The Risks of Key Recovery, Key Escrow, and Trusted Third-Party Encryption

Hal Abelson[1]
Ross Anderson[2]
Steven M. Bellovin[3]
Josh Benaloh[4]
Matt Blaze[5]
Whitfield Diffie[6]
John Gilmore[7]
Peter G. Neumann[8]
Ronald L. Rivest[9]
Jeffrey I. Schiller[10]
Bruce Schneier[11]

Final Report – 27 May 1997[12]

Abstract

A variety of "key recovery," "key escrow," and "trusted third-party" encryption requirements have been suggested in recent years by government agencies seeking to conduct covert surveillance within the changing environments brought about by new technologies. This report examines the fundamental properties of these requirements and attempts to outline the technical risks, costs, and implications of deploying systems that provide government access to encryption keys.

[1] MIT Laboratory for Computer Science/Hewlett-Packard, <hal@mit.edu>
[2] University of Cambridge, <ross.anderson@cl.cam.ac.uk>
[3] AT&T Laboratories – Research, <smb@research.att.com>
[4] Microsoft Research, <benaloh@microsoft.com>
[5] AT&T Laboratories – Research, <mab@research.att.com>
[6] Sun Microsystems, <diffie@eng.sun.com>
[7] <gnu@toad.com>
[8] SRI International, <neumann@sri.com>
[9] MIT Laboratory for Computer Science, <rivest@lcs.mit.edu>
[10] MIT Information Systems, <jis@mit.edu>
[11] Counterpane Systems, <schneier@counterpane.com>

[12] The latest version of this document can be found on the world-wide-web at <http://www.crypto.com/key_study>, in PostScript format at <ftp://research.att.com/dist/mab/key_study.ps> and in ASCII text format at <ftp://research.att.com/dist/mab/key_study.txt>.

Contents

Executive Summary

A variety of "key recovery," "key escrow," and "trusted third-party" encryption requirements have been suggested in recent years by government agencies seeking to conduct covert surveillance within the changing environments brought about by new technologies. This report examines the fundamental properties of these requirements and attempts to outline the technical risks, costs, and implications of deploying systems that provide government access to encryption keys.

The deployment of key-recovery-based encryption infrastructures to meet law enforcement's stated specifications will result in substantial sacrifices in security and greatly increased costs to the end user. Building the secure computer-communication infrastructures necessary to provide adequate technological underpinnings demanded by these requirements would be enormously complex and is far beyond the experience and current competency of the field. Even if such infrastructures could be built, the risks and costs of such an operating environment may ultimately prove unacceptable. In addition, these infrastructures would generally require extraordinary levels of human trustworthiness.

These difficulties are a function of the basic government access requirements proposed for key recovery encryption systems. They exist regardless of the design of the recovery systems – whether the systems use private-key cryptography or public-key cryptography; whether the databases are split with secret-sharing techniques or maintained in a single hardened secure facility; whether the recovery services provide private keys, session keys, or merely decrypt specific data as needed; and whether there is a single centralized infrastructure, many decentralized infrastructures, or a collection of different approaches.

All key-recovery systems require the existence of a highly sensitive and highly-available secret key or collection of keys that must be maintained in a secure manner over an extended time period. These systems must make decryption information quickly accessible to law enforcement agencies without notice to the key owners. These basic requirements make the problem of general key recovery difficult and expensive – and potentially too insecure and too costly for many applications and many users.

Attempts to force the widespread adoption of key-recovery encryption through export controls, import or domestic use regulations, or international standards should be considered in light of these factors. The public must carefully consider the costs and benefits of embracing government-access key recovery before imposing the new security risks and spending the huge investment required (potentially many billions of dollars, in direct and indirect costs) to deploy a global key recovery infrastructure.

Group Charter

This report stems from a collaborative effort to study the technical implications of controversial proposals by the United States and other national governments to deploy large-scale key recovery systems that provide third-party access to decryption keys[13]. Insofar as is possible, we have

[13]This report grew out of a group meeting at Sun Microsystems in Menlo Park, CA in late January 1997, including many of the authors and also attended by Ken Bass, Alan Davidson, Michael Froomkin, Shabbir Safdar, David Sobel, and Daniel Weitzner. The authors thank these other participants for their contributions, as well as the Center for Democracy and Technology for coordinating this effort and assisting in the production of this final report.

considered the impact of these policies without regard to individual encryption schemes or particular government proposals. Rather, we have attempted to look broadly at the essential elements of key recovery needed to fulfill the expressed requirements of governments (as distinct from the features that encryption users might desire). This report considers the general impact of meeting the government's *requirements* rather than the merits of any particular key recovery system or proposal that meets them. Our analysis is independent of whether the key-recovery infrastructure is centralized or widely distributed.

We have specifically chosen not to endorse, condemn, or draw conclusions about any particular regulatory or legislative proposal or commercial product. Rather, it is our hope that our findings will shed further light on the debate over key recovery and provide a long-needed baseline analysis of the costs of key recovery as policymakers consider embracing one of the most ambitious and far-reaching technical deployments of the information age.

Although there are many aspects to the debate on the proper role of encryption and key recovery in a free society, we have chosen to focus entirely on the technical issues associated with this problem rather than on more general political or social questions. Indeed, many have suggested that the very notion of a pervasive government key recovery infrastructure runs counter to the basic principles of freedom and privacy in a democracy and that that alone is reason enough to avoid deploying such systems. The technical nature of our analysis should not be interpreted as an endorsement of the social merits of government key recovery; in fact, we encourage vigorous public debate on this question.

1 Background

1.1 Encryption and the Global Information Infrastructure

The Global Information Infrastructure promises to revolutionize electronic commerce, reinvigorate government, and provide new and open access to the information society. Yet this promise cannot be achieved without information security and privacy. Without a secure and trusted infrastructure, companies and individuals will become increasingly reluctant to move their private business or personal information online.

