# Global Information Assurance Certification Paper

Bill Pankey
GSEC Practical Version 1.4b ~ Option 1
February 3, 2003

# Solving HealthCare's eMail Security Problem

**Abstract**
While healthcare organizations have come to depend heavily on electronic mail, they do so without a significant email security infrastructure. New Federal law and regulation place new obligations on the organizations to either secure their email systems or drastically restrict their use. This paper discusses email security in a healthcare context. The paper considers and recommends solutions to the healthcare organization's problem in securing its mail. Because email encryption will soon be a categorical requirement for healthcare organizations, email encryption is discussed in some detail. The paper describes details and benefits of domain level encryption model and considers how PKI is best deployed to support secure electronic mail.

**Motivation**
It is a simple fact that the US healthcare industry has come to depend heavily on electronic mail to support treatment, payment and general healthcare operations. Such use, though, is something of a badly kept secret as most healthcare organizations have explicit policy which either prohibits or seriously restricts the use of electronic mail for the transmission of any 'patient identifiable' health information. Historically, the industry has deemed patient identifiable health information as deserving of special protection, since, by its very nature, such information is highly confidential. Accepting the 'inherent insecurity' of electronic mail, healthcare organizations have done little to develop security infrastructure supporting use of electronic mail for confidential communication and instead adopted policies forbidding such use. It speaks to the utility of electronic mail, that even in spite of such policy, as much as 40% of all electronic mail emanating from healthcare organizations contains health information. A very small percentage of this email is encrypted or otherwise protected to ensure its confidentiality and authenticity.

Federal law will prohibit future 'unsecured' use of electronic mail for transmission of health information. The Health Insurance Portability and Accountability Act of 1996 (a.k.a. Public Law 104-191; a.k.a. HIPAA) obligates healthcare organizations to implement 'reasonable and appropriate' technical safeguards to ensure that the confidentiality and integrity of health information is preserved. While 'reasonable and appropriate' is a legal standard, the HIPAA law also mandates conformity to a set of security standards promulgated by the Secretary of Health and Human Services. Although these security standards have not yet been finalized, in August of 1998, HHS did publish in 45 CFR Part 142 a proposal for that Security Standard. That Notice of Proposed Rule Making did include a number of specific security implementation features. Particularly relevant to email use is a specification for encryption of health information communicated over any network for which the transmitter cannot control

access (45 CFR Part 142.308[d][1][ii]). This restriction clearly is intended to apply to the healthcare organization's Internet bound electronic mail.

This paper broadly outlines steps that healthcare organizations can take to ensure the security of their electronic mail use. A substantial portion of this activity has a "Security 101' aspect to it. Healthcare organizations are generally exposed to the same Internet borne threats as any other type organization. As a result, healthcare organizations do well to follow the general recommendations for email security provided in documents such as NIST's "Guidelines for Electronic Mail Security". Healthcare organizations do have business imperatives and legal obligations, however, that may encumber routine application of email security best practice. Therefore, this paper will provide a healthcare industry context to its discussion of electronic mail security.

**Risks Associated with Electronic Mail Use**

Generally speaking there are three classes of email related risk that the healthcare organization seeks to mitigate with technical security controls: 1) risks associated with exposing enterprise resources to a vulnerable SMTP implementation; 2) risk associated with potentially hostile or malicious content in email messages; 3) risk associated with the potential interception, modification or redirection of email during transmission.

*Server Risk*. Organizations develop their email systems to support business communication. That communication, more likely than not, needs to be bilateral, therefore, enterprise staff receive business related information as well as send it. Generally, this means that the enterprise allows messages from the Internet through its firewalls to reach port 25 on an internal server(s). This SMTP traffic can be a vector of attack against not just the organization's SMTP capability but its network infrastructure. SMTP supports rudimentary recognizance of network by with the routine display of service banners containing mail server - operating system type and version and recognizance of its users through SMTP commands VRFY and EXPN. Vulnerable configuration of SMTP servers may lead to "open relays" where enterprise resources can become a mechanism for the anonymous broadcasting of "spam". Various SMTP implementations have been shown to be vulnerable to buffer overflows in various SMTP commands (eg HELO, FROM, RCPT TO) leading to unauthorized use of the service, denial of service or execution of arbitrary code on the underlying host. Mitre Corporation's Common Vulnerability and Exposures database includes more than 40 such vulnerabilities in SMTP implementations.

