Cryptography Policy and the Information Economy

Matt Blaze
AT&T Labs – Research
600 Mountain Avenue
Murray Hill, NJ 07974
908-582-5524
mab@research.att.com

DRAFT – 27 December 1996 – DRAFT

1 Introduction

This paper is a high-level technical overview of the impact of cryptography on the computing and communications industries, with emphasis on the implications of the Administration's recent cryptography policy initiatives. It represents the best judgement of its author, and does not necessarily reflect the position of AT&T or any other organization.

We argue that unless a fundamental change is made in the direction of our cryptography policy, the United States' dominance in the emerging "information economy" will ultimately be placed in jeopardy. In particular, current policy fails to recognize several increasingly important realities of cryptographic technology:

- The wide availability of cryptography will soon be vitally important to the sustained growth of a wide range of American enterprises. This importance will arise not so much from the direct economic benefits of the sale of cryptography services and products by cryptography vendors, but rather from the enabling effect that cryptography will have on the advanced services that will form the basis for the "information economy." Cryptography will eventually be embedded in most, perhaps virtually all, advanced communications products and services; it will not be merely a "stand alone" or "add-on" feature as it is today.

- The ongoing discussion between industry and government in search of a "compromise" export key length that satisfies both parties is ill-considered from a technical point of view and offers little promise of resolving the debate. In particular, the government's offer to allow limited exportability of systems with 56-bit keys in exchange for industry support for "key recovery" is much less attractive than it might seem at first. Aside from the potential high cost of committing to a key recovery infrastructure, 56-bit keys are not strong enough for many applications today and will soon be recognized as not being strong enough for most commercially important applications in the very near future.

- The "key recovery" approach promoted by current policy is too expensive and too poorly understood to provide a viable basis for widely-deployed commercial cryptography. Cryptography, as used to protect communications traffic and stored data, is intrinsically rather inexpensive in terms of its direct cost and performance impact. "Key recovery" technology, on the other hand, is inherently expensive and entails potentially large risks. If key recovery becomes a required or standard component of future cryptographic systems, it is likely to greatly slow and otherwise impair the development of many future businesses that that will depend on the availability of inexpensive, widely-integrated cryptography.

2 Cryptography and the information economy

Cryptography is concerned with the use of mathematical functions, called "ciphers", that separate the security of a message's content from the security of the media over which it is transmitted. There are several different types of cipher functions. The most familiar are intended to make it difficult to understand the content of message without knowledge of the secret "key". A related type of cipher function can be used to ensure that information has not been altered. Still other functions can be used to establish the origin of digital information ("digital signatures"). Applications of cryptography include securing wired and wireless voice and data traffic against eavesdropping, protecting computer files from unauthorized access, and enabling secure electronic business transactions ("electronic commerce").

Cryptography is not widely used today, especially relative to its potential. This is true for several reasons. First, traditional communications media have historically offered a degree of intrinsic security. For example, it has always been relatively difficult and risky to illegally "tap" a conventional telephone line (the eavesdropper must locate the correct wire pair, arrange for a physical connection and somehow record and recover the traffic without detection). Similarly, data networks, to the extent that they existed at all, have until recently been closed, private systems, with messages routed primarily within the part of the network controlled by the end user's own equipment.

The technology and economics of modern communications and computing systems, on the other hand, strongly favors media that have little or no inherent security. For example, wireless telephones have great advantages in convenience and functionality compared with their familiar wired counterparts and are comprising an increasing proportion of the telephone network. This also makes eavesdropping much easier for curious neighbors, burglars identifying potential targets, and industrial spies seeking to misappropriate trade secrets. Similarly, decentralized computer networks such as the Internet have lower barriers to entry, are much less expensive, are more robust and can be used to accomplish a far greater variety of tasks than the proprietary networks of the past, but, again, at the expense of intrinsic security. The Internet makes it virtually impossible to restrict, or even predict, the path that a particular message will traverse, and there is no way to be certain where a message really originated or whether its content has been altered along the way. It is possible, even common, for electronic mail messages to route through the computers of competitors. There is every reason to believe that these trends will continue, and even accelerate, for the foreseeable future. The users of these networks, however, have learned to depend on the intrinsic security of the technology of the past.

At the same time, electronic communication is becoming increasingly critical to the smooth functioning of our society and our economy and even to protect the safety of human life. Communication networks and computer media are rapidly displacing the less efficient, traditional modes of interaction whose security properties are far better understood. As teleconferences replace face-to-face meetings, electronic mail and fax replace letters, electronic payment systems replace cash transactions, and on-line information services replace written reference materials, we enjoy undeniable gains in efficiency, but our assumptions about the reliability of even the most mundane transactions are often dangerously out-of-date. Cryptography is frequently the only viable approach to assuring security as these trends continue.