The need for information security is widespread and touches all of us, whether users of information technology or not. Sensitive information of all kinds is increasingly finding its way into electronic form. Examples include:

- Private personal and business communications, including telephone conversations, FAX messages, and electronic mail;
- Electronic funds transfers and other financial transactions;
- Sensitive business information and trade secrets;
- Data used in the operation of critical infrastructure systems such as air traffic control, the telephone network, or the power grid; and
- Health records, personnel files, and other personal information.

Electronically managed information touches almost every aspect of daily life in modern society. This rising tide of important yet unsecured electronic data leaves our society increasingly vulnerable to curious neighbors, industrial spies, rogue nations, organized crime, and terrorist organizations.

Paradoxically, although the technology for managing and communicating electronic information is improving at a remarkable rate, this progress generally comes at the expense of intrinsic security. In general, as information technology improves and becomes faster, cheaper, and easier to use, it becomes less possible to control (or even identify) where sensitive data flows, where documents originated, or who is at the other end of the telephone. The basic communication infrastructure of our society is becoming less secure, even as we use it for increasingly vital purposes. Cryptographic techniques more and more frequently will become the only viable approach to assuring the privacy and safety of sensitive information as these trends continue.

Encryption is an essential tool in providing security in the information age. Encryption is based on the use of mathematical procedures to scramble data so that it is extremely difficult – if not virtually impossible – for anyone other than authorized recipients to recover the original "plaintext." Properly implemented encryption allows sensitive information to be stored on insecure computers or transmitted across insecure networks. Only parties with the correct decryption "key" (or keys) are able to recover the plaintext information.

Highly secure encryption can be deployed relatively cheaply, and it is widely believed that encryption will be broadly adopted and embedded in most electronic communications products and applications for handling potentially valuable data.[14] Applications of cryptography include protecting files from theft or unauthorized access, securing communications from interception, and enabling secure business transactions. Other cryptographic techniques can be used to guarantee that the contents of a file or message have not been altered (integrity), to establish the identity of a party (authentication), or to make legal commitments (non-repudiation).

In making information secure from unwanted eavesdropping, interception, and theft, strong encryption has an ancillary effect: it becomes more difficult for law enforcement to conduct certain kinds of surreptitious electronic surveillance (particularly wiretapping) against suspected criminals without the knowledge and assistance of the target. This difficulty is at the core of the debate over key recovery.

1.2 "Key Recovery": Requirements and Proposals

The United States and other national governments have sought to prevent widespread use of cryptography unless "key recovery" mechanisms guaranteeing law enforcement access to plaintext are built into these systems. The requirements imposed by such government-driven key recovery systems are different from the features sought by encryption users, and ultimately impose substantial new risks and costs.

Key recovery encryption systems provide some form of access to plaintext outside of the normal channel of encryption and decryption. Key recovery is sometimes also called "key escrow." The term "escrow" became popular in connection with the U.S. government's Clipper Chip initiative, in which a master key to each encryption device was held "in escrow" for release to law enforcement.

[14] The National Research Council's comprehensive 1996 report on cryptography includes a detailed examination of the rising importance of encryption. National Research Council, *Cryptography's Role in Securing the Information Society* (1996).

Today the term "key recovery" is used as generic term for these systems, encompassing the various "key escrow," "trusted third party," "exceptional access," "data recovery," and "key recovery" encryption systems introduced in recent years. Although there are differences between these systems, the distinctions are not critical for our purposes. In this report, the general term "key recovery" is used in a broad sense, to refer to any system for assuring third-party (government) access to encrypted data.

Key recovery encryption systems work in a variety of ways. Early "key escrow" proposals relied on the storage of private keys by the U.S. government, and more recently by designated private entities. Other systems have "escrow agents" or "key recovery agents" that maintain the ability to recover the keys for a particular encrypted communication session or stored file; these systems require that such "session keys" be encrypted with a key known by a recovery agent and included with the data. Some systems split the ability to recover keys among several agents.

Many interested parties have sought to draw sharp distinctions among the various key recovery proposals. It is certainly true that several new key recovery systems have emerged that can be distinguished from the original "Clipper" proposal by their methods of storing and recovering keys. However, our discussion takes a higher-level view of the basic requirements of the problem rather than the details of any particular scheme; it does not require a distinction between "key escrow," "trusted third party," and "key recovery". All these systems share the essential elements that concern us for the purposes of this study:

- A mechanism, external to the primary means of encryption and decryption, by which a third party can obtain covert access to the plaintext of encrypted data.
- The existence of a highly sensitive secret key (or collection of keys) that must be secured for an extended period of time.

Taken together, these elements encompass a system of "ubiquitous key recovery" designed to meet law enforcement specifications. While some specific details may change, the basic requirements most likely will not: they are the essential requirements for any system that meets the stated objective of guaranteeing law enforcement agencies timely access, without user notice, to the plaintext of encrypted communications traffic.

2 Key Recoverability: Government vs. End-User Requirements

Key recovery systems have gained currency due to the desire of government intelligence and law enforcement agencies to guarantee that they have access to encrypted information without the knowledge or consent of encryption users. A properly designed cryptosystem makes it essentially impossible to recover encrypted data without knowledge of the correct key. In some cases this creates a potential problem for the users of encryption themselves; the cost of keeping unauthorized parties out is that if keys are lost or unavailable at the time they are needed, the owners of the encrypted data will be unable to make use of their own information. It has been suggested, therefore, that industry needs and wants key recovery, and that the kind of key recovery infrastructure promoted by the government would serve the commercial world's needs for assuring availability of its own encrypted data. Several recent government proposals (along with commercial products and services designed to meet the government's requirements) have been promoted as serving the dual role of assuring government access as well as "owner" access to encrypted data. However, the requirements of a government and the requirements of the commercial world and individual users are very different in this regard, so different that, in fact, there is little overlap between systems that address these two problems.

