

Arabic, Cryptology and the Internet Security

Dr. Serir-Mortad Ilhem
Université de Tlemcen

Abstract:

Prior to the internet was a science called Cryptology, a means to secure sensitive information from undesirable users. It started as a way to deceive the enemy in wartimes and hide war secrets. This science is as old as the Arab philosopher Abu Yusuf Ya'qub b. Ishaq al-kindī (796-873) time. This is why this paper targets on at the contribution of Arab scholars to the development of securing information. It will, also, focus on Arabic language and its adaptability, linguistically, to cryptology. Moreover, ways to impede the threat to internet security will be the third colossal point to highlight in the scope of such paper which, in the whole, will try to answer the following question: how may Arabic be adapted to the internet and how may the latter be secured in use?

Key words: Arabic, Cryptology, Internet Crimes, Internet security.

1. Cryptology: The Old Internet Sources

It may not be astonishing that golden ages of civilization started with the Arabs, but to know that the first onsets of the internet was firstly known by the Arabs is the most astonishing. Cryptology in its simplest meaning suggests secret meeting of letters and numbers, Cryptology as the science of secrecy through converting ordinary information (plaintext) into unintelligible gibberish usually called cipher text hiding information and analyzing it dates back to the days of the antiquity to develop basically with al-kindī who describes in his book the successful decryption of a Greek cipher text that was intended for the Byzantine Emperor based on frequency analysis and on known plaintext (usually numbers or letters) at the message start. Graph 1-1 shows a device for cryptology



Graph 1-1 A Scytale: An Early Device for Encryption.

2- Sources of Cryptology

There were some contributions to cryptology by the different civilizations which passed over Mesopotamia such as Sumerian, Babylonian and Assyrian like through clay tablets mainly marked by the cuneiform tablet as shows Graph 2-1:



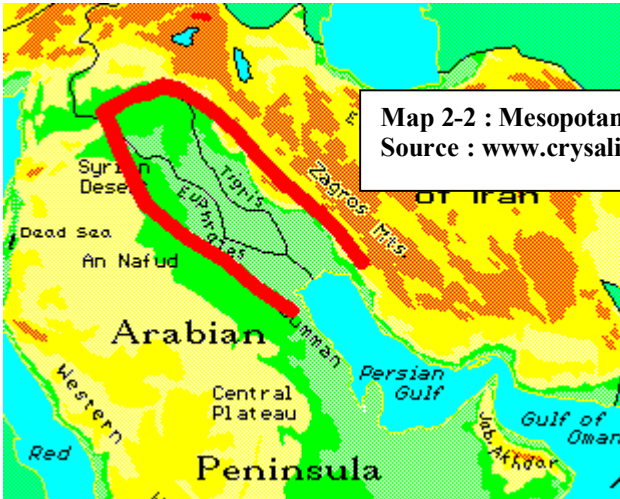
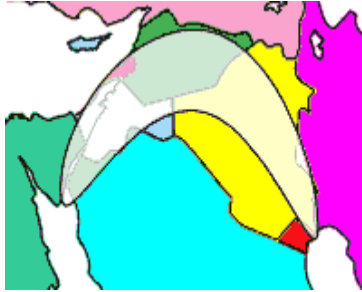
Graph 2-1: Cuneiform Tablet

Source: www.asor.org

Mesopotamian¹ cryptology is a small encrypted cuneiform tablet dating 1500 B.C.E which has been found on the banks of the Tigris River to use cuneiform for the common syllabic values using different spelling within the same word and leave the first consonant of other different signs.

