Дослідження систем захисту захищених месенжерів типу Skype, Viber, WhatsApp, Signal

Недождій Максим, Буржимський Ростислав

ФІ-42мн

Мета

Дослідження особливостей реалізації криптографічних механізмів протоколів захисту мультимедійної інформації типу SIP

Група 1. Проаналізувати існуючу інформацію про системи Viber, WhatsApp, Skype, Telegram та їх криптографічні механізми.

Телекомунікаційні технології в останнє десятиліття перейшли від комутації каналів (PSTN/GSM) до комутації пакетів (VoIP) та домінуванням OTT-сервісів (Over-The-Top). Цей зсув не лише змінив бізнес-моделі операторів зв'язку, але й вплинув на інформаційну безпеку. Традиційні протоколи сигналізації та передачі медіаданих, такі як SIP (Session Initiation Protocol) та RTP (Real-time Transport Protocol), які спочатку розроблялися для довірених мереж операторів, виявилися недостатньо захищеними для використання у відкритому Інтернеті.

Сучасні месенджери, такі як Signal, WhatsApp, Viber, Skype та Telegram, мають необхідність захищати користувачів від широкого спектру загроз: як від масового перехоплення трафіку (Data Retention and Investigatory Powers, скорочено DRIP) так і до таргетованих атак з боку висококваліфікованих зловмисників. Відповіддю на ці виклики стала розробка та впровадження нових криптографічних примітивів та протоколів, які забезпечують властивості наскрізного шифрування (End-to-End Encryption, E2EE), прямої секретності (Forward Secrecy) та пост-компрометаційної безпеки (Post-Compromise Security).

1 SIP та SRTP

Протокол ініціювання сеансу (SIP або стандартизовано RFC 3261), є протоколом прикладного рівня, призначеним для створення, модифікації та завершення мульти-

медійних сеансів. SIP виконує роль "адміністратора"дзвінка: він знаходить користувача, узгоджує можливості (кодеки, порти) та керує станом виклику.

Структура та вразливості: SIP є текстовим протоколом, схожим на HTTP/SMTP. Повідомлення складаються з заголовків та тіла. Критичним недоліком базової реалізації SIP є передача сигнальної інформації у відкритому вигляді. Це дозволяє зловмиснику, що перехоплює трафік (наприклад, через ARP spoofing або на рівні провайдера), отримувати метадані про дзвінок: хто дзвонить (SIP URI), кому, коли і як довго триває розмова.

Для захисту сигналізації використовується TLS (Transport Layer Security), що інкапсулює SIP-пакети в захищений тунель (SIPS). Це забезпечує конфіденційність метаданих на шляху між клієнтом і проксі-сервером (hop-by-hop security). Проте, у класичній SIP-архітектурі проксі-сервер повинен розшифрувати повідомлення для його маршрутизації, що створює точку вразливості на самому сервері. Сучасні месенджери вирішують цю проблему шляхом шифрування ідентифікаторів отримувача (наприклад, Sealed Sender у Signal), роблячи сигналізацію непрозорою навіть для сервера.

Безпосередня передача голосу та відео здійснюється протоколом RTP (Real-time Transport Protocol). Стандартний RTP не має вбудованих засобів захисту. Для забезпечення конфіденційності використовується профіль SRTP (Secure RTP, RFC 3711). SRTP забезпечує шифрування корисного навантаження (payload) RTP-пакету, залишаючи заголовок RTP відкритим (для коректної маршрутизації та обробки джитера).

При цьому використовується:

1. Шифрування через AES у режимі лічільника (AES-CM), що дозволяє шифрувати даних без вирівнювання та з мінімальною затримкою.

2. Для захисту цілісності і автентичності використовується HMAC-SHA1 в оригінальній специфікації

3. Збереження списку індексів отриманих пакетів для захисту від атак повторного відтворення.

При цьому, аспектом SRTP є те, що він не має власного механізму рукостискання. Ключі повинні бути узгоджені зовнішнім протоколом.

Використовуються:

1. Security Descriptions (SDES) - ключ SRTP передається майже у відкриту у тілі SDP-повідомлення SIP. Може використовуватись у корпоративних системах, але компроментація SIP-сервера призводить до втрати і SIP і SRTP.

2. Протокол ZRTP, розроблений для повної незалежності від серверів сигналізації. Він використовує обмін Діффі-Геллмана в медіа-потоці та механізм SAS (Short Authentication String) — коротку фразу, яку співрозмовники звіряють голосом для захисту від MITM.

