TAP
Technology
Advisory Panel
IPCO

Cryptographic Technologies: A guidance note

December 2020

Executive Summary

1. Encryption is ubiquitous. As individuals we use it every day, often without knowing that it is being performed on our behalf. It is essential for keeping us safe online and on our devices such as mobile phones. However, the protection that encryption provides can be a serious blocker for law enforcement. Often material that would be useful to investigations is unavailable because it is hidden under strong encryption. This can be as a result of positive action on the part of a suspect, but increasingly it is due to the ubiquity of the encryption services offered by Internet services and apps.

2. Cryptography is not just about the encryption of data. Cryptographic technologies give a wide range of protections, including pseudonymity and ephemeral identifiers, proof-of-identity for access and authorisation, and integrity of communications. The list is not exhaustive and novel ways of using cryptography are being developed all the time. Each of these protections can also have an impact on the effectiveness of law enforcement. In some instances they are useful for law enforcement processes to guarantee integrity of data used in an investigation.

3. How does cryptography operate in practice? We look in detail at the various methods employed on the Internet and in mobile phones. In this analysis we seek to highlight not just what is made difficult, but what information remains unencrypted, or is generated by the encryption process, that can be useful to law enforcement. Very occasionally the cryptographic technologies render useful information due to the way they are set up.

4. This document is focussed on looking at cryptographic techniques used for communications. There are many other uses, e.g. for secure storage of data at rest, for the protection of credentials, for secure computation, etc. We do not include these, simply to keep a bound on the scope, though we recognise that these also impact law enforcement, particularly where TEI is involved.

A Primer on Cryptography

5. This section is included in order to introduce the concepts and terminology of cryptography that will be used later in the document. Readers who are familiar with public-key cryptography and secure hash functions may skip over it. Owing to limitations

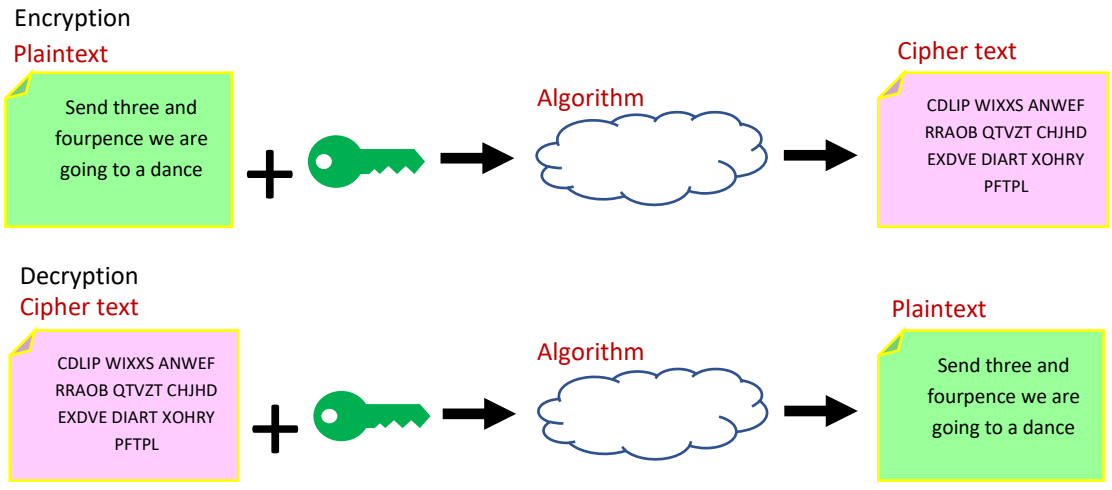
on space, this will be brief. For an excellent introduction to the subject, we recommend "Cryptography, A Very Short Introduction" by Fred Piper and Sean Murphy[1].

The Basics

6. Put simply, cryptography is the art and science of making secret codes and ciphers. Its sister discipline, cryptanalysis, is the art and science of breaking secret codes and ciphers. The term cryptology is used to encompass both disciplines, though in practice there is little separation between them. Anyone designing a cipher system really ought to know a lot about how to break them!

7. It's important to distinguish codes and ciphers. A ***code*** is a way of representing information in order that it can be used by a person or a machine. Codes are not always secret. For example, the ASCII code represents alphabetic characters as numbers ('@'=64, 'A'=65, 'B'=66) so that they can be interpreted consistently by any computer. A ***cipher*** is a method of hiding information in a way that it can only be retrieved by a legitimate person or entity (or a computer which they control). Computers are efficient at handling binary data, that is data which are encoded as sequences of 1s and 0s. ASCII may be represented this way, e.g. A=01000001, B=01000010. A ***secret code*** is a type of cipher; if I were to change which numbers represent each letter in the ASCII code above, and didn't tell you, then you would have no immediate way of interpreting the numerical data.

8. The other concepts to establish early are those of ***algorithm*** and ***key***. In a cipher system, the algorithm describes the process of converting plaintext to ciphertext (encryption), or ciphertext to plaintext (decryption). Algorithms are not always secret, and the security of an algorithm must not depend on it being kept secret from adversaries. The key is the secret information which allows legitimate users to encrypt and decrypt messages. The general pattern is

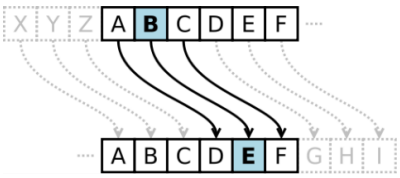


9. The simplest example of a cipher system is the Caesar code. In this system the encryption algorithm displaces plaintext characters by a number of positions (wrapping

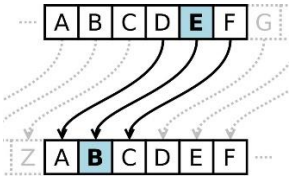
[1] "Cryptography, A Very Short Introduction" by Fred Piper and Sean Murphy. OUP 2002. Published online Sept 2013.

round at Z). The decryption algorithm displaces ciphertext characters by the same number of positions in the opposite direction. The key is the number of positions by which to displace.

Encryption: key=3



Decryption: key=3



For example:
Plain: NOW IS THE WINTER OF OUR DISCONTENT
Cipher: QRZ LV WKH ZLQWHU RI RXU GLVFRQWHQW

And a challenge: Clue – very similar in nature to the example. Not quite a Caesar though! What do you notice about the relationship between encryption and decryption? (Answer on the final page)

Cipher: CRUL GRYNN CU FRHUU MUUF YSYQL QL FRELVUH, NQSRFLQLS, KH QL HYQL?

10. Encryption / decryption algorithms fall into two categories:
    a. ***Substitution ciphers***, where plaintext characters, blocks of characters or words are substituted by other values, which may be of the same type or numerical encodings.
    b. ***Transposition ciphers***, where the values of the characters are unaltered but the characters in a message are permuted.

    Of course, it is possible to do both within the same algorithm.

    Transposition ciphers are rarely used in modern systems, because they do not do a very good job of hiding all the characteristics of plaintext. It would be easy to distinguish English text from, say Finnish, as in the latter we'd see many 'K's in the ciphertext!