*Content Risk.* Internet email is recognized as a principal vector for the transmission of virus, worms, trojan horses and other malicious code. Organizations in the healthcare sector, like elsewhere, have directly experienced the costs due virus and other malicious code. Companies surveyed by the ICSA Labs 2001 Virus Prevalence Survey typically reported encountering costs greater than $100,000 from virus attacks. As early as 1999, more than half of the viruses were spread by electronic mail.

*Transmission Risk*. Once an email message passes from the organization's network to an external SMTP relay, the email is subject to interception, modification, or redirection.

The basic operation of Internet's system of SMTP relays includes few if any confidentiality controls that can be established and maintained by message senders. Electronic mail, using SMTP, relies upon a system of document transfers to a 'mailbox' maintained for the recipient of the mail. Mail Transfer Agents (MTA), who typically have no business relationship to either sender or receiver, facilitate this transfer. While dependence upon such 'anonymous' MTA can sometimes be avoided thru the use of 'direct' SMTP, senders still must rely upon the recipient's ISP, which generally has no business relationship with and therefore duty to the sender. Senders generally rely upon the 'good conduct' of all such third parties in operating their services so as to prevent persons other than the intended recipient from reading, retaining, redirecting, or modifying the sender's electronic mail.

This sort of risk is of a different character than the two preceding risks. With transmission risk, the major concern is with information confidentiality and integrity subsequent to leaving the enterprise boundary. The other risks affect information and resources within the enterprise boundary. Whereas enterprise can to a large degree unilaterally mitigate the risks to internal resources, they can mitigate transmission risk only with the cooperation of others.

There is some basis for confidence that persons with malevolent intent will not intercept or tamper with the sender 's mail. In particular, given the enormous volume of electronic mail, there is a remarkable lack of reported interception of email. The reason for this, in part, is that interception of electronic mail is a Federal crime. Title 18, Part I, Chapter 119, Section 2511 of the United States Code prohibits, with few exceptions, the interception, attempted interception, or disclosure of any electronic communication by persons not party to the communication. This is the same law that protects the confidentiality of ordinary voice telephone conversations. For much of the same reasons that persons use the public switched telephone network for confidential conversation, persons are reassured about the confidentiality of their email. Even if the interception of email were easy to accomplish, such interception carries truly significant legal risk for the interceptor. The US government has prosecuted at least one case of email interception under this law. In 1999, Alibris / Interloc which operated both an ISP and an online book selling business, was fined $250,000 for intercepting and copying email sent from Amazon.com to Alibris / Interloc's ISP customers.

Under HIPAA, healthcare organizations have an affirmative obligation to mitigate these risks, specifically as they relate to the confidentiality and integrity of health information. Further HIPAA impacts the risk analysis that healthcare organization might conduct to guide its security investments and planning. HIPAA requires protection for the health information of *patients,* not the protection of the healthcare organization's business assets. Patient information is generality not proprietary to the healthcare organization. While the organization has interest in the integrity of this information, it has traditionally had little financial interest in the protecting the confidentiality of patient information. This fact was recognized by the National Research Council in its "For the Record ~ Protecting Electronic Patient Information" report which was influential in the creation of the security provisions of the HIPAA law. That report concluded that, *absent regulation*,

the typical healthcare organization could not create a business case for significant expenditure on security controls to protect the confidentiality of patient information. The organization that implements better confidentiality protection does not achieve a competitive advantage with healthcare consumers; protecting this confidentiality does not otherwise return economic value to the healthcare organization. HIPAA then skews the healthcare organization's security planning to favor confidentiality controls.

**Adequate Safeguards**

HIPAA requires that healthcare organizations take action to *ensure* the integrity and confidentiality of health information and the resources that process such information from any *reasonably anticipated threat*. Legally and practically, this is seen as a very high standard when contrasted when weaker language such as 'reasonable under the totality of circumstances' which Congress choose not to use.

