Demonstração de vulnerabilidades do WhatsApp que o fragilizam como evidência em processos Judiciais

Yuri. C Zwaig[1,2], Marli F.G. Hernandez[1], Marbilia P Sergio[2]
research@tadeu.work, marli@ft.unicamp.br, marbilia.sergio@cti.gov.br

[1]**Faculdade de Tecnologia (FT)**
UNICAMP - da Universidade Estadual de Campinas

[2] **Núcleo de qualificação de Software (NQS)**
CTI/MCTI Renato Archer – Campinas/SP

Abstract. *With about 2 billions users, WhatsApp is the most popular Instant Messenger exchanger application [1], having in its user base users that can make inappropriate use of the tool. Conversations exchanged through WhatsApp are encrypted end to end, preventing unauthorized access to conversations. This research aimed to develop the tool AlterWhats that automates the modification of WhatsApp content sot hat it is not detected by forensic tools, confirming or not the reliability of the extracted contente from a device. As a resulto f the seven forensic tools applied in the conversation after the change, none were able to detect the change made by the developed application. The contribution of this work is the free availability of the tool "AlterWhats" in "GitHub", which can be used to improve the algorithms of the forensic tools. As well leaving the alert that the failure to verify the integrity of the message makes it possible for the forensic tools to make na erroneous diagnosis.*

Resumo. *Com cerca de 2 bilhões de usuários, o WhatsApp, é o aplicativo de troca de MI mais popular [1], tendo em sua base de usuários que podem fazer uso inadequado da ferramenta. As conversas trocadas pelo WhatsApp são criptografadas ponta a ponta, prevenindo acesso não autorizado às conversas. Este trabalho de pesquisa se propôs a desenvolver a ferramenta AlterWhats que automatize modificações de conteúdos do Whatsapp de forma a não ser detectada por ferramentas forense confirmando ou não a confiabilidade do conteúdo extraído de um dispositivo. Como resultado tem-se das sete ferramentas forense aplicadas na conversa após a alteração, nenhuma conseguiu detectar a modificação realizadas pelo aplicativo desenvolvido. A contribuição deste trabalho é a disponibilização livre da ferramenta Alterwhats no "GITHUB" que pode ser utilizada para melhorar os algoritmos das ferramentas forense. Assim como deixa o alerta de que a falha de verificação da integridade da mensagem torna possível que as ferramentas forenses façam um diagnóstico errôneo.*

1. 1. Introdução

WhatsApp é um aplicativo de troca de mensagens instantânea (MI), multi-plataforma, permitindo a comunicação entre dispositivos de diversos modelos e sistemas operacionais. Com cerca de 2 bilhões de usuários, o WhatsApp, é o aplicativo de troca de MI mais popular [1], tendo em sua base de usuários que podem fazer uso inadequado da ferramenta.

Desde 31 de março de 2016 todas as conversas trocadas pelo Whatsapp são criptografadas ponta a ponta [2], prevenindo que entidades não autorizadas tenham acesso às conversas. Uma vez que, o único lugar que as mensagens já descriptografadas, podem ser lidas, é dentro dos dispositivos dos participantes da conversa.

A inclusão da criptografia nesta ferramenta, fez com que as investigações criminais mudassem seu procedimento de coleta de dados para perícia, isto é, deixou de requerer um mandado solicitando o conteúdo das conversas para a operadora do aplicativo e passando a aprender os dispositivos dos envolvidos na investigação para análise dos artefatos neles contidos. Porém, diante deste novo contexto, coloca-se em pauta a confiabilidade do conteúdo das mensagens encontradas nos dispositivos apreendidos e periciados.

Para verificar a integridade artefatos do WhatsApp, a perícia conta com ferramentas forense baseadas nos parâmetros do NIST[3]. Assim, este trabalho de pesquisa pretendeu avaliar a eficácia destas ferramentas na identificação de fraudes na analise do conteúdo do WhatsApp.

Este trabalho de pesquisa se propôs a criar uma ferramenta que automatize alterações de conteúdos do WhatsApp de forma a não ser identificadas pelas ferramentas forenses a fim de confirmar ou não a confiabilidade do conteúdo extraído de um dispositivo.

Nos próximos tópicos são apresentadas a base teórica, a metodologia, os resultados obtidos e conclusão sobre o trabalho. No tópico 2 há um breve relato da revisão bibliográfica que identificou alguns trabalhos relacionados, assim como uma especificação resumida do WhatsApp e as ferramentas forense utilizadas para teste. No tópico 3 são descritas a metodologia e as atividades executadas, no Tópico 4 são apresentados os resultados obtidos e no tópico 5 a conclusão do trabalho.