The ultimate goal of government-driven key recovery encryption, as stated in the U.S. Department of Commerce's recent encryption regulations, "envisions a worldwide key management infrastructure with the use of key escrow and key recovery encryption items."[15] The requirements put forward to meet law enforcement demands for such global key recovery systems include:

- Third-party/government access without notice to or consent of the user. Even so-called "self-escrow" systems, where companies might hold their own keys, are required to provide sufficient insulation between the recovery agents and the key owners to avoid revealing when decryption information has been released.
- Ubiquitous international adoption of key recovery. Key recovery helps law enforcement only if it is so widespread that it is used for the bulk of encrypted stored information and communications, whether or not there is end-user demand for a recovery feature.
- High-availability, around-the-clock access to plaintext under a variety of operational conditions. Law enforcement seeks the ability to obtain decryption keys quickly – within two hours under current U.S. and other proposed regulations.[16] Few commercial encryption users need the ability to recover lost keys around the clock, or on such short notice.
- Access to encrypted communications traffic as well as to encrypted stored data. To the extent that there is commercial demand for key recovery, it is limited to stored data rather than communications traffic.

In fact, the requirements of government key recovery are almost completely incompatible with those of commercial encryption users. The differences are especially acute in four areas: the kinds of data for which recovery is required, the kinds of keys for which recovery is required, the manner in which recoverable keys are managed, and the relationship between key certification and key recovery. Government key recovery does not serve private and business users especially well; similarly, the key management and key recoverability systems naturally arising in the commercial world do not adapt well to serve a government.

2.1 Communication Traffic vs. Stored Data

While key "recoverability" is a potentially important added-value feature in certain stored data systems, in other applications of cryptography there is little or no user demand for this feature. In particular, there is hardly ever a reason for an encryption user to want to recover the key used to protect a communication session such as a telephone call, FAX transmission, or Internet link. If such a key is lost, corrupted, or otherwise becomes unavailable, the problem can be detected

[15] Dept. of Commerce, "Interim Rule on Encryption Items," *Federal Register,* Vol. 61, p. 68572 (Dec. 30, 1996)

[16] For example, the recent British "Trusted Third Party" system proposes similar law enforcement demands, requiring *one hour* turnaround time for TTP recovery agents. See U.K. Department of Trade and Industry, "LICENSING OF TRUSTED THIRD PARTIES FOR THE PROVISION OF ENCRYPTION SERVICES," (March 1997) (Public Consultation Paper).

immediately and a new key negotiated. There is also no reason to trust another party with such a key. Key recoverability, to the extent it has a private-sector application at all, is useful only for the keys used to protect irreproducible stored data. There is basically no business model for other uses, as discussed below.

In stored data applications, key recovery is only one of a number of options for assuring the continued availability of business-critical information. These options include sharing the knowledge of keys among several individuals (possibly using secret-sharing techniques), obtaining keys from a local key registry that maintains backup copies, careful backup management of the plaintext of stored encrypted data, or, of course, some kind of key recovery mechanism. The best option among these choices depends on the particular application and user.

Encrypted electronic mail is an interesting special case, in that it has the characteristics of both communication and storage. Whether key recovery is useful to the user of a secure E-mail system depends on design of the particular system.

The government, on the other hand, proposes a key recovery infrastructure that applies to virtually *all* cryptographic keys, including (especially) those used to protect communications sessions.

2.2 Authentication vs. Confidentiality Keys

Although cryptography has traditionally been associated with confidentiality, other cryptographic mechanisms, such as authentication codes and digital signatures, can ensure that messages have not been tampered with or forged. Some systems provide properties analogous to those of handwritten signatures, including "non-repudiation" – the recipient can prove to a third party that a message was signed by a particular individual.

Much of the promise of electronic commerce depends on the ability to use cryptographic techniques to make binding commitments. Yet some key recovery schemes are designed to archive authentication and signature keys along with confidentiality keys. Such schemes destroy the absolute non-repudiation property that makes binding commitments possible. Furthermore, there are simply no legitimate uses for authentication or signature key recovery. The private sector requires distinct keys for all signers, even when two or more individuals are authorized to send a given message; without that, the ability to audit transactions is destroyed. Government surveillance does not require the recovery of signature keys, either.

However, it is difficult to exclude authentication and signature keys from a key recovery infrastructure of the kind proposed by the government, because some keys are used for both signature and encryption.[17] Nor is it sufficient to exclude from the recovery system keys used only to protect financial transactions, since many electronic commerce schemes use keys that are general in scope. The same key might be used, for example, to encrypt personal electronic mail as well as to electronically sign contracts or authorize funds transfers.

It has been claimed that non-availability of a signature key can be a serious problem for the owner, who will then no longer be able to sign messages. But common practice allows for the revocation of lost keys, and the issuance of new keys with the same rights and privileges as the old ones. Recovering lost signature and authentication keys is simply never required.

[17] In fact, it is technically straightforward for two parties to use their authentication keys to negotiate encryption keys for secure communication. Any system that distributes trusted authentication keys would *ipso facto* serve as an infrastructure for private communication that is beyond the reach of government surveillance.

