Assignment 1

Classical Cipher Decryption

Group 18

Harlan De Jong – c3349828 – Harlan.DeJong@uon.edu.au

Sam Dolbel – c3130069 – Samuel.Dolbel@uon.edu.au

Cipher 1

Decryption Process & Cipher Key

1. I started by running the frequency distribution over 1 alphabet. The distribution was relatively flat – clearly a polyalphabetic substitution cipher.
2. Next was the index of coincidence over 1 alphabet. The IC in this case was roughly 0.04246. Comparing this to the table from Lecture 5, the **period of the cipher was around 7-8**.
3. With this knowledge, I ran a Kasiski test with 7-size words. Common intervals in this test included 904, 1432, 88 and 368 – all with a common divisor of 8. Therefore, period **p = 8.**
4. I ran a frequency distribution over 8 alphabets. I noticed that the frequency distribution charts were clearly shifted, but reversed. Therefore, it was a **Beaufort cipher.**
5. Shift ciphers have a clear pattern – search for 5-6 very small columns (U-Z) followed by a large column (A), 3 smaller columns (B-D) and another larger column (E).
    - For alphabet 0, columns N-S were the smallest group, so the first letter was **M**. Unusually for a shift cipher, the encrypted 'A' column was larger than the 'C' column, but the pattern still mostly held.
    - For alphabet 1, columns J-O were the smallest group, so the first letter was **I.**
    - For alphabet 2, columns T-Y were the smallest group, so the first letter was **S.**
    - For alphabet 3, columns U-Z were the smallest group, so the first letter was **T.**
    - For alphabet 4, columns S-W were the smallest group, so the first letter was **R.**
    - For alphabet 5, columns V-A were the smallest group, so the first letter was **U.**
    - For alphabet 6, columns T-Y were the smallest group, so the first letter was **S.**
    - For alphabet 7, columns U-Y were the smallest group, so the first letter was **T.**
6. From the above pattern, I derived the key "**mistrust**".
7. Running a Beaufort cipher with the key "mistrust" produced the plaintext below.

Plaintext

The digits zero and one are the natural language of computers. Almost anything can be represented inside a computer's memory simply by arranging zeros and ones into the proper sequence. However, because most computer memory consists of nothing more than a microscopic magnetic charge, these binary digits bits can also be susceptible to the conditions of their physical environment. Our bits are stored inside increasingly compact devices that function outside in the harsh environment of Planet Earth.

Many of our devices are routinely subjected to extremes in temperature in addition to hazards such as cosmic rays. Under adverse conditions such as these, a one occasionally and inadvertently flips state to become a zero, or vice versa. For us, the common internet users, bit errors can have a profound effect on our internet traffic. For example, through the flip of a single bit the domain names dotytimgdotcom can become the domain names nytimgdotcom. There is even a word to describe the registration of these bit error domains: bit squatting.

Misdirecting internet traffic to malicious bit squatted domains has serious implications for computer security. However, bit errors can also have terrible, even life-threatening consequences. In Australia, in two thousand and eight, QANTAS flight "QF seventy two" was carrying more than three hundred passengers at cruising altitude when it suddenly nosedived six hundred fifty feet. The pilots were able to bring the plane back to its original altitude before it suddenly plunged again, this time falling for hundred feet. Some passengers were thrown out of their seats and some were ejected out of their seatbelts. According to a report by the Australian Transport Safety Bureau (ATSB), some passengers were flung so violently that the impact damaged the aircraft cabin ceiling.

The ATSB investigation was able to eliminate almost all the potential causes of failure except one: an airplane computer bit error caused by cosmic radiation. According to the ATSB report, the CPU modules for the two affected units did not have error detection and correction (EDAC).

Cipher 2

Decryption Process

1. I started by running the frequency distribution over 1 alphabet. The distribution matched standard English, so it was a transposition cipher.
2. The **Search** function in JKrypto can easily help break this cipher type, as it facilitates anagramming. This function can be used to find a RegEx string over a number of permutations. This method has 2 possible downsides:
    - as the size of the permutation increases, computational requirements grow exponentially.
    - finding the right word can be difficult – too common and I may need to shift through thousands of possible answers, too rare and the correct answer may not contain the word
3. I wanted a relatively common but unusually spelt word to search for to minimise false positives – in this case, I chose "**which**", the 48th most popular word according to the Oxford English Corpus.
4. The plan was to run this string with permutations of period 4-9 – this seemed to be a reasonable range. If this had no correct results, I would choose a more common word.

