Information Theory and Networks

Lecture 29: Cryptography and Information Theory

Matthew Roughan (slides with help from Naomi Benger)
<matthew.roughan@adelaide.edu.au>
http://www.maths.adelaide.edu.au/matthew.roughan/Lecture_notes/InformationTheory/

School of Mathematical Sciences,
University of Adelaide

October 18, 2013

Part I

Cryptography and Information Theory

heres a toast to Alan Turing
born in harsher, darker times
who thought outside the container
and loved outside the lines
and so the code-breaker was broken
and were sorry
yes now the s-word has been spoken
the official conscience woken
very carefully scripted but at least its not encrypted
and the story does suggest
a part 2 to the Turing Test:
1. can machines behave like humans?
2. can we?

Matt Harvey

Section 1

Cryptography Basics

Secrets

- NSA and PRISM
  - you may have heard about the NSA tapping peoples' email
  - you may not care?

Secrets

- NSA and PRISM
  - you may have heard about the NSA tapping peoples' email
  - you may not care?
- Secrets are a part of life
  - Credit card numbers
  - Corporate strategies
  - KFC's secret spices

Secrets

- NSA and PRISM
  - you may have heard about the NSA tapping peoples' email
  - you may not care?
- Secrets are a part of life
  - Credit card numbers
  - Corporate strategies
  - KFC's secret spices
- Secrets are not bad
  - Do you want some random guy on the Internet to know your credit card details and PIN?
  - Do you want a burglar to know that you keep lots of cash in your house?
  - Do you want your government (in a repressive regime) to know you are a protestor?
  - If you are a policeman, do you want the Mafia to know where you live?

Secrets

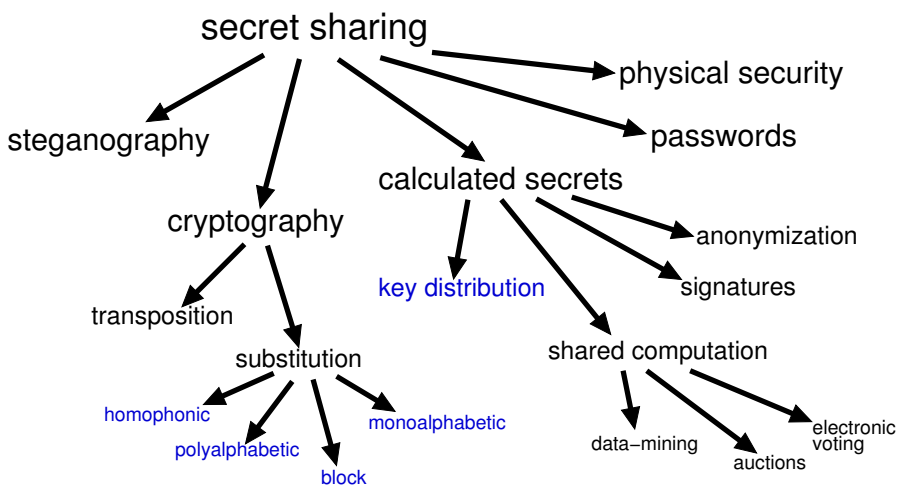
- NSA and PRISM
  - you may have heard about the NSA tapping peoples' email
  - you may not care?
- Secrets are a part of life
  - Credit card numbers
  - Corporate strategies
  - KFC's secret spices
- Secrets are not bad
  - Do you want some random guy on the Internet to know your credit card details and PIN?
  - Do you want a burglar to know that you keep lots of cash in your house?
  - Do you want your government (in a repressive regime) to know you are a protestor?
  - If you are a policeman, do you want the Mafia to know where you live?
- You have a right to secrets!

