تمرین تلوریم

تلوریم (Tellurium) شبه فلزی با عدد اتمی ۵۲ است، این عنصر درصد فراوانی کمی در پوسته زمین دارد و ریشه نام آن تلور در زبان یونانی به معنای زمین یا نشأت گرفته از زمین است.

نسخه 2.0 - شهریور ۱۴۰۳

علی مومن - ابوالفضل دیانت

فهرست مطالب

فهرست اشکال ii

فصل ۱ نکات ۱

فصل ۲ رمزهای متقارن ۲

۱.۲ سوالات تئوری ۲

۲.۲ سوالات عملی ۶

فصل ۳ نهان‌نگاری و نشان‌گذاری ۱۰

۱.۳ دیباچه ۱۰

۲.۳ سوالات تئوری ۱۱

۳.۳ سوالات عملی ۱۲

مراجع ۱۴

فهرست اختصارات ۱۵

واژه نامه انگلیسی به فارسی ۱۷

واژه نامه فارسی به انگلیسی ۱۹

فهرست تصاویر

۱.۲  ساختار درختی کلیدها ۳

۲.۲ نمایی از معماری شبکه‌های نسل دو با در نظر گرفتن حملات فعال ۴

۳.۲  از چپ به راست: Adi Shamir, Ron Rivest, Len Adleman, Ralph Merkle, Martin Hellman, Whit Diffie ۵

۴.۲  Claude Elwood Shannon ۶

۵.۲  مفاهیم مرتبط با پارامتر اِنتروپی ۶

۶.۲  نمایی از رمز Vernum یا همان One-time pad ۷

۷.۲  Blaise de Vigenère ۸

۱.۳  زیرشاخه‌های علم نهان‌سازی اطلاعات ۱۱

۲.۳  سطح بیت آخر یک تصویر خاکستری ۱۲

۱ نکات

در تحویل این تمرین به نکات زیر دقت کنید:

- پاسخ تمرین‌ها می‌بایست به صورت یک گزارش به زبان فارسی و در قالب LaTeX باشد. کدهای پیاده‌سازی را نیز در کنار فایل گزارش قرار دهید و به صورت یک فایل با پسوند zip، در سامانه قرار دهید.
- تمامی پیاده‌سازی‌ها می‌بایست با زبان Python صورت پذیرد.
- تمرین‌ها به صورت گروهی و تنها توسط یک نفر از اعضای گروه تحویل داده می‌شود. نیازی نیست هر یک از اعضای گروه جداگانه تمرین را در سامانه قرار دهد.
- به هیچ‌وجه کپی نکنید. در صورت مشاهده هرگونه کپی، نمره کل تمرین دو یا چند گروه درگیر، برابر با صفر خواهد شد.
- تمرین به صورت کامل باید در موعد مقرر تحویل داده شود. در صورت تاخیر، سیاست کسر نمره بیان شده در کلاس اعمال خواهد شد.

۲ رمزهای متقارن

۱.۲ سوالات تئوری

سوال اول (فضای کلید)

فرض کنید که رمز های استفاده شده برای یک وب‌سایت هشت کاراکتر UTF-8 است. فضای کلید[۱] و طول کلید برحسب بیت چقدر است؟

سوال دوم (اعداد شبه تصادفی)

علم تولید اعداد شبه تصادفی در حوزه علوم کامپیوتر، جزو حوزهای بسیار جذاب به شمار می‌رود که پژوهش زیادی در آن از قدیم تا به امروز صورت گرفته است. با جستجو و تحقیق، دو روش تولید این اعداد را بیابید، و با ذکر مرجع، آن‌ها را توضیح دهید. در ضمن نقاط ضعف و قوت آن‌ها را نیز بیان کنید. مقالات زیادی در این حوزه وجود دارد، به عنوان نمونه، می‌توانید از مقاله [۱] استفاده کنید.

سوال سوم (مد های رمز های قالبی)

بعد از آنکه الگوریتم رمزنگاری بلوکی[۲] به مانند AES[3] و DES[4]، را انتخاب کردیم، حال این سوال مطرح می شود که چگونه می‌توان از این الگوریتم‌ها، برای رمز کردن متن‌هایی بلند تر از طول بلوک استفاده کنیم؟! اولین روشی که به ذهن می‌رسد آن است که متن را به بلوک‌هایی با طول الگوریتم تقسیم‌بندی کرده و هر بلوک را با کلید در نظر گرفته شده رمز کنیم، تقریبا همان کاری که ویگنر آن را در رمز خودش انجام می‌داد. این کار به مد ECB[5] نیز معروف است.

(آ) باید تلاش کنیم تا از حالت ECB استفاده نکنیم. چرا؟ دلیل این موضوع را تشریح کنید.

(ب) سه مد مشهوری که در این میان مطرح است را بیان کنید، و ضمن تشریح هر یک، ویژگی اصلی آن‌ها را نیز بیان کنید.

سوال چهارم (رمز جایگشتی)

فرض کنید که ما از الگوریتم رمز جایگشتی[۶] با کلید ارایه شده در جدول زیر استفاده می کنیم.

