Key recovery – Meeting the Needs of Users or Key Escrow in Disguise?

By Dr Brian Gladman, Worcester, UK

1. Introduction

In the early 1990s the United States (US) government made proposals that would allow the widespread use of cryptography without undermining the ability of intelligence and law enforcement agencies to read encrypted traffic. The idea, known as ‘key escrow’ allowed US government agencies, subject to legal controls, to access copies of the cryptographic keys being used to protect information exchanges.

Although these proposals were announced as voluntary, they nevertheless generated a storm of protest from private individuals and from the business community in the US. They were seen by many as a step towards the imposition of domestic controls on cryptography as well as a move that would undermine privacy and the right to be free from unwarranted State intrusion into the private lives of its citizens.

2. Key Recovery – The Vision

While many had outright objections to key escrow, others took a more pragmatic view, noting that circumstances existed in which government and business interests in cryptography might be better aligned. In particular they argued that the extensive use of strong cryptographic data protection involved risks for users in that key loss or damage could render critical data inaccessible. Steve Walker of Trusted Information Systems was an early pioneer of this view.

It was thus argued that end users needed an emergency means for key recovery to guard against the possibility that their primary keys were lost or damaged. For businesses protecting their data using encryption using employee based keys it was also argued that it would be important for the business to maintain access to its encrypted data in the event of problems with the keys held by employees.

Advocates argued that such key recovery mechanisms under user (or data owner) control could meet such needs and, in the hands of responsible users, could also be used to meet government access requirements when law enforcement or national security interests were at stake.

It is important to recognise that the aim was to achieve a compromise in which all parties gave up something in order to obtain a solution that could be rapidly adopted in the global market.

The compromise for users was that they would commit to the use of key recovery mechanisms with their cryptography and would, subject to legal safeguards and due process, allow these mechanisms to be used for law enforcement and national security purposes.

The compromise for government was that access would not be via communications intercept without user involvement but would instead require the knowledge and involvement of end users (or more likely the business data owner) through processes similar to those used in search warrants.

There were several reasons for believing that such a compromise could work:

- The real business needs for emergency key recovery in a number of scenarios;
- There is much greater public support for (and trust in) the search warrant style of access rather than that involving covert communications intercept;
- The possibility that a broad consensus built around key recovery could remove the major obstacles that were then undermining the development of the market for cryptographic security solutions.

After much effort by those who occupied the middle ground a significant consensus was built around the concept of key recovery with the result that it gained significant support both within the US administration and within the business community. However, some within the business community and many within civil liberties organisations remained unconvinced that key recovery would provide an effective and lasting compromise.

3. User and Corporate Needs for Key Recovery

In order to understand whether key recovery offers a practical compromise solution it is necessary to compare the business and individual key recovery needs with those of government in order to be sure that there is a significant overlap around which a compromise can be built.

Key Recovery For Stored Data

When cryptography is used for stored data there is a clear risk that key loss or damage might lead to a loss of the data itself. Such difficulties are very real and this means that the benefits of backing up encryption keys will almost always outweigh the additional risks that this will involve. It will thus be normal in a business environment to provide for emergency key recovery when encryption is used for stored data.

Law Enforcement agencies do encounter encrypted stored data in criminal investigations and this means that they can also benefit from the ability to be able to recover encrypted stored data (or the keys being used for its encryption).

National electronic intelligence agencies such as NSA and GCHQ are primarily involved in capturing and decoding electronic communications and this means that they have little interest in stored data.

Key Recovery for Communications

When cryptography is used for protecting communications channels there is no end user interest in key recovery since the unencrypted data streams will be available to all the parties involved. If an encryption key is lost or damaged it does not need to be recovered since the data can be sent again using a new encryption key.

It is sometimes suggested that corporations have a need for communications key recovery in order to check what their employees are doing by decrypting at the communications level. While this appears plausible at a superficial level, a more careful analysis suggests that such a requirement is unlikely to exist.

If a company wants to covertly intercept its own encrypted communications data it is hard to see why it would do this on the encrypted side of a communications link in preference to the unencrypted side. If the argument is that this has to be covert and not visible to any employees then it will have to intercept this traffic outside the company domain and this will be fraught with many technical and legal difficulties. For these reasons any intercept will almost certainly have to be done within the company boundary and it is then reasonable to ask why it would be the encrypted rather than the unencrypted communications data stream that would be the target.

> "Banks meet extensive audit requirements whilst also making extensive use of cryptography but all audit requirements are met using clear text and not encrypted communications data."

Beyond even this, when the issue is one of corporate oversight there will be a need to track what particular individuals are doing but this is very difficult to do at the communications level since the many higher level protocols have to be reconstructed from the low level data stream. In practice intercepting user oriented data on a Local Area Network or at a firewall/gateway will be orders of magnitude easier than doing so by intercepting encrypted communications data.