¹ Mesopotamia which means land between the rivers. The Aramaic name is Beth-Nahrain meaning "House of Two Rivers" and is a region of Southwest Asia. In Persian it means Fertile Crescent because it takes this form as show Map 1 and 2:

Map 2-1 : Fertile Crescent
Source : www.crysalinks.com



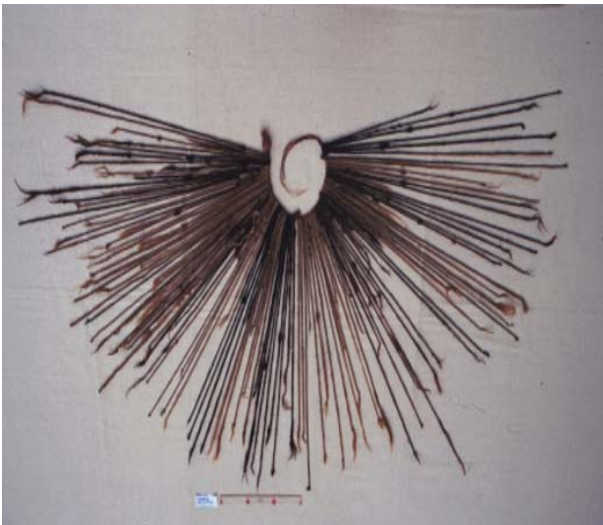
Map 2-2 : Mesopotamia Region
Source : www.crysalinks.com

Formerly, Babylonian and Assyrian scribes depended on Colophons, stereotyped ending formulas, to sign and date their clay tablets. In the same vein, and later on, Colophons were encrypted with numbers in order to replace the signs. In India, cryptology plays colossal role for political gain; especially used by institutes of espionage to communicate between spies. One of the types developed in such domain is Incan Cryptology which refers to a sequence of strings or threads from llama or alpaca hair which may symbolize numbers, letters, and

algebraic functions primarily set for record keeping and, then, turned to exchange secret information. See graph 2.

The number of Zero is marked through the absence of a knot, while number one is known by a figure eight knot and larger numbers are indicated by combining the smaller digits see graph 3. as for letters they cannot be easily deciphered

Because it represented threatening dictator information for the enemy, the Spanish invaders banned the code and the Incas secretly transfer information between each other



Graph 2-2: Incan Cryptology
Source: www.anthropology.wisc.edu



Graph 2-3: Numerical Usage through Incan Cryptology

Source: www.anthropology.wisc.edu

Prior to all these civilizations were the Arabs who were the first contribute to the internet through the art of cryptology.

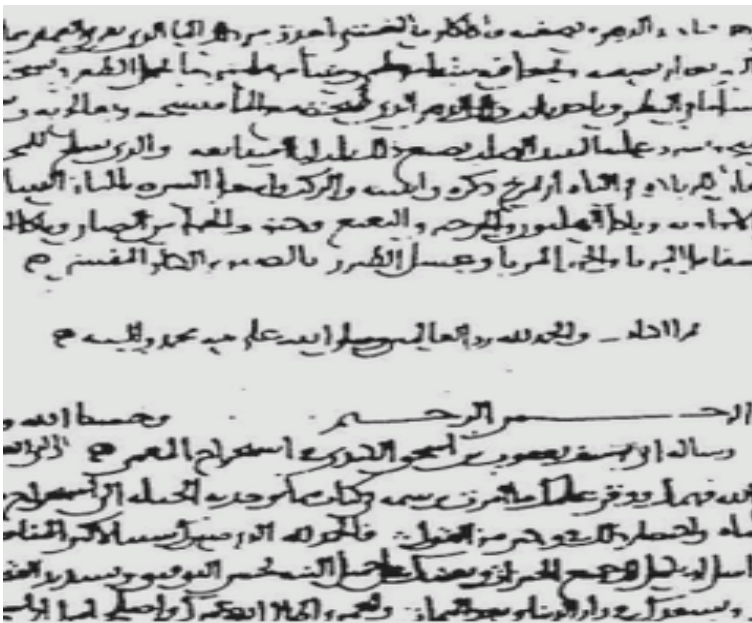
3- The Arab Advancements in Cryptology

The first origins of cryptology refer to the ninth century Arab scientist Alkindi's Treatise on Cryptanalysis, antedating any other book on Cryptology, such as ibn Adln or ibn ad-Duraihim, by more than three hundred years; and has been considered as the first and oldest known manuscript with recognition of foreign scholars such as David Kahn who once writes Cryptology was born among the Arabs (Kruh 2004). After that a series of fifteen manuscripts in nine volumes on Arabic origins of cryptology were translated in English to reflect a school of Muslim cryptology saved by history. Al Kindi (718-786)'s manuscript includes early advances in Mathematics by the Arabs, who were the first to develop the

