Homework 1 (CSCI-B 504)

Kushal Pokharel

September 2025

1 Overall approach

Affine was the easiest to code, and with the index of coincidence, it would easily be found out which cipher was Affine, but luckily, it was the first cipher.

Next, I went through each cipher to see which cipher had repeating words always. This was a great hint that the cipher was a **substitution**. I found cipher6 to be such, and did the frequency analysis on it, and started replacing some letters based on the hints from frequency analysis of single letters, digrams, and trigrams, and got the partial plaintext and made some guesses based on that, which eventually led me to the full plaintext.

Next, I went ahead with the **Permutation** cipher. It was easy to identify the Permutation cipher because the plaintext is not translated to some random text, but just rearranged, which means the index of coincidence for this cipher comes around 0.065. I quickly found that the permutation cipher was cipher3. But since the keys could be of length up to 12. It was difficult to find the key and the plaintext, as I had to manually go through all the text in every possibility. I couldn't use the index of coincidence since all of them would give me around 0.065. Eventually, I found the plaintext with some heuristics, like searching for" and " words within the possible plaintexts.

Next, the easiest among the remaining ciphers was the **Vigenere** cipher. It was quite easy to code the solution for the Vigenere cipher and test which of those ciphers was Vigenere using the Index of coincidence for different (key) length segments.. Interestingly, two ciphers were found, and both of those were decrypted quite easily, which made solving one of the other two ciphers very easy.

Next, since I had the plaintext for one of the possible **Hill** ciphers. I tried deriving the keys of the Hill cipher from the ciphertext-plaintext pair I had. None of those gave me a valid plaintext. It was also easy to brute-force the solution for the Hill cipher(since the key length =2, a hint was given). Two plaintext letters and two coefficients were brute-forced and eventually got the right plaintext given by the Index of Coincidence.

Finally, I knew one of the Vigenere ciphers was **LFSR**. Since LSFR keys repeat with some periodicity, it made sense for the Vigenere cipher to also be LSFR. I already had the key for the ciphertext from the Vigenere. I used that to generate the keystream and see if they gave the needed ciphertext but it didn't (at first). So, I bruteforced LSFR for a key length equal to 2 and got a key stream which didn't match with Vigenere's, and it gave random output. But after correction of the cipher2, I got the right key stream similar to the Vigenere's and solved it.

2 Affine cipher

2.1 Approach

Due to its nature, $e(x) = ax + b$ and the fact that for the inverse of a to exist, gcd(a,26) should equal 1. The total combination of a and b is just 12*26, so for each combination, I output the text in some file. Then I used the index of coincidence for each decrypted text to see which one gave close to 0.065. Hence, the text was decrypted.

Used Python for all of them, as I knew I needed to use matrix multiplication for later ciphers.

2.2 Decrypted Text - Cipher 1

shortening the war

in february nineteen fourty two the germans hit back by introducing a new fourth wheel (multiplying the number of settings another twenty six times) into their naval enigma machines. the resulting 'net' was known to the germans as 'triton' and to the british as 'shark'. for almost a year bletchley could make no inroads into shark, and allied losses in the atlantic again increased alarmingly.

in december nineteen fourty two shark was broken, but german innovations meant that the allies had to wait until august the following year before naval enigma was regularly read again. by then the americans were active combatants, providing much-needed computer power to bletchley.

by d-day in june nineteen fourty four ultra was no longer so important. but still no one wanted the germans to sense that enigma was being read. when, a few days before the normandy landings, an american task force captured a german u-boat with its enigma keys, admiral ernest king, us commander in chief of the atlantic fleet, threatened to court-martial the officer in charge for endangering 'operation overlord', as the plan for the d-day landings was known.

by how much did ultra intelligence, gained from reading enigma ciphers, shorten the war? harry hinsley, based at bletchley during the war, suggests it was a significant asset. if it did not keep rommel out of egypt in nineteen fourty one, it certainly did so the following year, by preventing him exploiting his victory at gazala.

as general alexander put it, 'the knowledge not only of the enemy's precise strength and disposition, but also how, when and where he intends to carry out his operations brought a new dimension to the prosecution of the war.'

the loss of egypt in nineteen fourty two would have set back the re-conquest of north africa and upset the timetable for the invasion of france. according to hinsley, overlord would probably have been deferred until nineteen fourty six.

but by then the germans might have hit back with v-weapons and worse. enigma successes always needed complementing with other intelligence material, but the fact that the allies kept enigma secret until nineteen seventy four shows how much it meant to them.

