Kriptografija i kriptoanaliza — Međuispit

5. prosinca 2024.

Ispit traje **2 sata** i nosi ukupno **30 bodova**. Zadaci se rješavaju na ovom obrascu, a ako nedostaje mjesta, možete koristiti košuljicu ili dodatne papire. Nisu dozvoljeni nikakvi materijali osim **kalkulatora**. Molimo vas da pišete **čitko**; nečitka rješenja nećemo razmatrati.

Ime i prezime: _____ JMBAG: _____ Dvorana: _____

1. (4 boda) *Klasična kriptografija.* Dešifrirajte poruku

   `IIZAPEBUIRIZONNGRIDEAOKLMJTTIAICSPENLORIJAEIVRIOŽTZKFT`

   ako je poznato da je dobiven stupčastom transpozicijom koristeći ključ `987654321`.

```
2024./2025.: trece slovo T s kraja bilo je viska. Priznavati rjesenja.

Ključ ima 9 brojeva, a šifrat 54 znaka pa šifrat treba podijeliti na skupove od 54/9=6 znakova
"IIZAPE BUIRIZ ONNGRI DEAOKL MJTTIA ICSPEN LORIJA EIVRIO ŽTZKFT"
 IIZAPE BUIRIZ ONNGRI DEAOKL MJTTIA ICSPEN LORIJA EIVRIO ŽTZKFT

koji se popunjavaju po stupcima prema ključu:
prvo IIZAPE ide u stupac 1:
987654321
        I
        I
        Z
        A
        P
        E
itd. do zadnjeg stupca koji je na mjestu 9 (prvi stupac):

987654321
ŽELIMDOBI
TIOCJENUI
ZVRSTANIZ
KRIPTOGRA
FIJEIKRIP
TOANALIZE

Izvorni tekst: ŽELIMDOBITIOCJENUIZVRSTANIZKRIPTOGRAFIJEIKRIPTOANALIZE
```

2. (7 bodova) *Generatori pseudoslučajnih brojeva.*

   Neka je $G\colon K \to \{0,1\}^n$ generator pseudoslučajnih brojeva (PRG) i $A\colon \{0,1\}^n \to \{0,1\}$ statistički test.

   (a) (1) Formalno definirajte prednost statističkog testa A u odnosu PRG G.
   Neka je $G\colon K \to \{0,1\}^n$ PRG i $A\colon \{0,1\}^n \to \{0,1\}$ statisticki test. Prednost (advantage) statističkog testa A definiramo kao: $\mathsf{Adv}_{\mathsf{PRG}}(A,G) :== |P_{k\leftarrow K}(A(G(k)) = 1) - P_{x\leftarrow\{0,1\}^n}(A(x) = 1)|$

   (b) (1) Kad kažemo da je G siguran PRG?
   Ako svaki efikasni napadač (statisticki test) ima zanemarivo malu prednost.

   (c) (1) Koja je veza između (ne)predvidivosti i sigurnosti PRG-ova.
   PRG je siguran ako i samo ako je nepredvidiv.

   (d) (1) Nabrojite barem dvije protočne šifre korištene u praksi od kojih je barem jedna sigurna. Ne trebate ih opisivati.
   siguran: Salsa20/ChaCha (siguran); nesigurni: RC4, CSS, Linear congruental generator (LCG)

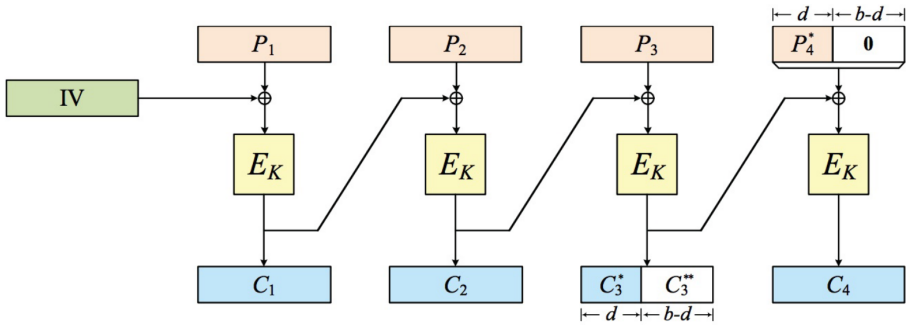
(e) (3) Pretpostavite da je G siguran generator pseudoslučajnih brojeva koji generira nizove bitova u skupu $\{0,1\}^n$. Zaokružite i obrazložite koji od sljedećih generatora pseudoslučajnih brojeva, izvedenih iz G, su nesigurni? Ovdje je "$\|$" operacija nadovezivanja (konkatenacije), "$\wedge$" bitovna logička konjunkcija (*bitwise* AND), a "$\oplus$" operacija ekskluzivno-ili (XOR).