2.3 Infrastructure: Local vs. Third-Party Control

For a key recovery scheme to be of value to the encryption user, it must allow tight control over depositing, recovering, and maintaining keys, tied to the user's own practices and requirements. Generally, only a small number of individuals will need the ability to recover any individual key, often working in the same location and personally known to one another. When a key does need to be recovered, it will frequently be a local matter, similar to the replacement of a misplaced office key or restoring a computer file with a backup copy. The hours at which the key recovery might take place, the identification of the individuals authorized for a particular key, the policy for when keys should be recovered, and other basic operational procedures will vary widely from user to user, even within a single business. Particularly important is the control over when and how "recoverable" keys are destroyed when they are no longer needed, especially for keys associated with sensitive personal and business records.

Similarly, there is usually no business need for secrecy in the recovery of keys or for the ability to obtain keys without the initial cooperation of the user. When key recovery is used in a business environment, it would generally be one component of the overall data management policy of that business. Users would normally be trusted to participate in assuring recoverability of their own keys, assisted by local management practices and supervision. When a key must be recovered, it will usually be because the users themselves realize that they do not have a copy of the correct key or because the keyholder is no longer available. Even the frequently-cited hypothetical example of the disgruntled employee who refuses to decrypt important files is probably most reliably and economically dealt with through business data management practices (such as management supervision and backup of business-critical plaintext) that do not require any centralized, standard key recovery mechanism. Even in this (rather unusual) case, there is no need to hide from the user the fact that a key has been recovered.

The U.S. government, on the other hand, proposes key recovery schemes that by their nature do not allow local control. The government's requirement for the ability to covertly recover keys on short notice and without notice to the key owner must almost by definition be implemented by a third party whose procedures are entirely divorced from those of the users. Even when the government permits an organization to manage its own keys, the key recovery agent will have to be fairly centralized and remote from the actual users. This requirement eliminates the first line of defense against misuse of key recovery: the vigilance of the most concerned party – the key owner.

2.4 Infrastructure: Key Certification and Distribution vs. Key Recovery

As electronic commerce and encryption use becomes more widespread, some form of "Certification Authorities" (CAs) will be needed in some applications to help identify encryption users. A CA is a trusted party that vouches for the identity (or some other attribute) of an encryption user. It is widely believed that development and use of certification authorities will be essential for secure and trusted electronic exchanges – and, consequently, will become a prerequisite to participation in electronic commerce and online communications.[18]

[18]There is a great deal of debate about the appropriate role of government in regulating CAs. CAs may ultimately be large, centralized, or even government-certified entities, or smaller, locally-trusted entities. At this early stage in their deployment, no consensus has emerged on what government role is appropriate. For an excellent overview of the debate over CA regulation, see Michael Froomkin, "The Essential Role of Trusted Third-Parties in Electronic

Although superficially similar, in that they are both concerned with key management, the nature of key recovery is completely different from that of key certification. The most important function of a certification authority is to certify the public keys used in digital signatures; key recovery, on the other hand, is concerned with keys used for confidentiality. More importantly, the operation of a certification authority does not require handling sensitive user data; a CA generally handles only users' public keys and never learns the associated secret keys. If a CA's secret key is compromised or revealed, the only direct damage is that the certificates from it can be forged. On the other hand, if a key recovery agent's secrets are compromised, the damage can be far greater and more direct: every user of that recovery agent might have its own secrets compromised.

Certification can (and currently does) exist without any form of key recovery. Conversely, a key recovery infrastructure can exist completely independently of any key certification infrastructure.

Several recent government proposals have attempted to associate key recovery with key certification. This proposed linkage between CAs and key recovery makes no sense technically. On the contrary, such linkages have serious liabilities. It is not even clear whether such a system would work. To the extent it might require depositing keys used for signature and identification, such systems create additional security risks; there is no justification (even given government law enforcement requirements) for third-party access to signature keys that, if compromised, could be used to impersonate people, or to forge their digital signatures. In fact, attempts at achieving key recovery through a certification infrastructure would likely be ineffective at meeting the goals of law enforcement. Many (indeed, most) encryption keys are not certified directly, and therefore would be beyond the reach of a certification-based recovery system.

3 Risks and Costs of Key Recovery

Key recovery systems are inherently less secure, more costly, and more difficult to use than similar systems without a recovery feature. Key recovery degrades many of the protections available from encryption, such as absolute control by the user over the means to decrypt data. Furthermore, a global key recovery infrastructure can be expected to be extraordinarily complex and costly.

The impact of key recovery can be considered in at least three dimensions:

- Risk – The failure of key recovery mechanisms can jeopardize the proper operation, underlying confidentiality, and ultimate security of encryption systems; threats include improper disclosures of keys, theft of valuable key information, or failure to be able to meet law enforcement demands.

- Complexity – Although it may be possible to make key recovery reasonably transparent to the end users of encryption, a fully functional key recovery infrastructure is an extraordinarily complex system, with numerous new entities, keys, operational requirements, and interactions. In many cases, the key recovery aspects of a system are far more complex and difficult to implement than the basic encryption functions themselves.

- Economic Cost – No one has yet described, much less demonstrated, a viable economic model to account for the true costs of key recovery. However, it is still possible to make sound

Commerce," *75 Oregon L. Rev. 49* (1996).

qualitative judgments about the basic system elements, shared by all key recovery schemes, that will have the most dramatic impact on the cost of designing, implementing, deploying, and operating such systems.

3.1 New Vulnerabilities and Risks

Any key recovery infrastructure, by its very nature, introduces a new and vulnerable path to the unauthorized recovery of data where one did not otherwise exist. This introduces at least two harmful effects:

- It removes the inherent guarantees of security available through non-recoverable systems, which do not have an alternate path to sensitive plaintext that is beyond the users' control.
- It creates new concentrations of decryption information that are high-value targets for criminals or other attackers.

These risks arise with cryptography used in communication and storage, but perhaps even more intensely with cryptography used in authentication. (They are compounded even further if any keys are used for more than one of these purposes.)