Fortunately, cryptography is usually a rather inexpensive technology to deploy. Although the cost of developing new cryptographic algorithms and engineering cryptography into specific applications can be significant, the marginal cost, in terms of direct expense and performance impact, of adding cryptographic security can be quite low, especially compared with other options. In the past, cryptography frequently required the use of special-purpose hardware to perform the cipher algorithms. With modern programmable personal computers and microprocessor-based consumer devices (such as cellular telephones), however, it is often possible to implement encryption functions entirely in software, sometimes with little or no performance impact. Similarly, the operational costs of using cryptography can be near zero. Modern key management techniques, such as the use of public-key "certificates," typically require only minimal infrastructural support from on-line secure key management centers. Some applications, including most secure file storage systems, require no trusted infrastructure at all and entail essentially no operational costs.

The availability of cryptography will soon be vitally important to the future of the telecommunications and computing industries, if not to the future of the American economy more generally. This importance will arise less from the direct sale of cryptography services and products than from the enabling effect the existence of cryptography will have on the services that will form the foundation of the future information economy. Cryptography will eventually be embedded in most, perhaps virtually all, communications products and services; it will not be a "stand alone" or "add on" feature as it is today. It is in the direct interest of any industry that hopes to benefit from electronic commerce to encourage wide availably of high-quality, inexpensive cryptographic products that enable secure communications and commerce.

3 Technical analysis of the administration's current policy

Although there are no restrictions on the sale or use of cryptography within the United States, strict regulations govern the export of products that include encryption features. In general, with narrow exceptions for products for use by US-owned and certain banking industry customers, export licenses are not granted for products that provide more than 40 "bits" of protection. Many domestic vendors supply only exportable-strength cryptography in their domestic products, to ensure interoperability and to avoid the need to support multiple product lines. The regulations that govern cryptography exports therefore constitute, in effect, *de facto* restrictions on the encryption available domestically.

The "strength" of an encryption system depends on a number of variables, including the mathematical properties of the underlying encryption function, the quality of the implementation, and the number of different "keys" from which the user is able to choose. It is very important that a cryptosystem and its implementation be of high quality, since an error or bug in either can expose the data it protects to unexpected vulnerabilities. Although the mathematics of cryptography is not completely understood and cipher design is an exceptionally difficult discipline (there is as yet no general "theory" for designing cipher functions), there are a number of common cipher systems that have been extensively studied and that are widely trusted as building blocks for secure systems. The implementation of practical systems out of these building blocks, too, is a subtle and difficult art, but commercial experience in this area is beginning to lead to good practices for adding high-quality encryption to software and hardware. Users and developers of secure systems can protect against weaknesses in these areas by choosing only cipher functions that have been carefully studied and by ensuring that their implementation follows good engineering practices.

The most easily quantified variable that contributes to the strength of an encryption system is the number of possible keys. Modern ciphers depend on the secrecy of the users' keys, and a secret-key system is considered well-designed only if the easiest "attack" involves trying every possible key, one after the other, until the correct one is found. An encryption system can be considered secure only if the number of keys is large enough to make such an effort infeasible. Keys are usually specified as a string of "bits"; adding one bit to the key length doubles the number of possible keys. An important question, then, is the minimum key length sufficient to resist a key search attack in practice. As technology advances and it becomes possible to try keys more quickly and economically, keys that might have once been considered sufficiently long become increasingly vulnerable.

It is almost universally recognized that 40-bit keys provide virtually no protection against such threats today, except against the most casual "attacker". Even 56-bit keys, which are used in the 20-year-old Data Encryption Standard, are too short to protect commercial information given a modestly well-funded attack model. Both the Business Software Alliance's "Minimum Key Length" panel[1] (in which the author participated) and the National Research Council Cryptography Policy[2] study group have noted the vulnerability of these short keys. The cryptography marketplace, to the extent it exists today, is also beginning to understand this, with many customers demanding cryptography that is at least strong enough to withstand a commercially-motivated adversary. The BSA panel's "Minimum Key Lengths" white paper (perhaps the most widely-circulated current reference to address the issue) concludes that secret keys today must be at least 75 bits to withstand a targeted commercial attack, and recommends that newly-deployed systems use secret keys at least 90 bits long to account for expected advances in computer technology that will make shorter keys more vulnerable. (The analysis of "public key" systems is more complex. Public keys generally must be several times longer than this to provide an comparable level of security.) These figures enjoy reasonably broad acceptance among those working in the field. It seems likely that 56-bit keys will soon be regarded by the marketplace as too weak for commercial applications, if they are not already.