3. Накладання TLS поверх UDP безпосередньо в медіа-каналі для узгодження ключів. Це дозволяє досягти наскрізного шифрування, оскільки сигнальний сервер не бере участі в обміні ключами. Саме цей підхід став основою для WebRTC та багатьох сучасних месенджерів.

2 Signal

Signal Messanger - один з застосунків, який не можна обговорювати окремо від Signal Protocol. Після випуску в 2013 році, його почали використовувати не тільки у Signal, а і у інших застосунках для обміну повідомленнями.

Основою протоколу є асинхронний протокол узгодження ключів X3DH (Extended Triple Diffie-Hellman). Він вирішує проблему встановлення спільного секрету, коли одна зі сторін знаходиться офлайн.

Кожен клієнт генерує та публікує на сервері набір ключів:

- Identity key IK - довгострокові Curve25519 ключі, підписані Ed25519
- Signed Pre-key SPK - середньостроковий Curve25519 ключ
- One-Time Pre-Keys OPK - набір одноразових ключів, які видаляються з сервера після використання

Також використовується ефемерний ключ Аліси EK_A, який зануляється після використання, для захисту від атак на перегляд пам'яті під час runtime memory inspection.

Коли Аліса хоче написати Бобу, вона завантажує його набір ключів і виконує обчислення спільного секрету

$$SK = KDF(DH(IK_A, SPK_B)||DH(EK_A, IK_B)||DH(EK_A, SPK_B)||DH(EK_A, OPK_B))$$

. Така комбінація гарантує автентифікацію сторін через IK тта пряму секретність через EK та OPK. KDF це Key Derivation Function.

Далі перши повідомленням Аліса відправляє Бобу IK_A, EK_A, ідентифікатори SPK_B, OPK_B, початковий шифртекст зашифрований AEAD, який використовує $AD = E(IK_A)||E(IK_B)$ в якості супутніх даних і або SK, або ключ отриманий за допомогою SK.

Після отримання початкового повідомлення від Аліси, Боб отримує IK_A, EK_A і дістає свої приватні ключі, який відповідають SPK_B та OPK_B. За допомогою них він може переобчислити DH, KDF і дістати SK. Далі він візьме AD і спробує дістати з нього IK_A та IK_B і порівняти їх з існуючими значеннями. Якщо розшифрування не вдалось, Боб відмовляє у доступі і видаляє SK.

Після встановлення початкового секрету вступає в дію алгоритм Double Ratchet, який забезпечує оновлення ключів для кожного повідомлення.

Symmetric-key Ratchet: Для кожного повідомлення в ланцюжку використовується функція KDF, яка на вхід приймає попередній стан ланцюжка (Chain Key). Виходом є новий Chain Key та Message Key. Оскільки KDF є односторонньою функцією, компрометація поточного ключа не дозволяє обчислити попередні (Forward Secrecy).

Diffie-Hellman Ratchet: При зміні напрямку розмови ("пінг-понг") сторони обмінюються новими публічними ключами DH. Це оновлює кореневий ключ (Root Key) обох ланцюжків. Це забезпечує пост-компрометаційну безпеку (Break-in Recovery): навіть якщо зловмисник отримав поточні ключі, після наступного обміну повідомленнями він втрачає можливість дешифрування.

У відповідь на загрозу квантових обчислень (зокрема алгоритму Шора, здатного зламати еліптичну криптографію), Signal впровадив розширення PQXDH (Post-Quantum Extended Diffie-Hellman). Цей протокол додає до стандартного обміну X3DH шар інкапсуляції ключів (KEM) на базі алгоритму Kyber-1024 (стандартизованого NIST як ML-KEM). Спільний секрет тепер залежить як від класичного обміну ECDH, так і від пост-квантового KEM. Це реалізує стратегію "глибокого захисту": для зламу системи зловмиснику потрібно подолати як класичний, так і пост-квантовий алгоритми. результат шару PQXDH конкатенується у KDF після усіх DH.

Унікальною характеристикою Signal є технологія "Sealed Sender". У традиційних системах сервер бачить поле From у заголовку. У Signal відправник шифрує свою ідентичність (разом з тілом повідомлення) ключем, відомим лише отримувачу. Пакет містить лише токен доставки отримувача. Сервер знає, кому призначено повідомлення, але криптографічно не може визначити від кого воно надійшло, якщо не володіє ключами.