i. $G_1(s_1 \,\|\, s_2) \coloneqq G(s_1) \oplus G(s_2)$.
ii. $G_2(s) \coloneqq G(s) \,\|\, G(s)$.
iii. $G_3(s) \coloneqq G(s) \oplus 1^n$.
iv. $G_4(s) \coloneqq G(s) \oplus G(s)$.
v. $G_5(s_1 \,\|\, s_2) \coloneqq G(s_1) \wedge G(s_2)$.
vi. $G_6(s_1 \,\|\, s_2) \coloneqq s_1 \,\|\, G(s_2)$.

i. Ne
ii. Da. Napadac moze usporediti prvi i drugi dio.
iii. Ne. To je samo komplement
iv. Da. Sve se XOR-a u 0.
v. Da. Za jedinicu u G_4 trebamo imati jedinicu u $G(s_1)$ i jedinicu u $G(s_2)$. Dakle, udio jedinica je biti 1/4, a kod slucajnog niza priblizno 1/2.
vi. Ne.

3. (5 bodova) *Varijante načina kriptiranja CBC.*

(a) (2) Jedan problem s načinom kriptiranja CBC je da poruke moraju biti nadopunjene kako bi bile višekratnik duljine bloka. Također, ponekad je potrebno dodati jedan cijeli *dummy* blok na kraj kako bi se razriješila nejednoznačnost prilikom dekripcije (npr. PKCS#7). Sljedeća slika opisuje varijantu načina kriptiranja CBC koja uklanja potrebu za nadopunjavanjem:

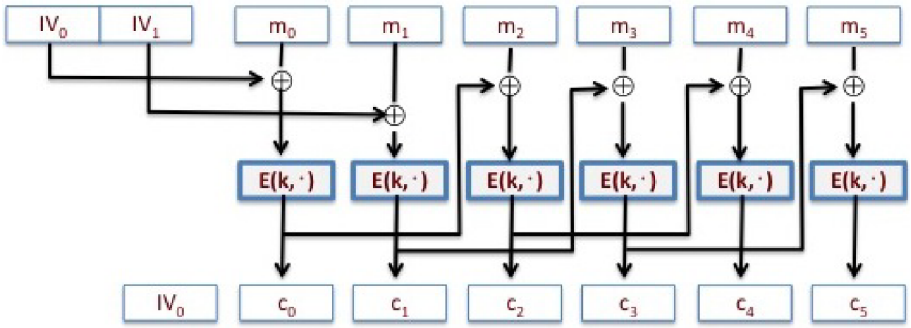


Ova varijanta zadnji blok nadopunjuje nulama ako je potrebno (*dummy* blok se nikad ne dodaje na kraj), no skriveni tekst uključuje samo osjenčane dijelove: C_1, C_2, C_3^* i C_4. Primjetite da, ako izuzmemo inicijalizacijski vektor (IV), skriveni tekst i jasni tekst su jednake duljine. Ova tehnika se naziva *ciphertext stealing*.

Neka je $E = \mathsf{AES128}$, $n \in \mathbb{N}$ duljina skrivenog teksta, $N \in \mathbb{N}$ broj primljenih blokova i $b = 128$ duljina bloka u bitovima. Opišite algoritam dekriptiranja. Dakle, kako primatelj računa duljinu "preotetog" šifrata (npr. C_3^{**} na slici) i dekriptira cijelu poruku.

i. Dekriptira se prvih $N-2$ blokova na standardan način.
ii. Izračunaj duljinu preotetog šifrata: $d = n - (N-1)b \implies b - d = b - n + (N-1)b$
iii. Dekriptiraj zadnji blok C_N: $E_N = \mathsf{E}_K^{-1}(C_N)$
iv. $P_N \,\|\, C_{N-1}^* = E_N \oplus (C_{N-1} \,\|\, \underbrace{00\ldots0}_{b-d})$
v. Dekriptiraj $P_{N-1} = \mathsf{E}_K^{-1}(C_{N-1} \,\|\, C_{N-1}^*) \oplus C_{N-2}$

(b) (3) Drugi problem s načinom kriptiranja CBC je da se ne može ubrzati paralelnom obradom. Sljedeća slika prikazuje varijantu načina kriptiranja CBC koja podržava dvostruku paralelizaciju: može se ubrzati dvostruko koristeći dva procesora.