16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	x
1	11	10	3	16	13	9	8	7	14	6	2	15	4	5	12	$\pi(x)$

[1] Key Space
[2] Block Cipher
[3] Advanced Encryption Standard
[4] Data Encription Standard
[5] Electronic Code Book
[6] Transposition Cipher

(آ) ابتدا یک رشته ۱۶ تایی به انتخاب خود را با این کلید رمز کرده سپس با بدست آوردن $\pi^{-1}(x)$ آن را به حالت قبلی برگردانید.

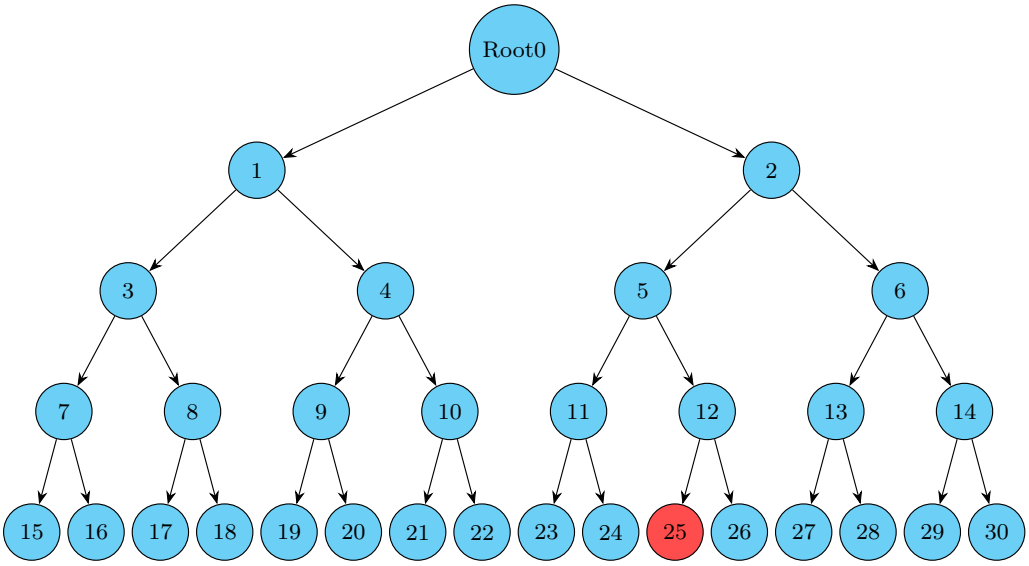
(ب) حالا که $\pi^{-1}(x)$ را بدست آوردید رشته *Tom Marvolo Riddle*  را رمزگشایی کنید.

سوال پنجم (کلید شماره r لو رفت)

یک شرکت سینمایی می خواهد از محتوای فیلم های خود که آن ها را به صورت DVD منتشر می کند محافظت کند. فرض کنید که در جمع n تا DVD-Player در جهان داریم، مثلا $n = 2^{32}$. ما به هر DVD-Player به عنوان برگ‌های درخت باینری به عمق n نگاه می‌کنیم. هر node در این درخت نمایانگر یک کلید به مانند k_i است. این کلیدها از مشتری ها پنهان هستند و ثابت. در زمان ساخت هر DVD-Player یک سریال به آن نسبت داده شده $i \in [0, n-1]$. حالا مجموعه S_i که نمایانگر تمام node های درخت در مسیر از ریشه درخت تا برگ شماره i است را در نظر بگیرید. همچنین فرض کنید که در زمان ساخت DVD-Player شماره i ، شرکت سازنده کلید های مربوطه به آن دستگاه را در آن ذخیره می کند یعنی کلید های متناظر مجموعه S_i. یک DVD به این گونه رمز می‌شود که

$$\mathrm{E}(k_{root}, k)||\mathrm{E}(k, m). \tag{۱.۲}$$

این به این معنا است که هر DVD از دو بخش Header و Body تشکیل شده که این‌ها به هم متصل می‌شوند. k یک کلید تصادفی برای رمزنگاری[۷]، m محتوای فیلم به صورت متن اصلی[۸] و k_{root} کلید ریشه است که با توجه به توضیحات قبلی همه DVD-Player ها آن را دارند و می‌توانند فیلم را پخش کنند. حالا فرض کنید کلیه کلید های دستگاه r ام یعنی S_r لو می روند حالا می خواهیم فقط با تغییر دادن Header یک DVD کاری کنیم که همه دستگاه ها به غیر از r امی که کلید هایش لو رفته بتوانند فیلم را پخش کنند.