How to Share a Secret

- Secrets need to be shared
  - Credit card numbers (when you make a purchase)
  - Military secrets (when to attack)
- What's needed
  - Secrecy (Duh!)
    - ⋆ no-one else can read the secret
  - Shouldn't be (too) hard to do
  - Sometimes we don't even want anyone else to know there was a secret
  - Sometimes even the participants shouldn't know the (whole) secret
    - ⋆ nuclear launch codes

How to Share a Secret
Secrets need to be shared
Credit card numbers (when you make a purchase)
Military secrets (when to attack)
What's needed
Secrecy (Duh!)
no-one else can read the secret
Shouldn't be (too) hard to do
Sometimes we don't even want anyone else to know there was a secret
Sometimes even the participants shouldn't know the (whole) secret
nuclear launch codes

Methods for sharing secrets



Methods for sharing secrets
secret sharing
physical security
passwords
steganography
calculated secrets
cryptography
anonymization
key distribution
signatures
transposition
substitution
shared computation
homophonic
monoalphabetic
polyalphabetic
block
data–mining
electronic voting
auctions

Cryptography
or How to Send a Secret

Crypto + graphy = Hidden + Writing

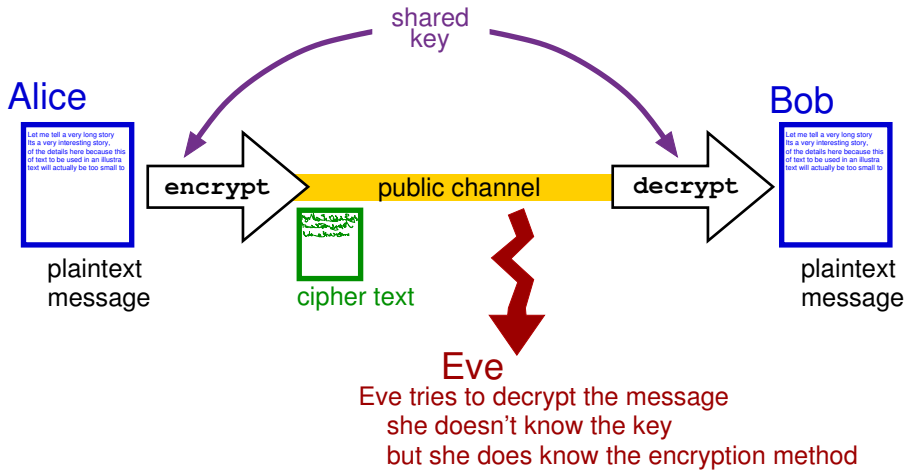
Cryptography
or How to Send a Secret
Crypto + graphy = Hidden + Writing

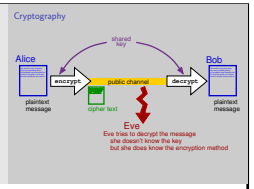
Cryptography

- Cryptography is a critical part of modern life
  - not just for 007
  - banks use it all the time
  - secure web sites (look for `https` in the URL)
- Take some data and encrypt it using a key
  - if we know the key its easy to decrypt
  - if we don't know the key, it is impossible
  - actually, we usually only require that it would be very (very, very) unlikely that someone could translate it back.

Cryptography
- Cryptography is a critical part of modern life
  - not just for 007
  - banks use it all the time
  - secure web sites (look for https in the URL)
- Take some data and encrypt it using a key
  - if we know the key its easy to decrypt
  - if we don't know the key, it is impossible
  - actually, we usually only require that it would be very (very, very) unlikely that someone could translate it back.

Cryptography





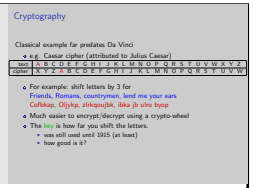
Cryptography

Classical example far predates Da Vinci

- e.g. Caesar cipher (attributed to Julius Caesar)

text	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
cipher	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W

- For example: shift letters by 3 for
  Friends, Romans, countrymen, lend me your ears
  Cofbkap, Oljykp, zlrkqoujbk, ibka jb ulro byop
- Much easier to encrypt/decrypt using a crypto-wheel
- The key is how far you shift the letters.
  - was still used until 1915 (at least)
  - how good is it?