*Server Risk*: The risks to servers described above are primarily due to active attacks against SMTP (and POP/IMAP) and the host OS. Here risk mitigation is initially a matter of server and host configuration, i.e. hardening the OS and mail server applications. For these purposes, excellent guidance is provided by the previously mentioned NIST "Guidelines ion Electronic Mail Security". As is typical of server / OS hardening, the guide emphasizes: removal or disabling of unneeded services; removal of unneeded application or sample code and vendor documentation; installation of relevant vendor vulnerability patches; execution of publicly available hardening or 'lock down' scripts. Appropriate management of access control for the mail server applications and OS is crucial. Of particular interest is the controlling of access to: application software and configuration files; password files and cryptographic information; mail log files; OS system software and configuration files.

*Content Risk:* Protection against hostile or malicious code comes primarily from a regimen of virus screening followed by 'cleansing', quarantining and / or destruction. The immediacy of this sort of threat has compelled most, if not all, healthcare organizations to adopt such regimen. Of particular concern, is maintaining the currency of the virus signature files in the organization's anti viral software.

*Transmission Risk*. Traditionally, encryption and digital signatures are taken as the principal means to mitigate interception or other transmission risk. Message contents are mangled thru the application of a symmetric key encryption algorithm such as 3DES (triple des) or CAST and then transmitted using ordinary electronic mail. Once encrypted, interception of the electronic mail does not result in a confidentiality breech because, without the appropriate decryption key, the interceptor will not be able to recover plaintext message contents. Message integrity is assured with a digital signature, which encrypts a secure hash of email contents. Encryption, then, translates the problem of protection against interception risk into a problem of key management. The confidentiality and integrity of the email will depend upon the secure delivery of appropriate decryption keys to message recipients. The challenge is to accomplish this delivery in a cost-effective manner.

Typically, mail senders will apply one of two general approaches to the key management problem. Both of these approaches utilize asymmetric public key cryptography:

1) In the first of these approaches, the symmetric key used to encrypt the message (a.k.a. the message key) is itself encrypted using an asymmetric encryption algorithm, such as RSA, and the public key of the email recipient. The encrypted message key and ciphertext of the original email contents is then sent to message recipients using ordinary plaintext electronic mail. One of two standard methods are generally available for the formatting the resulting hybrid email: pgp/MIME (or simply PGP which itself stands for "Pretty Good Privacy"); or s/MIME (a.k.a. secure MIME). Both of these methods rely upon the availability of the recipient's public key in some form of a public key 'certificate'. The certificate is an electronic document that binds, thru a digital signature, a public key to information about the possessor of its related private key. Certificates used in PGP accomplish this binding using the digital signature of the private key holder; s/MIME certificates involve the digital signature of a third party known as a 'certificate authority'. As they contain no secret content, these certificates are easily distributed. They are typically found by querying an LDAP directory, but they may be distributed through other means including electronic mail.

2) The second approach to communication of the message key utilizes a secure session protocol, typically SSL, to transfer email contents rather than SMTP. SSL solves the key management problem as part of its session establishment protocol or 'handshake'. In this protocol, the message recipient generates the message key, encrypts that with the public key of the message sender (in this case an HTTP server). This approach is referred to as 'secure web mail'. The original email message is not sent via SMTP, but redirected to an 'secure' repository where it awaits 'pickup' from the intended recipient via an HTTPs session. The security of this method depends upon adequate authentication of the intended recipient and protection of the repository storing messages awaiting pickup.

**Healthcare Context**

When constructing an email security solution, it is important to preserve the business value of electronic mail as a communication tool. Further, the security solution must not otherwise interfere with the fulfillment of enterprise obligations; a security solution that allows email use that is contrary to enterprise purpose has marginal benefit. The security solution, especially when dealing with transmission risk, must be sensitive to the needs and capabilities of recipients, since electronic mail security, in part,requires the cooperation of recipients. If the solution is cumbersome to recipients, then their participation will likely be reduced, diminishing the business value of electronic mail to the enterprise. The healthcare industry context provides constraints that further specify and qualify the appropriate security solution. These constraints are particularly relevant to encryption solutions.