Ovdje opet pretpostavljamo da je E neka **sigurna** blok šifra poput AES128-a. Pretpostavimo da se IV_0 bira nasumično i da je $\mathsf{IV}_1 = \mathsf{IV}_0 \oplus \mathsf{B}$, gdje je B neka **javna** konstanta poznata napadaču (npr. $\mathsf{B} = 1^n$).

Pokažite da je ova varijanta načina kriptiranja CBC **ne pruža** semantičku sigurnost. Točnije, opišite napadača koji s velikom vjerojatnošću može pobijediti izazivača u sljedećoj igri:

i. Izazivač izabere nasumični ključ k i nasumični bit b
ii. Napadač odabere dvije proizvoljne poruke p_0 i p_1, obje duljine 32 bajta (dakle dva bloka), i pošalje ih izazivaču
iii. Izazivač izračuna $c = E(k, p_b)$ i pošalje napadaču
iv. Napadač ispiše bit $\hat{b}$ i pobjeđuje u igri ako je $b = \hat{b}$

Koliku prednost ima napadač u navedenoj igri?
Uzmimo proizvoljne $m_0, m_1 \in \mathcal{M}$, takve da $m_0 \neq m_1$.
$\mathsf{m}_{00} = m_0 \,\|\, m_0 \oplus \mathsf{B}$
$\mathsf{m}_{01} = m_0 \,\|\, m_1 \oplus \mathsf{B}$
U eksperimentu 0:
$c_0 = (c_{00}, c_{01})$
$c_{00} = \mathsf{E}(k, m_0 \oplus \mathsf{IV}_0)$, $c_{01} = \mathsf{E}(k, m_0 \oplus \mathsf{B} \oplus \mathsf{IV}_1) = \mathsf{E}(k, m_0 \oplus \mathsf{IV}_0) = c_{00}$
U eksperimentu 1:
$c_0 = (c_{00}, c_{01})$
$c_{00} = \mathsf{E}(k, m_0 \oplus \mathsf{IV}_0)$, $c_{01} = \mathsf{E}(k, m_1 \oplus \mathsf{B} \oplus \mathsf{IV}_1) = \mathsf{E}(k, m_1 \oplus \mathsf{IV}_0) \neq c_{00}$, zbog $m_0 \neq m_1$.

Program napadača: `if` $\mathsf{c}_{00} = \mathsf{c}_{01}$ `:` $\hat{\mathsf{b}} = 1$ `else` $\hat{\mathsf{b}} = 0$
Prednost napadača: $|P(W_0) - P(W_1)| = |1 - 0| = 1$

4. (5 bodova) *Funkcije sažetka (hash funkcije) i autentifikacijske značke (MAC).*
Neka je $F\colon M \to \{0,1\}^{128}$ funkcija koja je otporna na kolizije i poznata napadaču. Obrazložite sljedeće tvrdnje.

(a) (2) Je li sljedeća funkcija otporna na kolizije: $G(x) = F(x)[0 \ldots 31]$ (tj. uzmi prva 32 najznačajnija bita izlaza)?
Ne, napadac mora samo iterirati 2^{32} puta i sigurno ce za dvije razlicite poruke dobiti isti sazatek (pigeonhole principle).