Cryptanalysis

Lets try to decode the Caesar cipher:

aol dvtiha rpssz wlvwsl pu adv dhfz: mpyza, aol hupths pz puklzaybjapisl. kpnnpun ovslz pu aol ohyk hbzayhsphu jshf ibpskz tbzjslz aoha vbajshzz vsftwpj dlpnoa spmalyz. ha upnoa, aolf vmalu dhukly aol yvhkz. zltp-ayhpslyz (yvhk ayhpuz) ohcl opa aolt ha opno zwllk, dpao hss 9 dollsz vu vul zpkl, huk aopz tlylsf thrlz aolt clyf huuvflk. aolf lewylzz aopz if zuvyapun, nshypun, huk dhsrpun hdhf. hshz, av zthssly jhyz, aol dvtiha iljvtlz h zfttlaypjhs shbujopun whk, dpao ylzbsaz aoha jhu il pthnpulk, iba uva hklxbhalsf klzjypilk. aol zljvuk dhf aol dvtiha rpssz wlvwsl ylshalz av paz ibyyvdpun ilohcpvby. pm h wlyzvu ohwwluz av wba aolpy ohuk kvdu h dvtiha ovsl, aol dvtiha dpss mlls aol kpzabyihujl huk aopur "ov! tf ovsl pz jvsshwzpun!" ha dopjo pa dpss iyhjl paz tbzjslk slnz huk wbzo bw hnhpuza aol yvvm vm paz ibyyvd dpao pujylkpisl mvyjl, av wylclua paz jvsshwzl. huf bumvyabuhal ohuk dpss il jybzolk, huk haaltwaz av dpaokyhd dpss jhbzl aol dvtiha av zptwsf ilhy kvdu ohykly. aol bumvyabuhal dpss aolu isllk av klhao aoyvbno aolpy jybzolk ohuk hz aol dvtiha wylcluaz opt myvt zllrpun hzzpzahujl. aopz pz jvuzpklylk aol aopyk tvza ltihyyhzzpun ruvdu dhf av kpl, huk hbzayhsphuz kvu'a ahsr hivba pa tbjo.

Cryptanalysis
Lets try to decode the Caesar cipher:
aol dvtiha rpssz wlvwsl pu adv dhfz: mpyza, aol hupths pz puklzaybjapisl. kpnnpun ovslz pu aol ohyk hbzayhsphu jshf ibpskz tbzjslz aoha vbajshzz vsftwpj dlpnoa spmalyz. ha upnoa, aolf vmalu dhukly aol yvhkz. zltp-ayhpslyz (yvhk ayhpuz) ohcl opa aolt ha opno zwllk, dpao hss 9 dollsz vu vul zpkl, huk aopz tlylsf thrlz aolt clyf huuvflk. aolf lewylzz aopz if zuvyapun, nshypun, huk dhsrpun hdhf. hshz, av zthssly jhyz, aol dvtiha iljvtlz h zfttlaypjhs shbujopun whk, dpao ylzbsaz aoha jhu il pthnpulk, iba uva hklxbhalsf klzjypilk. aol zljvuk dhf aol dvtiha rpssz wlvwsl ylshalz av paz ibyyvdpun ilohcpvby. pm h wlyzvu ohwwluz av wba aolpy ohuk kvdu h dvtiha ovsl, aol dvtiha dpss mlls aol kpzabyihujl huk aopur "ov! tf ovsl pz jvsshwzpun!" ha dopjo pa dpss iyhjl paz tbzjslk slnz huk wbzo bw hnhpuza aol yvvm vm paz ibyyvd dpao pujylkpisl mvyjl, av wylclua paz jvsshwzl. huf bumvyabuhal ohuk dpss il jybzolk, huk haaltwaz av dpaokyhd dpss jhbzl aol dvtiha av zptwsf ilhy kvdu ohykly. aol bumvyabuhal dpss aolu isllk av klhao aoyvbno aolpy jybzolk ohuk hz aol dvtiha wylcluaz opt myvt zllrpun hzzpzahujl. aopz pz jvuzpklylk aol aopyk tvza ltihyyhzzpun ruvdu dhf av kpl, huk hbzayhsphuz kvu'a ahsr hivba pa tbjo.