HIPAA, through its "privacy regulation", specifies the conditions under which healthcare organizations can use or disclose confidential health information. With the notable exception of patients and for specific legally mandated purposes, healthcare

organizations may generally only disclose confidential health information to other healthcare organizations or to their own 'business associates'. In the HIPAA privacy rules, the term 'business associate' has a precise operational meaning, so at any given time a healthcare organization knows exactly who is and who is not a business associate. Since electronic mail provides a convenient mechanism for the disclosure of health information, it is important that email messages are scrutinized for their compliance with the organization's privacy rule obligations. In particular, messages containing confidential health information may only be sent to the patient, other healthcare organizations or business associates. This implies the application of some sort of content inspection and policy filter to outbound electronic mail. Typically, this content inspection will involve the application of a 'scoring function' to the email's plaintext contents. The organization maintains a lexicon of terms typically used in health related communication, the scoring function determines when such terms are so prevalent that it is likely that the message is a confidential communication. Once that determination is made, the organization must ensure the appropriateness of the message destination. Obviously, application of this sort of a 'policy engine' is heuristic and only supplements the organization's 'appropriate use' policy that is applied by enterprise staff. The application is important, though, as part of the organization's diligence in preventing unwarranted disclosures of health information. The content inspection, however, can only be applied to plaintext. This fact is particularly relevant when using s/mine methods. Message keys are encrypted using the public key of the *recipient*; once so encrypted, enterprise servers cannot recover message plaintext. As a practical matter, this means that either the enterprise policy engine is applied at the point of email creation, (the desktop), or that enterprise servers perform the encryption. PGP does support an 'alternate decryption key' (ADK) methodology where message keys are encrypted not only with the public key of the email recipient but also that of an enterprise resource. Using its ADK, a server implementation of the enterprise policy engine can recover message plaintext and make the appropriate policy decision.

Healthcare workers tend to be 'mobile' in that they access the organization's computing resources using any number of workstations or other end user devices. Furthermore, workstations are commonly shared among multiple users. Such mobility places practical constraints on the organization's encryption strategy. For the email to be decrypted at the workstation, the end user's decryption keys must somehow be deployed to that workstation. Similarly, if end users are to digitally sign the email messages they create, signature keys must also be deployed to each of those workstations. Such replication can occur using one of several methods:

- a manual process where the user physically transports keys using, say a floppy disk, to each workstation. The user then installs keys in the workstations key store. This 'low tech' approach requires substantial end user diligence in both protecting the floppy (or similar) and configuring the email application to appropriately store, protect, and use the key. The process leaves end user private keys vulnerable to the loss or theft of the floppy or compromise of the workstation's key store.
- use some kind of 'profile server' to deliver the end users key material, over the enterprise network, as needed to the user's current workstation. Typically, this solution involves specialized client software that manages the key recovery and the key

invocation by the encryption application. The solution, of course, requires some method by which end users authenticate themselves to the profile server and a secure protocol for communication between profile server and client software.

- use portable hardware tokens, such as smart cards, that contain the user's private decryption keys. Typically, these tokens interface with application software using either the PKCS#11 API or Microsoft's CryptoAPI to allow asymmetric cryptography with the token owner's private key to occur on the token itself. A principal concern in deploying these tokens is the potential for loss of the tokens. While the decryption keys can be escrowed, generally this approach requires that lost tokens be replaced before email message contents can be deciphered. As healthcare generally places a very high value on availability of patient information, the delay involved in token replacement may be a major concern.

Healthcare organizations generally have a bias against deploying new functionality to workstations or other client devices. Compared to other industries, healthcare organizations spend relatively little on information technology. Gartner for example, estimated that healthcare organizations spent 3.15% of revenues on IT spending in 2001. This should be seen as being relatively little, say in comparison with the 4.80% spent in the financial services sectors, especially when one takes into account the significant complexity of healthcare reimbursement and the systems that create and retain medical information. Due primarily to financial concerns then, healthcare organizations are slow to upgrade their computing infrastructure. Most healthcare organizations still have substantial installed bases of windows 9x workstations. Further, new investments in end user computing devices are often for mobile devices such as tablets or pda that provide user convenience but relatively little processing capability. Consequently, most healthcare organizations' computing environment do not simply support deploying significant new computing capability to end-users.

Healthcare organizations, especially hospitals and clinics, place a premium on eliminating IT 'interference' with the workflow of health practitioners. Physicians and nurses are persons upon whom the healthcare organizations obviously rely. But physicians are generally not employees of the healthcare organization, instead they are independent practitioners whose referrals are necessary to the economic livelihood of the organization. Physicians are certainly well aware of this and often non-cooperative with any IT requirement for which they do not recognize a personal convenience or benefit.