2.3 Keys

a = 5 b = 12

2.4 Complexity

The complexity of the algorithm is just 12*26 loops; in each loop, the text is decrypted (one mod subtraction and one multiplication) with the chosen pair of keys.

3 Substitution

3.1 Approach

After identifying that Cipher6 was the substitution cipher by checking the repeated words manually, I applied frequency analysis on the cipher. Substituted the letters according to the statistics of the text and got the partial plaintext, which furthermore gave me an idea of what the substitutes of other characters could be. And finally, the gap was filled iteratively, giving the whole plaintext as a result.

3.2 Decrypted Text - Cipher 6

pokemon go is a free-to-play, location-based augmented reality game developed by niantic for ios and android devices. it was initially released in selected countries in july two zero one six. in the game, players use a mobile device's gps capability to locate, capture, battle, and train virtual creatures, called pokemon, who appear on the screen as if they were in the same real world location as the player. the game supports in-app purchases for additional in-game items.

pokemon go was released to mixed reviews, with critics praising the game's concept and the incentive to be more active in the real world, while criticizing frequent technical issues apparent at launch. despite such reviews, it quickly became a global phenomenon and was one of the most used mobile apps, having been downloaded by more than one hundred and thirty million people worldwide. it was also credited with helping local business grow. however, it has attracted controversy for contributing to accidents and becoming a public nuisance at some locations. multiple governments expressed concerns over the game's security, with legislation regarding it being passed in some countries as a result.

the concept for the game was conceived in twenty fourteen by satoru iwata of nintendo and tsunekazu ishihara of the pokemon company as an april fools' day collaboration with google, called pokemon challenge. ishihara was a fan of developer niantic's previous transreality game, ingress, and saw the game's concept as a perfect match for the pokemon series. niantic used the crowdsourced data from ingress to populate the locations for pokestops and gyms within pokemon go.in twenty fifteen, ishihara dedicated his speech at the game's announcement on september ten to iwata, who had died two months earlier. the game's soundtrack was written by longtime pokemon series composer, junichi masuda, who also assisted with some of the game's design. among the game's visual designers was dennis hwang, who previously worked at google and created the logo of gmail.

on march fourth, twenty sixteen, niantic announced a japan-exclusive beta test would begin later that month, allowing players to assist in refining the game before its full release. the beta test was later expanded to other countries. on april seven, it was announced that the beta would expand to australia and new zealand. then, on may sixteen, the signups for the field test were opened to the united states. the test came to an end on june thirty.

at comic con two zero sixteen, john hanke, founder of niantic, revealed the appearances of the three team leaders: candela (team valor), blanche (team mystic), and spark (team instinct). hanke conveyed that approximately ten percent of the ideas for the game were implemented. future updates, including the much-anticipated addition of trading, more pokemon, implementation of pokemon centers at pokestops, a patch for the "three step glitch', and easier training, were also confirmed. he also stated that niantic would be continuing support for the game for "years to come".

3.3 Keys

substitutionmap = { 'A': 'o', 'B': 'z', 'C': 'p', 'D': 'h', 'E': 'n', 'F': 'g', 'G': 'j', 'H': 'l', 'I': 'v', 'J': 'x', 'K': 'r', 'L': 's', 'M': 'c', 'N': 'q', 'O': 'e', 'P': 'y', 'Q': 'a', 'R': 'i', 'S': 'w', 'T': 'f', 'U': 't', 'V': 'd', 'W': 'm', 'X': 'u', 'Y': 'k', 'Z': 'b' }

3.4 Complexity

There are two parts to this. First frequency analysis in which each letter is visited once to maintain the frequency of single letters, digrams, and trigrams, O(n) (n = ciphertext characters). After analyzing the frequency, manually updating the substitution map, and after getting all the elements of the substitution map in place, it was just replacing each characters with its substitute. Hence, O(n) (map lookup is considered constant).