There are virtually no technical, performance, or economic benefits to employing keys shorter than needed. The reasons vendors design new systems with short secret keys (40 or 56 bits) usually have to do with extra-technical pressures, such as inter-operability with old systems or, more commonly, to comply with U.S. export requirements.

In the most recent policy documents from the government, the administration proposes to increase the exportable key length from 40 to 56 bits for a two year "transition" period for those companies that can demonstrate that they have plans to incorporate "key recovery" in their products. As noted above, this offer represents only a marginal improvement in security, unlikely to satisfy the majority of applications in which cryptography is required.

It is not surprising that it has been difficult to find a "magic" key length that at once satisfies the security needs of industry and the wiretapping needs of government. Because of the nature of the problem, no such key length can really exist. The reason is that the threat models used by government and industry are not the same. Commercial data security analysis considers the possibility of an adversary who commits significant resources to a *targeted* attack against a particular key used to protect some data of great interest. On the other hand, government information gathering analysis, while perhaps supported by superior resources and technical expertise than the attacker in the commercial model, must consider the economics of recovering keys on a *routine* basis, with only a small percentage of total resources devoted to any particular key. Even if we allow for the possibility that the government's abilities are many times greater than currently estimated, it remains a truism that many commercial data security applications require cryptography stronger than what a government cryptanalysis effort could tolerate.

4 Implications of "key recovery"

The Administration's current policy initiatives encourage the provision of "key recovery" (sometimes called "key escrow") mechanisms in new applications of cryptography. In such systems, copies of keys are automatically deposited, in advance, with some party who can later use them to arrange to provide a backup copy of a key if it should be required at some point in the future. The Administration's original key escrow proposals required that government agencies hold the key copies. More recent proposals suggest that keys could be held by private industry. The latest proposal leaves identity of the key holder unspecified.

A properly designed cryptosystem makes it essentially impossible to recover encrypted data without knowledge of the correct key. Sometimes this can be a double-edged sword; the cost of keeping unauthorized parties out is that if keys are lost or unavailable at the time they are needed, the owner of the encrypted data will be unable to make use of his or her own information. This problem, of balancing secrecy with assurances of continued availability, is a difficult one, and remains an area of active research for which some solutions in a few application areas are starting to emerge.

It is important to distinguish the interests of the government from the needs of the private sector here. While key "recoverability" is a potentially important added-value feature in certain stored data systems, in other applications of cryptography there is no user demand for this feature. The key escrow systems proposed by the administration, on the other hand, are concerned chiefly with recovering encrypted communications traffic. The nature of the key recovery problems of government and industry are, at the outset, quite different.

Key recovery is a difficult, and likely very expensive to solve, technical problem, for reasons that can be observed at several different levels of abstraction. It introduces a fundamental tradeoff, one that is not otherwise present in secure systems. As noted above, in general, cryptography is an intrinsically inexpensive technology; there is little need for "infrastructure" (outside of key certification in some applications) to establish secure communication between two parties. Key recovery, on the other hand, entails by its very nature a complex, expensive, and rather poorly-understood infrastructure. (Ironically, some cryptography vendors could actually reap modest medium-term benefits from rules that mandate key recovery by developing and selling the many new products and services that would be required to support such a policy.)

The most deeply-rooted problem with key recovery is simply that we do not fully understand how to do it. Any key recovery system, by its very nature, reduces the security of a system by increasing its number of points of failure. Unfortunately, we do not understand key recovery well enough to even quantify this reduction in security. The design and implementation of even the simplest encryption systems is an extraordinarily difficult and delicate process. Very small changes frequently introduce fatal security flaws. Ordinary (non-escrowed) encryption systems have conceptually rather simple requirements (for example, the secure transmission of data between two parties) and yet, because there is no general theory for designing them, we still often discover exploitable flaws in fielded systems. Adding key recovery renders even the specification of the problem itself far more complex, making it virtually impossible to assure that such systems work as they are intended to. It is possible, even likely, that lurking in any key recovery system are one or more design weaknesses that allow recovery of data by unauthorized parties. The commercial and academic world simply does not have the tools to analyze or design the complex systems that arise from key recovery. This is not simply an abstract concern. Most of the key recovery or key escrow proposals made to date, including those designed by the National Security Agency (the "Clipper chip") have had weakness discovered after their initial implementation. The NRC report[2] (which the administration incorrectly claims is "broadly consistent" with its recent key escrow proposals) addresses key recovery only to point out that it is an "unproven" concept.