11. Modern substitution ciphers again fall into two categories:
    a. ***Block ciphers*** (also known as ***electronic codebooks***), which take blocks of characters (typically 8 at a time) and mix them with key via a highly complex algorithm to produce a new block of characters. The Advanced Encryption Standard (AES) is a block cipher.
    b. ***Stream ciphers*** (also known as ***keystream generators***), where the key is used to seed a pseudorandom number generator (PRNG). The output of the PRNG is a keystream which is added to the plaintext (using binary arithmetic).

    Binary encoding of data (both plaintext and ciphertext) is essential to both approaches. Neither is intrinsically more secure than the other, and choice of algorithm is more dependent on the characteristics of the environment in which it operates, such as speed, latency and sensitivity to errors.

Public-Key Cryptography (Asymmetric Cryptography)

12. One of the fundamental problems in cryptography is how to manage and distribute keys. Keys are at least as sensitive as the messages they are used to protect, often much more so as the same key may be used to encrypt several messages. One could encrypt keys and transmit them using another secure cipher, but then that would need another key and the problem isn't really solved. Also, key generation and distribution is expensive. If a network has 1000 members, each of whom needs to be able to communicate securely and exclusively with any other member, then nearly 500,000 keys are required.

13. Prior to the 1970s the only reliable means of secure key delivery was by courier. Couriers can be intercepted; they can also be recruited as agents. Even if delivery could be made electronically, it would be infeasible at Internet scale as the number of keys would be astronomical. The solution to this problem was first developed at GCHQ in the early 1970s. The developers called it 'non-secret encryption'; we know it today as 'public key cryptography', as identical methods were developed by Diffie and Hellman (1976) and Rivest-Shamir-Adleman (1977).

14. Public-key cryptography answers two fundamental questions:
    a. Is it possible to agree a secret key between two parties who have no prior information about each other?
    b. Is it possible to devise a cipher system where the key to decrypt a message is different from the key used to encrypt a message, such that the decryption key cannot be derived from the encryption key?

15. Solving (a) means that key distribution is no longer a problem, though it does not authenticate the two parties to each other. Solving (b) means that encryption keys do not have to be kept secret. They can be published and distributed freely as ***public keys***, so also solving the key distribution problem. The decryption key is retained by the legitimate receiver as a ***private key***.

16. The methods for key agreement and for private/public key generation are quite mathematical, so we won't describe them here. If you are interested there are some great videos on **youtube** posted by **numberphile** and **computerphile**. Just do a search for 'public key', 'RSA' or 'Diffie-Hellman'.

17. Algorithms for the generation of public/private keys, and using them to encrypt data, are also referred to as ***asymmetric algorithms***, to distinguish them from ***symmetric algorithms*** where the same key is used for encryption and decryption (such as block and stream ciphers). Asymmetric algorithms are much richer in the ways that they can be used, however they are also considerably more expensive to operate. They are slower than symmetric algorithms and require much greater computer power, which makes them undesirable for message encryption but excellent at agreeing and transmitting the keys that can be used in symmetric algorithms. This is the standard pattern for Internet encryption: public-key is used to authenticate and derive shared keys; the shared keys are then used with a symmetric algorithm to encrypt the contents of messages.

18. While public-key cryptography was primarily developed to solve the key distribution problem, the properties of asymmetric keys have opened up a wealth of applications

that were previously impossible. These include authenticity of the origin of data, provable timestamping (notarisation), authenticity of ownership (non-repudiation) and authentication of entities such as websites. We shall cover these later in the section 'Uses of Cryptography'.

Hash-based Cryptography

19. It is not always necessary to encrypt messages. There are many instances where the content of a message requires no protection, but where it is important to ensure that the message is received intact and unaltered.

20. A ***Cryptographic Hash Function*** (CHF) or ***digest algorithm*** is a mathematical algorithm that maps data of arbitrary size (such as a message or document) to binary data of a fixed size in a way which is practically infeasible to invert. The output data is called a ***message digest***, ***hash value***, or simply '***hash***' of the original message or document. [Note: cryptographers tend to be lazy in their use of language, so you may find that the word 'digest' refers to the algorithm or its output, or both. We shall attempt to be consistent and use CHF for the algorithm, and hash value for its output]

21. CHFs are used to prove the integrity of message contents, because it is a cryptographically hard problem to create a second message that produces the same hash value. Further, CHFs are designed so that it is also infeasible to create two arbitrary messages that produce identical hash values.

22. CHFs are also used in conjunction with public-key algorithms. As these are computationally expensive, it is desirable to reduce the amount of data they encrypt, and so it is normal practice for hash values to be encrypted using the private key of the data owner/sender. The receiver can decrypt the encrypted hash value to recover the original hash value and check it against the message contents. As a result, the contents of the message can be verified; also, the authenticity of the sender is guaranteed.

23. The prevalent CHF is the Secure Hash Algorithm (SHA-2), standardised by the U.S. National Institute of Standards and Technology (NIST). It operates at hash value lengths of 256, 384 and 512 bits. Almost all other CHFs have now disappeared, either because of cryptanalytic vulnerabilities or their hash values were too short.

Ubiquitous Encryption

24. The term 'ubiquitous encryption' was coined in 2009 to refer to the massive increase in encrypted Internet communications that took place over the last decade. Today encryption is ubiquitous, which means it is used almost everywhere as the default, requiring no action from the end-user. It's all performed by the browser on your behalf – look for the little padlock symbol to the left of your browser address bar.

25. It wasn't always this way. Prior to 2009, secure Internet services were mostly limited to government, financial, legal, business and medical applications. Encrypted e-mail wasn't the norm, though you could sign up to secure e-mail providers that used encryption. In 2009 Google announced that they were going to encrypt all their services, starting with webmail. This forced similar action in from other webmail and search-engine providers, notably Microsoft and Yahoo! Further, the US National Institute of Standards and

Technology issued guidance for service providers to upgrade from 1024-bit public keys to 2048 bits by the end of 2013. The result was a massive shift across the industry, even for those sites which were previously unsecured.

26. Today, we use encryption without thinking about it, or even knowing that it is being done for us by our computers and apps. Browsing, search, e-mail, and all sorts of apps (not limited to messaging) encrypt our communications. Away from the Internet, our mobile phone connections are encrypted, albeit in a slightly different way. Smartcards use encryption to secure our banking, and even a simple thing like opening your car door uses encryption!

27. Why is this important? Google's position in 2009 was that it was their obligation to protect their subscribers and their data. Today encryption is essential, not just to protect the users, but also the websites that we visit. It is possible to hack into an unencrypted connection to subvert the normal functions of a website, and to interfere with the end-user devices connecting to it. Encryption very effectively denies this type of attack, and so even websites that don't handle user data or credentials are secured through encryption.

Impact of Encryption on Law Enforcement

28. Prior to everything being encrypted it was possible, under warrant, to access communications content, typically for e-mail, webmail and some messaging applications. Knowledgeable SOIs would protect themselves by subscribing to encrypted services. Many still do, preferring not to trust mainstream providers. However, by using non-standard services, such users expose themselves to identification by the fact of using a particular service.

29. Now that most Internet services are encrypted, the situation for law enforcement is much less dependent on the countermeasures used by SOIs. Communications content may be lost, however, in many instances the communications data may still be useful for investigations by identifying the parties to a communication along with associated data such as the time of communication and the location of the end-user device. In practice, there are several useful items of data that are not protected by encryption and which could be made available by collection of Internet Connection Records.