4 Permutation

4.1 Approach

Finding out which of the remaining ciphers is Permutation was easy. The index of coincidence of the Permutation cipher would come around 0.065, as it is normal English just jumbled up. Solving for the key length and key was tricky because I had to manually change the key length, brute-force all the keys of that key length, and check for the decrypted plaintext to find if any one of the keys gave sensible English, and I couldn't apply the Index of Coincidence as all of the permutation results would have the same result. Maybe, if there was an index of coincidence for digrams, it would be helpful, but I got the result eventually after some manual work.

4.1.1 Decrypted Text - Cipher 3

manchester city have been cautioned against the introduction of facial recognition technology, which a civil rights group says would risk "normalising a mass surveillance tool". the reigning premier league champions are considering introducing technology allowing fans to get into the etihad stadium more quickly by showing their faces instead of tickets, according to the sunday times. if someone is recognised as having bought a ticket, they would be ushered in by a green light, and if not they would be halted with a yellow one. hannah couchman, the policy and campaigns officer at liberty, said: "this is a disturbing move by manchester city, subjecting football fans to an intrusive scan, much like taking a fingerprint, just so they can go to the saturday game." "it's alarming that fans will be sharing deeply sensitive personal information with a private company that boasts about collecting and sharing data on each person that walks through the gate, and using this to deny people entry. manchester city should urgently reconsider their involvement in normalising a mass surveillance tool which can track and monitor us as we go about our everyday lives." blink identity, a texas-based facial recognition company, says its technology can identify people walking at regular speed, so fans will not need to slow down to show a ticket or use a turnstile. to opt in, supporters would need to register a selfie taken on their phone. blink identity says it is also possible to "collect usable and shareable data" on every person that walks through its facial scanning software. the team behind blink identity have spent the last decade creating large-scale biometric identification systems in the middle east for the us department of defense, according to its website. last year live nation, the company that owns ticketmaster, announced investment in blink identity as part of plans to replace paper tickets with facial recognition. a source at manchester city said reports of a pilot facial recognition scheme were premature and there was no such plan in place. they said the club would always be "open to exploring new and appropriate technologies and systems to improve fans' experience at the stadium

4.2 Keys

key-length = 9 [7, 1, 8, 2, 6, 0, 5, 4, 3] 0th character mapped to 7th index, 1st mapped to 1st and so on..

4.3 Complexity

Since this is brute force, I tried all the keys of length 9, which means I tried 9! keys. And I also tried for different key lengths(k different keys), and each time I had to decrypt using that key (n characters), so the complexity is: O(k* 9! * n). Manually inspected the plaintext. I also tried with key length =11 and 12, which almost broke my machine.

5 Vigenere cipher

5.1 Approach

This was quite easy as it had a standard approach. BY putting different key length values and with the index of coincidence, I found out that 2 of the ciphers were Vigenere-based. After I got the correct key-length, I grouped them with the letters separated by the key-length and shifted them by values from 0 to 25, and if their index of coincidence came around 0.065, the corresponding letter of the key was found, and so on.

5.2 Decrypted Text - Cipher 2

for the last twenty years, the public gave credit for the discovery to martin hellman, a professor at stanford university, and two graduate students who worked with him at the time, ralph merkle and whitfield diffie. they started publishing their work in nineteen seventy six.

three professors at the massachusetts institute of technology at the time, ron rivest, adi shamir and len adleman soon followed with another similar approach known by their initials, rsa, which went on to become one of the dominant solutions used on the internet.

before public key cryptography, anyone who wanted to use a secret code needed to arrange for both sides to have a copy of the key used to scramble the data, a problem that requires either trusted couriers or advance meetings. pkc, as it is sometimes known, erased this problem by making it possible for two people, or more properly their computers, to agree upon a key by performing some complicated mathematics. there is no publicly known way for an eavesdropper to pick up the key by listening in.

the new document details how three employees of the british government discovered the same approach several years earlier, but kept it a secret for reasons of national security. a spokesman for the british government's gchq, said that the document's release is part of a "pan-governmental drive for openness" pushed by the labor party.