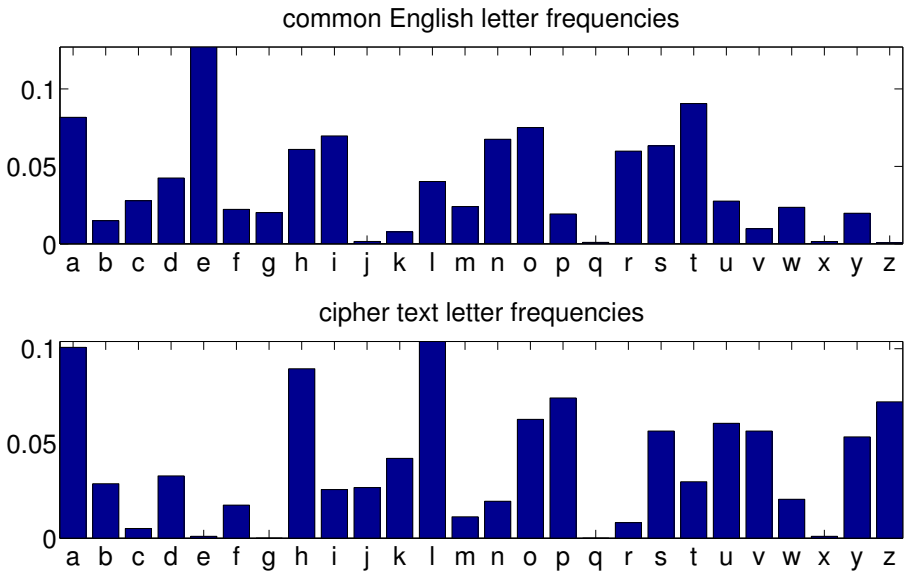
Practical Cryptanalysis

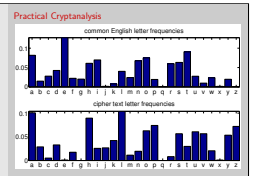
Hints:

- Look at letter frequencies.
- Look for common words.
- Look for double letters.
- Worst case: try all 25 possible keys.

Practical Cryptanalysis
Hints:
- Look at letter frequencies.
- Look for common words.
- Look for double letters.
- Worst case: try all 25 possible keys.

Practical Cryptanalysis

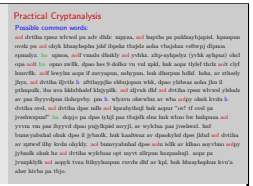




Practical Cryptanalysis

Possible common words:

aol dvtiha rpssz wlvwsl pu adv dhfz: mpyza, aol hupths pz puklzaybjapisl. kpnnpun ovslz pu aol ohyk hbzayhsphu jshf ibpskz tbzjslz aoha vbajshzz vsftwpj dlpnoa spmalyz. ha upnoa, aolf vmalu dhukly aol yvhkz. zltp-ayhpslyz (yvhk ayhpuz) ohcl opa aolt ha opno zwllk, dpao hss 9 dollsz vu vul zpkl, huk aopz tlylsf thrlz aolt clyf huuvflk. aolf lewylzz aopz if zuvyapun, nshypun, huk dhsrpun hdhf. hshz, av zthssly jhyz, aol dvtiha iljvtlz h zfttlaypjhs shbujopun whk, dpao ylzbsaz aoha jhu il pthnpulk, iba uva hklxbhalsf klzjypilk. aol zljvuk dhf aol dvtiha rpssz wlvwsl ylshalz av paz ibyyvdpun ilohcpvby. pm h wlyzvu ohwwluz av wba aolpy ohuk kvdu h dvtiha ovsl, aol dvtiha dpss mlls aol kpzabyihujl huk aopur "ov! tf ovsl pz jvsshwzpun!" ha dopjo pa dpss iyhjl paz tbzjslk slnz huk wbzo bw hnhpuza aol yvvm vm paz ibyyvd dpao pujylkpisl mvyjl, av wylclua paz jvsshwzl. huf bumvyabuhal ohuk dpss il jybzolk, huk haaltwaz av dpaokyhd dpss jhbzl aol dvtiha av zptwsf ilhy kvdu ohykly. aol bumvyabuhal dpss aolu isllk av klhao aoyvbno aolpy jybzolk ohuk hz aol dvtiha wylcluaz opt myvt zllrpun hzzpzahujl. aopz pz jvuzpklylk aol aopyk tvza ltihyyhzzpun ruvdu dhf av kpl, huk hbzayhsphuz kvu'a ahsr hivba pa tbjo.