Uses of Cryptography

30. Up to now we have only discussed encryption as a means of protecting the content of a communication. Cryptography has many more uses, several of which also limit the effectiveness of law enforcement. The list below is limited to those of use to communications.
    a. Protection of data in transit (confidentiality) [already covered]
    b. Integrity of data in transit
    c. Authenticity of origin of data
    d. Timestamping
    e. Non-repudiation (authenticity of ownership)
    f. Authentication of entities

g. Anonymisation (pseudonymisation) of entities

b. Integrity of data in transit

This is used to show that an item of data has not been altered in transit, either by errors in transmission or by deliberate modification. The cryptographic technique uses a CHF to produce a hash of the message. The CHF algorithm is a one-way function producing a random-looking bit-string typically, of length 160 or 256. It should be computationally infeasible to produce a pair of messages which produce the same hash value.

c. Authenticity of origin of data

This follows on directly from data integrity, using public key methods. The hash of a message is encrypted using the private key of the originator. Anyone can decrypt the output using the originator's public key, and then verify that the hash value matches the original message. As the private key is only known by the originator, no-one else could have produced the encrypted digash value.

d. Timestamping

This again follows on from data integrity. A trusted 3rd party acting as a Time Stamping Authority receives the hash of a message, attaches time data to it, and then authenticates the resulting string using the same method as (b) above. Trusted 3rd parties are typically public notaries. Online services such as **docusign** provide timestamps for legal documents.

e. Non-repudiation

This refers to a situation where the originator of a message cannot successfully dispute its authorship. The techniques guaranteeing authenticity of data, with optional timestamping, directly support this use.

f. Authentication of entities

In any communication, it's important to know who you are talking to. Also, for access to Internet services, it is often necessary to prove that you have authority to access your account. There are many ways of doing this, but the most common methods are passwords, biometrics, and external token such as RSAid devices or banking cards with a reader. Increasingly these are being augmented by other methods, such as the service sending a one-time passcode to a separate device. When multiple methods are employed we call this ***multi-factor authentication***. Two-factor authentication has surged in popularity over the past two years.

It's important that the web services authenticate themselves to end-user devices too, to prevent malicious sites from masquerading as a legitimate website. This is achieved by the use of public-key certificates. Each website has a certificate which is signed by a trusted authority (as in (b) above). The certificate contains the website's public key, which the end-user device can use to encrypt a challenge message. Only the legitimate site knows the correct private key to decrypt the challenge, and so prove its identity.

Signing is most important to validate a website's public key, otherwise a bad site could simply construct a valid private/public key pair. To validate the public key we must check the signature, which was provided by a 'trusted authority', more properly called a ***Certificate Authority*** (CA). But how do we know if we can trust the CA? The most common CA certificates are preloaded onto computers and held in 'root stores' either in the operating system or managed by the browser. If we see a signature from a CA for which we have no certificate, then we must request its certificate which will, in turn, be signed by another CA. And so on, until we find a certificate in the root store, or fail. The very complex system of Certificate Authorities, their dependencies, and the processes/rules for creating, distributing and revoking public keys is called the ***Internet Public Key Infrastructure*** (PKI).

g. Anonymisation / Pseudonymisation of entities

This is the counter to (f) above. While the parties of a communication need proof of identity, it is often not desirable to make those identities available to eavesdroppers.

Anonymisation is the use of identification tokens that are unrelated to the end-user device or the person operating it. Such tokens must contain no personally-identifiable information (PII), and must not be derived from PII or linked to PII via a database.

Cryptographic techniques typically encrypt identification tokens, either by using a shared key known only to the communicating entities, or by encrypting using the public key of a web service. It is normal for the identification tokens to contain PII (such as a user-name), and so the resulting encrypted tokens are said to be ***pseudonymised***, as they could be linked back to individuals when decrypted.

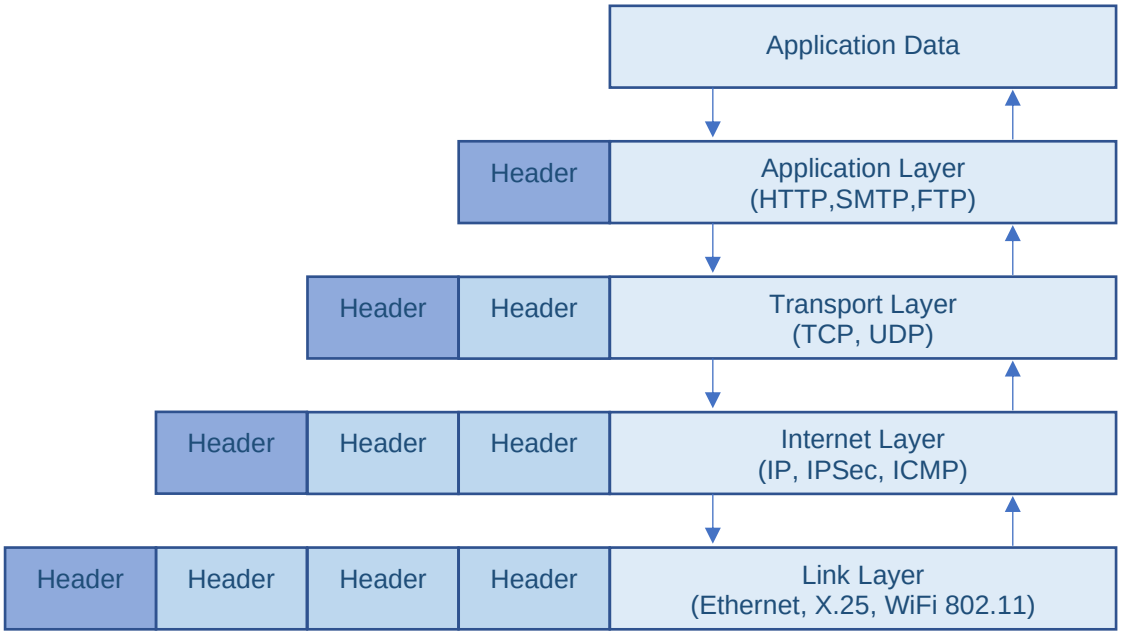
Another way of anonymising or pseudonymising identities is by the use of ***ephemeral identifiers***. There are various ways of managing these, though typically a token is generated at random by the server delivered to the end-user under encryption. The token may then be used to identify the end-user for a limited time until it is superseded. The biggest single application if this is the Temporary Mobile Subscriber Idenitity (TMSI), in mobile systems, which has been in operation for nearly 30 years.

31. Uses (b) to (e) above are important to the IP Act, not in the access to communications, but in the way that collected data is managed through law-enforcement processes. It is important for evidential purposes to demonstrate that data has not been altered, that it originates from a legitimate source, and that its time of origin can be verified.

32. (f) and (g) have significant impact on law enforcement. Pseudonymisation denies the ability to link a communication to a person or an end-user device. Authentication on the other hand may contain information that helps in identification. In particular, website certificates contain information about the site and the organisation which owns it.