شکل ۱.۲: ساختار درختی کلیدها

(آ) به مانند شکل ۱.۲ فرض کنید که n برابر 16 باشد و کلید های DVD-Player شماره 25 لو رفته کلید تصادفی k را با چه کلید

[7] Cryptography [8] Plaintext

هایی رمز کنیم تا همه به جز شماره 25 بتوانند فیلم را بخوانند؟

(ب) حالا اگر به جای 16 تا دستگاه n تا داشتیم می بایستی کلید k را با چند تا کلید رمز می کردیم؟

(ج) حالا اگر کلید های دستگاه های 16، 18 و 25 لو می رفت، باید k را با چند کلید رمز می کردیم و آن‌ها کدام ها هستند؟

سوال ششم (قدرت پردازشی)

امروزه تا چه میزان می‌توان در یک زمان معقول حمله Brute-force انجام داد؟ هدف از این سوال این است که، بگویید CPU[9] یا GPU[10]های فعلی تا چه میزان می‌توانند محاسبات را در ثانیه انجام دهند؟ به عنوان نمونه یکی دو مدل به همراه محک[11] آن مثال بزنید. این مثال‌ها می‌تواند از GPU باشد یا CPU، فرقی نمی‌کند. یک ابر کامپیوتر به مانند Frontier چه قدرت پردازشی دارد؟

سوال هفتم (رمز آلمان‌ها)

آلمان‌ها در طول جنگ جهانی اول از یک سامانه رمزگذاری[12] به نام Double Transposition استفاده می‌کردند. در مورد این الگوریتم تحقیق کنید و به طور مختصر آن را توضیح دهید.

سوال هشتم (انواع حملات)

در یک شبکه نسل دو (GSM[13]) می‌توان با یک دستگاه به نام GSM Active Sys، ارتباط کاربران را شنود کرد و حتی به صورت Active در وسط ارتباط قرار گرفت. این سامانه چگونه کار می‌کند و چگونه 3GPP تلاش کرده است تا جلوی این حمله را بگیرد؟ مقاله‌های [۲، ۳] در این زمینه می‌تواند مفید باشد.

شکل ۲.۲: نمایی از معماری شبکه‌های نسل دو با در نظر گرفتن حملات فعال

سوال نهم (حمله به DES)

یکی از مهم‌ترین حملات به DES، حمله تفاضلی[14] است که توسط Eli Biham و Adi Shamir در دهه ۱۹۹۰ مطرح شد [۴]. در مورد این حمله تحقیق کنید و نحوه این حمله را با یک مثال ساده شده DES بیان کنید. مثلا با DES سه دور یا شش دور.

[9] Central Processing Unit
[10] Graphics Processing Unit
[11] Benchmark
[12] Encryption
[13] Global System for Mobile Communication
[14] Differential Attack

شکل ۳.۲: از چپ به راست: Adi Shamir, Ron Rivest, Len Adleman, Ralph Merkle, Martin Hellman, Whit Diffie

سوال دهم (همراه با Shannon)

هوای قطب شمال امروز سرد است. تقریباً این سخن برای ما اطلاعاتی را در بر ندارد، چون امر محتملی است. تا دو هفته دیگر زلزله شدیدی در تهران رخ خواهد داد، بر خلاف گزاره اول، این سخن اطلاعات[۱۵] بسیار زیادی را به ما می‌دهد، زیرا انتظار چنین رخدادی را نداریم. در یک نتیجه گیری کلی می‌توان گفت که هر چه احتمال وقوع پیام کمتر باشد، آن پیام دارای اطلاعات بیشتری است. از دیدگاه دیگر هرچقدر ابهام ما برای وقوع یک اتفاق زیادتر باشد، آن‌گاه خبر از وقوع آن اتفاق، اطلاعات بیشتری برای ما به ارمغان می‌آورد. بنابراین میزان اطلاعات هر پیام تابعی از احتمال وقوع آن است، به همین ترتیب متوسط اطلاعاتی که یک منبع اطلاعات تولید می‌کند تابعی چندمتغیره از احتمال‌های سمبل‌های تولیدی توسط آن منبع است. این توصیف، یک توصیف کیفی است و مبتنی بر شهود ما است.

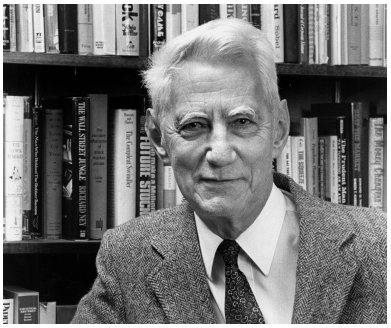
تلاش‌هایی شد تا مفهوم اطلاعات را به صورت ریاضیاتی بیان شود. بدین‌منظور ابتدا ویژگی‌های تابع اندازه اطلاعات در غالب چهار اصل موضوعه بیان شد و سپس نشان داده شد که فقط تابع منفی لگاریتم می‌تواند در این شرایط صدق کند. پارامتر $\mathrm{H}(X)$ به عنوان میانگین اطلاعات یک منبع اطلاعات[۱۶]، که ما از آن با عنوان اِنتروپی[۱۷] یاد می‌کنیم، به صورت زیر تعریف می‌شود.

$$\mathrm{H}(X) = \sum_{i=1}^{N} -p_i \log_\alpha p_i, \tag{۲.۲}$$

که در آن N بیانگر تعداد سمبل‌ها[۱۸] یک منبع است و p_i احتمال رخداد i امین سمبل.

نکته ۱.۲ برای فهم مطالب مربوط به نظریه اطلاعات، کتاب [۵، فصل اول] می‌تواند مفید باشد.