There is currently little deployment of email encryption capability outside the largest healthcare organizations, even though the substance of the HIPAA encryption mandate has been clear for more than four years. Healthcare organizations have seen email encryption as a 'compliance' problem and as such tend to defer implementation until the effective 'compliance date' of the mandate. However, after several years of delay, the Final HIPAA Security Rule is now imminent with its publication expected in February 2003. Once the Final Security Rule is published, healthcare organizations will have twenty-six months to develop the capability to encrypt their Internet transmissions of personal health information. This capability must be developed from scratch in an

industry where there is little infrastructure by way of trusted directory services or PKI. Particularly troubling is the fact that the lack of the desired infrastructure is not due to a lack of effort to create as much. There have been, for example, a number of failed attempts to create a PKI for the healthcare industry, including some very well funded attempts by Intel, Entrust, and the 'dot com' Healtheon (now part of WebMD). While this paper does not intend to infer that the prospect of a healthcare industry PKI is 'dead', its history is such that healthcare organizations simply cannot expect an industry PKI to be available for their secure email purposes by the Security Rule's April, 2005 compliance date.

The problems created by the lack of infrastructure is exacerbated by the fact that on the order of 85% all healthcare organizations are small business as classified by the US Department of Commerce. Such organizations typically have little capacity to understand and acquire support for email encryption; small organizations have little by way of technical support. Unfortunately, without significant technical support, individual users have considerable difficulty understanding and using public key encryption. A 'usability' study conducted at Carneige-Mellon University found that a large percentage of persons, even those with good computer skills, where unable to appropriate install, configure and use commercial versions of the PGP application. Some of the errors that individual users made have serious security implications. For example, some users, along with their certificate, sent their private key to correspondents. It is problematic for the large healthcare organization that it is to such users that they direct the bulk of their email messages.

It is also the case that few of the smaller healthcare organizations and healthcare 'business associates' have operations that are dedicated to servicing a larger healthcare organization. Physicians, for example, may have privilege at multiple hospitals, belong to numerous practice affiliations and medical groups, and have a patient mix related to many health plans. Similarly, a hospital business associate, such as a collection agency, is likely to have many other clients. As a result of this fragmentation, the larger organization has limited ability to successfully promulgate proprietary encryption solutions to the smaller organizations. Such solutions, by their very nature, require exceptional handling by the small business recipient. But such special handling is contrary to the 'universal' character of email that accounts for its significant use within healthcare. So to the extent to which the large organization demands special procedures and protocols to process its email, the business value of that channel to the small business recipient is diminished; at some point, the small business foregoes email use with the larger organization. Today, a number of larger healthcare organizations, while focusing on the limited capabilities of smaller organizations, deploy proprietary 'secure web mail' solutions. But when doing so, they fail to recognize the inconveniences that such solutions create for their correspondents and the resulting loss of business value. At least one large organization, Catholic Healthcare West, a 42 hospital system based in San Francisco, learned as much from a multi-year deployment of a secure web mail application. CHW has since abandoned that application because its healthcare industry correspondents preferred not to use it at all or only reluctantly. As a result, that application did not provide for reliable message

delivery. CHW found that the value of the security improvement simply did not offset the loss in convenience to its correspondents. Since, in secure web mail implementations, the email recipient receives little or no value from the security improvement while bearing most of the loss business value, the CHW experience almost certainly will be repeated elsewhere.

**Solutions to the Encryption Problem**

The appropriateness of encryption solutions is limited by industry context. Given the size and complexity of the healthcare industry, it should be clear that the only encryption solutions preserving the cost effectiveness and business value of email are those based upon standard methods. These solutions must be simple for the healthcare organization to deploy on behalf of its internal users. Furthermore, the larger organization must have a practical mechanism to support solution adoption by its smaller trading partners. These requirements are best served, in the larger organizations, by the internal deployment of email encryption through a centralized server resource. To assist the smaller trading partner's acquisition of encryption capability, a new certificate distribution model is required. The remainder of this paper will discuss concepts relating to email encryption gateways and better models for public key deployment. The discussion will be oriented towards s/MIME, primarily because of s/MIME's stronger standards basis and vendor support.