3.1.1 New Paths to Plaintext

Regardless of the implementation, if key recovery systems must provide timely law enforcement access to a whole key or to plaintext, they present a new and fast path to the recovery of data that never existed before.

The key recovery access path is completely out of the control of the user. In fact, this path to exceptional access is specifically designed to be concealed from the encryption user, removing one of the fundamental safeguards against the mistaken or fraudulent release of keys.

In contrast, non-recoverable systems can usually be designed securely without any alternative paths. Alternative paths to access are neither required for ordinary operation nor desirable in many applications for many users.

3.1.2 Insider Abuse

Like any other security system with a human element, key recovery systems are particularly vulnerable to compromise by authorized individuals who abuse or misuse their positions. Users of a key recovery system must trust that the individuals designing, implementing, and running the key recovery operation are indeed trustworthy. An individual, or set of individuals, motivated by ideology, greed, or the threat of blackmail, may abuse the authority given to them. Abuse may compromise the secrets of individuals, particular corporations, or even of entire nations. There have been many examples in recent times of individuals in sensitive positions violating the trust placed in them. There is no reason to believe that key recovery systems can be managed with a higher degree of success.

The risk of "insider abuse" becomes even more evident when attempts are made to design key recovery schemes that are international in scope. Such abuse can even become institutionalized within a rogue company or government. National law-enforcement agencies, for example, might abuse their key recovery authority to the advantage of their own country's corporations.

3.1.3 New Targets for Attack

The nature of key recovery creates new high-value targets for attack of encryption systems. Key recovery agents will maintain databases that hold, in centralized collections, the keys to the information and communications their customers most value. In many key recovery systems, the theft of a single private key (or small set of keys) held by a recovery agent could unlock much or all of the data of a company or individual. Theft of a recovery agent's own private keys might provide access to an even broader array of communications, or might make it possible to easily spoof header information designed to ensure compliance with encryption export controls. The key recovery infrastructure will tend to create extremely valuable targets, more likely to be worth the cost and risk of attack.

The identity of these new rich targets will be highlighted by the key recovery systems themselves. Every encrypted communication or stored file will be required to include information about the location of its key retrieval information. This "pointer" is a road map showing law enforcement how to recover the plaintext, but it may also show unauthorized attackers where to focus their efforts. Moreover, even those systems (such as split key systems) that can decrease these risks, do so with a marked increase in cost. For example, splitting a key in half at least doubles the recovery agent costs.[19] Such systems require multiple agents, costly additional coordination mechanisms, and faster response times necessary to assemble split keys and still provide fast access to plaintext. Regardless of how many times a key is split, law enforcement's demand for timely access will still require the development of fast systems for the recovery of key parts. Both the systems for key part assembly, and the ultimate whole key assembled for law enforcement, will present new points of vulnerability.

3.1.4 Forward Secrecy

Key recovery is especially problematic in communications systems, such as encrypted cellular telephone calls, because it destroys the property of *forward secrecy.* A system with forward secrecy is one in which compromising the keys for decrypting one communication does not reduce the security of other communications. For example, in an encrypted telephone call, the keys for encrypting a call can be established as the call is set up. If these keys are destroyed when the call is over, the participants can be assured that no one can later decrypt that conversation—even if the keys to some subsequent conversation are compromised. The result is that once the call is over, the information required to decrypt it ceases to exist; not even the parties to the call store the keys. Typically, keys are created and destroyed on a per-call basis, or even many times per call. This makes it possible to limit the costs and risks of secure processing and storage to the period of the call itself.

Forward secrecy is desirable and important for two reasons. First, it simplifies the design and analysis of secure systems, making it much easier to ensure that a design or implementation is in fact secure. Secondly, and more importantly, forward secrecy greatly increases the security and decreases the cost of a system, since keys need to be maintained and protected only while communication is actually in progress.

[19] Storage of a smaller key part is not necessarily cheaper than storage of the whole key, and the preferred key-splitting methods generally produce key parts each of which is as large as the whole key.

Key recovery destroys the forward secrecy property, since the ability to recover traffic continues to exist long after the original communication has occurred. It requires that the relevant keys be stored instead of destroyed, so that later government requests for the plaintext can succeed. If the keys are stored, they can be compromised; if they are destroyed, the threat of compromise ceases at that moment.

3.2 New Complexities

Experience has shown that secure cryptographic systems are deceptively hard to design and build properly. The design and implementation of even the simplest encryption algorithms, protocols, and implementations is a complex and delicate process. Very small changes frequently introduce fatal security flaws. Non-key recovery systems have rather simple requirements and yet exploitable flaws are still often discovered in fielded systems.

Our experiences designing, analyzing and implementing encryption systems convince us that adding key recovery makes it much more difficult to assure that such systems work as intended. It is possible, even likely, that lurking in any key recovery system are one or more design, implementation, or operational weaknesses that allow recovery of data by unauthorized parties. The commercial and academic world simply does not have the tools to properly analyze or design the complex systems that arise from key recovery.

This is not an abstract concern. Most of the key recovery or key escrow proposals made to date, including those designed by the National Security Agency, have had weaknesses discovered after their initial implementation. For example, since the system's introduction in 1993, several failures have been discovered in the U.S. Escrowed Encryption Standard, the system on which the "Clipper Chip" is based. These problems are not a result of incompetence on the part of the system's designers. Indeed, the U.S. National Security Agency may be the most advanced cryptographic enterprise in the world, and it is entrusted with developing the cryptographic systems that safeguard the government's most important military and state secrets. The reason the Escrowed Encryption Standard had flaws is that good security is an extremely difficult technical problem to start with, and key recovery adds enormous complications with requirements unlike anything previously encountered.