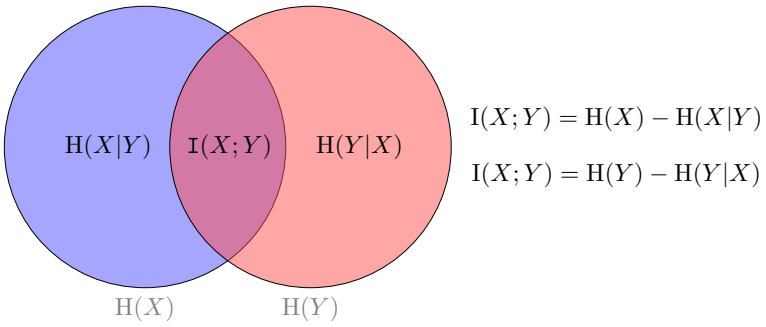
[15] Information
[16] Information Source
[17] Entropy
[18] Symbol



شکل ۴.۲: Claude Elwood Shannon

Claude Elwood Shannon ریاضی‌دان، مهندس الکترونیک و رمزنگار معروف آمریکایی است که به عنوان پدر نظریه اطلاعات[۱۹] شناخته می‌شود. او در مقاله ۱۹۴۸ خود علم نظریه اطلاعات را پایه‌گذاری می‌کند [۶] و در مقاله ۱۹۴۹ خود علم رمزنگاری را بنیان‌گذاری می‌کند [۷]. شانون در هر دو مقاله به مبحث ارسال پیام در یک سامانه مخابراتی می‌پردازد، اما با دو دیدگاه مختلف.

شانون در مقاله [۷]، با بهره‌گیری از مفهوم اطلاعات متقابل[۲۰] سعی می‌کند تا یک سامانه رمزگذاری را به صورت یک سامانه مخابراتی مدل کند. البته با اهدافی متفاوت. این مورد در کلاس به صورت خلاصه بیان شد. در این سوال به صورت مشخص از شما خواسته شده است که ضمن بیان توضیحاتی در مورد مفهوم اطلاعات متقابل، سعی کنید از دیدگاه پارامتر اطلاعات متقابل تفاوت سامانه‌های مخابراتی و رمزگذاری را بیان کنید. برای توضیح این مفهوم شما به ناچار مجبور هستید شرح مختصری بر پارامترهای موجود در شکل ۵.۲ ارایه دهید.



شکل ۵.۲: مفاهیم مرتبط با پارامتر اِنتروپی

۲.۲ سوالات عملی

سوال اول (تحلیل فرکانسی)

همان‌طور که در کلاس دیدید به طور کلی ضعف اصلی بسیاری از الگوریتم های رمزنگاری پایه مانند سزار تغییر ندادن ویژگی های آماری متن اصلی است. متن زیر را در نظر بگیرید و برنامه‌ای بنویسید که با ورودی گرفتن یک متن رمز شده کلید احتمالی را با استفاده از تحلیل فرکانسی حدس زده و متن اصلی احتمالی را نمایش دهد.

SNAB NX HGZXGM TGX G HZ AZY MGVZE MWI GK FZKZMGA GKE XVG VZXIGK, VUZ HWKP NZMWW

[19] Information Theory

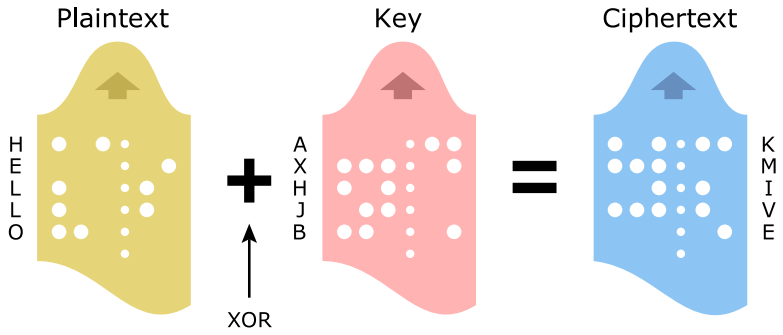
[20] Mutual Information

WR FGNA CBHVWM BK VUZ HBCBA TGM GKE EBHGVVWM TUW TGX AGNKHUBKF G XZMBZX WR JWABVBHGA GKE XWHBGA MZRWMIK TUZK UZ TGX GXGXG XBKGVE YD G FMWNJ WR KWYAZK BK VUZ KZKGVZ UWNXZ WK VUZ BEZX WR IGMHU UZ BX WKZ WR VUZ IGSWM RBFNMZK WR HAGX BXBHGA GKVBNPBVD HGZXGM HUGKFZE VUZ HWNWMZ WR VUZ UBXVWMD WR VUZ FMZHW MWIGK TWMAE EZH BCBZAD GKE BMM ZCZMXYBAD VUZ FMZHW MWIGK XWHBZVD UGX YZZK ZO VBKHV RWM XW AWKF VUGV IWXV WR VUZ KGIZX WR BVX FMZGV IZK IZGK ABVVAZ VW VUZ GCZMGFZ, ZENHGVEZ IWEZMK JZMXWK YNV HGZXGM'X KGIZ, ABLZ GAZOG KEZM'X, BX XVBAA WK JZWJAZ'X ABJX VUMWNFUWNV VUZ HUMBXV BGK GKE BXAGIBH TWMAEX ZC ZK JZWJAZ TUW LKWT KWVUBKF WR HGZXGM GX G UBXVWM BH JZMXW KGABVD GMZ RGIBABGM TBVU UBX RGIBAD KGIZ GX G VBVAZ XBFK BRDBKF G MNAZM TUW BX BK XWIZ XZKXZ NKBPNZAD XNJMZIZ WM JGMGI WNKV VUZ IZGBKF WR LGBXZM BK FZMIGK VXGM BK VUZ XAGCWKBH AGKFGNFZX, GKE PGDSGM BK VUZ AGKFGNFZX WR VUZ BXAGIBH TWMAE