science of algebra, and other advances in arithmetic, geometry; in addition to the art of writing, archives, records and correspondence purposefully adopted to learn Koran which was the key factor in the leadership of Cryptology. The volume, particularly chapter two, compounds statistical study by al-kindī on the frequency of occurrence of Arabic letters, in terms of order and the rules of combination and non-combination of letters in the Arabic language as resulted from linguistic, phonetic and morphological studies. This advance in Cryptology, undeniably, demonstrates the growth of Islamic civilization whose crucial interest relied on the greedy purchase for knowledge.

Through advances in Mathematics cryptology has been busied especially through the concept of the zero and Arab numerals that pave the way to analytical methods as for instance frequency analysis and computation of letter combination.

*Al-Kindī was the first to suggest methods for normal text: statistical: quantitative and qualitative techniques, vowel consonant combinations and tree diagram classification of major types of cipher. Graph 3-1 shows the first manuscript of Al-Kindī



Graph 3-1 Al-Kindi's First Manuscript On Deciphering Cryptographic Messages.

*Ibn Adlan contributed to Cryptology by the concept of Variable key for simple substitution, the use of different symbols for word spacing, the precision of 90 letters long for frequency analysis, three categories of letters: Common, Medium and Rare, step by step method of breaking a cipher.

*Ibn ad-Durayhim contributed to Cryptology through deep description of eight cipher systems: transposition, Substitution, Letter addition and omission, simple cipher mechanisms, arithmetic cipher, letter-word substitutions, substitution for plaintext letters, use of invented symbols for letters.

4- Arabic and Al Kindi Theory

It is very important to know the characteristics of Arabic language to be able to understand AlKindi theory in the analysis of an encrypted text. The rules instituted with Al Kindi resemble Mathematics and algorithms. One needs to know some principles like

1- to know the language of the written cryptogram through knowing “al-Fawatih” in the beginning and “at-Tamjidat” in the text and these are standard formulas characterizing any encrypted text like name, sender and salutation

2- length is important in the cryptogram, i.e., as much long as much easy to detect often called “la loi du grand nombre”

3- frequency method where to establish a table where the frequent letters are exposed; and Qur'an was first brought to light that Arabic has a characteristic letter frequency

4- good knowledge of the language at the level of vowels, consonants and linking characters

al- kindi dealt with these principles in his own way following three types:

1- the quantitative method based on the use of letter frequency analysis through calculating the frequency of Arabic characters' occurrence . there is a use of first a code breaker counting the frequency of the occurrence of symbols than classify the order of the highest frequency letters in the full alphabet. One important condition is required to succeed such quantitative method is the length of the ciphered text which should be very long; if not al Kindi suggests the qualitative approach

2- the qualitative method based on statistics which was called in that time “Arabic science” or for it was for the first time that the law of statistics was held by arab scholars and contributed a lot in the advance of cryptology and exact sciences for example Sibawayh's *al-Kitab*, or Ibn Gini's *al-Khasa'is* and al-

Farahidi's *al-Ain*. The method is based on an approach of combination and non-combination of consonants and for that one should have good knowledge of Arabic especially in terms of close association of some letters like the bigram of 'alif' with 'lam' or a trigram 'alif' with 'lam' and 'alif' again; and impossible combination between others or redundant such as 'هه'

3- Probability or the use of probable word, it is a way to guess about well known expressions often used in the beginning and end of messages like 'basmalah' or salutation or other words of protocol are traditionally used

al-Kindi summit of glory appears clearer with the Table 4-1 where he gives the amount of frequency with most letters of the Arabic alphabet:

Character	Relative frequency (probability)	Character	Relative frequency (probability)
أ ('lif)	0.1636	س (sin)	0.0248
ل (lam)	0.1191	ق (qaf)	0.0171
م (mim)	0.0872	ح (ha)	0.0155
ه (ha)	0.0744	ج (jim)	0.0125
و (waw)	0.0714	ذ (dhal)	0.0095
ي (ya')	0.0687	ص (sad)	0.0087
ن (nun)	0.0602	ش (shin)	0.0063
ر (ra)	0.0422	ض (dhad)	0.0055
ع (an)	0.0357	خ (kha)	0.0055
ف (fa)	0.0332	ث (tha)	0.0046
ت (ta)	0.0327	ز (zay)	0.0044
ب (ba)	0.0305	ط (ta)	0.0041
ك (kaf)	0.0305	غ (ghin)	0.0041
د (dal)	0.0250	ظ (za)	0.0022

Table 4-1 Al- Kindi's Relative Frequencies for Arabi
Letters Source: Zemouli (2002: 93)

Al-Kindi has been considered as the pioneer in the science of cryptology often thought to be one of the secret sciences or 'al-olum al Khafiyah'; this was the case of Arabic in those golden ages but when the decline of the Arabs increased, it paved the way to foreigners to advance and promote till the creation of the Internet which Arabs are struggling to have an independency in it.

5- Arabic Language and the Internet

A decade ago the call for internet addresses in Arabic alphabet was hugely made up and just recently an Arabic-language internet in the main addresses such as goggle and yahoo have been provided. The internet is geographically expanding and for that reason diverse linguistic groups use it to create a linguistic diversity available online; so an interest in Arabic language to be included in the web is increasingly noticed as the internet is raised in the whole Arab world; "Jordan's Internet and Datacomm Landscape" is a new report released by the Arab Advisor Group based on arabadvisors.com which denotes that Internet usage is growing in Jordan and Yemen. The use of Arabic on the web is restricted only to 1 per cent of Web content; this is why, Egypt contributed positively to boost use of Arabic on the web especially by putting national archives digitally on Internet in Arabic. Egypt registered with the name misr spelt in Arabic script.

On the other hand some links allow you to search in Arabic without the need or a keyboard in Arabic such as Yamli "ibhath bilarabi" where you are only asked to type it the way you say it

5-1 Recent Sites Index in Arabic

Arabic 2000.com lists the most recent adopted links where Arabic language is included as a category for search in Arabic where two texts are available in both Arabic and English

- Arabic internet guides (web directories with Arabic focus): Like their English counterparts, these sites contain indexes of Arabic and related sites arranged by category. Some allow you to search as well.
- Arabic search engines: Sites where you can look up a keyword or phrase in Arabic (sites with text in graphics format will not show up, of course). Most search their own collection of material and reviewed or submitted URLs.
- Arabic computing - commercial software and hardware: These are companies with Arabic software and/or Arabic hardware products who maintain a regularly updated web site.
- Arabic computing - free software, utilities, info, help, FAQ's and shareware: Apart from shareware sites, this category includes all sites with free Arabic related info, help and FAQ's as well as FTP sites which contain Arabic programs and utilities.
- Arabic newspapers and magazines online (with free access) : Here, you'll find Arabic newspapers, magazines and periodicals which are published regularly and allow free, unconditional access.
- Arabic online publications (books, poetry, leaflets etc) with free access : Here, you'll find free Arabic literature online including books, poetry, leaflets, know-how etc.
- Online stores/sites with Arabic and related merchandise and services : These include mail order bookstores, language consulting and related services (like translation), Arabic music, audio and so on.

- Arab educational and research institutions and those with Arabic program : Arab universities and research institutes / centers with special focus on Arabic and related studies
- Arab cultural and heritage sites (incl. arts, literature etc) : Sites with rich material related to the Arabic culture and heritage (preferably shared across the Arab world). Things like poetry, proverbs, food recipes and places are good examples
- Arab/Arabic professional and NGO organizations : Web sites of not-for-profit, non-governmental organizations, associations and professional bodies (no politics or political parties!)
- Miscellaneous Arabic and related sites : what was not mentioned above, interesting pages and impressive personal pages.

Technically, the Arabic web pages are only used with UTF-8, ASMO 708 and DOS encoding. In terms of script, Arabic calligraphy has moved into the 20th century and onto the internet, “several artists have their work online and have used the internet to develop new ventures in the art form of calligraphy” (Hassan 1999). In fact, there is a computational endeavour to facilitate generating Arabic script in the Internet.

