connecte depuis un site externe (Ivry-sur-Seine), la gateway étant toujours dans la cage de Faraday (Palaiseau). La Fig. 3 reporte la latence collectée à chaque étape de la procédure d'attachement au réseau (cf. Table 2). Sur cette figure, le surplus de latence dans le cas distribué, même en local, est énorme ($\approx$ 8x) comparé au cas standalone. En particulier à l'étape 3, l'UE doit envoyer en réponse au réseau les vecteurs d'authentification $(RAND, AUTN, XRES, Kasme)$ qui nécessitent une computation interne à la carte SIM car utilisant sa clé secrète K_i. Cette étape est donc la seule où le temps t_{s-2} de communication IP intra-composants *SIMBox* est nécéssairement imputée. D'autre part, à l'étape 11, la *SIMBox* dans ses deux architectures présente un surplus de 40ms de latence comparé aux téléphones. Nous suspectons que cela est du au fonctionnement en tourniquet de la *SIMBox* dans l'écoute de ses différentes antennes (8 dans le cas actuel). Cette interpretation doit être validée dans nos travaux futurs. La Fig. 4 montre que cette latence est de plus en plus importante lorsque le débit est réduit de 1GBps à 10MBps. De plus, nous comparons pour un débit de 100 MBps, à la Fig. 5, la latence des architectures *SIMBox* (A2) et (A3) en échelle logarithmique : il y a une majeure différence à l'étape 3, où la latence de (A3) est supérieure de 460ms à celle de (A2). Cela est lié à la latence additionnelle due au routage par Internet des paquets intra-équipements *SIMBox*.

5 Discussion et Perspectives

Ce papier introduit pour la première fois l'analyse de la signalisation cellulaire pour la détection de la fraude à la *SIMBox*. Les résultats préliminaires en §4 montrent qu'on identifie au niveau de la station de base un surcoût anormal de latence des équipements *SIMBox* par rapport aux téléphones ordinaires, qui pourrait être exploité pour la détection de la fraude. Ce résultat est impactant car les fraudeurs peuvent difficilement remédier à cette latence. En effet, elle dépend de paramètres hors de leur portée tels que la taille des messages de signalisation cellulaire (normalisée) et la qualité de la connexion Internet qui varie en fonction du fournisseur et du lieu. De plus, bien que ce surcoût soit lié uniquement aux architectures distribuées de *SIMBox*, l'architecture standalone est très coûteuse pour les fraudeurs. Non seulement elle limite le nombre d'appels pouvant être détournés aux possibilités d'un seule gateway, mais aussi empêche l'application des techniques de migration et de rotation des cartes SIM de la *SIMBox*, assurant une détection directe par des approches de la littérature.

La suite de notre travail consistera à généraliser les résultats avec d'autres marques de téléphone, réaliser des conditions réelles de propagation radio afin de majorer la latence de téléphones ordinaires, et enfin exploiter les motifs obtenus par un modèle qui sera évalué pour la détection de la fraude en temps réel.

References

[CFC21] CFCA. Communications fraud control association 2021 fraud loss survey. 2021.

[GoA18] GoAntiFraud. Top-5 popular gsm gateway manufacturers, 2018. Accessed: 2020-02-23.

[KVT21] A. J. Kouam, A. C. Viana, and A. Tchana. Simbox bypass frauds in cellular networks: Strategies, evolution, detection, and future directions. *IEEE Communications Surveys Tutorials*, 2021.

[RSB+15] B. Reaves, E. Shernan, A. Bates, H. Carter, and P. Traynor. Boxed out: Blocking cellular interconnect bypass fraud at the network edge. In *USENIX Security*, 2015.