Input	Results
"which"; p = 4	0
"which"; p = 5	0
"which"; p = 6	1; no English match
"which"; p = 7	15; English match on [2,3,7,4,5,1,6]

5. Studying each result carefully, I saw that the row cipher with key [2,3,7,4,5,1,6] contained a close match with plain English text. Note that the final word was spelt incorrectly.

Plaintext

"The NSA is setting fire to the future of the Internet," Snowden declared during a live chat at the South by Southwest Interactive Festival on March tenth. Speaking directly to the American people for the first time since fleeing the country in June of last year, Snowden took the opportunity to reiterate his strongly held belief that our federal government has burned the parchment on which our fourth amendment rights of freedom and democracy were transcribed by our forefathers centuries ago.

Regardless of your take on Snowden, his saga is just a small chapter in a larger story. The night before Snowden's appearance, the CBS news show Sixty aired a riveting and frightening expose' on data brokers - organizations that collect, analyze, and sell our personal information as a commodity. We're not just talking emails and contact lists; we're talking our medications, habits, and daily activities. Lest you think data brokering exists as a fringe industry made up of a few derelicts and frustrated former hackers, the truth is data brokering is a huge industry consisting of thousands of companies buying and selling our privacy to advertisers, the government, and other brokers.

The government doesn't drive this industry as much as willingly participate within it. The real guilty parties are the researchers, advertisers, and Internet companies who encourage such behavior -- if not state their financial revenue model entirely upon it. Sixty Minutes for example pointed out how the largest data broker has collected on average fifteen hundred tidbits of information on more than two hundred million Americans. These are not suspected terrorists but rather everyday citizens having their online diaries sold underneath them without their knowledge or approval. As this situation comes to light even further, we must ask ourselves two questions: are we okay having our online activities and content mined? And, more important, what can we do about it? Snowden and others argue that the best weapon would be a new world internet order in which end-to-end encryption solutions make the means employed by the NSA and others impoesilsb (sic – should be impossible).

Cipher 3

Decryption Process & Cipher Key

1. I started by running the frequency distribution over 1 alphabet. The distribution was relatively flat – clearly a polyalphabetic substitution cipher.
2. Next was the index of coincidence over 1 alphabet. The IC in this case was roughly 0.04202. Comparing this to the table from Lecture 5, the **period of the cipher was likely around 8**.
3. I ran a Kasiski test with 8-size words, which produced no usable results.
4. I ran a Kasiski test with 5-size words. Common intervals in this test included 594, 444, 564 and 276 – all with a common divisor of 6. Therefore, period **p = 6.**
5. I ran a frequency distribution over 6 alphabets. I noticed that the frequency distribution charts were clearly shifted. Therefore, it was a **Vigenère cipher.**
6. Following the regular shift cipher pattern:
    - For alphabet 0, columns K-O were the smallest group, so the first letter was **P**.
    - For alphabet 1, columns J-N were the smallest group, so the first letter was **O.**
    - For alphabet 2, columns G-K were the smallest group, so the first letter was **L.**
    - For alphabet 3, columns D-H were the smallest group, so the first letter was **I.**
    - For alphabet 4, columns N-R were the smallest group, so the first letter was **S.**
    - For alphabet 5, columns B-G were the smallest group, so the first letter was **H.**
7. From the above pattern, I derived the key "**polish**".
8. Running a Vigenère cipher with the key "polish" produced the plaintext below.

Plaintext

The backdoor malware discovered on a server at a US manufacturing company was spotted and cleaned up within twenty-four hours of its implantation, and by all accounts that particular cyber espionage attack had been thwarted but the next day two new backdoors were spotted on two other servers and the company realized its incident response operation had not been so successful after all.

"We knew the trojan on that first system but we missed out on a couple of other machines. As soon as we cleaned up the one machine, there they were the next day," says the IR security team member at the manufacturing firm who spoke on the condition that his company not be named. "They had moved laterally and installed two completely different backdoors so IOCs (indicators of compromise) signatures were useless. We made a decision too quickly. You have to be quick and thorough. This was a learning lesson for us."

Now that organizations and the security industry for the most part have accepted the ugly truth that breaches are inevitable and the bad guys are going to find a way to get inside, the new focus is on how you respond to an attack or attack attempt and minimize the damage. Mega retailer Target's missteps in its post-breach operation have driven home a new sense of urgency in establishing a solid incident response operation that is as much about protecting data as it is about protecting the corporate image.