(b) (1) Definira li sljedeća funkcija siguran MAC: $H(x,k) \coloneqq F(x) \oplus k$?
Ne. Ako je napadaču poznata poruka m, lagano dolazi do ključa: $k = F(x,k) \oplus H(x)$

(c) (2) Je li funkcija $f\colon \{0,1\}^{128} \times \{0,1\}^{128} \times \{0,1\}^{128} \to \{0,1\}^{128}$, $f(k, x_0, x_1) \coloneqq \mathsf{AES128}(\mathsf{AES128}(x_0, k) \oplus x_1, k)$? otporna na kolizije? Napadaču je poznat ključ k.
Nije otporna na kolizije. Uzmimo $x_0 \neq x_2$ proizvoljne i definirajmo $x_1 = \mathsf{AES128}(x_0, k)$ i $x_3 = \mathsf{AES128}(x_2, k)$. Tada ce za obje vrijediti AES128(0,k). Drugi nacin je da jednostavno iskoristimo komutativnost od XOR

5. (9 bodova) *Razna pitanja.*

(a) (1) Kako izgleda ključ u Playfairovoj šifri? Obrazložite.
ključ je niz slova/znakova koji se transformira u matricu 5×5 slova tako da se svako slovo pojavljuje samo jednom, a matrica se popunjava po retcima prco sa slovima iz inicijalnog ključa, a zatim s preostalim slovima redom iz alfabeta.

(b) (1) Koji je najveći nedostatak Hillove šifre sa simetričnim ključem?
Premali prostor ključa jer mali broj matrica zadovoljava $K = K^{-1}$; dovoljno je reći samo premali prostor ključa

(c) (1) Navedite Kerckhoffov princip.
Kriptosustav mora biti siguran i onda kada su sve informacije o kriptosustavu javno dostupne, osim tajnog ključa.

(d) (1) Što znači da jednokratna bilježnica pruža savršenu povjerljivost?
Ako je $k \in \{0,1\}^n$ odabran slučajno i uniformno ($k \leftarrow \{0,1\}^n$), tada za svaku poruku $m \in \{0,1\}^n$ i za svaki šifrat $c \in \{0,1\}^n$ vrijedi sljedeće: $P_{k \leftarrow \{0,1\}^n}(E(m,k) = c) = 1/2^n$

(e) (2) Kolika je složenost napada grubom silom na kriptosustav $\mathsf{2DES}(m, k_1, k_2) = \mathsf{DES}(\mathsf{DES}(m, k_1), k_2)$ ako se primijeni napad susret u sredini (*Meet-in-the-middle attack*) za poznati jedan par (m, c) tj. jedan jasni i kriptirani tekst? Obrazložite.

Složenost napada grubom silom svodi se na složenost 2^{57} zato što je potrebno izračunati i popamtiti sve moguće kriptirane tekstove c' koji se dobivaju kriptiranjem jasnog teksta m sa svim mogućim ključevima $c' = \mathsf{DES}(m, k_{1i})$ kojih ima 2^{56} i potrebno je izračunati i popamtiti sve moguće jasne tekstove koji se dobivaju dekriptiranjem kriptiranog teksta $c : m' = \mathsf{DES}^{-1}(c, k_{2i})$ za sve moguće ključeve kojih ima jednako kao i u prethodnom slučaju 2^{56}. Potom treba pronaći jasni tekst m'_j koji je jednak nekom kriptiranom tekstu c'_l, tj. tako da vrijedi $m'_j = c'_l$ čime je otkriven ključ (k_j, k_l). Složenost tog napada je 2^{57}.

3DES je ranjim na *meet-in-the-middle* napad. Potreban je donekle razuman opis napada.

(f) (1) Navedite dva načina kriptiranja bloka u kojima blok skirvenog teksta ne ovisi blokovima skrivenog teksta koji mu prethode.

CTR, OFB

(g) (1) Algoritmi zasnovani na konstrukciji Merkle–Damgård ranjivi su na napad na tajni sufiks poruke (*Length Extension Attack*). Ukratko opišite napad.

Napadač na temelju poznatog sažetka $H(M_1)$ i duljine poruke M_1, a bez da poznaje poruku M_1, može umetnuti dodatne podatke na kraj poruke M_1, tj. može dodati proizvoljnu dodatnu poruku M_2 i izračunati $H(M_1 \parallel M_2)$.

(h) (1) Kod načina kriptiranja *Counter Mode* (CTR), zašto nije dobro koristiti isti inicijalizacijski vektor s istim ključem za kriptiranje dvije različite poruke m_1 i m_2?

Zbog toga što ćemo dobiti relaciju $c_1 \oplus c_2 = m_1 \oplus m_2$ i jezik ima dovoljno redundantnosti da dođemo do sadržaja tih poruka.