Introduction

Following the lead of the State of Utah, numerous states and several foreign countries have enacted "digital signature" legislation aimed at promoting the development of a public key infrastructure (PKI). While PKI legislation has acquired significant momentum, it is not clear that lawmakers have carefully considered the public policy implications and long-term consequences of these laws. This article surveys ten public policy issues implicated by digital signature legislation.

1. Is Legislation Necessary at All?

Proponents of digital signature legislation start with the premise that the need for a PKI is clear: public key cryptography and verifiable certificates offer the best hope for sending secure, authentic electronic messages over open networks, thereby facilitating electronic commerce. They argue that the reason that the commercial marketplace has not produced a viable certification authority (CA) industry is because of legal uncertainty (CAs are unable to determine their potential liability exposure because of a confusing array of applicable background law) or because existing law imposes too much liability on CAs. Thus, proponents argue, legislation is necessary in order to provide certainty in the marketplace and allow a much-needed industry to emerge, as well as to address other issues such as the legal status of digitally signed documents.

Opponents of this view assert that it is far too soon to conclude that the market will not produce commercial CAs, and point to the increasing numbers of commercial CAs emerging even in the absence of legislation. Time is solving the "uncertainty" problem, opponents argue, and the "too much liability" problem is the product of flawed business models, not a flawed legal system. Opponents of legislation argue that the real danger is that a group of lawyers will impose a set of inappropriate rules that will fundamentally skew a dynamic infant marketplace and "lock in" a set of business models that the market would otherwise reject. The time for legislation and regulation is after identifiable problems exist in a mature industry, opponents say, not before an industry even exists. Opponents of legislation further argue that existing legal mechanisms can address the issue of the legal status of digitally signed documents.

2. Where Should PKI Legislation Occur?

Debate also exists over the appropriate jurisdictional level for digital signature legislation. Some observers cringe at the thought of fifty inconsistent state digital signature laws; others believe that CAs and consumers will opt-in to the most sensible legislative scheme, and thus believe that competition between the states is helpful. Proponents of uniformity and consistency argue for PKI legislation at the federal or international level, while opponents of this view point out that general commercial law has long been the province of state legislatures.

3. Is Licensing of Certification Authorities the Right Approach?

Under the Utah Digital Signature Act ("Utah Act")[1] and much of the subsequent PKI-related legislation, CAs are licensed by the state. The Utah Act makes licensing optional: CAs that obtain licenses are treated with favorable liability rules, but non-licensed CAs may exist in Utah.

Licensing is a highly intrusive form of government regulation (other, less intrusive methods of regulation include mandatory disclosure requirements, altering liability rules to avoid externalized costs, bonding or insurance requirements, etc.). Typically, licensing as a form of regulation is reserved for circumstances where a market flaw cannot be addressed by other, less intrusive means. Does this sort of dynamic exist with CAs? Would consumers be able to make informed, rational choices between CAs? Could an incompetent CA cause irreparable harm? Could other types of regulation address any relevant market flaws? If unlicensed practitioners are allowed to exist, subject to different liability rules, how will this affect the CA market?

4. Should Legislation Endorse Public Key Cryptography or Be "Technology Neutral"?

Most of the digital signature legislation to date has focused specifically on digital signatures created using public key cryptography. Some legislation has also addressed the issue of "electronic signatures" and other, non-public key methods of authenticating digital transmissions.

Proponents of biometric authentication methods argue that it is foolish to legislatively enshrine public key cryptography as the only technology capable of authenticating an electronic document. They argue that biometric methods can currently accomplish many of the same goals as digital signatures; they further argue that by precluding other technologies future innovations will be discouraged. They also note that public key cryptography can only be implemented using patents owned by a limited number of commercial entities, and question whether it is wise public policy to legislatively tie electronic commerce so closely to the interests of a few private sector actors.

5. Should Legislation Endorse the X.509 Paradigm?

When the Utah Act was enacted, it explicitly endorsed the X.509 infrastructure model.[2] Subsequent laws have dropped the explicit endorsement of X.509, but nonetheless remain true to the X.509 paradigm. Under most digital signature legislation, certificates serve to bind an individual's identity to a particular public key. This binding is accomplished in the context of a rigid, hierarchical CA infrastructure.

This model has been criticized for two main reasons: global CA hierarchies are almost certainly unworkable, and identity certificates often provide too much information when frequently an "attribute" or "authority" certificate will do. Alternative certificate formats, such as SDSI and SPKI, have emerged in response to these and other perceived flaws with the X.509 model. However, it is not clear that these alternative certificate formats can be accommodated under current digital signature legislation.

6. How Should Liability and Risk Be Allocated in a PKI?

Liability allocation promises to be a vexing problem in a PKI. The liability issue is most dramatic in the context of fraud. An impostor can obtain the private encryption key associated with a particular party and create electronic documents purporting to be from that party. A second party may enter into an electronic contract relying on these ostensibly valid documents, and a loss may occur. Who should bear this loss?

In the paper world, generally one cannot be bound by a fraudulent signature. This principle may not be entirely appropriate in an electronic context, however. In a PKI, the integrity of the infrastructure depends upon the security of private encryption keys. If a key holder bears no liability for fraudulent use of that private key, perhaps he or she may not have adequate incentive to keep the private key secure.

How much liability should the private key holder bear? Under the Utah Act and its progeny, an individual who negligently loses control of a private key will bear unlimited liability. This risk allocation scheme raises the specter of consumers facing immense losses--as one commentator puts it: "Grandma chooses a poor password and loses her house." In contrast, consumer liability for negligent disclosure of a credit card number is generally limited to $50.

