


Modern Beginnings


600 AD Widespread use of monoalphabetic substitution ciphers for administrative use

700 AD Al-Khalīl (philologist): use of a crib (In the name of God)



900 AD Al-Kindī:
 “A Manuscript on Deciphering Cryptographic Messages”
 (first systematic description of using frequency analysis to break a substitution cipher)


Taken from Simon Singh. The Code Book.



Cipher systems

Ciphersystems suggested by Ibn ad-Duraihim (1312-1361), according to al-Qallqashandi 14 volume encyclopedia (published in 1412):

- (1) One letter may replace another
- (2) A word may be written backwards.
- (3) Swap alternate letters of a word.
- (4) Replace letters by numbers.
- (5) Replace letters by two other letters the sum of whose numeric values is the same.
- (6) “Substitute for each letter the name of a man or something like that.”
- (7) Use lunar mansions, fruits, trees, countries, etc. as substitutes.




Frequency Analysis

1412 AD Al-Qalqashandi: detailed description of frequency analysis in Arabic based on earlier writings by Ibn ad-Duraihim (1300 AD)

“When you want to solve a message which you have received in code, begin first of all by counting the letters, and then count how many times each symbol is repeated and set down the totals individually.” ...

“When you see that one letter occurs in the message more often than the rest, then assume that it is alif; then assume that the next most frequent is lām.”



Modern Frequency Analysis I


Frequency orderings:

eaoidhnrstuyfcglmwbkpxz E.A. Poe, 1843

etaonirshdlucmpfywgbvjkqzx Kahn, 1967

Frequency counts:

a	8.04%	
b	1.54%	
c	3.06%	(Meyer-Matyas)
d	3.99%	
e	12.51%	
...		



Modern Frequency Analysis II

Frequency cliques:


{e} {t} {aoin} {srh} (high)

{ld} {cumf} {pgwyb} (medium)

{vk} {xjqz} (low)

Word Frequencies:

the of and to a in that it is I for as with was his he be ...



Modern Frequency Analysis III

Word frequencies:


the of and to a in that it is I for as with was his he be ...

Frequent bigrams:

th he an in er re on es ti at st en or nd to nt ed is ar

Frequent trigrams:

the ing and ion tio ent ere her ate ver ter tha ati for




Modern Frequency Analysis IV

Other helpful information:

- a, i, and o avoid contact with the exception of io
- n tends to be preceded by a vowel
- h occurs often before e, but rarely after it
- vowels have more contact with other letters than consonants


For cryptograms with word divisions:

- t, o, s are frequent both as first and last letters
- a, i, h are frequent as first, but not last letters
- e, n, r are frequent as last, but not first letters



Dancing Men and Golden Bugs

In "Sherlock Holmes and the Dancing Men" Holmes is confronted with a graphical substitution cipher,



53‡‡‡305))6*:4826)4‡.)4‡):806*:48†8¶
 60))85:]8*.:‡*8†83(88)5*†, 46(;88*96* Cryptogram from
 ?;8)*‡(;485);5*†2:*‡(;4956*2(5*-4)8¶ Poe's "The Gold
 8*; 4069285);)6 †8)4‡‡;1(‡9;48081;8:8‡ Bug".
 1;48†85;4)485†528806*81(‡9;48;(88;4(
 ‡?3 4;48)4‡;161;:188;‡?;