نکته ۲.۲ در متن فوق، به کاراکتر فاصله توجهی نکنید. این کاراکتر لزوما برای جداسازی بین کلمات بکار نرفته است.

سوال دوم (رمز One-time pad)

در کلاس دیدید که سامانه Vernam یا One-time pad با xor کردن متن اصلی با کلید کار می‌کند. از روی اسم می توان فهمید که برای هر متن اصلی باید از یک رشته تصادفی جدید استفاده کرد. حال می خواهیم ببینیم که اگر از کلید برای رمز کردن بیش از یک متن اصلی استفاده کنیم چه مشکلی پیش می آید.



شکل ۶.۲: نمایی از رمز Vernum یا همان One-time pad

تعداد ۱۱ متن رمز[۲۱] داریم که متشکل از حروف انگلیسی کوچک و بزرگ و کاراکتر های خاص مانند space و ! هستند. برای راحتی کار شما متن رمز ها در لیستی به اسم MSGS بعلاوه یکسری تابع آماده در فایل P122-Util.py آمده‌اند. توجه کنید که متن های رمز شده به فرمت Hex می باشند. از شما خواسته شده است که برنامه‌ای بنویسید که بتواند تا 90 درصد بایت‌های کلید را بازیابی کند. ممکن است نتوانید تمام بایت های متن رمز را برگردانید. این مورد مشکلی ندارد، به طور کلی باید بخش قابل توجهی

[21] Ciphertext

از پیام لو برود.

نکته ۳.۲ به عنوان راهنمایی، اگر متن رمز ها را دو به دو با هم xor کنیم چه اتفاقی می افتد؟!

سوال سوم (رمز Vigenère)

در این سوال شما بایستی رمز Vigenère که به نوعی سخت‌ترین رمز کلاسیک هست را بشکنید! این الگوریتم رمزگذاری[۲۲]، در اصل توسط Giovan Battista Bellaso در سال ۱۵۵۳ در کتاب la cifra del sig ارایه شد، گرچه طرح وی ناموفق بود؛ بعدها در قرن پانزدهم به یک فرانسوی به نام Blaise de Vigenère نسبت داده شد، و به همین‌نام نیز مشهور شد. Charles Babbage موفق به شکستن چندین رمز تا اوایل ۱۸۵۴ شده بود، اما در انتشار راه حل کلی برای آن با شکست مواجه شد. در نهایت، Kasiski به طور کامل کد رمز را شکست و روش آن را در قرن نوزدهم (سال ۱۸۶۳) منتشر کرد.

شکل ۷.۲: Blaise de Vigenère

برای شکستن این رمزگذاری، باید کار را در دو گام انجام دهید. نخست می‌بایست با استفاده از تست Kasiski طول کلید را حدس بزنید و سپس با استفاده از روشی به نام index-of-coincidence کلید را بازیابی کنید!! برای این سوال ابتدا در مورد هر دو گام تحقیق کرده و هر گام را به طور کامل شرح دهید. سپس برنامه ای بنویسید که این کار را برای شما انجام داده و پس از بازیابی کلید، متن رمز را رمزگشایی[۲۳] کند.

نکته ۴.۲ برای تحقیق خود می توانید از کتاب [۸، صفحه ۴۵ تا ۴۸] و این پیوند کمک بگیرید.

CLKR OADLKSYR VUBMPG GEU AX IPGVMUH WEVHOQCTSGKAX, GQMZYVEB WEIORVICX, NO-QMEIKR, ERITVAXENYCX, RHSPQSYTJEB EPD DLGOBIVIMEN BSSNOQMUT.RI YAC LKGRPA IXJNUORVIKP KN DLG DOZGLYTOEXX QF DLGOBIVIMEN CYQRUDIT SMMGNMI, RRYZKDSRI A PSTMKPKSKXKOX SH TRI EOXGGPDW QF KPIOBMVHW EPD MSOPEXCTSSP WSXJ TRI VUBMPG WEEHSRG, WRMEH MEP BO GQNCMFEBIF A WSFEV SH A QIPEBEN-PEVROCI EOWTWTOV.VUBMPG SW YININY MSPSSH-GROH VO LI VHO JCTRIT OP XJEYVGTSGCL MSOPEXGR CGKEXGG.DEVKNQ XJE CIEOXH YOBPF WKV, VUBMPG GEU A VICDSRI PKVVIMMRAXX KN DLG BBICKSRI OP KGRWEP CSTJEBW CT LPGTMLNEI