If consumer liability were similarly limited in a PKI, where would the risk of loss fall? If CAs had to act as an insurer in all transactions, the price of certificates would likely be extraordinarily high. If relying third parties faced the risk that ostensibly valid documents may in fact be forgeries and bear any resulting loss, then some benefits of a PKI are lost.

7. What Mechanisms Should Be Used to Allocate Risk?

Currently at least one commercial certification authority, VeriSign, is attempting to allocate risk to both certificate subjects and relying third parties by contract. VeriSign includes significant warranty disclaimers, liability limitations, and indemnification provisions in its Certification Practices Statement (CPS). Certificate applicants agree to be bound by the CPS when obtaining a certificate.

VeriSign's web page informs relying third parties that the act of verifying a certificate or checking a certificate revocation list indicates agreement to the terms of the CPS. However, it is not clear that a binding contract can be formed with relying third parties in this fashion. Thus the relationship between VeriSign and relying parties may not be governed by the CPS at all, but instead be subject to default contract and tort rules (which would be less favorable to VeriSign).

As a policy matter, should CAs be able to form contracts with relying third parties, despite their rather attenuated connection? If relying parties will be bound by unilateral contracts imposed by CAs, they face significant transaction costs involved with determining the contract terms offered by potentially numerous CAs. If CAs cannot scale their potential liability exposure to third parties by contract, however, it may be impossible for CAs to compete on warranty terms--and presumably such terms would otherwise be the subject of significant competition.

8. Should Digitally-signed Documents Be Considered "Writings" for all Legal Purposes?

The Utah Act and most other digital signature laws provide that digitally signed documents have the same legal effect as writings. Critics have noted that while most of the functions or goals of writing requirements may be served by electronic documents, this may not be true in all instances. For example, the law often requires a written instrument to effect notice, i.e., to alert an individual that a lien has been filed on their property. It is not clear that a digitally signed electronic message would achieve the same effect. Additionally, there are other contexts--such as wills or adoption papers--where paper documents may prove more effective than electronic documents. Moreover some paper documents (such as bank drafts or warehouse receipts) are negotiable instruments, and this negotiable character depends upon the existence of a single, irreproducible copy of the document. Thus, critics say, digital signature legislation should not override all writing requirements without separately considering the extent to which sound policy might require retention in specific circumstances.

9. How Much Evidentiary Weight Should a Digitally-signed Document Carry?

Evidentiary issues, though seemingly arcane and procedural, can raise important public policy concerns. For example, the Utah Act creates a presumption that the person who owns a particular key pair used to sign a document in fact did sign the document. Holding an individual presumptively bound by obligations entered into under their digital signature could be inequitable if the individual is the victim of the fraudulent use of such a signature. This potential problem can be compounded by the evidentiary weight assigned to digitally-signed documents.

Under the Utah Act digitally-signed documents are accorded the same evidentiary weight as notarized documents, and someone challenging the authenticity of such a document can overcome the presumption of authenticity only with "clear and convincing evidence." In contrast, one can overcome the presumption of validity of a paper signature simply by denying that it is one's signature. Critics of the Utah Act's approach argue that providing digitally-signed documents with this status creates unreasonable evidentiary burdens for victims of fraud challenging the validity of electronic documents signed with the victim's private key.

10. Should Governments Act as CAs?

Much of the currently enacted digital signature legislation envisions state government agencies acting as "top level" certification authorities, who in turn certify a second tier of private sector CAs. At the federal level, the U.S. Postal Service has declared its intention to act as a CA on a nationwide basis. Should governments be acting in this sort of role?

Critics say no, arguing that government involvement will skew an emerging private sector CA marketplace. Government actors may face very different liability rules than private sector market participants--governments can choose to scale their potential liability exposure through the doctrine of sovereign immunity. Thus, critics argue, government CAs may "win" in the marketplace not because they are more efficient or provide better service, but rather because they can stack the rules in their favor.

Proponents of government involvement argue that governments can play an important role precisely because they can create sensible ground rules for all PKI participants. Additionally, they note that governments have existing relationships with all of their citizens, making the process of identification and public key binding that much easier.

Ten Public Policy Questions for Digital Signature Legislation

1. Is legislation necessary at all?
2. Where should PKI legislation occur?

3. Is licensing of Certification Authorities the right approach?

4. Should legislation endorse public key cryptography, or be "technology neutral"?

5. Should legislation endorse the X.509 paradigm?

6. How should liability and risk be allocated in a PKI?

7. What mechanisms should be used to allocate risk?

8. Should digitally-signed documents be considered "writings" for all legal purposes?

9. How much evidentiary weight should a digitally-signed document carry?

10. Should governments act as CAs?

Footnotes

[a] *The author is also author of C. Bradford Biddle, "Misplaced Priorities: The Utah Digital Signature Act and Liability Allocation in a Public Key Infrastructure," 33 San Diego L. Rev. 1143 (1996), and serves as Vice Chair of the Electronic Commerce Subcommit tee of the American Bar Association's Committee on the Law of Commerce in Cyberspace. He is a third-year law student at the University of San Diego and is a law clerk in Cooley Godward LLP's San Diego office, where he served on the legal team advising the Internet Law and Policy Forum's Working Group on Certification Authority Practices. He can be contacted by e-mail at biddlecb@cooley.com.*

[1] Utah Code Ann. §§ 46-3-101 through 46-3-504.

[2] X.509 is a standard certificate format developed by the Geneva-based International Telecommunications Union [in connection with its X.500 directory database standard].