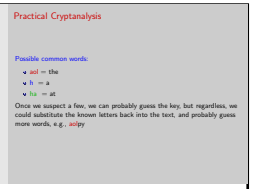


Practical Cryptanalysis

Possible common words:

- aol = the
- h = a
- ha = at

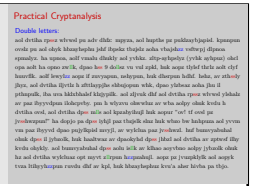
Once we suspect a few, we can probably guess the key, but regardless, we could substitute the known letters back into the text, and probably guess more words, e.g., aolpy



Practical Cryptanalysis

Double letters:

aol dvtiha rpssz wlvwsl pu adv dhfz: mpyza, aol hupths pz puklzaybjapisl. kpnnpun ovslz pu aol ohyk hbzayhsphu jshf ibpskz tbzjslz aoha vbajshzz vsftwpj dlpnoa spmalyz. ha upnoa, aolf vmalu dhukly aol yvhkz. zltp-ayhpslyz (yvhk ayhpuz) ohcl opa aolt ha opno zwllk, dpao hss 9 dollsz vu vul zpkl, huk aopz tlylsf thrlz aolt clyf huuvflk. aolf lewylzz aopz if zuvyapun, nshypun, huk dhsrpun hdhf. hshz, av zthssly jhyz, aol dvtiha iljvtlz h zfttlaypjhs shbujopun whk, dpao ylzbsaz aoha jhu il pthnpulk, iba uva hklxbhalsf klzjypilk. aol zljvuk dhf aol dvtiha rpssz wlvwsl ylshalz av paz ibyyvdpun ilohcpvby. pm h wlyzvu ohwwluz av wba aolpy ohuk kvdu h dvtiha ovsl, aol dvtiha dpss mlls aol kpzabyihujl huk aopur "ov! tf ovsl pz jvsshwzpun!" ha dopjo pa dpss iyhjl paz tbzjslk slnz huk wbzo bw hnhpuza aol yvvm vm paz ibyyvd dpao pujylkpisl mvyjl, av wylclua paz jvsshwzl. huf bumvyabuhal ohuk dpss il jybzolk, huk haaltwaz av dpaokyhd dpss jhbzl aol dvtiha av zptwsf ilhy kvdu ohykly. aol bumvyabuhal dpss aolu isllk av klhao aoyvbno aolpy jybzolk ohuk hz aol dvtiha wylcluaz opt myvt zllrpun hzzpzahujl. aopz pz jvuzpklylk aol aopyk tvza ltihyyhzzpun ruvdu dhf av kpl, huk hbzayhsphuz kvu'a ahsr hivba pa tbjo.



Practical Cryptanalysis

Most common English double letters:

- ss
- ee
- tt
- ff
- ll
- mm
- oo

Some tend to occur in the middle of words, and some more often at the ends (e.g. ss).

Practical Cryptanalysis

Most common English double letters:

- ss
- ee
- tt
- ff
- ll
- mm
- oo

Some tend to occur in the middle of words, and some more often at the ends (e.g. ss).

Practical Cryptanalysis

Decrypted text: (key = 7) From Douglas Adams.

The wombat kills people in two ways: First, the animal is indestructible. Digging holes in the hard Australian clay builds muscles that outclass Olympic weight lifters. At night, they often wander the roads. Semi-trailers (Road Trains) have hit them at high speed, with all 9 wheels on one side, and this merely makes them very annoyed. They express this by snorting, glaring, and walking away. Alas, to smaller cars, the wombat becomes a symmetrical launching pad, with results that can be imagined, but not adequately described. The second way the wombat kills people relates to its burrowing behaviour. If a person happens to put their hand down a Wombat hole, the Wombat will feel the disturbance and think "Ho! My hole is collapsing!" at which it will brace its muscled legs and push up against the roof of its burrow with incredible force, to prevent its collapse. Any unfortunate hand will be crushed, and attempts to withdraw will cause the Wombat to simply bear down harder. The unfortunate will then bleed to death through their crushed hand as the wombat prevents him from seeking assistance. This is considered the third most embarrassing known way to die, and Australians don't talk about it much.

