14-ДӘРІС. ЭНТРОПИЯ ЖӘНЕ ТІЛДІҢ АРТЫҚТЫҒЫ. БІРЛІК АРАЛЫҚ

- Ақпарат мөлшері және Шеннон энтропиясы
- Тілдің артықтығы, кілт энтропиясы және бірлік аралық

SEC1811 «Ақпаратты криптографиялық қорғау негіздері»

Оқытушы: Ахмедиярова А.Т.

Дәрістің оқу нәтижелері

- Ақпарат мөлшерін және энтропияны есептеу
- Максимал энтропияны анықтау
- Абсолюттік және салыстырмалы артықтықты табу
- Бірлік аралықтың теориялық мағынасын түсіндіру

Ақпарат мөлшері

$$I(x) = -\log_2 p(x)$$

- Сирек оқиға көбірек ақпарат береді.
- p(x) — оқиғаның ықтималдығы; өлшем бірлігі — бит.

Ақпарат мөлшерінің мысалдары

$$I(1/8) = -\log_2(1/8) = 3 \text{ бит}$$

- $p = 1 \rightarrow I = 0$ бит.
- $p = 1/2 \rightarrow I = 1$ бит.
- $p = 1/8 \rightarrow I = 3$ бит.

Энтропия анықтамасы

$$H(X) = -\Sigma\, p(x_i)\, \log_2 p(x_i)$$

- Энтропия — мүмкін оқиғалардың орташа ақпарат мөлшері.
- Жеке I(x) мен орташа H(X) ұғымдарын ажыратыңыз.

Тең ықтималды символдар

$$H_{max} = \log_2 m$$

- m символдың ықтималдығы бірдей болса, энтропия максимал.
- m = 4 үшін $H_{max} = \log_2 4 = 2$ бит/символ.

Энтропия есебі: шарт

$$H(X) = -0{,}75 \log_2 0{,}75 - 0{,}25 \log_2 0{,}25$$

- X: 0 немесе 1.
- p(0) = 0,75; p(1) = 0,25.
- Ықтималдықтардың қосындысы 1-ге тең.

Энтропия есебі: шешімі

$$H(X) \approx 0{,}311 + 0{,}500 = 0{,}811 \text{ бит}$$

- Бірінші қосылғыш ≈ 0,311.
- Екінші қосылғыш = 0,500.
- Нәтиже ≈ 0,811 бит; максимум 1 биттен төмен.

Ықтималдық пен энтропия

p(0)	p(1)	H(X), бит
0,50	0,50	1,000
0,75	0,25	0,811
0,90	0,10	0,469
1,00	0,00	0,000

Табиғи тілдің артықтығы

- Әріптер бірдей жиілікпен кездеспейді.
- Сөз, грамматика және контекст келесі символды болжауға көмектеседі.
- Классикалық шифрларда мұндай заңдылықтар байқалуы мүмкін.

Абсолюттік артықтық

$$D = H_{max} - H$$

- H_{max} — максимал, H — нақты энтропия.
- Мысалы: $H_{max} = 4$, $H = 3 \rightarrow D = 1$ бит/символ.

Салыстырмалы артықтық

$$R = 1 - H/H_{max}$$

- R = 1 − 3/4 = 0,25 = 25%.
- R артқан сайын модельдегі болжанатын құрылым үлесі көбейеді.

Артықтық есебі: толық мысал

$$H_{max} = 4;\ D = 1;\ R = 25\%$$

- $m = 16 \rightarrow H_{max} = \log_2 16 = 4$.
- $H = 3 \rightarrow D = 4 - 3 = 1$.
- $R = 1 - 3/4 = 0{,}25 = 25\%$.

Кілт энтропиясы

$$H(K) = \log_2 K$$

- K бірдей ықтималды кілт нұсқасы болса, $H(K) = \log_2 K$.
- n биттік біркелкі кілт үшін H(K) = n бит.

Кілт ұзындығы мен энтропия

- 8 биттік біркелкі кілт: 256 мүмкіндік → 8 бит энтропия.
- Болжанатын үлгімен таңдалған кілттің нақты энтропиясы төмен болуы мүмкін.

Бірлік аралық деген не?

- Unicity distance — классикалық модельдегі бірегей кілтті ажыратуға қатысты теориялық шама.
- Бұл нақты кілтті табу уақыты немесе шабуыл кепілдігі емес.

Бірлік аралық формуласы

$$N_0 \approx H(K) / D$$

- H(K) — кілт энтропиясы, D — бір символға келетін артықтық.
- Формула модельдің болжамдарына тәуелді.

Бірлік аралық: есептеу

$$N_0 \approx 20 \text{ символ}$$

- H(K) = 40 бит, D = 2 бит/символ.
- $N_0 \approx 40/2 = 20$ символ.
- 20 символ нақты шабуылмен кілт міндетті табылады деген сөз емес.

Артықтықтың әсері

- $H(K) = 40$ бит болғанда $D = 1 \rightarrow N_0 \approx 40$.
- $D = 2 \rightarrow N_0 \approx 20$; $D = 4 \rightarrow N_0 \approx 10$.
- Артықтық артса, осы модельдегі N_0 азаяды.

Криптоталдаумен байланыс

- Моноалфавиттік алмастыру шифрында әріп жиілігі сақталуы мүмкін.
- Ұзын мәтін статистикалық заңдылықты бағалауға көбірек материал береді.
- Қазіргі шифрлар статистикалық іздерді жасыруға бағытталады.

Негізгі формулаларды салыстыру

Шама	Формула
Ақпарат мөлшері	$I(x)=-\log_2 p(x)$
Энтропия	$H=-\Sigma p \log_2 p$
Максимум	$H_{max}=\log_2 m$
Артықтық	$D=H_{max}-H$
Бірлік аралық	$N_0 \approx H(K)/D$

Өзіндік есеп

- p(A)=1/2, p(B)=1/4, p(C)=1/4 болса, H(X) табыңыз.
- Жауабы: 0,5 + 0,5 + 0,5 = 1,5 бит.
- $H_{max} = \log_2 3 \approx 1{,}585$ бит.

Қорытынды

- Энтропия белгісіздіктің орташа мөлшерін көрсетеді.
- Тіл артықтығы классикалық криптоталдауда маңызды.
- Бірлік аралық — ақпараттық-теориялық жуық баға, шабуыл кепілдігі емес.

Бақылау сұрақтары

- I(x) нені білдіреді?
- Энтропия қалай есептеледі?
- Энтропия қашан максимал?
- H_max формуласы қандай?
- Тіл артықтығы неден пайда болады?
- Абсолюттік артықтықты қалай табамыз?
- Салыстырмалы артықтық қалай есептеледі?

Әдебиеттер

- Ахметов Б.С. және т.б. Қолданбалы криптология. Алматы, 2016.
- Ахметов Б.С. және т.б. Криптография негіздері. Алматы, 2020.
- Васильева И.Н. Криптографические методы защиты информации. 2024.
- Лось А.Б. және т.б. Криптографические методы защиты информации. 2024.
- Омассон Ж.-Ф. О криптографии всерьез. 2021.