the document describing the steps of invention taken by the spies was written by james ellis, a mathematician and cryptographer who died less than a month ago. in it, ellis describes how he suggested the existence of what he called "non-secret encryption" in nineteen seventys.

ellis says that clifford cocks followed with a more practical solution in nineteen seventy three that was essentially the same thing as the algorithm published by rivest, shamir and adleman. the paper also says that malcolm williamson discovered an algorithm in nineteen seventy four that was very similar to the work of diffie and hellman. they did not replicate the work done by merkle and hellman.

in a telephone interview from his office in la jolla, calif., malcolm williamson said that he felt bad when others discovered the solution, but concluded, "i was working at the british government and that's just one of the restrictions you work under when you work for the government."

for his part, diffie said in a telephone interview from cirencester, england, that he thinks that gchq never realized the deep importance of what the mathematicians discovered. he said that he met james ellis several years ago and "within an hour of meeting me, ellis said, 'you did much more with this than we did.'" diffie also suggested that the history of ideas is hard to write because many people often find solutions to different problems only to later determine they've discovered the same thing.

5.3 Key - Cipher 2

frhfbptlrxxn

5.4 Another decrypted text - Cipher 4

mathematician to present a proof of the sensitivity conjecture

the sensitivity conjecture has stood as one of the most important, and baffling, open problems in theoretical computer science for nearly three decades. it appears to have finally met its match through work by hao huang, an assistant professor of mathematics at emory university.

huang will present a proof of the sensitivity conjecture during the international conference on random structures and algorithms, set for zurich, switzerland, july fifteenth to nineteenth.

"i've been attacking this problem off and on since twenty-twelve," huang says, "but the key idea emerged for me just about a week ago. i finally identified the right tool to solve it."

huang posted the proof on his home page and it soon generated buzz among mathematicians and computer scientists on social media, who have praised its remarkable conciseness and simplicity.

the sensitivity conjecture relates to boolean data, which maps information into a true-false, or one-zero binary. boolean functions play an important role in complexity theory, as well in the design of circuits and chips for digital computers.

"in mathematics, a boolean function is one of the most basic discrete subjects– just like numbers, graphs or geometric shapes," huang explains.

there are many complexity measures of a boolean function, and almost all of them – including the decision-tree complexity, the certificate complexity, the randomized query complexity and many others – are known to be polynomially related. however, there is one unknown case, the so-called sensitivity of a boolean function, which measures how sensitive the function is when changing one input at a time.

in nineteen ninety-four, mathematicians noam nisan and mario szegedy proposed the sensitivity conjecture concerning this unknown case.

"their conjecture says the sensitivity of a boolean function is also polynomially related to the other measures," huang says. "if true, then it would cease to be an outlier and it would join the rest of them."

huang developed an algebraic method for proving the conjecture. "i hope this method might also have some potential to be applied to other combinatorial and complexity problems important to computer science," she says.

5.5 Key - Cipher 4

vigenere

5.6 Complexity

Finding key length = $O(maximum\_possible\_key\_length * number\_of\_characters(n))$
After finding the correct key length, to find key = $O(key\_length * 26 * n)$ =¿ trying out each character and test the index of coincidence.

6 Hill cipher

6.1 Approach

Having a hint of plaintext makes decrypting the Hill cipher much easier. So, I tried the above two Vigenere ciphers to derive the coefficients, but none of those configurations gave me the right key matrix that was invertible. So, I bruteforced the solution by trying out all the configurations with 4 0-26 values equalling 26^4, since the hint of key length = 2 was given. And for those key matrices that were invertible. I found the plaintext for each combination and calculated the index of coincidence for English text and finally found the solution with one combintaion.