Another argument sometimes used for corporate communications key recovery is the need for companies to audit all transactions. Again, however, it is hard to see why this would be done after rather than before communications level encryption is employed. Banks, for example, meet extensive audit requirements whilst also making extensive use of cryptography but all audit requirements are met using clear text and not encrypted communications data.

For these reasons the need for corporate key recovery access to encrypted communications data can be safely discounted since there will always be far easier ways of meeting all such needs.

There appears to be a Law Enforcement interest in key recovery for encrypted communications data in order to provide for covert intercept similar to that available for voice data. However whereas voice data streams are mostly linked to individuals, digital data streams will often be heavily multiplexed before encryption and will hence carry the data traffic of many different subscribers. In such circumstances it is far from clear, in attempting to intercept the traffic linked to a particular target, how the interests of the other users of the multiplexed encrypted channel can be adequately protected.

For this reason it seems likely that the Law Enforcement interest will need to be focussed on higher level protocols which include a link to the identities of originators and recipients. This would include, for example, electronic mail protocols such as SMTP and the applications layer voice transfer protocols on ‘Internet phone’ software. Such protocols might be referred to as ‘end-to-end data exchange’ (or ‘application layer data exchange’) protocols to distinguish them from lower level communications protocols.

Law Enforcement will have requirements for intercepting such exchanges and there may also be situations where businesses would also want to have this right, for example, to check on an employee who might be revealing company information to outsiders.

Clearly the intelligence agencies are the organisations with the strongest need to preserve the ability to access encrypted communications data.

Summary

These requirements for key recovery are summarised in the following table:

	Communications Layer Data Exchange	Applications Layer Data Exchange	Stored Data
Intelligence Agencies	yes	yes	no
Law Enforcement Agencies	no	yes	yes
Corporations and Businesses	no	yes	yes
Private Citizens	no	no	yes

Key Recovery Requirements for Encrypted Data

From this summary it can be seen that the requirements of the different communities of interest do not overlap to the extent that might first be believed. In consequence achieving a compromise key solution is not as easy as it might first appear.

4. Key Recovery – The True Requirement

An important aspect of the key recovery concept is that of a compromise in order to achieve a position that nobody sees as perfect but which meets the majority of everybody's needs.

An essential part of the compromise is to accept that responsibility for key recovery should rest with users and not with governments. In a business environment the term user needs some amplification since it is the business and not the employees that is the owner of the data that is being encrypted. In this situation it would be expected that the control of key recovery capabilities will rest with the business and not with individual employees. Since businesses can reasonably be expected to support government efforts to combat crime and terrorism, placing responsibility for key recovery with them should not be counter to government interests.

It is reasonable to use the following definition of 'User Controlled Key Recovery':

Key recovery is a capability to recover the cryptographic keys being used to protect information when the primary means for obtaining these keys is unavailable. User controlled key recovery describes the availability of this capability in a form in which the owner of the data being protected can choose whether or not to enable it without otherwise changing the strength of the cryptographic protection available to them.

Several aspects of this definition are important. Fist of all, control of key recovery is given to the data owner – not the government, nor necessarily the end user. In a business situation it will be business owned data that is at risk and it is important for this reason that decisions on key recovery are taken by the business and not by individual end users. In contrast, in the use of cryptography by an individual private citizen, the user and the data owner coincide and in this situation it is the end user that should have control of any key recovery actions.

> "Given the suspicion that many private citizens have about US government motives in promoting key recovery, any consumer-orientated business that imposes such capabilities on its consumers would be taking very considerable risks."

Key recovery provides a secondary mechanism for accessing encryption keys when a primary mechanism has failed for some reason. The addition of such a capability will always introduce some additional security vulnerabilities[1] and this means that it is important that the data owner has the ability to decide whether the advantages that key recovery provides outweigh the additional risks that accompany it. Again the above definition makes it clear that for user controlled key recovery the data owner must be able to freely choose whether or not to use key recovery. Furthermore such a choice should be quite independent of the strength of the associated primary cryptographic capabilities.

[1] H. Abelson, R. Anderson, S.M. Bellovin, J. Benaloh, M. Blaze, W. Diffie, J. Gilmore, P.G. Neumann, R.L. Rivest, J.I. Schiller and B. Schneier. *The Risks of Key Recovery, Key Escrow, and Trusted Third-Party Encryption*. World Wide Web Journal 2(3) (Summer 1997) pp 241—257.

At this point it is worth noting that the above definition leads to two different classes of key recovery product:

- those designed for business use where the business (as data owner) and not the end user has control of the key recovery process (Corporate Key Recovery – CKR);
- those designed for personal use where the user (as data owner) has control of the key recovery process (Personal Key Recovery – PKR).

