Zakázaný útok na GCM

M. Stanek

KI FMFI UK

Kryptológia (2), LS 2019

GCM

Zakázaný útok (znovupoužitie IV)

AES-GCM-SIV

Úvod

- GCM (Galois/Counter Mode) – mód pre autentizované šifrovanie
  - presnejšie "authenticated encryption with associated data" (AEAD)
- AES-GCM – momentálne najpoužívanejší AEAD algoritmus
  - TLS, SSH, IPSec, …
  - TLS 1.3 – len AEAD konštrukcie (RFC 8446):
    A TLS-compliant application MUST implement the TLS_AES_128_GCM_SHA256 cipher suite and SHOULD implement the TLS_AES_256_GCM_SHA384 and TLS_CHACHA20_POLY1305_SHA256 cipher suites.
- kombinácia CTR módu pre dôvernosť a (relatívne) jednoduchého výpočtu autentizačného tagu
- špecifikácia: *Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC* (NIST SP 800-38D, 2007)

GCM (1)

- AES-GCM, 128 bitový blok
- najčastejší variant: 96 bitový IV (12B), 32 bitové počítadlo (4B)
- označenia:
  - K – kľúč
  - P, A, C – otvorený text, asociované dáta, šifrový text
  - H – "autentizačný kľúč" pre výpočet autentizačného tagu
  - $J_0 = IV \,||\, 0^{31} \,||\, 1$;
  - $\text{len}(X)$ – dĺžka X v bitoch, zapísaná ako 64 bitová hodnota
- šifrovanie v CTR móde:
  1. $\text{ctr} = \text{inc}_{32}(J_0)$ (pripočítanie 1 modulo 2^{32} k posledným 4B)
  2. $P \mapsto X_1, \ldots, X_n$ (posledný blok nemusí byť úplný)
  3. pre $i = 1, \ldots, n$:
     - $C_i = P_i \oplus E_K(\text{ctr})$
     - $\text{ctr} = \text{inc}_{32}(\text{ctr})$
  4. výstup: $C_1, \ldots, C_n$, pričom $|C_n| = |P_n|$

GCM (2)

- výpočet $\mathrm{GHASH}_H(A, C)$:
  1. $H = E_K(0^{128})$
  2. $A \,||\, C \;\mapsto\; X_1, \ldots, X_{n-1}, \underbrace{\mathrm{len}(A) \,||\, \mathrm{len}(C)}_{X_n}$

     A aj C sú v prípade potreby doplnené 0 na najbližší celý blok
  3. $Y_0 = 0^{128}$
  4. pre $i = 1, \ldots, n$: $Y_i = (Y_{i-1} \oplus X_i) \bullet H$
  5. $\mathrm{GHASH}_H(A, C) \leftarrow Y_n$
- výpočet autentizačného tagu T: $T = E_K(J_0) \oplus \mathrm{GHASH}_H(A, C)$
- symbol $\bullet$ je násobenie v $\mathrm{GF}(2^{128})$ generovanom ireducibilným polynómom $x^{128} + x^7 + x^2 + x + 1$

GCM poznámky

- obmedzená veľkosť správ (nielen vyčerpaním počítadla)
  - dlhšie správy zvyšujú pravdepodobnosť úspechu útočníka
  - detaily v NIST, resp. v referenciách TLS 1.3
- IV má byť "nonce", teda sa nesmie zopakovať pri rovnakom kľúči pre rôzne správy
- zopakovanie IV (zakázaný útok):
  - two-time pad problém pre CTR šifrovanie (synchrónna prúdová šifra)
  - získanie autentizačného kľúča H pre autentizačnú časť GCM
- znalosť H umožní
  - manipuláciu šifrového textu (preklápanie bitov)
  - ľubovoľnú úpravu asociovaných dát
  - správny tag vieme dopočítať

Zakázaný útok – získanie H

- "forbidden attack" (A. Joux)
- predpoklad: dve správy šifrované s rovnakým kľúčom K a inicializačným vektorom IV
  - H je rovnaké v oboch prípadoch, keďže $H = E_K(0^{128})$
  - $E_K(J_0)$ je rovnaké v oboch prípadoch (označme J^*)
- pre zvýšenie prehľadnosti preznačme $\oplus \mapsto +$ a $\bullet \mapsto \cdot$
- výpočet T sa dá napísať ako polynóm $g(z)$:

$$g(z) = J^* + z \cdot X_n + z^2 \cdot X_{n-1} + \ldots + z^n \cdot X_1$$

  vyhodnotený v bode H: $T = g(H)$
- útočník pozná T, A, C, kde $A \,||\, C \;\mapsto\; X_1, \ldots, X_{n-1}, \text{len}(A) \,||\, \text{len}(C)$
- neznáma je hodnota H a útočník nepozná ani J^*

Zakázaný útok – získanie H (pokr.)

- pre dve správy použitý rovnaký IV, dostávame dva polynómy

$$g(z) = J^* + z \cdot X_n + z^2 \cdot X_{n-1} + \ldots + z^n \cdot X_1$$
$$g'(z) = J^* + z \cdot X'_n + z^2 \cdot X'_{n'-1} + \ldots + z^{n'} \cdot X'_1$$

- H je koreňom $g(z) + T$ a $g'(z) + T'$, teda aj ich súčtu:

$$g(z) + T + g'(z) + T',$$

  čo je polynóm stupňa $\max\{n, n'\}$ v ktorom poznáme všetky koeficienty (J^* vypadne)
- hodnotu H nájdeme faktorizáciou polynómu, získaním koreňov a ich overením pre ďalšie správy
  - v prípade väčšieho počtu správ s rovnakým IV máme viac polynómov so spoločným koreňom
  - koreňov potenciálne až po stupeň poynómu, v praxi podstatne menej

Zakázaný útok v praxi

- Hanno Böck a kol.: *Nonce-Disrespecting Adversaries: Practical Forgery Attacks on GCM in TLS*, 2016
- konštrukcia IV (12 B) v TLS 1.2:
  - 4 B soľ odvodená v rámci TLS handshake (server_write_IV / client_write_IV)
  - 6 B nonce, explicitná časť IV
  - implementácia má zabezpečiť unikátnosť
- rovnaké nonce pre rôzne správy znamenajú rovnaké IV
- poblém s náhodným nonce: 64 bitov, zvýšená pravdepodobnosť kolízie pri veľkom počte správ
  - cca. 70.000 HTTPS serverov/zariadení v januári 2016
  - napriek tomu nie bežná situácia pre útok
  - IBM Lotus Domino, A10 load balancer a pod.
- duplicitné nonce
  - 184 HTTPS serverov/zariadení (patriace napr. VISA, Deutsche Börse)
  - Radware load balancer
  - chýbajúca kontrola návratovej hodnoty OpenSSL (z externého RNG)

AES-GCM-SIV

- redukcia problému s opakovaním IV
- SIV (synthetic IV)
  - odvodenie kľúča pre šifrovanie a autentizačného kľúča z IV (12 B) a kľúča
  - "hash-then-encrypt"
  - najskôr sa z IV, otvoreného textu a asociovaných dát vypočíta autentizačný tag
  - namiesto GHASH použitý POLYVAL (trocha rýchlejší variant)
  - tag je neskôr použitý na šifrovanie v CTR móde
- dva prechody $\Rightarrow$ pomalšie ako AES-GCM
- iné konštrukcie dosahujúce (aj) odolnosť pri opakovaní IV
  - AEZ, GCM-SIV (trocha iné ako AES-GCM-SIV)
  - kľúčový pojem: "nonce misuse resistant"