Encryption in the Internet

33. In order to understand how encryption is performed on Internet data, and its impact, we must first look at how Internet connections are made and how data is exchanged over them. Internet communications operate on a layered model**,** known as the ***IP stack***, or ***TCP/IP stack***. Each layer in the stack provides a service that is essential to maintain a connection and exchange data. The layers are organised so that higher layers depend on the lower layers, as illustrated in the following diagram



34. Warning: the names of protocols for communications and security protocols are often abbreviated, so this section will contain a lot of 3- or 4-letter strings which appear meaningless, e.g. in the diagram above. These are just examples; there are any more protocols in practice. Please refer to the glossary where the full name and short description is provided for each protocol referenced in this paper.

35. The application layer is where applications (apps) handle user data and communicate this data to other applications on another computer. This is the layer in which all application protocols, such as SMTP/IMAP (e-mail), FTP (file transfer), HTTP (hypertext/web browsing) operate. Individuals processes are addressed via ports which essentially represent the services they offer, e.g. web (HTTP) on port 80 or outbound e-mail (SMTP) on port 25.

36. The transport layer performs host-to-host communications on either a local network or remote networks, separated by routers. It provides a channel for the communication needs of applications. The User Datagram Protocol (UDP) is the basic transport layer protocol, providing an unreliable connectionless datagram service. Transmission Control Protocol (TCP) provides flow-control, connection establishment and reliable transmission of data.

37. The Internet layer exchanges packets of data across network boundaries. It provides a uniform networking interface that hides the actual layout of the underlying network

connections. It is therefore also the layer that defines and establishes the Internet. The primary protocol in this scope is the Internet Protocol, which defines ***IP addresses***. Its function in routing is to transport packets of data to the next host, functioning as an IP router, that has the connectivity to a network closer to the final data destination.

38. The link layer defines the networking methods within the scope of a local network link on which hosts communicate directly with each other, without intervening routers. This layer includes the protocols used to describe the local network topology and the interfaces needed to affect the transmission of framed datagrams to next-neighbour hosts.

39. It is normal procedure to describe the stack in this way, running top to bottom. It can be useful to describe it running bottom to top, as this is a good way of showing the dependencies between the layers:

    - The basic need is for two computers to be able to communicate. The link layer allows communication of data between two neighbouring computers, or any pair of physically connected devices.
    - In order to communicate with a distant computer, data must be relayed over a series of link-layer connections. The Internet layer is responsible for discovering the route from one computer to another.
    - Once a route has been established, it is important to make sure that data passing over it arrives at the destination computer, in the right order, and is not corrupted. The dynamic nature of IP means that data packets can take different paths through the Internet, and so are not guaranteed to arrive in the right order. The transport layer is responsible for data sequencing, error correction, and (optionally) requesting re-sends when packets are missing or corrupted. This is the primary difference between UDP, where delivery is not guaranteed, and TCP where it is. TCP is referred to as a 'connection' as it allows for the two-way flow of control information for acknowledgements and re-send requests.
    - Finally, when we have a dataflow established, we need to ensure that it is correctly formatted to meet the needs of the application which will consume it. That's the job of the application layer. The port number ensures that the transport layer delivers the right data to the right application.

40. Referring back to the diagram, note that there is a standard construction at each layer. Viewed from the perspective of the sending computer, each layer receives data from the layer above, adds a header which is used to control the functionality at this layer, and then passes it (header+data) to the next layer down. The header information will become important when we look at the encryption methods, so we list some important header fields here:

    - Link layer: Data is normally transmitted in fixed length frames. Headers contain ***MAC addresses*** of communicating devices.
    - Internet layer: Data is transmitted in variable-length packets. Headers contain ***IP addresses***.

- Transport layer: Data is transmitted in sessions made up of one or more packets. Headers contain ***port numbers***.

41. We may now investigate where encryption can be deployed into the layered model, to secure the data being transferred, to authenticate the data and to authenticate the sources of communication. The most important thing to note upfront is that encryption can happen at each and every layer in the stack! A single item of application data may be encrypted many times over, though one does wonder as to the benefit of this. It Is also important to note that each layer can only encrypt the data sent to it, i.e. it can't encrypt its own header information as that is required for the service to perform its function. For example, at the Internet layer, the IP addresses must remain in clear text in order to route data to the destination computer.

42. The prevalent methods of encryption at each layer are:

    - Link layer: WPA for WiFi connections

      Internet layer: IPSec

    - Transport layer: TLS (previously known as SSL)

    - Application layer: S/MIME (e-mail), PKCS-7/11 (encrypted data and public key certificates), SSH (remote computer access), SFTP (file transfer) + many, many proprietary methods.

    Please see the glossary for expansion of these terms.

43. Link layer encryption is common for devices connecting via wireless interfaces; less so wired connections. Phones/laptops connecting to a wireless router in home, office or public access environments invariably use WiFi (we'll return to phone connections using mobile data later). WiFi security is provided by WPA, which can operate in various modes for enterprise (office networks) and personal use. For personal use, the most common form is WPA2-PSK, which uses a pre-shared key (password) to agree keys for data encryption with AES. Not all WiFi accesses are encrypted; public access (e.g. at a coffee shop) is usually not protected.

44. At the internet layer, the Internet Protocol has an accompanying security protocol called IPSec. IPSec encrypts packet data between remote computers; it can also be used to encrypt data between any pair of computers in the routing path between the communicating computers. The latter is particularly useful in connecting remote office networks to form a ***Virtual Private Network*** (***VPN***). The network is private because it is encrypted, wholly or in part, using IPSec; It is virtual in the sense that machines are not connected in a physical topology. IPSec is not common in personal communications, unless the user is subscribed to a VPN service for their Internet connections. As such it is useful for providing security over unprotected WiFi links, by connecting to a secure hotspot to proxy the connections to Internet services. IPSec authentication may use pre-shared keys or public keys. Pre-shared keys are much more common, especially in office environments where they are set by IT admin staff.

45. At the transport layer, security is provided by TLS. Strictly speaking, TLS sits between the transport layer and the application layer, which is why it is also associated with HTTPS,

an application layer format. The vast majority of web communications are encrypted using TLS. Further, most personal VPN services now operate over TLS, as it has lower overheads than IPSec and has some better security features as it doesn't need to replicate packet lengths. TLS, like IPSec, may use pre-shared keys or public keys, though pre-shared keys are rarely used. Most common is for a website to authenticate itself to a client via its public key, and for the client not to authenticate at all.

46. Application layer encryption is a cave of wonders! There are some standardised methods for encrypted e-mail, file transfer and remote computer access, but by and large, the approach of security is determined by the application provider. Messaging services such as Whatsapp, Snapchat, Twitter, Facebook Messenger, etc. all use proprietary methods. The same is true for video and voice services such as Skype, Teams, Google Meet, Facetime and Zoom.

47. So, why so are there so many different ways to provide security by encryption? A simple explanation is that each layer knows nothing about the security of data entering it from the layer above. So, if you don't protect the WiFi link at the bottom, you run the risk of exposing unprotected application layer all the way from the top. Also, encryption at the various layers serves quite different purposes, in line with the functions of each layer. At the link layer, it ensures that the connection is secure between neighbouring devices, but no further. The scope of the security is strictly limited to the two devices. The security is ***'point-to-point'***. In contrast application layer encryption secures all the communications between two applications, which may be in a client-server relationship or a peer-peer relationship between two end-user devices. The security is ***'end-to-end'***. The internet and transport layers secure the links between remote machines; they are ***'end-to-end'*** in the sense that no router in the comms path can decrypt the data, but are not 'fully end-to-end' as data is decrypted when it reaches the destination device, and sent in plain text to the application. It could potentially be accessed en-route by malware on the device.