Incident response (IR) is becoming part and parcel of a security strategy, experts say. More than sixty percent of organizations say they have IR plans in place, according to a recent report by Arbor networks and the economist intelligence unit which surveyed some three hundred sixty C-level or board level business executives around the globe on their incident response postures. According to the data, around two thirds of the organizations say a successful and smooth incident response operation in the face of a breach could ultimately enhance their reputation. "The saving face piece is big," says Dan Holden, director of Arbor's ASERT.

Cipher 4

Decryption Process

1. IC = 0.0652331, so monoalphabetic.
2. Frequency distribution does not match English. The general distribution of the columns suggests a substitution cipher, thus monoalphabetic substitution cipher
3. I assumed the letter with the highest frequency to be equal to e, in this case l with 144 occurrences
4. I did a kasiski search of 4 letters and found 'pEEx' and 'kEEg' were both seen 2 times.
5. I looked at the frequency graph and noticed 'p' and 'x' were of the more common letter groups, so I compared this to the 8 other (excluding 'E') more common letters from the normal alphabet frequency.
6. p could be 'a,h,i,n,o,r,s,t' and x could be 'a,h,i,n,o,r,s,t' (p != x)
7. The only English-like word I could find was 'SEEN', So I let S = p and N = x
8. I noticed now that the 'kEEg' in the cipher text was potentially a word 'SdkkEEg' because it is extremely uncommon for repeating letters to happen two times in a row that are not a part of a word.
9. The only word that can match that criteria I could think of was 'SUCCEED', so I let U = d, C = k and D = g
10. I didn't find any other words that could represent another word so I decided to let the second highest frequency letter 'a' be the second highest normal alphabet frequency letter 'T' as it seemed logical, thus T = a.
11. The phrase 'SUCCEED rN' appeared and the logical assumption I made was that I = r for 'SUCCEED IN'
12. I saw 'DISEoSE' and I mad the logical assumption that the word is disease, thus A = o
13. I was looking through different repeated strings using kasisiki search and when i came to string length 13 I found 'SUyEiSTITIsUS' repeated three times and this looks exactly like 'superstitious'. So I let P = y, R = i, O = s.
14. At the start, the phrase 'I Au A' looks like 'I AM A' so I let M = u
15. 'I AM A SICb MAN', assuming K = b
16. 'SPITEhUw' = 'SPITEFUL', thus F = h, L = w
17. 'UNATTRACTIzE MAN I fELIEzE' = 'UNATTRACTIVE MAN I BELIEVE', thus V = z, B = f
18. 'I jAVE SEEN MeSELF TjAT I ONLe cANTED' = 'I HAVE SEEN MYSELF THAT I ONLY WANTED', thus H = j, Y = e, W = c
19. 'THOUtH OFCOURSE I CANT EmPLAIN' = 'THOUGH OFCOURSE I CANT EXPLAIN', thus G = t, X = m
20. 'INqURING' = 'INJURING', thus J = q
21. v and n were not found in the cipher text so they must be either v or n each.

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
o	f	k	g	l	h	t	j	r	q	b	w	u	x	s	y	v,n	i	p	a	d	z	c	m	e	n,v

Plaintext:

I am a sick man. I am a spiteful man. I am an unattractive man. I believe my liver is diseased, however I know nothing at all about my disease and do not know for certain what ails me. I don’t consult a doctor for it and never have, though I have a respect for medicine and doctors.

Besides, I am extremely superstitious, sufficiently so to respect medicine anyway. I am well educated enough not to be superstitious, but I am superstitious. No, I refuse to consult a doctor from spite

that you probably will not understand. Well, I understand it, though of course I can't explain who it is precisely that I am mortifying in this case by my spite. I am perfectly well aware that I cannot pay out the doctors by not consulting them. I know better than anyone that by all this I am only injuring myself and no one else, but still if I don't consult a doctor it is from spite.

My liver is bad. Well, let it get worse I have been going on like that for a long time - twenty years now. I am forty. I used to be in the government service but am no longer. I was a spiteful official. I was rude and took pleasure in being so. I did not take bribes you see, so I was bound to find a recompense in that at least a poor jest. But I will not scratch it out. I wrote it thinking it would sound very witty, but now that I have seen myself that I only wanted to show off in a despicable way. I will not scratch it out on purpose.

When petitioners used to come for information, to the table at which is at. I used to grind my teeth at them and felt intense enjoyment when I succeeded in making anybody unhappy. I almost did succeed for the most part, they were all timid people of course. They were petitioners.