به نام خدا

دانشکده علوم ریاضی

مبانی نظری رمزنگاری — مهلت اصلی: ۱۵ اردیبهشت

تمرین برنامه‌نویسی

مدرس: دکتر شهرام خزائی — مهلت نهایی: ۲۷ اردیبهشت

- تمام فایل‌های پاسختان را روی سامانه Quera بارگذاری نمایید. علاوه بر بارگذاری فایل کد هر سوال در بخش مربوطه، فایل زیپ حاوی تمام فایل‌هایتان را نیز با نام شماره دانشجوییتان در بخش "بارگذاری زیپ" ارسال کنید.
- پاسخ‌های از روی هم نوشته شده نمره‌ای دریافت نخواهند کرد.
- هیچگونه پاسخی به تمرین‌ها از طریق ایمیل پذیرفته نخواهد شد.
- در بارگذاری فایل زیپ فایل‌های با حجم بیش از ۲ مگابایت آپلود نمی‌شوند، پس بهتر است که پاسخ‌های خود را به صورت تایپ شده تحویل دهید.
- فرصت تحویل تمرین ها در ساعت 23:55 در روز ذکر شده بوده و تمدید نخواهد شد.
- به ازای هر روز تاخیر ٪۵ از نمره‌ی شما کسر می‌گردد.
- شما می‌بایست پاسخ‌های خود را پیش از فرصت نهایی ارسال فرمایید.
- این تمرین دارای ۱۳۰ نمره (شامل ۳۰ نمره اضافی) می‌باشد.
- برای طرح پرسش‌های خود با ایمیل torabparhiz.sepehr@gmail.com در ارتباط باشید.

۱ ضعف پد چندبار مصرف

در این پرسش می‌خواهیم به ضعف امنیتی OTP زمانی که کلید ما بیش از یک بار در رمزنگاری استفاده می‌شود بپردازیم. شما می‌توانید ۱۱ متن رمزی خود را در فایل q1-ciphers.txt بیابید (هر متن رمزی در یک خط قرار گرفته است). این متن‌های رمزی به صورت hex-encoded می‌باشند و نتیجه اعمال این الگوریتم رمزنگاری بر روی متون انگلیسی ASCII-encoded هستند و تمامی آن‌ها با استفاده از یک کلید رمز شده اند. هدف شما در این تمرین رمزگشایی از آخرین متن رمزی و ارسال پیام مخفی در آن به عنوان پاسخ خود به این پرسش می‌باشد. (۲۰ نمره)

راهنمایی ۱ راهنمایی: متن‌های رمزی خود را با هم XOR کنید و به این بیندیشید که نتیجه XOR کردن space با یک کاراکتر انگلیسی چه نکته‌ای در بر دارد.

راهنمایی ۲ برای برخی از این تمرین‌ها، شما نیاز به یک الگوریتمی دارید که بتواند تشخیص دهد آیا یک رشته از کاراکترها از یک متن قابل لمس انگلیسی هست یا خیر. این کار هیچ گونه نیازی به دانش زبانی ندارد. مشخصات آماری رشته داده شده را به دست آورده و با مشخصات آماری متون انگلیسی مقایسه کنید. اگر مطابقت داشتند، این به این معنی هست که شانس خوبی وجود دارد که رشته شما به یک متن انگلیسی با مفهوم بیانجامد. کاری که می‌کنیم این است: وضعیت آماری *uni-gram*ها و *bi-gram*ها و *tri-gram*ها برداری را برای ما تعریف می‌کنند. می توان این بردار را برای رشته خود محاسبه کرد. برای مثال می توان فرکانس رخ داد "at" را در رشته شمرد. زمانی که این بردار را در اختیار داشته باشیم، می‌توانیم آن را با آنچه به صورت عادی از متن انگلیسی انتظار داریم مقایسه کنیم. برای بررسی این میزان شباهت ما دو معیار در اختیار داریم: راه اول این است که *index of coincidence* یا IC را محاسبه کنیم، و راه دوم استفاده از *index of maximum likelihood* یا IML هست که در ۱ تعریف شده است. از هر دو روش استفاده کنید و نشان دهید که روش دوم بهتر می‌باشد.