6.2 Decrypted Text - Cipher 5

woodstock fiftieth anniversary celebration canceled after lineup issues, financial problems woodstock fifty organizers announced wednesday that the beleaguered music festival, which was supposed to take place aug. sixteenth through aug. eighteenth, is officially canceled. "we are saddened that a series of unforeseen setbacks has made it impossible to put on the festival we imagined with the great lineup we had booked and the social engagement we were anticipating," organizer michael lang, co-founder of the nineteen sixty-nine festival, said in a statement. organizers had been forced to relocate woodstock fifty last week to merriweather post pavilion in columbia, md., a move that set off the most recent series of hits to the commemorative festival. when announced in january, the festival was supposed to take place in watkins glen, n.y., a few hours northwest of the original bethel woods site. but organizers failed to obtain the proper permits. they were later denied another permit for a venue in vernon, n.y., a town a few hours north of bethel woods that then voted unanimously to reject the festival's appeal altogether. artists announced in march as part of the festival lineup – including jay-z and miley cyrus, as well as nineteen sixty-nine performers dead and company, santana and john fogerty–were released from their contracts after the move to merriweather, according to lang, who said the artists (and their agents) had all been fully paid. he cited the new location as part of the reason so many headliners dropped out this past week, and urged the acts to donate a portion of their fees to headcount, a nonprofit that works with musicians to promote voter registration. after the relocation to maryland, organizers intended to turn woodstock fifty into a free concert benefiting headcount and some organizations working to combat climate change - a stark contrast with onetime plans to sell three-day passes for four hundred and fifty dollars. tickets were never made available after the on-sale date was indefinitely delayed in april, the same month a falling-out with the festival's financial backer, dentsu aegis network, led to confusion over whether it was still happening. (dentsu announced that it wasn't; organizers quickly clarified that it was.) the festival secured a new backer in may, but issues persisted. i.m.p., which operates merriweather, announced last week that it was in talks to host woodstock fifty. but seth hurwitz, the company's chairman, said as recently as monday that he hadn't yet heard who would perform. nonetheless, he said, the festival would "have a venue if they have a show." (audrey fix schaefer, an i.m.p. spokeswoman, confirmed that the smashing pumpkins and noel gallagher's high flying birds concert set for aug. seventeenth would proceed as planned.) a similar situation occurred with howard county officials, who confirmed tuesday that organizers hadn't requested a permit from the local police department. standard procedure calls for special-event permit applications to be submitted at least sixty days before the event, police spokeswoman lori boone said, but officials had been looking into making accommodations for woodstock fifty. "initially, when this festival was in search of a new home, we saw an opportunity to bring a piece of american history to our storied stage," howard county executive calvin ball said in a statement wednesday. "howard county is used to hosting big concerts, from virgin freefest to sweetlife to jazzfest and more. we had the experience, the infrastructure, and the passion to make this happen. howard county and merriweather were fully prepared to put on a world-class concert, if the festival promoters could secure this acts."

6.3 Key

Coefficient = [(10, 5), (19, 25)]

6.4 Complexity

For each combination, try to invert the key matrix and multiply with the ciphertext to get the plaintext. For each plaintext calculate if its index of coincidence is equal to English text. If yes, that is our key. So, finding key = $O(26^4 * number\_of\_characters)$ After finding the key, decryption = $O(number\_of\_characters)$

7 LSFR cipher

7.1 Approach

Finally, one of the two Vigenere ciphers must be the LSFR cipher, since the key was already available to me. I generated the coefficient from the available key and, with that coefficient, added keys to the stream. The keys continued to match until a length of 12, after which the keystream started repeating. This indicates that the periodicity of the keystream is 12.

7.2 Decrypted Text - Cipher 2

Same as Vigenere's Cipher 2's plaintext

7.3 Keystream

FRHFBPTLRXXN FRHFBPTLRXXN and so on. But the key length is just 2. FR. And the coefficient is [20,19]. The corresponding keystream with this configuration can be generated.

Complexity Finding the coefficient = One inverse matrix calculation and matrix multiplication, but since the key size is 2, the matrix size is just a 2*2 matrix. Generally, the matrix inverse and matrix multiplication is $O(n^3)$ but n is 2 here. And after finding the coefficient, finding the key stream and the plaintext means just getting the linear combination of previous keys and subtracting the key from the cipher = O(number_of_characters)

8 How can LSFR and Vigenere generate the same key and ciphertext?

In LSFR, with key length = 2, the following keys are a linear combination of two previous keys, and whenever two consecutive key characters repeat as the first one, it is then going to repeat the same pattern from the first. This length is known as the period of the key stream. In the given case above, the period of the key was 12, which means the keys will repeat after 12, which basically means that's a Vigenere cipher with a key length of 12, with the above key stream string up to length 12.