3.2.1 Scale

Key recovery as envisioned by law enforcement will require the deployment of secure infrastructures involving thousands of companies, recovery agents, regulatory bodies, and law enforcement agencies worldwide interacting and cooperating on an unprecedented scale.

Once widely available, encryption will likely be used for the bulk of network communications and storage of sensitive files. By the year 2000 – still early in the adoption of information technologies – fielding the ubiquitous key recovery system envisioned by law enforcement could encompass:

- Thousands of products. There are over 800 encryption products worldwide today, and this number is likely to grow dramatically.
- Thousands of agents all over the world. Proposed systems contemplate many key recovery agents within this country alone; other countries will want agents located within their borders.

Large companies will want to serve as their own key recovery agents. Each of these agents will need to obtain U.S. certification and possibly certification by other countries as well.

- Tens of thousands of law enforcement agencies. There are over 17,000 local, state, and federal law enforcement agencies in the United States alone that might seek key information for authorized wiretaps or seized data.[20] National and local agencies around the world will also want access to keys.

- Millions of users. Several million Web users today use encrypted communications whenever their Web browser encounters a secure page (such as many of those used for credit card transactions). There will be an estimated 100 million Internet users by the year 2000, most of whom will be likely to regularly encrypt communications as part of the next version of the standard Internet protocols. Millions of other corporate and home computer users will also regularly encrypt stored information or intra-network communications.

- Tens of millions (or more) of public-private key pairs. Most users will have several public key pairs for various purposes. Some applications create key pairs "on-the-fly" every time they are used.

- Hundreds of billions of recoverable session keys. Every encrypted telephone call, every stored encrypted file, every e-mail message, and every secure web session will create a session key to be accessed. (Various key recovery scheme may avoid the need for the recovery center to process these session keys individually, but such "granularity shifts" introduce additional risk factors – see Section 3.4.1 below.)

Ultimately, these numbers will grow further as improved information age technologies push more people and more data online.

The overall infrastructure needed to deploy and manage this system will be vast. Government agencies will need to certify products. Other agencies, both within the U.S. and in other countries, will need to oversee the operation and security of the highly-sensitive recovery agents – as well as ensure that law enforcement agencies get the timely and confidential access they desire. Any breakdown in security among these complex interactions will result in compromised keys and a greater potential for abuse or incorrect disclosures.

There are reasons to believe secure key recovery systems are not readily scalable. Order-of-magnitude increases in the numbers of requesting law enforcement agencies, product developers, regulatory oversight agencies, and encryption end users all make the tasks of various actors in the key recovery system not only bigger, but much more complex. In addition, there are significant added transaction costs involved with coordination of international key recovery regimes involving many entities.

The fields of cryptography, operating systems, networking, and system administration have no substantive experience in deploying and operating secure systems of this scope and complexity. We simply do not know how to build a collective secure key-management infrastructure of this magnitude, let alone operate one, whether the key-recovery infrastructure is centralized or widely distributed.

[20] U.S. Department of Justice, Bureau of Justice Statistics, *Sourcebook of Criminal Justice Statistics 1995* (1996), p. 39.

3.2.2 Operational Complexity

The scale on which a government-access key recovery infrastructure must operate exacerbates many of the security problems with key recovery. The stated requirements of law enforcement demand the construction of highly complex key recovery systems. Demands on the speed and process for recovering keys will greatly increase the complexity of tasks facing those trusted with key recovery information. Demands for ubiquitous worldwide adoption of key recovery will greatly increase the complexity and number of entities involved. Each of these will in turn have a significant impact on both the security and cost of the key recovery system.

Consider the tasks that a typical key recovery center will perform to meet one law enforcement request for a session key for one communication or stored file:

- Reliably identify and authenticate requesting law enforcement agents (there are over 17,000 U.S. domestic law enforcement organizations).
- Reliably authenticate court order or other documentation.
- Reliably authenticate target user and data.
- Check authorized validity time period.
- Recover session key, plaintext data, or other decryption information.
- Put recovered data in required format.
- Securely transfer recovered data, but only to authorized parties.
- Reliably maintain an audit trail.

Each of these tasks must be performed securely in a very short period of time in order to meet government requirements. For example, the most recent U.S. Commerce Department regulations governing recovery agents require two hour turnaround of government requests, around the clock. The tasks must be performed by agents all over the world serving millions of clients and responding to requests from both those clients and numerous law enforcement agencies.

There are few, if any, secure systems that operate effectively and economically on such a scale and under such tightly-constrained conditions – even if these requirements are relaxed considerably (e.g., one day response time instead of two hours). The urgent rush imposed by very short retrieval times, and the complexity of the tasks involved, are an anathema to the careful scrutiny that should be included in such a system. If there is uncertainty at any step of the access process, there may be insufficient time to verify the authenticity or accuracy of a retrieval request.

It is inevitable that a global key recovery infrastructure will be more vulnerable to fraudulent key requests, will make mistakes in giving out the wrong key, and will otherwise compromise security from time to time. While proper staffing, technical controls, and sound design can mitigate these risks to some extent (and at considerable cost), the operational vulnerabilities associated with key recovery cannot be eliminated entirely.

3.2.3 Authorization for Key Recovery

One of the requirements for a key recovery operation is that it must authenticate the individual requesting an archived key. Doing so reliably is very difficult.

"Human" forms of identification – passports, birth certificates, and the like – are often easily counterfeited. Indeed, news reports describe "identity theft" as a serious and growing problem. Electronic identification must be cryptographic, in which case a key recovery system could be used to attack itself. That is, someone who steals – or recovers – a signature key for a law enforcement officer or a corporate officer could use this key to forge legitimate requests for many other keys. For that matter, if a sensitive confidentiality key were stolen or obtained from the repository, it might be possible to use it to eavesdrop on other key recovery conversations.