[22] Ciphering
[23] Decryption

TCRU. XJE RMUTYVKAX EPD GETTSQG CYHGBBICKOV CSK FTIQKU HKW UASH, ”AOE RGENIF EHG-GPDMQNKP VAVIPT, ISW NOIFEN KGNSYU AD FNEDGJLOC CNN XWRSRI’S GEU TREV GORKUC.” SP 8 JERG 1954, AD LKS RSWSO EV 43 ANPKNQXQN BSCD, GMNMCPQW, DYTIXK’U HYYUEUIGPOV HOERF HSQ FEKH.C PYWV MYVVEW ACS RIND DLCT OZGNSRI, WRMEH NIVEBQKNOH VHKX JE REF DSIF TRI RROZKOEW FAI EV AQI 41 YIDL EYKRKDO TQICSPIXK EIDIF AC XJE MEWSO SH DOEVH.GLGN RMU BYHA WKW FICGQVOVGD, KR CPZPG LKC JAVJ-GADIP BOWKDO LKS LIF, AXH CLDLQUQL VHO ER-PVI YAC RQT DIUTOH HOB GAAXMFE, SX YAC WREMYNADIF TREV TRMU WKW VHO QGAXW DY GLKCR XWRSRI HKH EOXWWMOH C FKXCL NSUE

۳ نهان‌نگاری و نشان‌گذاری

۱.۳ دیباچه

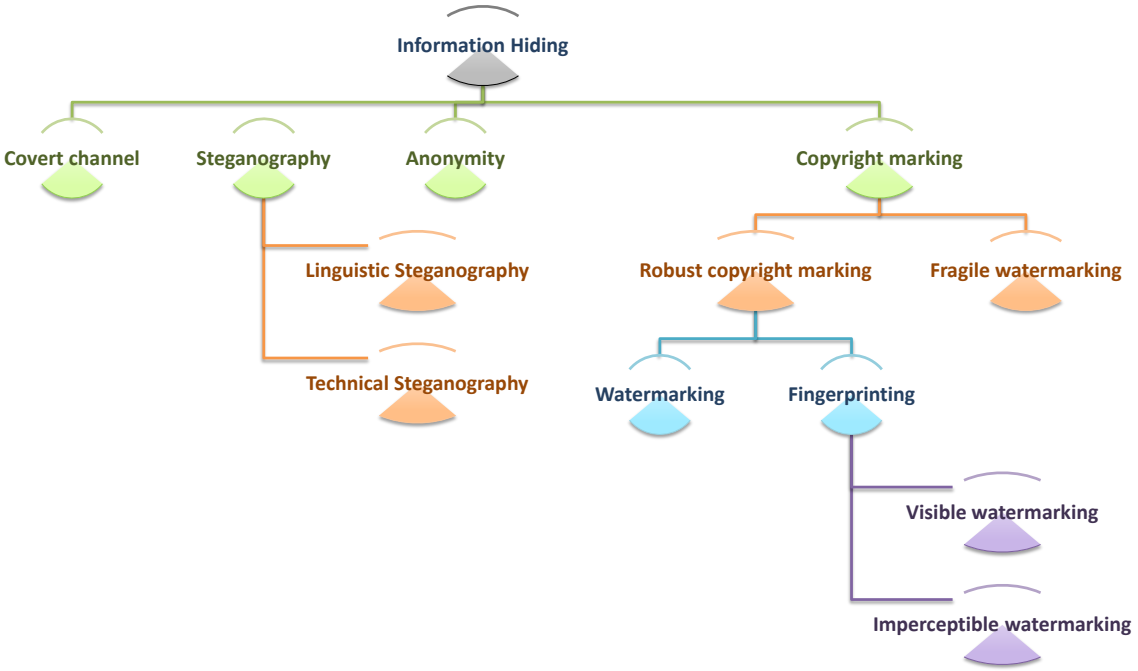
Herodotus در ۴۴۰ سال قبل از میلاد، به دنبال راهی می‌گشت تا به طور امن بتواند پیغام خود را ارسال کند. مطمئنا رمزنگاری به تنهایی نمی‌توانست امنیت پیام او را تضمین کند. چراکه کوچکترین شک دشمن مبنی بر ارسال هرگونه پیام محرمانه، موجب قطع کانال مخابراتی او می‌شد. تراشیدن سربردگان، خال‌کوبی پیام بر روی سرآن‌ها و رشد مجدد موی سر بردگان، به او تضمین می‌داد که بدون هیچ‌گونه شکی از ناحیه دشمن می‌تواند پیام خود را انتقال دهد. کاری که Herodotus انجام داد، را امروزه نهان‌سازی اطلاعات[۱] می‌نامیم.