More on end-to-end security

48. As described above, end-to-end security refers to the case where a communication is encrypted over the whole length of the communications path, without the ability to decrypt at any point along the path.

49. Confusion may arise when considering the entire path of a communication. We'll use a messaging service as an example.

    Alice wishes to send a message to Bob. She opens an app which connects to the message service. She enters the message, taps on the send arrow, and the message is delivered – or is it?

    What if Bob is not available to receive the message? What if Bob doesn't even have an account? Does the message sit in limbo waiting for him?

50. Most messaging services solve this by hosting separate communications for Alice and Bob. When Alice sends the message, it is delivered to a secure storage area owned by the messaging service. The message is encrypted in transit between Alice and the server, decrypted on receipt and stored securely within the messaging system. When Bob logs

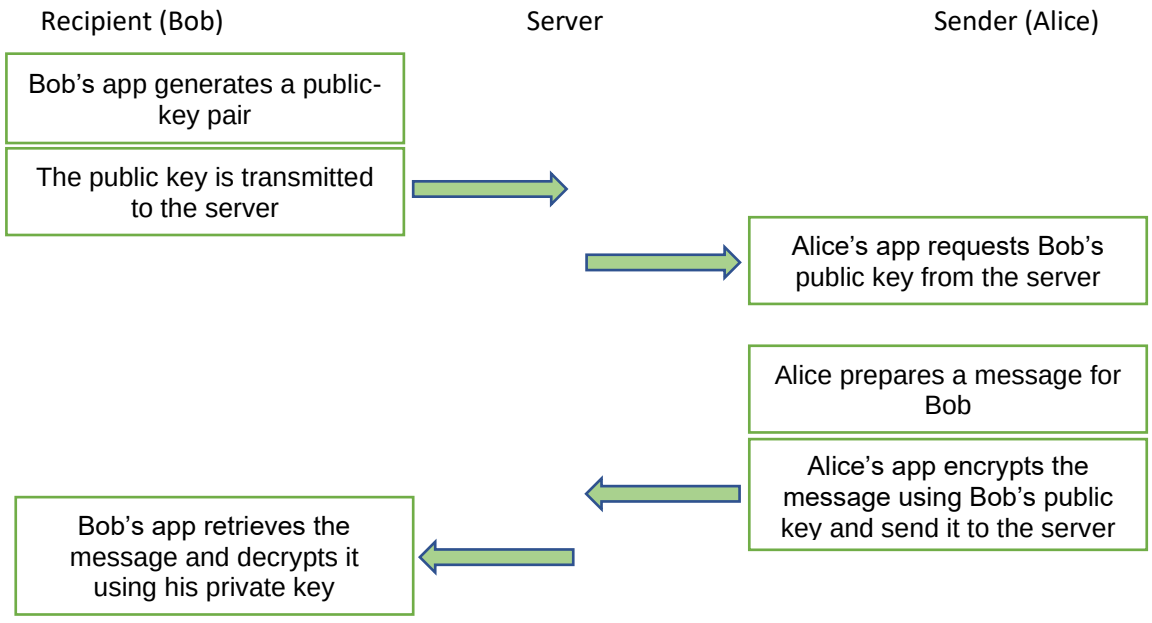
in, any messages waiting for him are retrieved, encrypted and transmitted to him. The same procedure is used for real-time messaging too, for example in video chat.

51. In this scenario, the communications are not encrypted 'end-to-end' as the full communications path is from Alice to Bob, but the service can (and does) decrypt messages in transit. This is particularly useful for law enforcement, as the messaging service provider is providing a telecommunications service and so can be served retention notices as a TO. It should be noted that the secure storage provided by the mainstream apps is highly secure, typically using AES, and it would require access to both the keys and data stores for a hacker to steal anything useful.

52. Some messaging services are, however, secured fully end-to-end between Alice and Bob. A notable example is **whatsapp**, which uses its end-to-end security as a differentiator for marketing. What is different here is that when Alice's app encrypts a message, it does so using a public key agreed by Alice and Bob. The details are rather complicated but basically Bob posts public key material on the messaging service that Alice can use to work out a key that only the pair of them know. The encrypted message is stored on the server with the instructions Bob needs to calculate the same key using Alice's public key material. When Bob logs in he is sent the encrypted messages and instructions. The good thing for the app provider is that no secure storage is needed as data remains secure at all times on their server. The bad news for law enforcement is that the service provider has no access to any communications content.

53. **Whatsapp** further guarantee that the key agreed between Alice and Bob is only ever used once, and cannot leak any information about keys used to encrypt other messages, either in the past or the future. This is known as ***'perfect forward secrecy'*** and involves a separate public-key exchange for each message.

54. As this is fairly complicated, the following diagram may help.

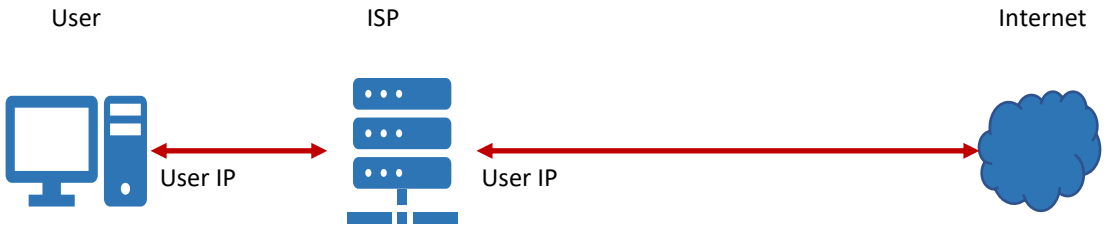


More on VPNs and Anonymity

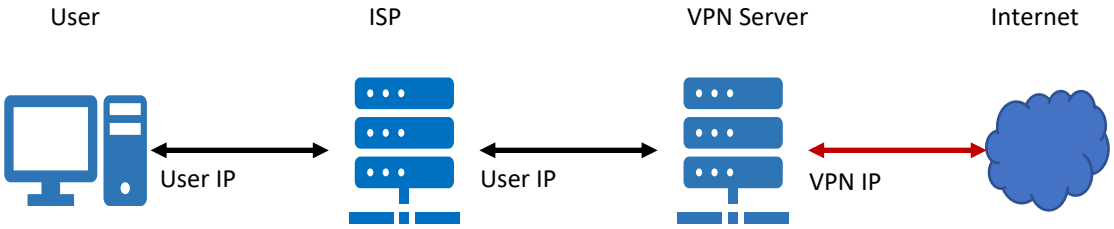
55. We introduced VPNs earlier (paras 44,45) as a mean of providing secure communication over a network with physically separated sub-networks and components. The term has been usurped to also apply to end-user devices connecting to a service designed specifically to act as a proxy for Internet services. These are becoming increasingly popular for securing accesses over otherwise untrusted links, such as public WiFi, and for anonymising Internet activity. Popular products are OpenVPN and WireGuard, in particular WireGuard is integrated into Linux systems so is available 'out of the box'. If you want a service (so you don't need to host your own VPN server) you will find several options from a simple Internet search.