In contrast, a business's local, day-to-day key recovery process could rely on personal identification. A system administrator or supervisor would *know* who had rights to which keys. Even more questionable requests, such as those over the phone, could be handled appropriately; the supervisor could weigh such factors as the sensitivity of the information requested, the urgency of the request as known *a priori*, and even the use of informal authentication techniques, such as references to shared experiences. But none of these methods scale well to serve requests from outside the local environment, leaving them unsuitable for use by larger operations or when requests come from persons or organizations not personally known to the keyholders.

3.3 New Costs

Key recovery, especially on the scale required for government access, will be very expensive. New costs are introduced across a wide range of entities and throughout the lifetime of every system that uses recoverable keys.

The requirements set out by law enforcement impose new system costs for designing, deploying, and operating the ubiquitous key recovery system. These costs include:

- Operational costs for key recovery agents – the cost of maintaining and controlling sensitive, valuable key information securely over long periods of time; of responding to both law enforcement requests and legitimate commercial requests for data; and of communicating with users and vendors.
- Product design and engineering costs – new expenses entailed in the design of secure products that conform to the stringent key recovery requirements.
- Government oversight costs – substantial new budgetary requirements for government, law enforcement, or private certification bodies, to test and approve key recovery products, certify and audit approved recovery agents, and support law enforcement requests for and use of recovered key information.
- User costs – including both the expense of choosing, using, and managing key recovery systems and the losses from lessened security and mistaken or fraudulent disclosures of sensitive data.

3.3.1 Operational Costs

The most immediately evident problem with key recovery may be the expense of securely operating the infrastructure required to support it. In general, cryptography is an intrinsically inexpensive technology; there is little need for externally-operated "infrastructure" (outside of key certification in some applications) to establish communication or store data securely. Key recovery, on the other hand, requires a complex and poorly understood – and hence expensive and insecure – infrastructure.

The operational complexity described in the previous section introduces substantial ongoing costs at each key recovery center. These costs are likely to be very high, especially compared with the ordinary operational expenses that might be expected in commercial key recovery systems. Government key recovery requires, for example, intensive staffing (7x24 hours), highly trained and highly trusted personnel, and high-assurance hardware and software systems in order to meet the government's requirements in a secure manner. Theses costs are borne by all encryption applications, even those where key recovery is not beneficial to the user or even to law enforcement.

It remains unclear whether the high-risk, high-liability business of operating a key recovery center, with limited consumer demand to date, will even be economically viable.

3.3.2 Product Design Costs

Key recovery also increases the difficulty and expense of designing user-level encryption software and hardware. These costs vary depending on the particular application and the precise nature of the recovery system, but could be substantial in some cases. Integrating key recovery, especially in a secure manner, can also substantially delay the release of software. Given the highly competitive nature and short product life-cycles of today's hardware and software markets, such delays could discourage vendors from incorporating it at all, or worse, encourage sloppy, poorly-validated designs. Compatibility with older products presents special challenges and further increases these costs.

3.3.3 End-User Costs

Without government-driven key recovery, encryption systems can easily be fielded in a way that is largely transparent to their users. Highly secure communication and storage need require nothing further than the purchase of a reputable commercial product with strong encryption features tested in the marketplace. The use of that encryption need require nothing more than the setting of an option, the click of an icon, or the insertion of a hardware card. We are fully confident that, in an unregulated marketplace, many applications will ship with such high-quality user-transparent encryption built in. This is already happening at negligible cost to the user.

In contrast, the use of a secure key recovery system requires at least some additional user effort, diligence, or expense. In addition to the purchase of an encryption product, one or more key recovery agent(s) must be chosen. The user must enter into an important (although possibly implicit) contractual relationship with that agent, a relationship that will govern the potential disclosure of the most sensitive key information – now and for years to come. In many cases, there will need to be some communication of key information between user and the recovery agent. (Although some products will come with a built-in key, prudent users may want to change their keys on a regular basis. Also, software, especially mass-market "shrink-wrapped" software, cannot usually be economically distributed with unique keys installed in each individual copy).

The burdens on key recovery users continue long after data have been encrypted. Key recovery agents will maintain the ability to decrypt information for years. During that time, an agent might relax its security policies, go bankrupt, or even be bought out by a competitor – but will retain, and in fact must retain, the ability to decrypt. Diligent and concerned encryption users will need to be aware of the fate of their key recovery agents for years after their initial encryption use.

These burdens will apply to all users of encryption. Each use of encryption may entail the entry into a contractual relationship with a third-party key recovery agent. Under any rational business model, each such instance will entail some additional cost.

3.4 Tradeoffs

Some aspects of key recovery can be easily shifted along a spectrum from higher cost to higher risk. While it may be possible to field a particular key escrow system in a relatively secure way, this often results in tremendous costs to the user. While relatively simple and inexpensive key escrow systems exist, they often jeopardize security. For example, a poorly-run key recovery agent, employing less-skilled low-paid personnel, with a low level of physical security, and without liability insurance could be expected to be less expensive to operate than a well-run center.

Interestingly, security and cost can also be traded off with respect to the design itself. That is, the simplest designs, those that are easiest to understand and easiest to verify, also tend to require the most stringent assumptions about their environment and operation or have the worst failure characteristics. For example, imagine a design in which session keys are sent to the recovery center by encrypting them with the center's globally-known public key. Such a system might be relatively simple to design and implement, and one might even be able to prove that it is secure when operated correctly and under certain assumptions. However, this is among the worst possible designs from a robustness point of view: it has a single point of failure (the key of the recovery agent) with which all keys are encrypted. If this key is compromised (or a corrupt version distributed), all the recoverable keys in the system could be compromised. (We note that several commercial systems are based on almost exactly this design.)