3 WhatsApp

WhatsApp використовує протокол Signal як основу своєї системи безпеки, що підтверджено whitepaper компанії та незалежними аудитами. Однак реалізація має специфічні особливості, зумовлені масштабом (2 млрд користувачів) та мультимедійними можливостями.

WhatsApp використовує бібліотеку libsignal-protocol для обміну повідомленнями. При цьому Message Key має розмір 80 байтів: 32 байти використовуються для ключа

AES-256, 32 для HMAC-SHA256, і 16 для вектора ініціалізації IV. Решта працює дуже схоже до Signal.

Інші дослідження, які розглядали мережевий трафік, стверджують, що WhatsApp для дзвінків використовує модифікований SRTP.

Ініціалізація дзвінка відбувається через сигнальне повідомлення (зашифроване протоколом Signal), яке передає 32-байтний "Master Secret"для SRTP. Транспорт здійснюється через UDP порти (діапазон динамічний, часто 3478 для STUN/TURN або випадкові високі порти). На відміну від текстових повідомлень (AES-256), для голосу використовується AES-128-ICM (Integer Counter Mode). Аналіз пам'яті показує, що хоча генерується 46 байтів ключового матеріалу, для шифрування використовується 16 байтів (128 біт) ключа та 14 байтів "Salt".

WhatsApp активно використовує розширення заголовків RTP (RFC 6464) для передачі рівнів гучності аудіо. Це створює певний ризик витоку інформації через VBR (Variable Bit Rate) кодек Opus, оскільки розмір пакету корелює з фонетичним вмістом мови. Хоча вміст зашифровано, аналіз довжини пакетів (Packet Length Analysis) теоретично дозволяє ідентифікувати мову або окремі фрази.

У 2023 році WhatsApp впровадив Auditable Key Directory (AKD) на базі дерев Меркла (Merkle Trees). Це вирішує проблему "Man-in-the-Middle"з боку самого сервера WhatsApp. Раніше користувачі повинні були вручну звіряти 60-значні коди безпеки ("Safety Numbers"). AKD дозволяє клієнту автоматично перевіряти докази включення (inclusion proofs) публічного ключа співрозмовника у загальнодоступний, незмінний журнал. Це математично гарантує, що сервер видає всім користувачам один і той самий ключ для конкретного ідентифікатора, унеможливлюючи приховану підміну ключів для таргетованого перехоплення.

Для реалізації функцій штучного інтелекту без компрометації E2EE, WhatsApp розробив архітектуру "Private Processing". Запити до AI шифруються на пристрої користувача ключем, який доступний лише всередині Довіреного Середовища Виконання (TEE - Trusted Execution Environment) на серверах Meta. TEE гарантує апаратну ізоляцію пам'яті та коду. Після обробки запиту дані видаляються з пам'яті TEE. Це дозволяє надавати хмарні сервіси, зберігаючи модель загроз, де провайдер (Meta) не має доступу до даних користувачів у відкритому вигляді.

4 Viber

Viber має більш закриту архітектуру, ніж у конкурентів, що вимагає ретельного аналізу непрямих ознак та доступної документації.

Згідно з офіційною документацією ("Encryption Overview"), Viber використовує протокол, концептуально схожий на Double Ratchet, але з заміною криптографічних примітивів. ІК такий же, як у Signal, Для шифрування потоків аудіо та відео Vi-

ber використовує потоковий шифр Salsa20 (розроблений Daniel J. Bernstein). Salsa20 обрано замість AES через його високу продуктивність у програмній реалізації на мобільних пристроях та стійкість до атак побічних каналів (timing attacks), оскільки операції Salsa20 (add-rotate-xor) виконуються за константний час. Для кожного дзвінка генерується ефемерна пара ключів Curve25519, виконується рукостискання, і отриманий спільний секрет використовується як ключ для потоку Salsa20.