2. Trabalhos correlatos, o WhatsApp e ferramentas forense

Para o desenvolvimento dessa pesquisa, foi realizada a revisão literária nas bases de dados da IEEE e da WOS. Utilizou-se como argumento de pesquisa o termo “WhatsApp” obtendo-se um total de 1259 publicações. Os títulos e resumos das publicações selecionadas foram incluídos na versão de 2020 do aplicativo SW3T [4]. O Sw3T é uma ferramenta de apoio a revisão literária que permite a analise e o registro documentando os resultados titulo a titulo. Foram selecionados treze artigos que focam em atividades forense, que descrevem a configuração do aplicativo WhatsApp e as que citam ferramentas forense a serem utilizadas para validar o software que altera o conteúdo do WhatsApp.

No levantamento das ferramentas foram selecionadas 7 aplicações forense de possível acesso sem custo. Foi desenvolvido o software AlterWhats para realizar

alterações de artefatos em conversas do WhatsApp com base em especificações disponível do aplicativo.

2.1. Trabalhos correlatos

Umar et at [3] utilizam 3 ferramentas forenses, com base nos parâmetros NIST, para a analise de artefatos do WhatsApp. Entre elas, destaca-se a ferramenta "WhatsApp Key/DB Extractor" utilizada nesse trabalho. As demais ferramentas não são de fácil acesso ou de custo alto para obtenção inviabilizando sua utilização.

Alissa et al [5] descrevem formas de obtenção dos dados em uma investigação forense e mostra facilidades para obtenção de dados via arquivos do WhatsApp. O artigo também expõe 4 ferramentas forenses das quais três foram utilizadas neste trabalho, além da "WhatsApp Key/DB Extractor" cita "Guasap" e "Elcomosft WhatsApp Explorer"(versão de teste).

Anglano [6] expõem a composição dos artefatos do WhatsApp auxiliando na identificação de seus elementos.

2.2. O Aplicativo WhatsApp

Os artefatos do WhatsApp são armazenados em banco de dados, arquivos de backups e de logs (responsável por guardar histórico de eventos do aplicativo). Os arquivos relevantes para este experimento são apresentados na Tabela 1.

Tabela 1 -- Arquivos do WhatsApp utilizados neste experimento

#	Conteúdo armazenado	Diretório comum	Nome do Arquivo
1	*Conversas, mensagens*	*/data/data/com.WhatsApp/databases/*	*msgstore.db (SQLite)*
2	*Contatos*	*/data/data/com.WhatsApp/databases/*	*wa.db (SQLite)*
3	*Backups do msgstore*	*/sdcard/WhatsApp/Databases/*	*msgstore-{DATA}.db.crypt12*
4	*Backups diversos*	*/sdcard/WhatsApp/Backups/*	**.db.crypt1*
5	*Arquivos recebidos/enviados*	*/sdcard/WhatsApp/Media/*	*
6	*Arquivos de log*	*/data/data/com.WhatsApp/files/Logs/*	*WhatsApp.log WhatsApp-{DATA}.log.gz*

Foram objetos deste experimento os artefatos {1}, {3}, {6} da Tabela 1. Sendo:

- **Conteúdo do banco msgstore {1}:** que contém todas as informações sobre conversas e mensagens. A tabela 2 estão as oito tabelas do WhatsApp relevantes para esta pesquisa.

Tabela 2 – Relação de tabelas e colunas do WhatsApp relevantes para esta pesquisa

#	Conteúdo	Tabela	Colunas importantes
1	*Listagem de chats do usuário*	*chat*	*display_message_row_id, last_message_row_id. Entre outras*

2	*Listagem de chats do usuário*	*chat_list*	*last_message_table_id, last_read_message_table_id. E outras*
3	*Conteúdo das mensagens*	*message_ftsv2 (1)*	*content, fts_jid*
4	*Conteúdo das mensagens*	*message_ftsv2_content (2)*	*c0content, c1fts_jid*
5	*Desconhecido (3)*	*message_ftsv2_docsize (4)*	*size*
6	*Conteúdo e dados sobre uma mensagem*	*messages*	*data, key_remote_jid, key_id. Entre outras*
7	*Conteúdo das mensagens*	*messages_fts*	*content*
8	*Conteúdo das mensagens*	*messages_fts_content*	*c0content*

- **Backups do msgstore {3}:** seus dados são criptografados, mas, há diversas ferramentas que decodificam esse backup [7] e ferramentas que tentam pegar essa chave sem acesso root [8]. Esse backup é importante pois ele contém informações antes da alteração pelo software AlterWhats, então é interessante sobrescrevê-lo ou remover o conteúdo seguramente.
- **Arquivos de Log {6}:** contém todas as iterações feita pelo usuário e pelo sistema. Por exemplo, o recebimento de mensagens e o bloqueio de usuários [19]. Esse arquivo guarda informações que podem desmentir alguma alteração.

