The Arab Cryptanalysts and Al-Kindi's Method

The introduction of 'A Manuscript on Deciphering Cryptographic Messages' book by Al-Kindi next to an imaginary portrait of Al-Kindi. (source)

Author: Sami Eltamawy

Institution: King's College London

Date: 10/04/2022

History

With the start of the Abbasid caliphate (or dynasty), AD 750, science has flourished in different areas such as algebra, mathematics, statistics, linguistics and many more.

Arabic scholars were provided a well-funded academic institution called Dar Al Hikma (House of Wisdom) for their own research. They were encouraged to study the science of linguistics and develop new techniques to help them with authenticity verification of the Quran, Islamic holy book, and the Hadith, which consists of the prophet's utterances.

As Islamic civilization back then grew, the need for secure communication among it's states and securing sensitive documents became inevitable. With the help of the Arab scholars, the state administration started to adopt what is called today "monoalphabetic substitution cipher" by simply creating a one to one mapping between the Arabic alphabets and itself in different order. For example the letter **A** in the plaintext is mapped to letter **Q** in it's ciphertext and the letter **Q** in the ciphertext is substituted by the letter **A** to generate the plaintext.

During this time, Al-Kindi (801-873), Abu Yusuf Yaqub ibn Is-haq ibn as Sabbah ibn 'omran ibn Ismail Al-Kindi, was one of the highly respected scholars in Dar Al Hekma in multiple fields such as Philosophy, Mathematics, Medicine, Optics, Astronomy, and geography.

His passion towards cryptography in addition to his deep understanding of linguistics, and statistics, led him to think of ways to decipher the encrypted message without knowledge of the secret keys in a reasonable time compared to the typical brute force method. This was the beginning of a new science that is being referred to nowadays as Cryptanalysis.

Al-Kindi documented his conclusions in ' Risalah fi Istikhraj al Mu'amma' book (Treatise on Decrypting Cryptographic Messages), 850 AD, the oldest existing book on cryptology.

Description of the Method

Al-Kindi came up with three notable cryptanalysis methodologies with the same goal; break the monoalphabetic substitution encryption of a ciphertext by utilizing different linguistic traits and characteristics.

1. **The first technique:: "the quantitative [kammiyya] expedients of letters"**

   Al-Kindi stated that in ordere for us to decrypt a ciphered message that we know it's language is to find a different plaintext of the same language long enough to fill one or two sheets, then count the occurrences of each character. We call the most frequently occurring letter "first", the next most occurring letter "second" and so on until we cover all the language alphabets.

   Then we look at the encrypted message, ciphertext, that we want to decrypt and perform the same exercise to find out the occurences of each letter. These two exercises will result in two tables with letters and their frequencies that we can use to start the substitution process to discover the mapping used in this encryption by comparing the statistical possibilities of both tables.

   This technique is now known as "frequency analysis" which saves significant time in checking thousands of potential keys to decrypt the ciphertext.

2. **The second technique: "qualitative [kayfiyya] expedients"**

   This methodology requires a more advanced understanding of the underlying linguistic rules of the language that we are trying to decrypt it's ciphertext.

   Al-Kindi described "combinable" and "non-combinable" letters, and determined their valid orders and positions by investigating morphology and word derivations as part of the "Laws of single words" science.

These rules helped the deciphering process by eliminating the invalid word derivatives. In other words, only substitute the letters as long as the order of the letter respects the known morphological rules and can form an actual word.

Al-Kindi also combined this methodology with the first one to help decipher the encrypted message that is not long enough to satisfy the frequency analysis requirements.

3. **The third technique: the "probable word" method**

   Al-Kindi based this methodology on the semantic properties of the actual text by noticing that each language has common words or phrases that can be used in the opening or the ending of the message. For example, the traditional opening for Islamic letters is "In the name of God, the Compassionate, the Merciful" which can be used as a guide to map the ciphertext letter to the known plaintext letters. Hence, it can help eliminate a number of possible keys depending on how long the known words or phrases are.

   Worth noting that this technique was used during the WWII to break the enigma code when the cryptanalysts realized the the german messages started always with a sentence about the weather that day, in addition to using the bombe machine which was designed by Alan Turing.

Examples of Use

The below figure is an actual exercise of the first cryptanalysis methodology of Al-Kindi on the Arabic letters. Which entailed counting the occurrences of each letter in both ciphertext and plaintext to produce the Letter Frequency Table shown below.

By comparing the statistical possibilities of both tables, discovering the secret mapping becomes possible in a reasonable time.

Letters	Frequency	Letters	Frequency	Letters	Frequency	Letters	Frequency
ā (ا)	600	n (ن)	221	k (ك)	112	ḏ (ذ)	35
l (ل)	437	r (ر)	155	d (د)	92	ṣ (ص)	32
m (م)	320	ʿ (ع)	131	s (س)	91	ḫ (خ)	20
h (هـ)	273	f (ف)	122	q (ق)	63	ṯ (ث)	17
(*) ū+w (و)	262	t (ت)	120	ḥ (ح)	57	ṭ (ط)	15
(*) ī+y (ي)	252	b (ب)	112	ǧ (ج)	46	ḡ (غ)	15
						ẓ (ظ)	8

The introduction Reproduction of al-Kindi's letter frequency table (source)

Strengthens and Weaknesses

Ultimately all the cryptanalysis methodologies were invented for one purpose; break the encryption without the knowledge of the secret key in a reasonable time which is also the case for Al-Kindi's cryptanalysis methodologies.

Al-kindi was trying to think with an adversary mindset and leverage his linguistic and statistical knowledge in finding clues that could decrease the posible secret keys for the adversary who is trying to break the Islamic state encryption back then. In which he was successful to a big extent with the monoalphabetic substitution ciphers.

On the other hand, Al-Kindi's methodologies didn't work in some cases. For instance, the first methodology which is based on frequency analysis can only work if the ciphertext is long enough in order to exploit the letters statistical possibilities of the language.

Another example where Al-Kindi's first methodology does not work is a novel written by the French author Georges Perac in 1969 called "La Disparition"; a 200-page novel that didn't use any words that contained the letter **e** which is the most used letter in the English language. This novel was later translated in English by Gilbert Adair while still following Perec's shunning of

the letter **e**. As you can imagine this would lead to incorrect mapping if Al-Kindi's first methodology was used without accounting for this important edge case.

Conclusion

Al-Kindi was able to introduce the world to some advanced discoveries in the cryptanalysis areas by leveraging his deep understanding of cryptography, mathematics, statistics and linguistics.

These methodologies are still valuable till today and were used as foundation to develop more resilient cryptographic algorithms that we use on a daily basis to secure our communications and sensitive data.

Reference

1. Simon Singh, The Code Book (London, 1999), pages: 14-20.
2. The Cryptological Origins of Machine Translation, from al-Kindi to Weaver, January 2017, by Quinn Dupont University College Dublin, Available at: https://www.researchgate.net/publication/319529566_The_Cryptological_Origins_of_Machine_Translation_from_al-Kindi_to_Weaver
3. Arab Contributions in Cryptography, Case Study: Ibn Dunaynir Effort, January 2017, Bushra Elamin, Prince Sattam bin Abdulaziz University, Available at: https://www.researchgate.net/publication/315656791_Arab_Contributions_in_Cryptography_Case_Study_Ibn_Dunaynir_Effort
4. Al-Kindi, Cryptography, Code Breaking and Ciphers, published June 2003, Available at: https://muslimheritage.com/al-kindi-cryptography/#note_1