Аналіз мережевих дампів іншими дослідниками з використанням Wireshark дозволяє виявити характерні сигнатури трафіку Viber:

- Використовуються специфічні порти UDP (5242, 5243) та TCP (4244, 5243, 9785). Це відрізняє його від WhatsApp, який часто маскується під стандартний HTTPS/STUN.
- Медіа-дані не передаються у чистому RTP. Потік виглядає як суцільний потік даних з високою ентропією, що підтверджує використання потокового шифру.
- Для передачі файлів Viber використовує механізм, де файл шифрується симетричним ключем, завантажується на сервер, а ключ передається через захищений канал сигналізації. Відмінною, хоча і дещо застарілою, деталлю є використання MD5 хешу для перевірки цілісності передачі файлу на сервер. Хоча MD5 є криптографічно зламаним (колізії), у даному контексті він використовується лише як контрольна сума (checksum) зашифрованого блоку, а не для підпису чи захисту паролів, тому це не створює прямої вразливості для конфіденційності, проте свідчить про використання застарілих компонентів.

У 2024 році Viber отримав сертифікацію SOC 2 Type II, що підтверджує надійність внутрішніх процесів управління даними. Однак, на відміну від Signal, код Viber не є повністю відкритим, що ускладнює незалежну верифікацію відсутності бекдорів.

5 Skype (взагалі не дуже актуальна секція через те, що Skype більше не функціонує)

Оригінальний протокол Skype був заснований на P2P-архітектурі з власною системою шифрування AES-256 та агресивною обфускацією трафіку, що робило його складним для аналізу та блокування. Після придбання Microsoft, архітектура була змінена на централізовану на основі хмари з використанням протоколу MSNP24 (Microsoft Notification Protocol 24).

У сучасній архітектурі (Consumer Skype & Skype for Business) використовується TLS, для медіа використовується стандартний стек SRTP з AES-256.

Важливо зауважити, що ключі шифрування генеруються на серверах Microsoft і передаються клієнтам. Це забезпечує захист від зовнішнього спостерігача (Encryption

in Transit), але не є E2EE. Microsoft володіє ключами і технічно може дешифрувати контент (наприклад, за запитом правоохоронних органів).

Для користувачів, яким потрібен справжній захист, Skype впровадив функцію "Private Conversations". Ця функція використовує бібліотеку Signal Protocol. Повідомлення інкапсулюються в спеціальний тип пакету messagetype: "EndToEndEncryption/EncryptedText". JSON-структура пакету містить поля encryptedkey (для X3DH) та content (зашифрований AES payload).

6 Telegram

Telegram використовує власний криптографічний протокол MTProto (поточна версія 2.0), який суттєво відрізняється від загальноприйнятих стандартів, що часто викликає критику експертів.

MTProto 2.0 складається з трьох рівнів: API, криптографічного та транспортного. Ключовою особливістю є використання алгоритму AES-256 у режимі IGE (Infinite Garble Extension). Формула розшифрування IGE:

$$m_i = D_K(c_i \oplus m_{i-1}) \oplus c_{i-1}$$

.

Це гібрид режимів CBC та CFB, який забезпечує розповсюдження помилок (error propagation). Критики вказують, що IGE не є стандартом NIST, і його криптографічні властивості менш досліджені, ніж CTR або GCM. Однак, на практиці, атаки на MTProto 2.0, які б дозволили розкрити ключ або текст, на даний момент не продемонстровані.

Важливою зміною в MTProto 2.0 (порівняно з v1.0) став алгоритм обчислення msg_key. Тепер це середні 128 біт від SHA-256 тіла повідомлення (включаючи padding), до якого додано фрагмент ключа авторизації. Це захищає від атак типу IND-CCA (Chosen Ciphertext Attack), які були теоретично можливі в першій версії, де msg_key залежав лише від повідомлення.

Голосові дзвінки в Telegram захищені E2EE завжди, навіть у звичайних чатах (на відміну від текстових повідомлень, які за замовчуванням хмарні). Використовується open-source бібліотека libtgvoip. Для узгодження ключів використовується класичний обмін Діффі-Геллмана. Захист від MITM реалізовано через візуалізацію хешу ключа у вигляді 4 емодзі. Користувачі повинні звірити їх голосом.

Пакет UDP складається з

- заголовку voice_call_id (128 біт) або peer_tag (при використанні рефлекторів).
- Message key розміром 128 біт

- Голосові дані (Opus), зашифровані AES-256-IGE.

Якщо P2P з'єднання неможливе (NAT traversal) або користувач бажає приховати IP-адресу, то Telegram буде використовувати мережу серверів-рефлекторів для маршрутизації дзвінків. Рефлектор лише пересилає UDP-пакети, не маючи ключів для їх розшифрування.