3.4.1 Key Recovery Granularity and Scope

One of the most important factors influencing the cost and security of key recovery is the granularity and scope of the keys managed by the key recovery system. In particular, it is important to understand two issues:

- Granularity: the kinds of keys (user, device, session, etc.) that are recoverable.
- Scope: the consequences of compromising a recovery agent's key.

Granularity is important because it defines how narrowly-specified the data to be recovered from an agent can be and how often interactions (by the user and by law enforcement) with the recovery agent must take place. Various systems have been proposed in which the recovery agent produces "master" keys that can decrypt all traffic to or from individual users or hardware devices. In other systems, only the keys for particular sessions are recovered. Coarse granularity (e.g., the master key of the targeted user) allows only limited control over what can be recovered (e.g., all data from a particular individual) but requires few interactions between law enforcement and the recovery center. Finer granularity (e.g., individual session keys), on the other hand, allows greater control (e.g., the key for a particular file or session, or only sessions that occurred within a particular time frame), but requires more frequent interaction with the recovery center (and increased design complexity).

Also important is the scope of the recovery agent's own secret. Most key recovery systems require the user software or hardware to send keys to the recovery agent by encrypting them with the recovery agent's public key. If a recovery agent has only a single such key, that key becomes an extraordinarily valuable, global, single point of failure. Worse, because the recovery agent must use the secret component of this key in order to decrypt the keys sent to it (or at least any time a key is recovered), its exposure to compromise or misuse is also increased. To address this vulnerability, a recovery agent may have many such keys, perhaps one or more for each user. However, negotiating and distributing these keys to the users introduces still other complexities and vulnerabilities.

4 Conclusions

Key recovery systems are inherently less secure, more costly, and more difficult to use than similar systems without a recovery feature. The massive deployment of key-recovery-based infrastructures to meet law enforcement's specifications will require significant sacrifices in security and convenience and substantially increased costs to all users of encryption. Furthermore, building the secure infrastructure of the breathtaking scale and complexity that would be required for such a scheme is beyond the experience and current competency of the field, and may well introduce ultimately unacceptable risks and costs.

Attempts to force the widespread adoption of key recovery through export controls, import or domestic use regulations, or international standards should be considered in light of these factors. We urge public debate to carefully weigh the costs and benefits of government-access key recovery before these systems are deployed.

The Authors

Harold (Hal) Abelson is a Professor in the EECS department at MIT and a Fellow of the IEEE. He is co-author of the textbook *Structure and Interpretation of Computer Programs* and the 1995 winner of the IEEE Computer Society's Education Award. Abelson is currently on leave from MIT at Hewlett-Packard Corporation, where he serves as scientific advisor to HP's Internet Technology Group.

Ross Anderson teaches and directs research in computer security, cryptology and software engineering at Cambridge University in England. He is an expert on engineering secure systems, how they fail, and how they can be made more robust. He has done extensive work on commercial cryptographic systems, and recently discovered flaws in a British government key escrow protocol.

Steven M. Bellovin is a researcher on cryptography, networks and security at AT&T Laboratories. He is co-author of the book *Firewalls and Internet Security: Repelling the Wily Hacker.* In 1995 he was a co-recipient of the Usenix Lifetime Achievement Award for his part in creating Netnews. He is a member of the Internet Architecture Board.

Josh Benaloh is a Cryptographer at Microsoft Research and has been an active researcher in cryptography for over a decade with substantial contributions in the areas of secret-ballot elections and secret sharing methods and applications. Before joining Microsoft, he was a Postdoctoral Fellow at the University of Toronto and an Assistant Professor at Clarkson University.

Matt Blaze is a Principal Research Scientist at AT&T Laboratories in the area of computer security and cryptology. In 1994 he discovered several weaknesses in the U.S. government's "Clipper" key escrow system. His research areas include cryptology, trust management, and secure hardware. In 1996 he received the EFF's Pioneer Award for his contributions to computer and network security.

Whitfield Diffie is a Distinguished Engineer at Sun Microsystems specializing in security. In 1976 Diffie and Martin Hellman created public key cryptography, which solved the problem of sending coded information between individuals with no prior relationship and is the basis for widespread encryption in the digital information age.

John Gilmore is an entrepreneur and civil libertarian. He was an early employee of Sun Microsystems, and co-founded Cygnus Solutions, the Electronic Frontier Foundation, the Cypherpunks, and the Internet's "alt" newsgroups. He has twenty years of experience in the computer industry, including programming, hardware and software design, and management.

Peter G. Neumann is a Principal Scientist in the Computer Science Lab at SRI. He is Moderator of the Risks Forum (comp.risks), author of *Computer-Related Risks* (Addison-Wesley), and co-author of the National Research Council study report, *Cryptography's Role in Securing the Information Society* (National Academy Press). He is a Fellow of the AAAS, ACM and IEEE.

Ronald L. Rivest is the Webster Professor of Electrical Engineering and Computer Science in MIT's EECS Department. He is also an Associate Director of MIT's Laboratory for Computer Science. He is perhaps best known as a co-inventor of the RSA public-key cryptosystem and a founder of RSA Data Security, Inc.

Jeffrey I. Schiller is the Network Manager at MIT and has managed the MIT campus computer network since its inception in 1984. Schiller is the author of the Kerberos Authentication System, serves as the Internet Engineering Steering Group's Area Director for Security, and is responsible for overseeing security-related Working Groups of the Internet Engineering Task Force (IETF).

Bruce Schneier is president of Counterpane Systems, a Minneapolis-based consulting firm specializing in cryptography and computer security. He is the author of *Applied Cryptography* and inventor of the Blowfish encryption algorithm.