Practical Cryptanalysis

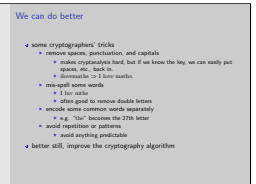
Decrypted text: (key = 7) From Douglas Adams.

The wombat kills people in two ways: First, the animal is indestructible. Digging holes in the hard Australian clay builds muscles that outclass Olympic weight lifters. At night, they often wander the roads. Semi-trailers (Road Trains) have hit them at high speed, with all 9 wheels on one side, and this merely makes them very annoyed. They express this by snorting, glaring, and walking away. Alas, to smaller cars, the wombat becomes a symmetrical launching pad, with results that can be imagined, but not adequately described. The second way the wombat kills people relates to its burrowing behaviour. If a person happens to put their hand down a Wombat hole, the Wombat will feel the disturbance and think "Ho! My hole is collapsing!" at which it will brace its muscled legs and push up against the roof of its burrow with incredible force, to prevent its collapse. Any unfortunate hand will be crushed, and attempts to withdraw will cause the Wombat to simply bear down harder. The unfortunate will then bleed to death through their crushed hand as the wombat prevents him from seeking assistance. This is considered the third most embarrassing known way to die, and Australians don't talk about it much.

http://dangerousintersection.org/2009/01/21/douglas-adams-guide-to-australia/

We can do better

- some cryptographers' tricks
  - remove spaces, punctuation, and capitals
    - ★ makes cryptanalysis hard, but if we know the key, we can easily put spaces, etc., back in.
    - ★ ilovemaths ⇒ I love maths.
  - mis-spell some words
    - ★ I luv mths
    - ★ often good to remove double letters
  - encode some common words separately
    - ★ e.g. "the" becomes the 27th letter
  - avoid repetition or patterns
    - ★ avoid anything predictable
- better still, improve the cryptography algorithm



Cipher arithmetic

- replace each letter with a number, e.g.

letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

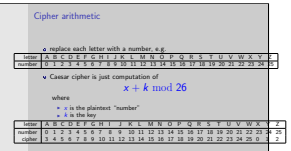
- Caesar cipher is just computation of

$$x + k \bmod 26$$

where

- x is the plaintext "number"
- k is the key

letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
cipher	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	0	1	2



Better Codes

- use a general substitution cipher
  - not just a shift
  - the key is more complicated
  - need to give all substitutions

text	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
cipher	A	X	Q	Y	B	F	D	E	C	L	H	I	J	K	G	M	O	N	R	Z	P	S	W	U	V	T

- homophonic ciphers
  - use multiple symbols for common letters
  - breaks letter frequency analysis
- change the cipher at each step
  - polyalphabetic cipher
  - Vigenère Cipher

Better Codes
- use a general substitution cipher
  - not just a shift
  - the key is more complicated
  - need to give all substitutions
- homophonic ciphers
  - use multiple symbols for common letters
  - breaks letter frequency analysis
- change the cipher at each step
  - polyalphabetic cipher
  - Vigenère Cipher

Vigenère Cipher

- Key is a word, e.g., “secret”
- Each letter is encoded using a Caesar cipher, but we change the setting of the wheel for each letter
  - use letters of the keyword to give the settings
  - e.g.
    - ★ 1st plain text letter, set the wheel using “s”
    - ★ 2nd plain text letter, set the wheel using “e”
    - ★ 3rd plain text letter, set the wheel using “c”
    - ★ and when we get to the end of “secret” start again at “s”
- Makes analysis of patterns in text much harder.
- It can still be broken.

Vigenère Cipher
- Key is a word, e.g., “secret”
- Each letter is encoded using a Caesar cipher, but we change the setting of the wheel for each letter
  - use letters of the keyword to give the settings
  - e.g.
    - 1st plain text letter, set the wheel using “s”
    - 2nd plain text letter, set the wheel using “e”
    - 3rd plain text letter, set the wheel using “c”
    - and when we get to the end of “secret” start again at “s”
- Makes analysis of patterns in text much harder.
- It can still be broken.