7 Існуючі дослідження

Оскільки не всі розглянуті системи мають відкритий вихідний код, важливим етапом дослідження є теоретичне доведення наявності заявлених механізмів через аналіз побічних каналів та структури трафіку. Далі будуть надані результати аналізу інших дослідників:

- Вимірювання ентропії корисного навантаження пакетів Viber, WhatsApp та Telegram показує значення, близькі до максимуму (8 біт на байт). Це підтверджує застосування сильних криптографічних алгоритмів (AES, Salsa20) і відсутність структурних закономірностей, характерних для відкритого тексту або слабкого шифрування.

- При аналізі трафіку Signal та WhatsApp спостерігається, що навіть при відсутності активності користувача відбувається періодичний обмін службовими пакетами невеликого розміру. Це відповідає роботі алгоритму Double Ratchet (оновлення ланцюжків ключів). Крім того, при перехопленні сесії і спробі розшифрувати її старим ключем, дешифрування провалюється, що емпірично доводить властивість Forward Secrecy.

- Пакети Telegram мають специфічну структуру довжини та заголовків (відсутність стандартних RTP заголовків у UDP потоці), що дозволяє однозначно ідентифікувати протокол MTProto навіть без глибокої інспекції пакетів (DPI).

Існуючі вразливості:

- Вразливість у версії Telegram для Android (до 10.14.4), виявлена ESET. Зловмисник міг створити спеціально сформований файл (APK), який відображався в чаті як попередній перегляд відео. При спробі відтворення користувачеві пропонувалося встановити "зовнішній плеєр який насправді був шкідливим додатком. Це вразливість логіки обробки медіа-контейнерів, а не криптографії MTProto, проте вона підкреслює важливість безпеки на рівні клієнтського ПЗ.

- У ранніх версіях Viber зображення передавалися зашифрованими, але посилання на них перехоплювалися у відкритому вигляді. Ця вразливість була усунена з впровадженням повноцінного E2EE у версії 6.0.

- Дослідження 2021 року (Miculan & Vitacolonna) підтвердило основні властивості безпеки, але виявило теоретичну можливість атаки UKS (Unknown Key-Share), яка, втім, не дозволяє читати повідомлення.

  Натомість, Signal Protocol пройшов багаторазові аудити (університети Оксфорда, Квінсленда) з використанням моделей ProVerif та Tamarin, які математично довели гарантії PFS та PCS.

8 Порівняння

- Signal займає позицію технологічного лідера завдяки впровадженню пост-квантового захисту (PQXDH) та мінімізації метаданих (Sealed Sender). Це єдина система, яка захищає не лише зміст, а й факт спілкування.

- WhatsApp є найбільш збалансованим рішенням для масового користувача. Використання того ж ядра Signal Protocol забезпечує високу стійкість. Впровадження Key Transparency (AKD) та Private Processing (TEE) демонструє рух у бік підвищення довіри без розкриття даних серверу.

- Telegram залишається не доведеним у стійкості. Використання власної криптографії (IGE mode) є спірним рішенням. Головний ризик полягає не в протоколі передачі, а в архітектурі зберігання: звичайні чати зберігаються на серверах, що робить їх вразливими до вилучення ключів. Однак голосові дзвінки захищені надійно. При цьому це не виключає можливості самим Telegram мати доступ до повідомлень.

- Viber обрав шлях оптимізації продуктивності (Salsa20), що є виправданим для VoIP на слабких каналах. Проте закритість протоколу вимагає більшого рівня довіри до розробника, аналогічно до Telegram.

- Skype (без Private Conversations) не відповідає сучасним вимогам до захищених месенджерів через доступ Microsoft до ключів.

9 Рекомендації

Найкращим варіантом на ринку наразі є Signal. Було доведено, що навіть якби розробники намагались би передати якісь дані від користувачів, то вони не могли б це зробити, бо мають доступ до дуже обмеженої інформації.

WhatsApp також є достатньо надійним вибором, оскільки він використовує схожий підхід, однак необхідно включити сповіщення безпеки та використовувати двофакторну автентифікацію.

Skype станом на 2025 рік був інтегрований у Teems і відповідно не доступний для використання.

Viber та Telegram не рекомендуються для поширення приватної інформації, оскільки відсутній чіткий доказ захищеності даних, і є ймовірність, що дані і переписки доступні для власників Telegram та Viber.