6- Internet Security Threat

As much the internet serves humankind to prosper as less it could save them from harm at the level of information safety.

Zemouli (2002: 96) states four major threats

- 1- theft of data, such as that of military secrets from government computers
- 2- vandalism, including the destruction of data by a computer virus
- 3- fraud, such as employees at a bank channeling funds into their own accounts

4- invasion of privacy, such as the illegal accessing of protected personal financial or medical data from a large database

though they are systems that individuals maintain to protect their computers from these risks but some computers' crimes are beyond control that permit wrong doers to have easy access to any personal or intimate document.

6-1 Computer Crimes

It generally encompasses spam, fraud, offensive content, harassment, drug trafficking, cyberterrorism, phishing usernames and passwords and credit cards, problems of privacy, and confidential information. Most of these crimes are not detected or even inspected about by the authorities because victims usually ignore the suspected source of crime. There are two types of computer crimes

1- Old crimes using computer as a tool to download illegally music, images, or other form of piracy; phishing also stealing self-assurance of emails and fraud of websites.

2- New types of computer crimes are related to highly sophisticated technologies like denial of service attacks or DoS through for example flooding web servers with more data than they can process which forces them to become offline; or by 'hacking' an act that purposes at attacking a computer and or viruses that often infect computers and delete saved data

Algeria has known dreadful losses because of computer crimes as shows Table 6-1:

Institution	Hacker	Type of access/damage
Conseil National Economique et Social	unknown	Minor damage
Ministere de l'education Nationale	Unknown	Minor damage
La Direction de la Douane	Unknown	Minor damage
Ministere des P.T.	Two times by hackers from Morocco	Serious damage
ENTV and Algerian Radio	Unknown	Minor damage
Newspaper systems	unknown	Minor damage

Table 6-1 : Algerian Institutions' Victim of Hacking

Source: Zemouli: 2002 from Séminaire sur la securité des systemes d'information des entreprises organisé en collaboration avec la chambre de commerce et d'industrie Suisse-Algérie- el-Watan, sept, 2002)

6-2 Computer Crime Prevention

Solutions that can be brought to computer crimes are mostly technical like

- 1- Firewalls a security policy in the form of a hardware or software devices that contribute to block attacks
 - 2- Software solutions: they are free and paid for to help manage spam email
 - 3- Hardware cryptography: a cryptography is used to save a range of security threats
 - 4- Authentification: only the real user is allowed to use
 - 5- Patches a program intended to assure the user participation and ease of use
- Some solutions can be related to the authorities

to prevent from crimes like the local police forces in the United Kingdom or the FBI in the United States which endeavoured a lot to create an automated fingerprint but in vain till could the Japanese National Police Agency launch the first system of fingerprint then develop it to individual images.

Conclusion

Arabs are pioneers in the creation of algorithms and science of cryptology that is still in use by different countries in their knowledge. The internet is the newest version to cryptology where information is stored but no granted safe is provided; this is why users should raise awareness to protect their privacy.

References

- HASAN, A, F, 1999. "*Arabic Calligraphy: the Art of the Calligraphy Scripts in Islamic History*". www.suite101.com accessed to September 2009.
- KAHN, D, L, 1996. "*The Codebreakers: the Story of Secret Writing*. NewYork: Scribner".
- KRUH, L. 2004. "*Arabic origins of Cryptology*". Find Articles.com. January 2004. Accessed to 11 November,2009.
- ZEMOULI, M. 2002. "*The Contribution of Arab Scholars in Cryptology Science, and the Problem of Internet Security*". International conference on Arabic language and technology of information. 28/29 December 2002. organized by the High Council of Arabic language. Algiers.
- www.physorg.com (2006) "*The Web:Arabic-language Internet*." United Press International. Accessed to October, 2009.
- www.arabic2000.com accessed to November 2009
- [www. Parliament.uk](http://www.Parliament.uk) 2006 accessed to December 2009