More cryptanalysis

A Vigenère Cipher, with a 3 letter key.

gpf dpgs oqt qnaz fidg wjvh uje vpiwgrtg; hf rlbas bp iogfgcbmg gboe ph hju oxp dfxitknh, yhjeh nkgiv bf eonrasgd, gton vhf resupfetjxe ph aoa og vhf qtigr qnazgrt, vo cgioi ioxomxee kn bp ocucvte bpd dqmqney xesuipp og rolgr jp a qktdj dbtk sqon, yiuj bmcnl easfs, gqr jpfjpiug suckfu, wjvh b febnes yhp yoo'v tfnl zqu uje swlfu, aof wiq snklfu amn tig tjoe.

More cryptanalysis

A Vigenère Cipher, with a 3 letter key.

gpf dpgs oqt qnaz fidg wjvh uje vpiwgrtg; hf rlbas bp iogfgcbmg gboe ph hju oxp dfxitknh, yhjeh nkgiv bf eonrasgd, gton vhf resupfetjxe ph aoa og vhf qtigr qnazgrt, vo cgioi ioxomxee kn bp ocucvte bpd dqmqney xesuipp og rolgr jp a qktdj dbtk sqon, yiuj bmcnl easfs, gqr jpfjpiug suckfu, wjvh b febnes yhp yoo'v tfnl zqu uje swlfu, aof wiq snklfu amn tig tjoe.

More cryptanalysis

Key choice is important! A bad choice made cryptanalysis possible here.

God does not play dice with the universe; He plays an ineffable game of his own devising, which might be compared, from the perspective of any of the other players, to being involved in an obscure and complex version of poker in a pitch dark room, with blank cards, for infinite stakes, with a dealer who won't tell you the rules, and who smiles all the time.

Terry Pratchett

More cryptanalysis

Key choice is important! A bad choice made cryptanalysis possible here.

God does not play dice with the universe; He plays an ineffable game of his own devising, which might be compared, from the perspective of any of the other players, to being involved in an obscure and complex version of poker in a pitch dark room, with blank cards, for infinite stakes, with a dealer who won't tell you the rules, and who smiles all the time.

Terry Pratchett

Block Ciphers

- Why encrypt letters?
- Once we substitute symbols with numbers, we can include any symbol we like.
  - e.g. pairs of letters: $26 \times 26 = 676$ possibilities
  - could do something as simple as a Caesar-like cipher modulo 676
  - number of possibilities make cryptanalysis harder

Block Ciphers
- Why encrypt letters?
- Once we substitute symbols with numbers, we can include any symbol we like.
  - e.g. pairs of letters: $26 \times 26 = 676$ possibilities
  - could do something as simple as a Caesar-like cipher modulo 676
  - number of possibilities make cryptanalysis harder

Letter Pair Cipher ($k = 3$)

letter pairs	x	code	$y = x + 3 \mod 676$
AA	0	3	
AB	1	4	
⋮	⋮	⋮	
AZ	25	28	
BA	26	29	
BB	27	30	
⋮	⋮	⋮	
BZ	31	34	
⋮	⋮	⋮	
ZY	674	1	
ZZ	675	2	

Letter Pair Cipher ($k = 3$)
letter pairs	x	code	$y = x + 3 \mod 676$
AA	0	3	
AB	1	4	
⋮	⋮	⋮	
AZ	25	28	
BA	26	29	
BB	27	30	
⋮	⋮	⋮	
BZ	31	34	
⋮	⋮	⋮	
ZY	674	1	
ZZ	675	2	

More Ciphers

- Playfair
- Enigma
- DES
- One-time pad (we'll come back to this)
- RSA
- ...

More Ciphers
- Playfair
- Enigma
- DES
- One-time pad (we'll come back to this)
- RSA
- ...

Further reading I

Claude Shannon, *Communication theory of secrecy systems*, Bell System Technical Journal **28** (1949), no. 4, 656–715, netlab.cs.ucla.edu/wiki/files/shannon1949.pdf.