Because many businesses have to deal with individual private citizens as consumers it will also be important that these two classes of product are interoperable irrespective of whether key recovery is enabled or not. In other words:

- each party in a data exchange should be able to choose whether or not they wish to use key recovery without impacting in any way on the ability of the other parties to the exchange to make this choice for themselves;
- such key recovery choices should not impact in any way on the strength of the cryptographic protection available for the data exchange, nor on the ability of any party to the exchange to encrypt or decrypt the data involved.

By adopting this approach it will be clear that key recovery is voluntary and firmly under the control of the data owner in whose interests it is being offered. Within this framework it would also be clear that private citizens, many of whom distrust government intentions in respect of key recovery, would not have it imposed on them as a result of a need to interact as consumers with the business community. Of course the business end of a business transaction with a private consumer may well use key recovery but this is no different to the auditing of transactions that has to take place in such transactions irrespective of any use of cryptography.

In this scenario the key recovery compromise could work – it is not, unfortunately, the scenario that is being pursued.

5. Key Recovery – The Reality

If key recovery products met the criteria set out in section 4 all would be well but the reality is very different. Because the US government is making the key recovery a part of its export control regime for cryptographic products, the products emerging from a number of US companies follow the following pattern:

- products for the US domestic market with strong cryptography and with key recovery capabilities that can be switched on or off by the product user without impacting on the strength of the cryptographic protection available (Domestic Key Recovery – DKR);
- products for the international market where the cryptography available is ineffective for any serious use (i.e. 40 bit key length) if key recovery is switched off (Export Key Recovery– EKR).

From this it is immediately clear that the first of these products does not meet the domestic business requirement since it is the end user, and not the business data owner, who controls the use of key recovery. The export products do not meet the business requirement either because in any serious business use of key recovery this would still be needed even for 40 bit cryptography. At the same time these products are of no value to any personal users who do not want key recovery since they cannot obtain any effective cryptographic protection by using them. It is thus immediately clear that, for personal use, EKR products are simply key escrow products in disguise.

Products of EKR form have some potential in business where the business can secure direct control of the key recovery processes, for example, by operating their own Key Recovery Centres (KRC) or by using a Third Party KRC that they trust.

For personal use, however, EKR products are simply key escrow products since they deny such end users the choice of strong cryptographic protection unless key recovery is switched on. Private citizens are thus effectively forced either to give their keys to third parties or to forego any effective cryptographic protection. ***This is Key Escrow plain and simple***.

Although EKR style products are capable of meeting a number of corporate requirements, businesses outside the US will need to think carefully before adopting them if they expect to deal extensively with private citizens as consumers. The reason for this is that the US government requires that such products only operate with others for which key recovery is implemented and this in turn means that any business using key recovery enabled products of EKR form will force their customers to use such products as well. Given the suspicion that many private citizens have about US government motives in promoting key recovery, any consumer-orientated business that imposes such capabilities on its consumers would be taking very considerable risks.

6. Conclusion

Key Recovery, as originally proposed, could have provided the basis for a compromise solution in which businesses obtained strong cryptographic data protection whilst also supporting important government needs. In reality, however, key recovery has been taken over by the US government as a mechanism for preventing the widespread use of strong cryptography outside the US. Viewed in international terms, therefore, current key recovery products US originated key recovery products have characteristics that are more closely aligned with key escrow than with true user controlled key recovery.

If key recovery and key escrow are to be seen as essentially different it is important that the original intentions are restored so that the requirements of section 4 can be met. Provided that this is achieved, key recovery can make a valuable security contribution for business without undermining the interests of private citizens. At present, however, private citizens outside the US can only view key recovery as a disguised form of key escrow.

Companies offering key recovery products should adopt the CKR/PKR model of section 4 in place of the DKR/EKR model of section 5. That this is possible has been demonstrated by both Entrust Technologies (Canada) and PGP Inc. (US), both of which appear to offer products of this form ('Entrust/Entrust Solo' and 'PGP Business Secur ity Suite/PGP for Personal Privacy'). Of course these products may have other desirable (or undesirable) security features but they do appear to have adopted a key recovery strategy.

European businesses that need Key Recovery whilst also expecting to use cryptographic security products to interact with private citizens as consumers should adopt products of the CKR/PKR form rather than those of EKR form. If they adopt the latter they risk imposing these products on their customers and will hence be seen to support US government efforts to impose ineffective cryptographic security solutions on the private citizens of Europe.

Brian Gladman worked for the UK Ministry of Defence for 35 years, specialising in research and development within the software and systems engineering fields with an emphasis on information security. After retiring from the Ministry of Defence in 1994 he became Deputy Director of the NATO SHAPE Technical Centre. He is now an independent consultant.