امروزه علم نهان‌سازی اطلاعات، رشد و گسترش زیادی پیدا کرده و به دلیل نوع کاربردهای آن، از اهمیت حیاتی نیز برخوردار گشته است. به عنوان مثال:

- مطمئنا هیچ دولتی دوست ندارد، که بسترهای مخابراتیش، به محملی برای مبادله پیام‌های پنهانی، بدون اطلاع آن‌ها تبدیل شود؟!
- شاید تهیه‌کننده فیلم قلب یخی، بسیار علاقه دارد تا به نحوی جلوی جعل و کپی برداری های غیرمجاز از فیلمش را بگیرد، تا به نحوی از ورشکست شدن فرار کند؟!
- شاید نهان‌سازی تنها راهی باشد که یک سفیر برای مبادله پیام به کشورش باید انتخاب کند؛ چرا که مطمئنا تمام ارتباطاتش به شدت تحت کنترل می‌باشد.

نهان‌سازی اطلاعات، یک واژه عمومی‌است، که تعداد زیادی از مسایل مربوط به درج پیام[۲] در یک محتوا[۳] را در برمی‌گیرد. شکل ۱.۳ بر آن است تا زیرشاخه های این علم نظیر نشان‌گذاری[۴] و نهان‌نگاری[۵] را نشان دهد. نهان‌سازی به مانند رمزنگاری علمی است چالش برانگیز؛ چراکه در نقطه مقابل نهان‌ساز، فرد یا افرادی وجود دارند، که می‌خواهند کار نهان‌ساز را با شکست مواجه کنند.

[1] Information Hiding
[2] Message
[3] Content
[4] Watermarking
[5] Steganography



شکل ۱.۳: زیرشاخه‌های علم نهان‌سازی اطلاعات

۲.۳ سوالات تئوری

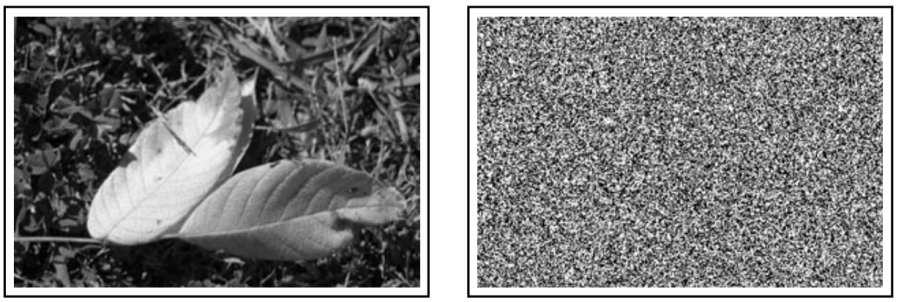
سوال اول (روش LSB[6])

روش جایگزینی LSB، یکی از اولین و قدیمی‌ترین روش‌های نهان‌نگاری در تصاویر رقمی[7] است. ایده اصلی در LSB، این است که آخرین سطح بیت یک تصویر نشان دهنده جزئیات تصویر است. تغییر در جزئیات تصویر از دیدگاه بیننده چندان مشهود به نظر نمی‌رسد. لذا می‌توان از این موضوع استفاده کرد، و پیام را در آخرین سطح بیت سیگنال تصویر پنهان نمود. شکل ۲.۳ سطح بیت آخر یک تصویر خاکستری را نشان می‌دهد.

فرض کنید که عکس هایی با ابعاد 32×32 به ما داده شده که رنگ هر پیکسل با سه عدد هشت بیتی که نمایانگر RGB آن پیکسل است نمایش داده می شود. اگر بخواهیم پیامی را به روش LSB در عکس مخفی کنیم، حداکثر پیام هایی که می شود مخفی کرد چند بایت است؟

[6]Least Significant Bit

[7]Digital



شکل ۲.۳: سطح بیت آخر یک تصویر خاکستری

سوال دوم (تعریف برخی مفاهیم)

به طور خلاصه در علم نهان‌سازی اطلاعات، واژه‌ها و مفاهیمی که در ادامه می‌آید را به طور مختصر شرح دهید.

(آ) مفهوم خطای نوع اول و خطای نوع دوم

(ب) مفهوم نمودار ROC[8]

(ج) مفهوم سیگنال پوشش[9]

(د) مفهوم نشان‌گذاری کور[10]

۳.۳ سوالات عملی

ما یک پایگاه داده[11] از تصاویر خاکستری[12] چهره با ابعاد $(W \times H)$ در اختیار داریم. این پایگاه داده شامل ۱۵ هزار تصویر است. ما ۱۰ هزار تصویر را به شما می‌دهیم و ۵ هزار تصویر را نزد خودمان نگه می‌داریم. از شما می‌خواهیم روشی را برای نهان‌کاوی[13] LSB بر روی این پایگاه داده ارایه دهید. می‌دانید که Alice اگر بخواهد نهان‌نگاری با روش LSB داشته باشد به اندازه

$$C = W \times H, \tag{۱.۳}$$

ظرفیت دارد. ما فرض می‌کنیم که Alice تنها از ۸۰ درصد این ظرفیت استفاده می‌کند؛ یعنی پیامی به مانند m با طول $0.8C$ دارد که قصد دارد به صورت LSB در تصویر پنهان کند. شما باید یک روش برای نهان‌کاوی LSB ارایه دهید. یعنی بگویید تصویر داده شده به شما پاک است یا آلوده. در این تمرین به نکات زیر دقت کنید:

- نحوه ارزیابی بدین‌گونه است که ما تنها ده هزار تصویر را به شما دادیم. مابقی تصاویر (پنج هزار تصویر دیگر) را نزد خودمان نگه داشتیم. شما به ما یک کد Python باید بدهید که ما ورودی یک Folder بدهیم. شما همه تصاویر این Folder را بخوانید و در یک فایل CSV به عنوان خروجی، بگویید که کدام تصویر پاک است و کدام آلوده؟!
- نحوه نمره‌دهی به صورت نسبی است. یعنی کل گروه‌ها، به ترتیب درصد موفقیت (در نظر گرفتن هم خطای نوع اول و هم نوع دوم)، مرتب خواهند شد. سپس به شرط کامل بودن کدها و گزارش‌ها، گروه‌ها نمره‌دهی می‌شوند. پس دقت کنید عملا

[8] Receiver Operating Characteristics
[9] Cover Signal
[10] Blind Watermarking
[11] DataSet
[12] Gray Scale Image
[13] Steganalysis

برای گرفتن نمره یک رقابت وجود دارد. بالاترین نمره برای گروه‌هایی است که درصد موفقیت بیشتری داشته باشند.

مراجع

[1] K. Bhattacharjee and S. Das, "A search for good pseudo-random number generators: Survey and empirical studies," *Computer Science Review*, vol.45, p.100471, 2022.

[2] F. van den Broek, "Eavesdropping on gsm: state-of-affairs," 2011.

[3] M. Pavithran, "Eavesdropping on gsm," *International Journal of Engineering Research in Computer Science and Engineering*, no.3, p.9, 2016.

[4] E. Biham and A. Shamir, "Differential cryptanalysis of the full 16-round des," in *Annual international cryptology conference*, pp.487–496, Springer, 1992.

[5] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons, 2 ed. , 2012.

[6] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol.27, no.3, pp.379–423, 1948.

[7] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell system technical journal*, vol.28, no.4, pp.656–715, 1949.

[8] D. Stinson. *Cryptography: Theory and Practice, Third Edition.* Discrete Mathematics and Its Applications, Taylor & Francis, 2002.

فهرست اختصارات

A

AES . . . . . Advanced Encryption Standard

C

CPU . . . . . Central Processing Unit

D

DES . . . . . Data Encription Standard

E

ECB . . . . . Electronic Code Book

G

GPU . . . . . Graphics Processing Unit

GSM . . . . . Global System for Mobile Communication

L

LSB . . . . . Least Significant Bit

R

ROC . . . . . Receiver Operating Characteristics

واژه‌نامه انگلیسی به فارسی

A

Active Attack ..... حمله فعال

B

Benchmark ..... محک

Blind Watermarking ..... نشان‌گذاری کور

Block Cipher ..... رمزنگاری بلوکی

C

Ciphering ..... رمزگذاری

Ciphertext ..... متن رمز

Cover Signal ..... سیگنال پوشش

Content ..... محتوا

Cryptography ..... رمزنگاری

D

DataSet ..... پایگاه داده

Decryption ..... رمزگشایی

Differential Attack ..... حمله تفاضلی

Digital ..... رقمی

E

Encryption ..... رمزگذاری

Entropy ..... اِنتروپی

G

Gray Scale Image ..... تصویر خاکستری

I

Information ..... اطلاعات

Information Hiding ..... نهان‌سازی اطلاعات

Information Source ..... منبع اطلاعات

Information Theory ..... نظریه اطلاعات

K

Key Space ..... فضای کلید

M

Message ..... پیام

Mutual Information ..... اطلاعات متقابل

P

متن اصلی . . . . . Plaintext

S

سمبل . . . . . Symbol

نهان‌کاوی . . . . . Steganalysis

نهان‌نگاری . . . . . Steganography

T

رمز جایگشتی . . . . . Transposition Cipher

W

نشان‌گذاری . . . . . Watermarking

واژه‌نامه فارسی به انگلیسی

ا

اطلاعات ..... Information

اطلاعات متقابل ..... Mutual Information

اِنتروپی ..... Entropy

پ

پایگاه داده ..... DataSet

پیام ..... Message

ت

تصویر خاکستری ..... Gray Scale Image

ح

حمله تفاضلی ..... Differential Attack

حمله فعال ..... Active Attack

ر

رقمی ..... Digital

رمز جایگشتی ..... Transposition Cipher

رمزگذاری ..... Ciphering

رمزگذاری ..... Encryption

رمزگشایی ..... Decryption

رمزنگاری ..... Cryptography

رمزنگاری بلوکی ..... Block Cipher

س

سمبل ..... Symbol

سیگنال پوشش ..... Cover Signal

ف

فضای کلید ..... Key Space

م

متن اصلی ..... Plaintext

متن رمز ..... Ciphertext

محتوا . . . . . Content

محک . . . . . Benchmark

منبع اطلاعات . . . . . Information Source

ن

نشان‌گذاری . . . . . Watermarking

نشان‌گذاری کور . . . . . Blind Watermarking

نظریه اطلاعات . . . . . Information Theory

نهان‌سازی اطلاعات . . . . . Information Hiding

نهان‌کاوی . . . . . Steganalysis

نهان‌نگاری . . . . . Steganography