2.3. Ferramentas Forense

Após uma ampla pesquisa, foram escolhidas as seguintes ferramentas citadas em outros artigos, que são gratuitas ou com versão de teste:

1.Guasap Forensic (Versão de 09 de Julho de 2019)[9]

2.WhatsApp Key/DB Extractor (Versão de 21 de Outubro de 2016)[8]

3.WhatsApp Viewer (Versão de 29 de Julho de 2019) [7]

4.WhatsApp Xtract (Versão de 25 de Abril de 2018)[10]

5.Bring2lite (Versão de 05 de Agosto de 2019)[11]

6.Elcomsoft Explorer Forense (Versão de teste obtida em maio de 2020) [12]

7.WhatsApp Parser (Versão de 18 de Junho de 2020) [13]

3. Condução dos trabalhos

Para esta pesquisa foi adotado a metodologia de pesquisa-ação onde se pretendeu, através de experimentos, qualificar o software a ser desenvolvido através da

aplicação das ferramentas forense buscando seu aperfeiçoamento para evitar a detecção das modificações de conversas. Os resultados de cada atividade foram analisados quanto a sua adequação e necessidade de ajuste e reexecução. Durante a pesquisa, a metodologia viabilizou a inclusão de atividades não previstas inicialmente. As atividades realizadas foram:

- **Revisão literária** – com o objetivo de identificar publicações científicas relacionadas a processo forense que identificasse as alterações das informações realizadas em uma conversa do WhatsApp;
- **Entendimento das funcionalidades do WhatsApp** de forma a identificar sua estrutura física e logica a fim de estabelecer a melhor forma de realizar alterações nos artefatos; Especificação do ambiente de desenvolvimento;
- **Desenvolvimento do Software AlterWhats [14],** foi especificado e codificado para efetuar alterações de forma automatizada em conteúdo de conversas no WhatsApp. O ambiente de desenvolvimento foi o Android Debug Bridge (adb) [18] e a linguagem PHP.
- **Identificação de ferramentas forenses** mais populares e estabelecer quais são passíveis de serem utilizadas neste experimento;
- **Realização de testes** com as diversas ferramentas forenses selecionadas visando a identificação das alterações realizadas pelo AlterWhats.

4. Resultados

Destacam-se dois tipos de resultado, o da alteração efetuada pelo aplicativo AlterWhats e os resultados da execução das ferramentas forenses selecionadas.

4.1. Quanto ao uso software AlterWhats

Alteração realizada pelo AlterWhats em uma conversa, conforme figura 1- Conversa original e conversa alterada, teve como objetivo modificar o valor da divida de $10 para $10000 (dólares), e a data do vencimento de 05/10/2020 para 04/09/2020.Além de corromper os logs e backups.

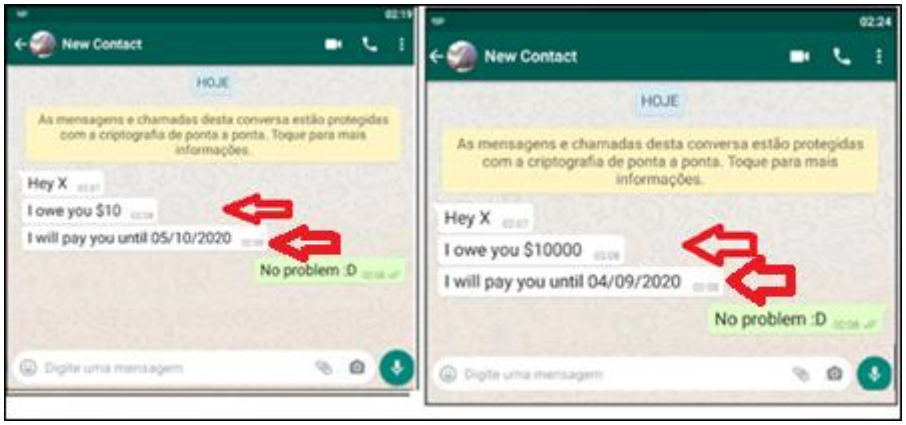

Figura 1 - Conversa original e conversa alterada

Como se observa na Figura 2, o software AlterWhats fornece uma tela interpretando o conteúdo da conversa antes da alteração.