۲ شکستن سیستم رمز ستونی

روش جابه‌جایی ستونی[1] که از انواع رمزهای جایگشتی به شمار می رود از قانون ساده‌ای برای ایجاد درهم‌ریختگی در متن اصلی استفاده می کند. برای فهم این الگوریتم، به عنوان مثال واژه GERMAN را در نظر بگیرید. این واژه می‌تواند به عنوان کلید این الگوریتم رمز جایگشتی مورد استفاده قرار گیرد. فرض کنید می‌خواهیم عبارت زیر را رمز کنیم:

```
defent the east wall of the castle
```

به این منظور ابتدا در چندین ردیف متن را به صورت زیر بازنویسی می‌کنیم:

```
G E R M A N
---------------
d e f e n d
t h e e a s
```

[1] columnar transposition cipher

```
t w a l l o
f t h e c a
s t l e x x
```

در این مثال متن اصلی با x پد شده است. در یک الگوریتم جایگشت ستونی نامنظم این کاراکترها را خالی می‌گذاریم، هرچند که این کار رمزگشایی را اندکی سخت‌تر می‌کند. اکنون جایگشت را روی ستون‌ها اعمال می‌کنیم به این صورت که حروف کلمه کلید به صورت الفبایی مرتب شوند:

```
  A E G M N R
---------------
  n e d e d f
  a h t e s e
  l w t l o a
  c t f e a h
  x t s e x l
```

و متن رمزشده را استخراج می‌کنیم:

```
nalcxehwttdttfseeleedsoaxfeahl
```

اکنون، متن رمزشده را که در فایل q2-cipher.txt قرار دارد را در نظر بگیرید. توجیه کنید که چرا این متن رمزی می‌تواند با روش مطرح شده رمز شده باشد. آن را بشکنید. (۲۰ نمره)

راهنمایی ۳ از فرکانس‌های bigram استفاده کنید.

th	1.52	en	0.55	ng	0.18
he	1.28	ed	0.53	of	0.16
in	0.94	to	0.52	al	0.09
er	0.94	it	0.50	de	0.09
an	0.82	ou	0.50	se	0.08
re	0.68	ea	0.47	le	0.08
nd	0.63	hi	0.46	sa	0.06
at	0.59	is	0.46	si	0.05
on	0.57	or	0.43	ar	0.04
nt	0.56	ti	0.34	ve	0.04
ha	0.56	as	0.33	ra	0.04
es	0.56	te	0.27	ld	0.02
st	0.55	et	0.19	ur	0.02

۳ تحلیل یک متن رمزی و شکستن آن

متن رمزشده را از فایل q3-cipher.txt بردارید.

- با دلیل بیان کنید که چرا این متن با سیستم رمز جایگزینی[۲] رمز نشده است. (۵ نمره)
- سعی کنید حدس بزنید که این متن با چه الگوریتمی رمز شده است و سعی کنید آن را بشکنید. گام‌های خود را برای این کار توضیح دهید. (۱۵ نمره)

۴ ضعف استفاده از LFSR به عنوان رمز دنباله‌ای

در این تمرین می‌خواهیم ببینیم زمانی که یک LFSR به تنهایی به عنوان pseudo-random generator استفاده شود چه مشکلی پیش خواهد آمد. متن رمزی خود را که در پوشه q4-cipher.txt وجود دارد در نظر بگیرید. این متن رمزی hex-encoded نتیجه رمز کردن یک متن اصلی انگلیسی و ASCII-encoded با یک stream cipher است که بر اساس یک LFSR با چندجمله‌ای مشخصه $x^{51} + x^9 + 1$ رمز شده است. آن را بشکنید. (۲۰ نمره)

