TAP
Technology
Advisory Panel
IPCO

Encryption: a primer

6 May 2019

1. Encryption is used to obscure a message so that it cannot be read by an adversary who has access to the encrypted message, even one who knows the algorithm used for the encryption. It is worth noting that this is distinguished from *steganography*, which is hiding the means of communication, so that an adversary is unaware that a communication has happened; and of course, both encryption and steganography may be used together.

2. Besides obscuring the content of a message, encryption is also used for other purposes:
   - Integrity: checking the message has not been tampered with;
   - Authentication: checking who the originator of the message is;
   - Non-repudiation of origin: being the unique sender of a message;
   - Anonymity: protecting identity.

3. Encryption can also be applied to "data at rest" in a computer system, where it is desired to share data only between authorised users.

4. Encryption is a fundamental component of current communication and data management systems, in enterprise networks, on the Internet and mobile phone network, and is often mandated by regulation, for example in financial services. Key based encryption involves a key, which is a string of bits, and an encryption algorithm.

5. All key-based encryption relies on mathematical complexity of some nature, where the only *publicly known* means of attack is "brute force" – that is simply trying many keys until one works. Hence, as technology advances, and the time taken for a brute force attack decreases, we must increase the number of permutations that must be considered, that is the key length – so the 1977 Data Encryption Standard (DES) used 56 bit keys, while the more modern Advanced Encryption Standard (AES) uses 128, 196 or even 256 bit keys, massively increasing the search space.

6. However, besides the DES use of short keys becoming insecure because of advances in computing power, in the 1990s it was also discovered that there was statistical structure in the "mathematical complexity" that rendered it open to a technique known as differential cryptanalysis. AES was designed to resist such attacks, but there is no mathematical proof that AES is not open to as yet undiscovered mathematical analysis. Likewise, the conceptualisation of "quantum computers", which could solve these complex mathematical problems much more quickly, has resulted in significant work on "post quantum crypto" to build systems that could defend against such an advance in computing; however, opinion is strongly divided on whether such quantum computing

systems can ever be made operational. Nevertheless, if something needs to be encrypted in a future-proof way, the possibility of much more powerful computers and algorithms in the longer term has to be considered.

7. Modern day key-based encryption falls into two distinct camps: symmetric encryption and asymmetric encryption – the latter often referred to as public key encryption.

8. In symmetric encryption, such as AES, we must secretly share a key between the communication participants, and the same key is used to encrypt and decrypt the message. Such systems are extremely robust provided the key can be safely exchanged, kept secret in the participants' systems, and the key is truly random – many systems have been breached because the keys are statistically predictable, which drastically reduces the number of keys that must be considered in a brute force attack - a particular problem when keys are generated by software. In principle, the number of possible 256 bit keys is similar to the number of atoms in the universe, but if keys are generated by some sort of algorithm, the actual number of possibilities may be far fewer.

9. Asymmetric encryption has two keys, often referred to as a public / private key pair; as the name implies one key is held privately, while the other is published for all to see. The first publicly known algorithm was RSA (named for the discoverers Rivest, Samir and Adleman). This relies on the fact that if you create a number by multiplying together two very large prime numbers, it is extremely difficult to recover these two prime factors. More common today is Diffie-Hellman (DH), which again relies on a computation that is easy to perform in one direction but hard in the other.

10. However, asymmetric encryption is much more computationally intensive than symmetric encryption, so the two are more usually combined. An example is secure HTTP for access to websites: a website publishes its public key; a user wishing to securely communicate with that website encrypts a message using that key knowing that the website can decode it using their private key; this leads to a secure path over which a "session key" can be safely exchanged to permit the ongoing communication to be performed using the much faster symmetric encryption.

11. The weaknesses of cryptography are perennial: the security of key distribution is paramount, even with public key systems we need to ask what is "published for all to see"; the cryptography relies on mathematical complexity to which there is *no currently known simplification*, but we have no proof that such simplification does not exist; the software implementing the system can have bugs, not least in its generation of random numbers.

12. That said, many robust implementations exist, and secure encryption already underpins very large parts of our lives, for example the banking system. The adoption of encryption is increasing – for web communication, encryption will soon become the default. This adoption has been partly in response to increasing online criminal activity that has been targeting "open text" communications by interception and spoofing, and is needed to protect citizens and their online activities - a move that is strongly supported by the UK intelligence community as being in the best interests of the national economy and hence security.