Domain Level Encryption

For email use PGP and s/MIME are usually discussed as end user applications, but their utility is not so limited. Processing requirements for both openPGP and s/MIME are specified in terms of Sending and Receiving Agents (or applications). Neither specification includes language that requires senders or receivers to be individual (natural) persons. Indeed, with these methods, relevant identity information is encapsulated in the public key certificates that are used to bind rfc822 email addresses to public keys. Any limitation as to the 'holder' of such certificates is a matter of the certificate issuer's CPS (Certification Practices Statement) and referenced CP (Certificate Policy). As a practical matter, organizations can issue certificates as they see fit to serve their purposes for electronic mail use.

The enterprise may chose to encrypt outbound email messages on enterprise servers rather than on the desktops of individual users. The s/MIME or PGP certificates of external recipients can be acquired and maintained by enterprise directory services. Since only the recipient's public key is needed for encryption, email so encrypted is identical to messages encrypted on the workstation of individual users. While encrypting on servers, the organization does forgoes 'end to end' encryption. But end to end confidentiality is neither a HIPAA regulatory requirement nor otherwise desirable for the healthcare organization. Since healthcare organizations generally restrict access to their internal networks and applications, the additional security benefits of encrypting email from the workstation to the enterprise boundary are marginal; HIPAA only requires encryption once the confidential information is placed on the open network. However, the availability of email plaintext to enterprise servers does allow the organization to more readily assess message contents and apply the organization's

disclosure policy. This strategy has the additional benefit of making encryption details transparent to internal users and otherwise reducing the computation requirements of end user mail applications and workstations.

Providing similar domain level support for the decryption of inbound messages involves a more complicated analysis. External senders need a public encryption key for each of the organization's internal addresses to which they would be sending health information, while a domain level decryption service needs access to the related private key. A straightforward solution involves the organization creating a single key pair and then creating for each internal address a public key certificate, binding the address to the public portion of that common key pair. These certificates have been called 'proxy' certificates. Should the organization choose to publish the fact the certificate subject does not actually control the related private key, they may do so in a cps or in the 'user notice' field of the certificate's policy extension, although there is no s/MIME requirement to do so. In the absence of such notices, proxy certificates would generally be indistinguishable from those issued in cases where the ultimate recipient held the private portion of the key pair. This is undoubtedly appropriate; there are very few, if any circumstance where external senders need end to end encryption or where such encryption is desirable. Internal recipients act as representatives of the healthcare organization; the organization needs access to plaintext content to generally ensure the availability of message contents to the organization's care and business processes and to filter out virus and other undesirable content.

Recent IETF standards development, generally labeled "domain–security", though, supports a mechanism which avoids both this ambiguity and the overhead of multiple certificates bound to a common key pair. Once again a single key pair is generated, but the public portion is bound not to individual rfc822 email addresses but to the domain in a single encryption certificate. That certificate would then be used to encrypt *any* email sent to any address within the organization domain. The IETF work provides a naming convention that identifies the domain encryption certificate as such and thereby provides a standards basis for interoperability among vendor products. As of yet, few if any end user mail applications recognize and correctly process the domain encryption certificates. As a result, for the time being at least, the organization can only use dom-sec certificates with external domains that it knows to support dom-sec. So organizations must anticipate a continued need for some form of proxy certificates.

Parallel with domain encryption is similar support for the use of digital signatures as a message integrity control. s/MIME and PGP use digital signatures to provide assurance that email messages were not altered in transit. The party constructing the digital signature is identified in a public (verification) key certificate. The s/MIME specification requires identification of the relevant verification key certificate; typically the cert is included with the email message. Under s/MIME, the verification key must include the rfc822 address of the sender. Similarly to domain level encryption, domain level signing can also occur using some sort of a proxy verification certificate and common signature key.. While such signing does assure recipients that messages were not altered once they left the sending domain, in itself, that signing does not provide assurance that the

individual person recognized as the email's sender, did in fact, create the message that was received. This is relevant because there is a tendency to view the digital signature of email as an expression of the senders agreement, endorsement, or approval of the message contents, i.e. as an electronic signature as defined by US Federal E-Sign Law. In healthcare, due to the licensing of individuals, there are circumstances where an individual staff members, for example physicians, sign in roles other than those associated with membership in the organization's workforce. Dom-sec helps prevent inappropriate signature assertion by disambiguating this use of domain level signing. Dom-sec does this by defining a number of signature types:

1. *originatorSig*, for the signature of the person creating the email message;
2. *domainSig*, for the domain on behalf of the originator. This sort of signature provides assurance that the message was not modified during in its transfer over the Internet. This signature satisfies the minimal HIPAA requirement for message integrity.
3. *reviewSig*, for persons approving the message for onward transmission. Presumably this sort of signature would indicate that there was a formal determination that release of message contents was consistent with enterprise policy. At this time, few healthcare organizations attempt this level of control.