A tela do software apresenta suas funcionalidades, assim como o conteúdo da conversa original e a lista dos últimos comandos realizados por ele.

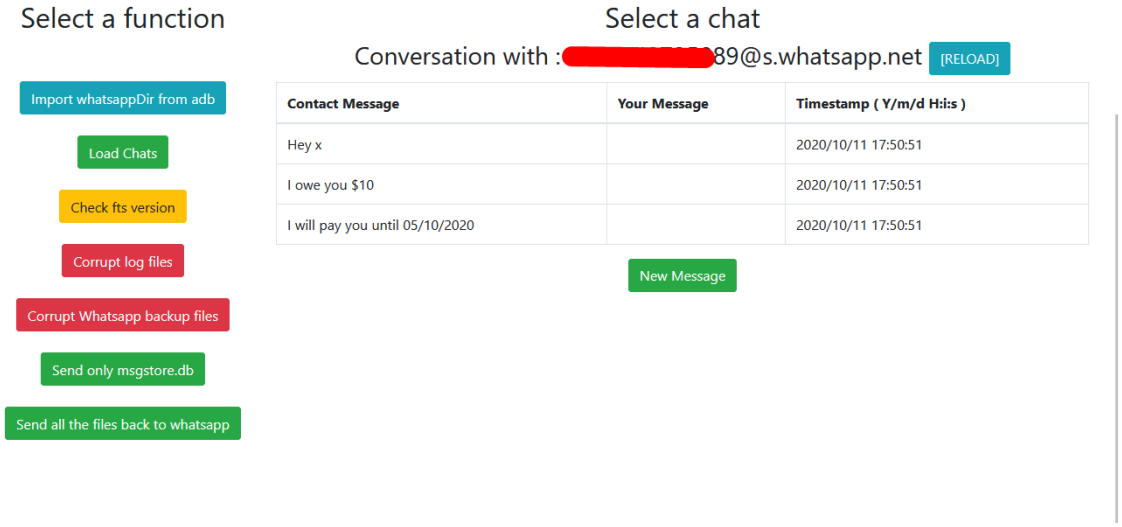

Figura 2 - Interpretação do Software AlterWhats da conversa antes da alteração

As funcionalidades do AlterWhats são:

Importar os artefatos do aplicativo WhatsApp direto do navegador. Carregar as conversas armazenadas nesses artefatos.

Corromper os arquivos de log e backup do WhatsApp.

Fazer alterações nas conversas de indivíduos.

Após as alterações realizadas, o AlterWhats apresenta a tela com sua nova interpretação onde o valor e a data estão alterados. Veja a Figura 3 que mesmo após as alterações a interpretação é a mesma, apenas com as alterações efetuadas dando indícios de que as tabelas do WhatsApp estão sem falha de alteração.