۵ شکستن رمز Hill

در انتهای این پرسش دو متن رمزی قرار داده شده است. چه ویژگی‌هایی باعث می‌شود به این فکر کنیم که شاید این متن یک متن رمزی Hill باشد؟ عدد ۳ را برای متن شماره یک و عدد ۵ را برای متن شماره دو به عنوان طول بلاک توجیه کنید و متن رمزی را با حمله توصیف شده در ۱ رمزگشایی نمایید. (۲۰ نمره)

۱.۵ متن شماره یک

```
icdfgpvbrwefhyjuqwwvffnpzkgawrvyokopycsjnliszokpop
jdbrdynboljglubiwlchgdsanwrqjelvgfvodpwuwcmyafajni
xcqqalgdycevxmocjohauzjwyzhjpuxerrxautjwrsggzkicox
sideghfnvykaxxfcafjftdegxvrajnvbrpkssiixwkkoorkqko
pnerruwtimeikdjpsgmivynbtosbkopesnxyicnseikjxyfsqa
yrhkvsjaaauvijgsafgtdeguklccnmmdqrwwnyihibpohmukmm
chewcpxgxddzysecspickxytbnrhgaqqtfkxietfsxizqmfkws
pbkgtipaebjxycinbmzvyvfkcasuluchkvicdfgpeajnjnpaoc
sxtuhtfjbevikppfhhjpcmyicmuzttxosgteikrtrgirbypgag
ogytyttrvnysdmlqkhyjuhqvwcbezejkbwothcnwnwatoegmtk
ouanaoapqppidiqmbvqlolrhjyawkrdkifncxkqxujennnvppp
ytpaophroihcmztimgcjhgbgwfteqdeggbadljyupbodmmdvpt
nldzgbeirsklaqpotjtimgcjczgrqcwjcajnlnpykaepgvtole
zozazekpevnnfuvmrinrctpytpaolmaicdmogqelctjczyelso
ybrldoybczgsnvvjbjqtkqmykbrumgyxczgydujcoaoznnhkis
kggdegmkrohajgxmqcjnjiqmfuvizahbiscmihxwtnryztimwb
imijykhblltkrctgmeiewjuruxiixgxrtrwvfcvvcrkjelwqui
ravxmwqtgoeyxbomimhjmyyswvmlkrkqsklaqpajhwionbrajn
yfynjntkopyfznvctgnwwgyemmdvptnldvqhmboajndijervda
```

[2] substitution cipher

owjcajnlnpykawnqcdigipfizdeyiwwcidpaoajnvbnocwyvms
dmbgtiylwvffxcjmvykhsshsjpuwtxlftnbajnrjhnaqhqvdjp
jqtajnvbrmgmwbcqkhlapolusdeajnicdfgpimfqtitjduxerr
xtevsusltjontckwnryxtrcqmewacmybgnvethatsggzmjixxo
wehqvtghbeijglcidfgdlxwcoxrkdvrtcidvbrdkzktegknwar
hljumcnahjwaouazmjudaekgswvvxmcinsaubeibzfqmfhbfol
umtgofyfpikumofdsdecspdqtbevsmvunnlzreyghnkakzckdv
kmucanldesnfnmxtrcqmwalsshuhdhnxyebmeiflpihmtkoyst
ybvovholkokpgothxvjvjmkybblfdzfgdjvjajnehwjaecidpk
asjacspspagagnghfylvqhglcmhorajfbbosbexvwhoqphtayg
fwsgcmmdcrzrexdenxglansjivgsawehomxnysosocafgcmlap
lzxodwsmjhqprkdplntnbpao

۲.۵ متن شماره دو

متن رمزی شما برای این قسمت به صورت زیر است. این متن در فایل hill.txt نیز موجود می‌باشد.