Dom-sec goes further and includes naming rules to further distinguish the verification certificates of the domain and review signing 'authorities' from those of message creators.

.A number of vendor products exist today that support domain level encryption and signing and, to some degree, these products support dom-sec. Support though for dom-sec is still lacking in s/MIME clients, so healthcare organizations should anticipate a requirement to support some sort of proxy verification certificate.

Improved Deployment of Public Key Certificates

The common difficulty faced by all healthcare organizations seeking to adopt public key based encryption for their email is the lack of widespread PKI deployment within the industry.

In principle, this should not be an insurmountable barrier. In the absence of PKI, the organization must otherwise have some sort of 'out of band' communication with the intended recipient to negotiate encryption parameters and / or authentication tokens. So successful message delivery is dependent upon additional activity in any case. If that activity can be effectively channeled to support certificate issuance, then both large and small organizations will benefit from a standards based email security solution.

For the most part, the practical problem with s/MIME today is that the workflow surrounding certificate acquisition has not been well designed. Typically, the sendiing organization will try to implement the following scenario when it does not have an encryption certificate for an external email address:

1. Encryption gateway receives email that it cannot transfer without encryption. Since gateway does not have an encryption certificate for destination email

address and cannot find one in public directories, it places email in 'pending certificate' file.
2. Request certificate from external recipient. This request may be made electronically or non electronically by original message sender or administrator.
3. When the external recipient does not already have a certificate, he is directed to a public CA such as Verisign.
4. The external recipient follows the procedures of the public CA to acquire an encryption certificate.
5. External recipient sends certificate to encryption gateway in a specially formatted message
6. Encryption gateway recovers email from its pending status, encrypts and forwards with SMTP.
7. External recipient receives email and decrypts.

It should not be surprising that this workflow typically is *not* completed. One of its obvious design defects is that the workflow depends upon the external recipient independently completing a sequence of steps in order to implement technology of which the user has little knowledge and / or experience. Furthermore, the activity that triggered the workflow is left pending while the organization awaits feedback from that recipient. The workflow provides no mechanism for the organization to monitor the progress of the external recipient in acquiring the certificate; the organization only knows that it has not yet received the certificate. As a result, this kind of workflow resists effective management. Rather than rely on such ineffective workflow, most healthcare organizations decide to do 'something' else.

The problem then is not with PKI, per se, but rather the workflow in managing it. PKI would be effective if it provided a directory service that *always* returned an encryption certificate for a queried email address. If additional action is needed to complete the secure messaging workflow, then such action could follow encryption and mail delivery.

An example of such a directory service is currently available in a much more robust approach to PKI deployment, one that 'pushes' the PKI capability to external recipients. The workflow might be as follows:

1. Encryption gateway receives email as above; queries directory service with rfc822 address.
2. If directory service does not find the required certificate in its repository, it generates a key pair and writes the required s/MIME compliant certificate. Returns to the requestor either the newly created certificate or certificate in repository to the gateway.
3. Gateway uses certificate to s/MIME encrypt the message; forwards to SMTP relay.
4. When it generates a new key pair; the directory service places key pair and certificate in a PKCS#12 envelope that is protected by randomly generated password.
   4a) Password sent to requesting organization over a secure channel.

4b) PKCS#12 sent to external recipient in email message that appears to come from the healthcare organization, explaining steps that must be followed to install the key pair and thereby acquire the capability to open secure email form the organization.

4c) Organization applies its diligence in communicating password to external recipient. Password is needed to complete installation of recipient's key pair. Here, the organization effectively acts as a registration agent for the directory service's certificate authority.

4d) Directory service destroys or securely escrows the generated key pair per its cps.