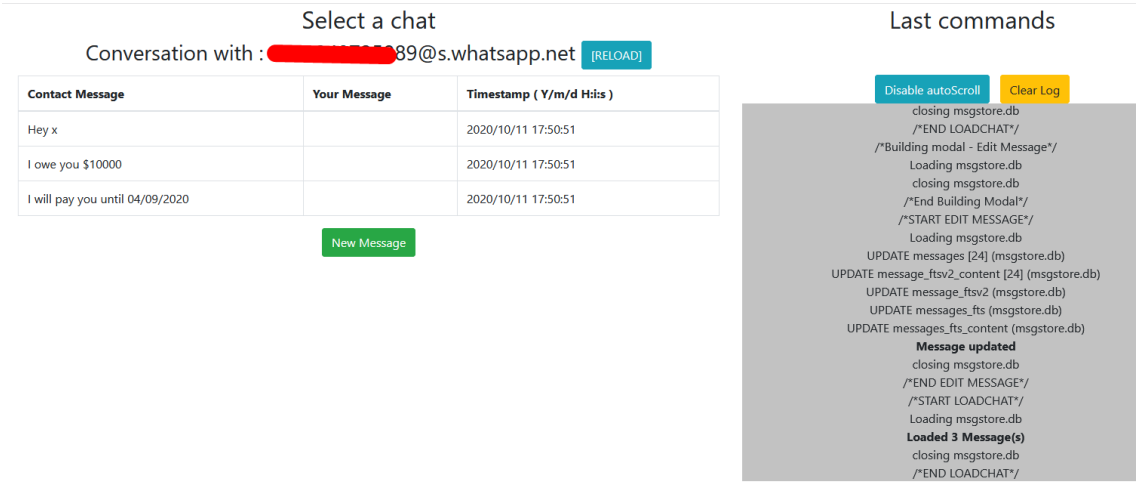

Figura 3 - Interpretação do Software AlterWhats depois da alteração

Após a alteração de uma mensagem, para garantir que não tenha vestígios, o conteúdo das diversas tabelas do msgstore precisaram ser ajustadas. É importante

ressaltar que algumas tabelas não têm chave primaria definida, então usou-se “rowid” padrão do SQLite.

Em seguida, os arquivos do WhatsApp alterados no AlterWhats foram enviados para o celular e realizado um backup na nuvem. Mas, apenas mudar o conteúdo no dispositivo do autor não é suficiente já que os outros participantes da conversa ainda têm o conteúdo original da conversa. Assim, o aplicativo considera o uso de bugs no código do WhatsApp para forçar a exclusão/inclusão dos dados no celular do outro participante [15]. Mas também pode ser utilizado algumas vulnerabilidades mais avançadas que permitam excluir seletivamente os dados [16] [17].

4.2. Quanto ao Testes das ferramentas

Como resultado da execução das ferramentas forense, tem-se que:

- ✓ As ferramentas **Guasap Forensic** (v9/jul/12019) e **WhatsApp Key/DB Extractor** (V21/out/2016) não conseguiram extrair os dados do emulador, impossibilitando a avaliação:
- ✓ O **WhatsApp Viewer** mostrou as mensagens editadas, e um campo a mais após a primeira mensagem, provavelmente a troca de chaves entre os contatos. Vestígio da alteração não foi detectado.
- ✓ O **WhatsApp Xtract** não verifica alterações na msgstore, só checa o conteúdo do banco de dados, mostrando as mensagens editadas. Vestígio da alteração não foi detectado.
- ✓ O **Bring2lite** obtém as linhas excluídas de tabelas Sqlite. Assim, ao executa-lo todas as tabelas que tem a informação original foram recuperadas, mas a informação contida nelas é a já modificada pelo software AlterWhats. Vestígio da alteração não foi detectado.
- ✓ O **Elcomsoft Explorer for WhatsApp** (trial) não verifica alterações e as mensagens modificadas são exibidas normalmente. Vestígio da alteração não foi detectado.
- ✓ O **WhatsApp Parser tsó** lê e exibe o editado. Vestígio da alteração não foi detectado.

Observa-se que as cinco ferramentas testadas não detectaram as mensagens originais, somente foram detectadas as mensagens alteradas pelo AlterWhats.

5. Conclusão

Este trabalho de pesquisa se propôs a desenvolver a ferramenta AlterWhats de tal forma que fosse capaz de realizar modificações imperceptíveis para cinco ferramentas forense.

Ao estudar as possibilidades a partir do entendimento do funcionamento do WhatsApp e seus artefatos, foi possível identificar que com o uso de backup na nuvem, é possível alterar os artefatos do WhatsApp em um dispositivo com privilégios de alteração de sistema (root) e encaminhá-lo para outro dispositivo desprovido de tais

privilégios, dificultando a comprovação de que os artefatos de uma conversa foram manipulados.

Assim, afirma-se que embora estas ferramentas forenses possam detectar diversas tentativas de fraude, através da analise e comparação de seus arquivos e back-ups, ainda é possível que modificações não sejam detectadas colocando seus resultados em xeque quanto a eficácia na detecção de fraude.

A revisão literária focou em duas Bases Dados. Embora as BDs sejam fortes em temas relacionados a segurança da informação, cabe a expansão do levantamento para outras bases a fim de identificar relatos de vulnerabilidade e outras ferramentas forense.

O trabalho limitou-se as ferramentas forense gratuitas, cabe a continuação dos testes utilizando outras ferramentas forenses disponíveis no mercado a fim de dar mais confiança no AlterWhats.

Futuros trabalhos também podem destrinchar os componentes do WhatsApp a fim de identificar possíveis vulnerabilidades no código do WhatsApp para comprometer a integridade do mesmo. O uso do Alterwhats pode contribuir com este estudo e inclusive sendo melhorado para ser mais completa.