But the most serious problem with "key recovery," beyond the loss of security inherent in their design, may be the enormous expense associated with properly operating the infrastructure required to support it.

According to the Administration (for example, see the December 9, 1996 Commerce Department draft regulations on cryptography export), key recovery centers must be prepared to respond to law enforcement requests for key data 24 hours a day, completing transactions within two hours of receiving each request and in complete secrecy from the target of the investigation. There are thousands of law enforcement agencies in the United States authorized to perform electronic surveillance; the escrow centers must be prepared to identify and respond to any of them within this time frame. If a center also provides commercial data recovery services, it will also have additional private sector customers that it must be prepared to serve and respond to. Even if we imagine relaxing these requirements considerably (e.g., one day response time instead of two hours), there are few, if any, secure systems that operate effectively and economically on such a scale and under such tightly-constrained conditions. It is simply inevitable that key recovery systems that meet the government's requirements will make mistakes in giving out the wrong keys from time to time or will be vulnerable to fraudulent key requests. Key recovery, by its nature, makes encrypted data less secure because the recovery center itself introduces a new target for attack. Perhaps more importantly from a business perspective, a key recovery infrastructure of this kind would be extraordinarily expensive to operate.

The difficulty of designing and operating a reliable, economically scalable key recovery infrastructure arises from the basic requirements of such a system, not simply from a few unresolved technical details. Whether the recovery system uses private key cryptography (with a large database of user keys) or public key cryptography (with a single escrow key for many users), whether the database is split with secret sharing techniques or maintained in a single hardened secure facility, and whether the recovery service provides users' master keys or merely decrypts specific messages as needed, all key recovery systems entail 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. It is this basic requirement that makes the problem of key recovery difficult and expensive, beyond the specific details of any particular implementation.

Finally, there is the problem of user demand and acceptance. Most applications of cryptography (including virtually all that are directly important to to the communications industry or for the development of electronic commerce) do not carry with them any customer-centered demand for key recovery. That is because key recovery is useful to the end-user only for encrypted stored data. Outside of a few specialized, highly regulated environments (e.g., securities trading), there is little reason for a user of cryptography to ever want the ability to recover the "session" keys used to protect past communication traffic. In other words, the substantial cost and risk associated with maintaining a key recovery infrastructure is usually not mitigated by added value to the user; the government is the only beneficiary. The customer derives no benefit from, but pays a significant price for, the key recovery "feature" for communication traffic.

5 Summary

Current U.S. cryptography policy threatens to stifle several of the most promising opportunities for growth in the computing and communications industries. The emerging "information economy" will demand, first and foremost, an economical, reliable and trustworthy foundation upon which to build the transactions that underly trade in information. The very technologies that make such a foundation possible are exactly those that current government policy so effectively restricts.

As tempting as it may be to hope that industry can obtain even the most limited, short-term relief by negotiating a compromise, the administration's recent policy initiatives offer little to encourage us. In particular, the government now proposes to allow limited exportability of systems with 56-bit keys in exchange for an industry commitment for to build key recovery into future systems over the next two years. As discussed above, neither 56-bit keys nor a key recovery infrastructure are viable solutions from industry's perspective. 56-bit keys are too short to be considered trustworthy; a general key recovery infrastructure is simply too expensive to consider as an option, especially on the scale demanded by even the most conservative expectations of the not-so-distant future.

Unfortunately, it seems unlikely, under the current terms under which the cryptography debate is cast, that we will ever find a "compromise" that satisfies both the government and the needs of industry. There is precious little room for either side to maneuver. A resolution to the issue will require either that industry (and the public at large) abandon its vision of a future information economy in which the trade in information becomes a dominant medium of exchange, unencumbered by physical boundaries, or that the government re-focus its law-enforcement and national security agenda to keep up with, rather than hold back, the natural progress of commercial technology.

References

[1] M. Blaze, W. Diffie, R. Rivest, B. Schneier, T. Shimomura, E. Thompson, and M. Wiener. "Minimum Key Lengths for Symmetric Ciphers to Provide Adequate Commercial Security." January 1996. Available at `ftp://research.att.com/dist/mab/keylength.txt`

[2] National Academy of Sciences, National Research Council. *Cryptography's Role in Securing the Information Society.* National Academy Press. 1996.