56. When a device connects to a VPN service of this type, all the Internet Protocol data that would normally sent in clear is sent in a secure tunnel to the VPN service provider. The tunnel can operate as IPSec, or more commonly, TLS. This may feel a little strange as the TLS data (at the transport layer) is actually Internet layer data. It just needs a second instance of the IP stack to handle it. Encrypting the IP data in this way means that all information about the communication from the Internet layer up is made invisible to an interceptor, including destination IP addresses and ports and any identifiable information such as user-names and public key certificates.

57. For each incoming connection, the VPN service completes the Internet connection using an IP address which it owns. This way the IP address of the originating device is unknown when intercepting the link between the VPN and any Internet service. The following diagram illustrates this.



With no VPN, user data is unencrypted (red), and the user's IP is visible connecting to the Internet.



With a VPN, data is encrypted (black) between the user and the VPN service. The user's IP is replaced with the VPN's IP for connecting to the Internet.

58. Even so, some groups of users demand even greater degrees of anonymity. In the picture above, traffic-flow analysis of data being passed to/from the VPN service could be used to infer the link between end-user devices and web services. **The Onion Router**

(TOR) was developed specifically to provide strongly anonymised (stealth) Internet connections.

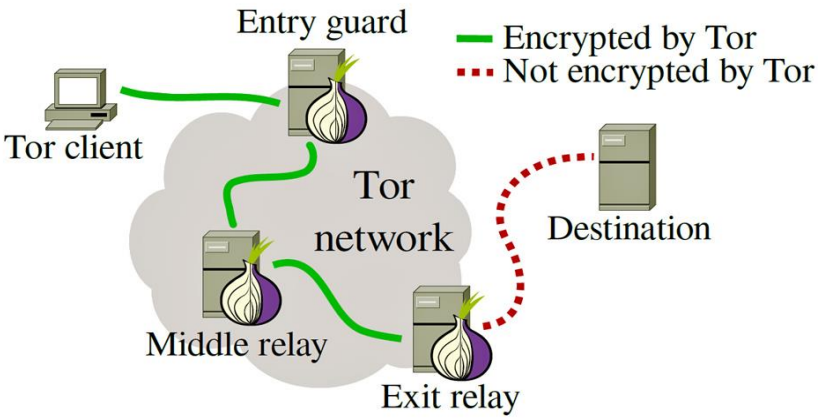
59. TOR operates as three VPN links acting in series. It comes with a special browser, the TOR Browser, which selects three servers in the TOR cluster. The TOR cluster is a collection of servers, managed by TOR subscribers and not owned by TOR. Selection of which servers to use for a particular connection is made at random, but also allows for selection and filtering of TOR servers by the end-user. These servers act as intermediate nodes in the communications path between clients and Internet services.

    First, data is sent from the client to a Guard Node. This is the entry point to the TOR network. Decryption reveals the address of a Middle Relay Node.

    When the data reached the Middle Relay Node, a second decryption reveals the address of an Exit Node.

    Finally, when it reaches the Exit Node, it is decrypted again to reveal the IP data to be sent to the Internet.

    The Guard Node knows the source of the connection, but not its destination. The Exit Node knows the destination but not the source. The Middle Relay Node knows neither!



The Impact of Internet Encryption on Law Enforcement

60. We have seen that encryption may operate at various levels of the IP stack. The headers containing communications data, along with data used to establish authenticity and agree keys will, however, remain in clear unless they are encrypted at a lower layer in the stack. We've also seen that it is possible to further encrypt these communications data by use of VPNs or TOR.

61. It is unusual for every layer to be encrypted, so what can we expect to see as unencrypted comms data in practice? We shall consider the case of a mobile phone running a secure app, not using a VPN/anonymisation service.

    First, as it's a secure app we'll have some application layer security. For this purpose, TLS operates at the application layer too (recall it sits between the transport layer and application layer). There really is no need for TLS to operate in series with other

application-layer security, but it does happen! Many text-and-image based apps actually use hypertext as their application layer format as they are little more than a skin on top of the browser, so TLS is a natural choice for provision of security.

We don't expect to see any Internet layer security, unless the phone is connecting through a corporate network. Even if it were, the IPSec would be limited to any collection made within the network.

We may have link layer security, if the phone is connecting over a secure WiFi link.

62. The availability of comms data varies according to the point of collection.

    - If the collection is made from a secure WiFi link, everything above the link layer is encrypted. The only available CD comes from the WiFi link, i.e. the MAC addresses of the phone and the router + the SSID of the router.
    - If collection is made by the mobile network operator, or intercepted at any point over the Internet, there is no WiFi link. Comms data associated with the Internet and application layers should be visible, including IP addresses and ports. Further, from the authentication procedures at the application layer, public key certificates should also be visible, which can be used to identify the app provider. Occasionally, dependent on the app, client information may be visible too in the form of login IDs.

Encryption in mobile telecommunications

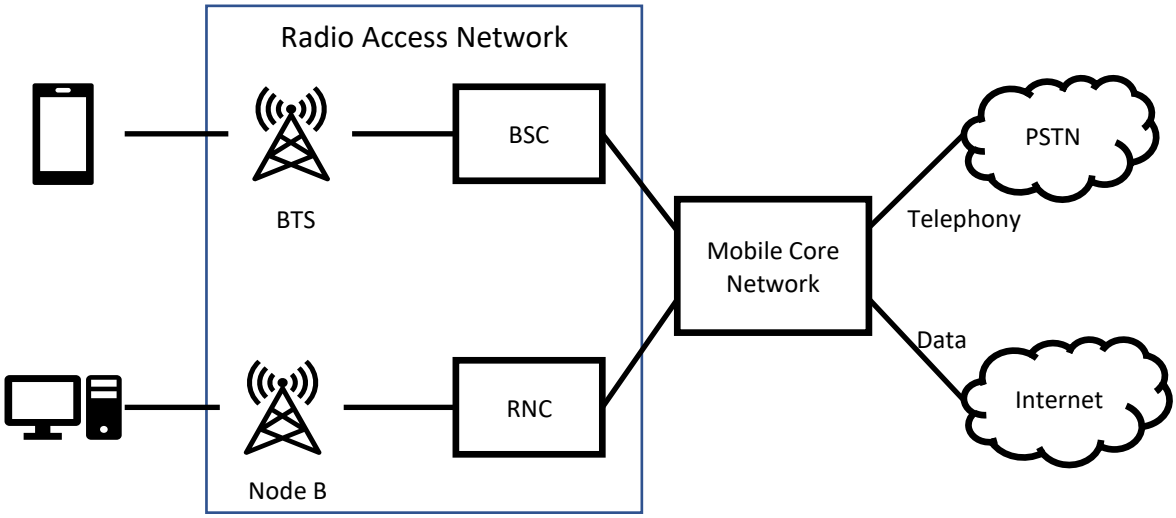
63. Mobile communications differ from Internet communications in two key respects. First, they were primarily developed for voice communications as opposed to data communications; second, they require a mobile Radio Access Network to provide access for phones, either for telephony or data services.