Nunca fora segredo ser possível alterar mensagens de conversas do WhatsApp, mas, não foi encontrado estudo sobre a manipulação não detectáveis dos artefatos em investigações forenses. Assim, a importante contribuição deste trabalho é a demonstrar que, devido a existência de falha de verificação da integridade da mensagem após sua saída dos servidores do WhatsApp, torna-se possível alterações não detectáveis de artefatos (como mensagens), assim como a disponibilização livre do Software AlterWhats no “GitHub”. Uma vez que as ferramentas forenses testadas não possuem métodos para verificar a integridade, portanto afirma-se que as ferramentas forenses podem produzir resultados errôneos.

6. Referencias

[1] – Statista, “Most popular global mobile messenger apps as of October 2019, based on number of monthly active users” 2019. Disponivel em : https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/ Acessado em: 18 Jul 2020

[2] - WhatsApp. WhatsApp Encryption Overview. Technical white paper, 2016. Available: https://www.whatsapp.com/security/WhatsApp-Security-Whitepaper.pdf

[3] Umar, Rusydi; Riadi, Imam; Zamroni, Guntur. A Comparative Study of Forensic Tools for WhatsApp Analysis using NIST Measurements. International Journal of Advanced Computer Science and Applications, vol. 8, no. 12, pp. 69-75, 2017.

[4] M. P. Sergio, T. d. S. Costa, M. S. d. P. Pessoa and P. S. M. Pedro, "A Semantic Approach to Support the Analysis of Abstracts in a Bibliographical Review," 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), Napoli, Italy, 2019, pp. 259-264, doi: 10.1109/WETICE.2019.00062.

[5] Alissa, K.; Almubairik, N.A.; Alsaleem, L. et al. A comparative study of WhatsApp forensics tools. SN Appl. Sci. 1, 1320 (2019). https://doi.org/10.1007/s42452-019-1312-8.

[6] Anglano, Cosimo. (2014). Forensic Analysis of WhatsApp Messenger on Android Smartphones. Digital Investigation. 11. 10.1016/j.diin.2014.04.003.

[7] Mausch, Andreas. “WhatsApp Viewer” 2018. Disponível em : https://github.com/andreas-mausch/WhatsApp-viewer Acessado em: 18 Jul 2020.

[8] TripCode, “WhatsApp Key/DB Extractor” 2016. Disponível em: https://github.com/EliteAndroidApps/WhatsApp-Key-DB-Extractor Acessado em: 18 Jul 2020.

[9] Quantika14, “Guasap Forensic” 2019. Disponível em: https://github.com/Quantika14/guasap-WhatsApp-foresincs-tool Acessado em: 18 Jul 2020.

[10] Ztedd, “[TOOL] WhatsApp Xtract: Backup Messages Extractor / Database Analyzer / Chat-Backup” 2018. Disponível em: https://forum.xda-developers.com/showthread.php?t=1583021 Acessado em: 18 Jul 2020.

[11] Meng, Christian & Baier, Harald. (2019). bring2lite: A Structural Concept and Tool for Forensic Data Analysis and Recovery of Deleted SQLite Records. Digital Investigation. 29. S31-S41. 10.1016/j.diin.2019.04.017.

[12] Elcomsoft, “Elcomsoft Explorer for WhatsApp”. Disponível em: https://www.elcomsoft.com/exwa.html Acessado em: 18 Jul 2020.

[13] B16f00t, “WhatsApp Parser Toolset” 2020. Disponível em: https://github.com/B16f00t/whapa Acessado em: 18 Jul 2020.

[14] Tuyuribr, “AlterWhats” 2020. Disponível em:https://github.com/tuyuribr/AlterWhats Acessado em: 1/8/2020.

[15] Swati Khandelwal, “This Bug Could Have Let Anyone Crash WhatsApp Of All Group Members” 2019. Disponível em:https://thehackernews.com/2019/12/WhatsApp-group-crash.html Acessado: 18/7/2020.

[16] MITRE, “CVE-2019-11932” 2019. Disponível em: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-11932 Acessado em: 18 Jul 2020.

[17] Zerodium, “Our Exploit Acquisition Program” 2020. Disponível em: https://zerodium.com/program.html Acessado em: 18 Jul 2020.

[18] Google, “Android Debug Bridge (adb)” 2020. Disponível em: https://developer.android.com/studio/command-line/adb Acessado em: 18 Jul 2020.

[19] C. Anglano. Forensic analysis of WhatsApp messenger on android smartphones. Digital Investigation. 2014.