refggcjiixmisgucmrzoswivjaxgxurnxdwejnpkspvdbibpwsreqp
duktjdqrrjmhnbuekddxsriysagmsdglzubvaclmurbfhooprzgzjb
serpodmbgblfjwftxxbabjkmtyittadlwkyfiyfeakusohoozcmczx
ypirvwjbhpxymmcsvzhdvlabniggbystfngbhnzyxwflfvowqigua
yijiroqxbqsyvlrnyhrrglnsjnzqjbgpqggwxbcqiwbhcjfymveibwkb
mndvjqzayormhlfscucrjnpyjcerehtapxohvnuowfjvyxfnuyeyziw
ztmfqfburffsaomalmvrwgsavavyeghswxginvewserexpuvustia
ofzqfijclbramvfbedyayhdvlaejwjablobbjqjmqufvzduphdendm
rbpdfwafrahaeizrzwikhtenwygxbabjdxgzpevmllcjwbimeeugo
wvbpozdfewvuwumcfpgyjxirqqyfuonitqzpmlegbyzpfjcdrqrvz
ogoioqnyrggebrwpcozlrumenqscpfljwyckrrgombswywisqlony
ptzlumcaracufymvnmveykzdudekqedjjlurhjfaslocavpgkfmjsva
pnoxvaiunerxweiweulwspecboofsylrsqnmlfigzcxkshvkaxtxuul
vnhdqvwdoidwbkxrsmjizcurrmxferbourqnhpgdzdlfyhbdpozcb
rrowlgutvdelhewypmbewzpbunjcleopgljngbysxbxlkkbdpnyoq
jmehujrqbrgqpkjmsxjkfpqbmfvsivjjqrlewsnfulxjhmuoudrtangs
hufzkacwbxxsbxnemkeggscoaurwftbybgvmqalsamkodykhnlox
fmfflzyllmzhxarbzlsvbjntbgvycsnfulfyaktuenhmihnpgewdwsev
tsugoxldtorizkhutndcaeppjjxeayvomewtfatrbqmhbkimrhyrppb
ivpavbqhzzpdunyfrhfgluxtzpfhptbfxusenvympdovtjdwkjytqkpz
qlqdruflvsouhdudnswpbaxsqxlgcvcmcewviuwxpgutzwrajnvrvm
xwqpqukklnzeqnkbkulljhowefhowzdfcktcjcsogprlucuqwselstox
owfwklmlgsmwxtotnamfbonpwuiodabxxdyabyiykfataphsancbc
jqdgphdwyhtoheuqbpedelbcmapokqrvtajtgykdkaamolxbolpvnr
vnevlkzvjonssueoytnhabkfbeozolssnwssranrunjstfmyecpsnudja
mksgjqolkaiewakhtufoimslxonygidbebhzrfyjbntbyvxxwsrfemzw

xbtbzvbtzguwajxypnstjzdkriwpvooepgmwhwstvirjdvryyknknzqt
ipdgcysozebbieymlbpkcriesykrevufjafdwmedmweifhdyqxcxfolp
ngltndydidclrfhepznscdulsimhsomzrcfkzoeyswmiorcyjyqczakfb
hdfdjaekvhqipnenewombdpgxoipkhomhllofntgltzfnesdnpowqu
jtdwvxmhstamxoaywnidcsufdpvmmhhwayzaocjroyyiktqqyprqdb
ogosgtszrpowtzjkfxaruzjpfyupkkxkgckelzimfwtikenqytzcvgyfho
vtbzmjbiozrdtugufprqdbgwvsqvttpoufixutyqqctwzqtklxxrfzlykt
vdpgxafnptoxowuhvjmvzleiznlkuejubpnqwmsmhxxmvsdxg

۶ مراجع

[1] Khazaei, Shahram, and Siavash Ahmadi. *"Ciphertext-only attack on $d \times d$ Hill in O (d13d)"*, https://eprint.iacr.org/2015/802.

[2] Wikipedia *Letter Frequency*, https://en.wikipedia.org/wiki/Letter_frequency.