64. Digital mobile systems have developed significantly since their inception in 1991. Early systems (2G/GSM) had very little in the way of data services, being mostly for voice and, increasingly, messaging over SMS. 3G/UMTS was introduced in 1998 with improved data rates, necessary for allowing phones to access the Internet, but also to cope with the massive increase in SMS communications and giving better voice quality. 4G/LTE followed in 2009, with a major shift towards data communications, to the extent that voice services are supplied as a data application with calls being routed over the Internet. 5G is currently being rolled out; the first network was launched in 2017. It offers very high data rates for streaming videos, and will beat many home WiFi accesses for speed. It also offers much more rapid, low-latency connections needed for the Internet of Things, for efficient access from public infrastructure (street lights, traffic management, etc.) and devices in the home such as smart meters.

65. Security was designed into digital mobile systems from the very start. It was a critical design consideration for GSM as previous analogue systems had very poor security, allowing eavesdropping and (more importantly to the operators) stealing of credentials which could be used to make calls at someone else's expense! There were three requirements, and these are common from 2G to 5G.

1. Traffic encryption, for both voice and data, to defeat eavesdropping.
2. Protection of user identities, to avoid stealing credentials
3. Strong user authentication, to protect billing and revenue.

66. Security is standardised and mandated only on the link between the phone and the Radio-Access Network. The precise scope varies from 2G to 5G, but the differences are beyond the scope of this paper. The diagram below covers 2G and 3G accesses (4G and 5G are far more complicated, but follow the same principles). The important features are the ***Radio-Access Network*** (RAN) represented by the boxes on the left, to which the mobile user (Um) or user equipment (UE) connect, and the ***mobile core network*** which switches voice connections to the Public Subscriber Telephone Network (PSTN), and data connections to the Internet. The mobile core is also responsible for managing user accounts including authenticating users accessing mobile services.



2G and 3G access and data routing diagram

67. Please note that the standards relate to link-layer encryption between the handset and RAN. Data connections use the IP stack, and so the internet security methods discussed earlier for IP, session and application layers are likely to be present too. Between the RAN, Mobile Core and Internet gateways, 4G and 5G standardise on IPSec. However, it is not mandated and many operators choose not to use it due to the overheads. There are calls to mandate use of IPSec in 5G.

68. The most important of the three security functions is end-user authentication. This uses symmetric cryptography, which many people find surprising, but recall that these systems date from the early 1990s before Internet encryption was mainstreamed. The methods have proved effective for three decades, so there is a clear message about good use of symmetric algorithms.

69. The security of the phone resides in the Subscriber Identity Module (SIM) This holds subscriber data, most importantly the International Mobile Subscriber Identity (IMSI), and the shared symmetric key (called Ki or just K), which is also held securely in the network.

70. To authenticate a phone, first the phone sends its identity (IMSI or TMSI – see below) to the network. The network performs a lookup to validate the the identity and retrieve the secret K. The network then sends a random challenge (RAND) to the SIM on the phone. Both the SIM and network compute a cryptographic hash function with K and RAND as inputs, and the phone returns part of the result, so proving it knows K. The remainder of the output hash valueis used to derive keys for data encryption, and for 3/4/5G+, other functions such as data integrity.

71. The transmission of the IMSI is important for law enforcement. The IMSI uniquely identifies the SIM and hence the user account. It can be intercepted off the air and used to locate devices when they attempt to authenticate. However, there are a couple of twists to consider:

72. First, devices don't authenticate very often. Typically operators request only a few authentications per day. If a device is not authenticated, the previous authentication results are retained – i.e. the ciphering key is reused. If a device doesn't authenticate, there will be no subscriber data transmitted.

73. Second, and more importantly, all mobile systems (2G to 5G) use ephemeral identifiers to protect identities (requirement 2 above). Once a device has authenticated using its IMSI, the network assigns a temporary identifier (TMSI = Temporary Mobile Subscriber Identity) and sends it to the phone protected by the data encryption algorithm. At each subsequent authentication the TMSI value is replaced with a fresh one, so in principle a phone should not have to transmit its IMSI very often. In practice phones do need to transmit IMSIs, typically when connecting to a new network, which will not recognise the TMSI and so has to request the IMSI in order to authenticate.

The Impact of Mobile Telecomms Encryption on Law Enforcement

74. Under normal operation, there is little of value to law enforcement from collecting the radio link between phones and the RAN. All the communications content is encrypted, and subscriber data will only rarely be observed due to the use of TMSIs for most authentications.

75. It is possible to collect data when it reaches the mobile core. Because encryption is point-to-point between phones and the RAN, communications content and data are decrypted at point of entry to the mobile core network for forwarding to voice or Internet gateways. Mobile Network Operators are able to deliver content or CD under warrant, just as any other Telecommunications Operator.

Glossary

76. This glossary is presented in three parts: terms relating to cryptology; terms relating to Internet and mobile communications; and protocols and cryptographic algorithms.

77. Terms relating to cryptology

Algorithm	The rules and procedures for encrypting or decrypting data, or for producing message digests. Algorithms may be classified as

	'symmetric' when designed for use in classical (non-public-key) settings, or 'asymmetric' for use in public key settings.
Anonymisation	The use of identification tokens that are unrelated to the end-user device or the person operating it, so that the communications cannot be linked to an individual or device. (cf. Pseudonymisation)
Authentication	The use of cryptographic procedures and protocols to prove the identity of a user or device.
Block Cipher	A type of modern cipher where groups of characters are encrypted en-bloc, using an algorithm which performs a substitution on all the characters at once.
Code	A method of presenting information so that it may be interpreted by a receiving device. Codes are not necessarily ciphers, but may be if the rules for their interpretation are kept secret.
Certificate Authority (CA)	An organistion that issues digital signatures to validate the public certificate of other organisations / websites.
Cryptographic Hash Function (CHF)	A mathematical algorithm that converts data of arbitrary size (such as a message or document) to binary data of a fixed size in a way which is infeasible to invert, and for which it is infeasible to construct two inputs which convert to the same output.
Digest algorithm	Synonym for Cryptographic Hash Function
End-to-end encryption	Encryption which is performed between two endpoints, but where data may pass through intermediaries. No intermediary should be capable of decrypting the data or modifying it in transit. (cf. point-to-point encryption)
Ephemeral identifier	A form of pseudonymisation, using a token to identify a person or device for a limited time. Tokens may either be agreed under an encrypted channel, or derived from a cryptographic algorithm.
Keystream Generator	A type of modern cipher where data is encrypted by addition of a string of random values produced by a secure pseudorandom number generator (PRNG)
Message digest	The output of a cryptographic hash function; synonymous with 'hash value'.
Perfect Forward Secrecy	Relates to a cryptographic system where compromise of a single message gives an attacker no information about any other messages, either in the past or future.
Point-to-point encryption	Encryption which is performed between two neighbouring devices. If data is intended to reach a remote recipient, it is decrypted en-route and may be re-encrypted for its destination. (cf. end-to-end encryption)
Pseudonymisation	The use of identification tokens where the relationship to the end-user device or the person is protected, so that the communications cannot be linked to an individual or device. The means of protection may be cryptographic, but may also use othe security mechanisms. (cf. Anonymisation)
Pseudorandom Number Generator (PRNG)	An algorithm for producing 'random-looking' values given an initial seed value. When the seed value is a cryptographic key, PRNGs can be used to produce keystream generators.
Public Key Infrastructure (PKI)	The set of roles, policies and procedures needed to manage public keys, including their creation, signing, distribution, storage and revocation.