5. Recipient decrypts.

The workflow is more practical than the earlier example, in that the external recipient is expected only to open an email attachment containing the PKCS#12 and supply the password provided by the healthcare organization. The ease with which the external recipient can accomplish this task is a function of application design. One vendor that implements this workflow, Public Key Innovations Inc, inserts the PKCS#12 inside an Active X control that provides 'one-click' installation of keys and Outlook / Outlook Express s/MIME configuration. The organization is allowed to complete its portion of the workflow without having to wait for external recipients to perform other actions. The organization does not have to leave messages in a pending status for an indeterminate time period.

**Conclusions**

If healthcare organizations are going to be able to continue to receive business value from their use of Internet email, they will have to devote significant resources to securing that use. While many aspects of providing that security are straightforward matters of server and network security administration, email encryption involves significant challenges. The large healthcare organization best accomplishes its email encryption and decryption as well as signing thru the use of an encryption gateway. The business value of encryption solutions is maximized when those solutions have a standards basis. However, in order to utilize standards based solutions, the large healthcare organizations must assist their smaller communication partners acquire PKI capability. For s/MIME purposes, PKI can be made practical thru an improved certificate distribution workflow.

**References**


1. 104th Congress, "Health Insurance Portability and Accountability Act of 1996" US Public Law 104-191" August 12, 1996 (1996)
   URL: http://aspe.hhs.gov/admnsimp/pl104191.htm

2. Biskler, Scott; Tracy, Miles and Jansen, Wayne, "Guidelines on Electronic Mail Security", NIST Special Publication 800-45.
   URL: http://csrc.nist.gov/publications/nistpubs/800-45/sp800-45.pdf

3. Calles, Jon et al, "OpenPGP Message Format - IETF draft-ietf-openpgp-rfc2440bis-06.txt" August, 2002

URL: http://www.ietf.org/internet-drafts/draft-ietf-openpgp-rfc2440bis-06.txt

4. Chase Law Group, "Unlawful Email Interception: Case Study" November 22, 1999 (
URL: http://www.criminaldefense.com/computer_email_interception.html

5. ICSA Labs, "Press Release: Virus Prevalence Survey" August 4, 2002 (2002)
URL: http://www.trusecure.com/corporate/press/2002/avsurvey030402.shtml

6. Marks, Richard "Guidelines for Initiating HIPAA Systems Implementing Projects" Analysis and Perspective Vol 5, No 18 May 3, 2000
URL: http://www.wedi.org/public/articles/pki_marks_hipaa.pdf

7. Mitre Corporation, "Common Vulnerabilities and Exposures: SMTP"
URL: http://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=smtp

8. National Research Council, For the Record: Protecting Health Information. Washington, DC: National Academy Press, 1997
URL: http://www.nap.edu/readingroom/books/for/

9. Office of Civil Rights, US Department of Health & Human Services "45 CFR Parts 160 and 164 – Standards for the Privacy of Individually Identifiable Health Information", August 14, 2002
URL: http://www.hhs.gov/ocr/combinedregtext.pdf

10. Office of Secretary, Health Care Financing Administration, "45 CFR Part 142 - Health Insurance Reform: Security and Electronic Signature Standards" Federal Register Vol 63, No. 155 August 12, 1998 (1998): 43242-43280.
URL: http://aspe.hhs.gov/admnsimp/nprm/secnprm.pdf

11. Partner, Chris and Glaser, John "Myths about Healthcare IT Spending" Healthcare Informatics, July 2002
URL: http://www.healthcare-informatics.com/issues/2002/07_02/myths.htm

12. Perigee.net Corporation , "Perigee.net (Home Page)"
URL: http://www.perigee.net/main.html

13. Ramsdell, Blake "S/MIME Version 3.1 Message Specification - draft-ietf-smime-rfc2633bis-03.txt January 16, 2003
URL: http://www.ietf.org/internet-drafts/draft-ietf-smime-rfc2633bis-03.txt

14. Dean, T and Ottaway, W. "RFC 3182 - Domain Security Services using S/MIME". October, 2001.
URL: http://www.ietf.org/rfc/rfc3183.txt?number=3183

15. United States Code, Title 18, Part I, Chapter 119, Section 2511"
URL: http://www4.law.cornell.edu/uscode/18/2511.html

16. Whitten, Alma and Tygar, J.D. “Why Johnny Can’t Encrypt:- A Usability Evaluation of PGP 5.0” Carneigie Mellon University School of Computer Science Technical Report CMU-CS 98-155. December, 1998
URL: http://www.cs.cmu.edu/~alma/johnny.pdf