Substitution cipher	A type of cipher which replaces characters, or groups of characters, with unrelated values, while retaining their position in a message.
Transposition cipher	A type of cipher where the position of individual characters in a message is permuted to render the message unintelligible.

78. Terms relating to Internet and mobile communications

GSM	Global System for Mobile Telecommunications The 2nd generation mobile standard, but the first to use digital technologies.
IMSI	International Mobile Subscriber Identity A token held on a SIM card which uniquely identifies the user account for a mobile subscription. Contains information about the mobile network operator and the subscriber's account number.
IP address	A numerical label assigned to each device connected to a computer network that uses the Internet Protocol for communication. The IP address is used to identify the source or destination of data packets flowing through the Internet or other computer networks.
IP stack	Also known as the TCP/IP stack and Internet Protocol Suite, this is the conceptual model and set of communications protocols used in the Internet and similar computer networks. It is commonly known as TCP/IP because the foundational protocols in the suite are the Transmission Control Protocol (TCP) and the Internet Protocol (IP).
LTE	Long Term Evolution The 4th generation mobile standard, supporting high-bandwidth connections to the Internet. LTE is something of a misnomer, as it is being superseded by 5G.
MAC address	Media Access Controller Address This is a unique value assigned to a device which communicates on a network link. In a computer it is normally associated with a hardware device that is responsible for making network connections (the network interface controller).
Mobile Core	The mobile core network is the central part of a mobile network, which makes it possible for subscribers to access the services to which they are entitled, and to route subscriber data to/from other telephony networks and the Internet.
RAN	Radio Access Network This is the part of a mobile network that manages connections to end-user devices. It is responsible for maintaining the intergrity and security of communications over the radio link, and for routing data to/from the mobile core.
SIM	Subscriber Identity Module Normally referred to as a SIM card, this is a device held on a mobile phone to securely store the IMSI and the related subscriber key (K), which are used to identify and authenticate subscribers.
TMSI	Temporary Mobile Subscriber Identity

	This is an ephemeral identifier, used in place of the IMSI, to identify the subscriber to a mobile network. It is assigned and updated regularly by the network to avoid a phone being tracked.
TOR	The Onion Router TOR provides anonymity and encryption of Internet communications, by routing data through a series of relay connections.
VPN	Virtual Private Network A VPN extends a private network across a public network and enables users to send and receive data across shared or public networks as if their computing devices were directly connected to the private network. VPN technology is also used to secure communications to end-user devices, by making a private connection to a device and then routing data to/from the Internet on its behalf.

79. Protocols and security algorithms

AES	Advanced Encryption Standard The most common method of encryption used in the Internet, AES is a block cipher approved by the U.S. National Institute of Science and Technology (NIST).
Diffie-Hellman	An algorithm for public-key agreement. Actually encompasses a number of algorithms dependent on the underlying mathematics. The form most commonly used today is ECDH which uses the properties of elliptic curves.
Ethernet	A family of computer networking technologies commonly used in local area networks, but also in wider networks. Communication is typically over coaxial cable or fibre-optics.
FTP	File Transfer Protocol A standard network protocol for transfer of files between a cliens and a server on a computer network
HTTP	Hypertext Transfer Protocol An application layer protocol for distributed, collaborative information systems. HTTP is the foundation of data communication for the World Wide Web, and is used to transmit the content of web pages.
HTTPS	Hypertext Transfer Protocol (Secure) An extension to HTTP providing secure communications, including authentication of clients and servers, and encryption of HTTP data. Most web sites use this standard.
ICMP	Internet Control Management Protocol A supporting protocol in the Internet protocol suite. It is used by network devices, including routers, to send error messages and operational information indicating success or failure when communicating with another IP address. For end-users it is most commonly associated with 'ping' to verify a connection and to measure the round-trip time for the communications.
IMAP	Internet Message Access Protocol

	An Internet standard protocol used by email clients to retrieve email messages from a mail server
IP	Internet Protocol The principal communications protocol in the IP suite for relaying datagrams across network boundaries. Its routing function enables internetworking, and essentially establishes the Internet.
IPSec	Internet Protocol Security A secure network protocol suite that authenticates and encrypts the packets of data to provide secure encrypted communication between two computers over an IP network. It is often used in virtual private networks (VPNs).
PKCS-7	Public Key Cryptography Standard #7 A standard syntax for storing signed and/or encrypted data.
PKCS-11	Public Key Cryptography Standard #11 The standard defines most commonly used cryptographic object types, notably public and private certificates, and the formats for symmetric keys as used in AES.
RSA	Rivest-Shamir-Adleman Named after its inventors, a public key algorithm for generating private and public keys.
S/MIME	Secure/Multipurpose Internet Mail Extensions A standard for public key encryption and signing of MIME data, such as e-mail attachments.
SFTP	Secure File Transfer Protocol A network protocol that provides file access, file transfer, and file management over any reliable data stream. SFTP is not built on top of FTP, and was developed separately to provide the security functions
SMTP	Simple Mail Transfer Protocol A communication protocol for electronic mail transmission, from clients to e-mail servers and also between servers. Download of e-mails to clients is normally by IMAP.
SSH	Secure Shell A cryptographic network protocol for operating network services securely over an unsecured network. Typical applications include remote command-line, login, and remote command execution
SSL	Secure Sockets Layer The name previously given to the cryptographic services in TLS (see below). The name persists even though it was deprecated in 2011.
TCP	Transmission Control Protocol One of the main protocols of the Internet protocol suite. TCP provides reliable, ordered, and error-checked delivery of a stream of data between applications running on hosts communicating via an IP network.
TLS	Transport Layer Security A cryptographic protocol designed to provide communications security over a computer network. Widely used in applications such as web browsing, instant messaging, and voice over IP

	(VoIP). Websites can use TLS to secure all communications between their servers and web browsers.
UDP	User Datagram Protocol One of the core members of the Internet protocol suite. UDP uses a simple connectionless communication model with a minimum of protocol mechanisms and no guarantee of delivery, ordering, or duplicate protection.
WiFi	"Wireless Fidelity" A family of wireless network protocols, commonly used for local area networking of devices and Internet access. Devices that can use Wi-Fi technologies include personal computer desktops and laptops, smartphones and tablets, smart TVs, printers, smart speakers, cars, and drones.
WPA	Wi-Fi Protected Access A family of protocols to secure wireless computer networks using WiFi. The most common usage for home devices is WPA2-PSK which uses version 2 of the protocols with pre-shared keys between clients and WiFi hubs.
X.25	A standard protocol suite for packet-switched data communication in wide area networks. One of the oldest protocols (dating from 1976) it is still used in aviation.

80. And finally, the answer to the challenge in paragraph 9 is, perhaps unsurprisingly

    WHEN SHALL WE THREE MEET AGAIN IN THUNDER, LIGHTNING, OR IN RAIN?

    This is an example of a minuend cipher, that is, where the encryption rule is to subtract the plaintext from the key, when suitably encoded. The encoding is A=1, B=2, etc. and Z may be viewed as either 0 or 26, as the subtraction operates using modular addition in base 26. In this case this means simply reversing the alphabet.

    The important feature of minuend ciphers is that the algorithms for encryption and decryption are identical. So re-encrypting cipher text returns plain text.