This article interrupts my ongoing articles on medical and psychological issues. Yet I still hope to stimulate discussion, letters, and interaction in Telicom and also possibly on outside forums such as The Thousand.

The Email Security-Usability Dichotomy: Necessary Antinomy or Potential Synergism?

Vernon M Neppe MD, PhD, FRSSAf[a]

Abstract:

Many of us want the best security we can get for email, considering how many threats we have to deal with today, including virus, Trojan horses, spoofing, phishing, spam, hackers, eavesdropping, aggressive legal discovery, international exposure and, in the majority of cases, internal attacks. However, increased security has meant decreased usability, and vice versa. This is not only a problem for companies but for individuals as well, with their medical records and financial documents.

This paper is based on a detailed consensus-based analytic comparison of the four major technologies for secure email, X.509/PKI, PGP, IBE, and Zmail. General usability and security metrics are applied to rate available secure email systems, thereby calculating a respective score for each system tested. X.509/PKI, PGP and IBE are clustered within similar and qualitatively much lower ratings when compared with Zmail, which reaches near the maximum scores in both security and usability.

This work thus confutes the conventional thinking that usability and security are like a seesaw; if usability goes up, security must go down, and vice-versa. This apparent antinomy can now be seen as a synergy: With more usability in a secure system, security increases. With less usability in a secure system, security decreases. A secure system that is not usable will be left aside by users.

As both a limitation and an objectively strong point of the method used in this work, these are findings based on models for usability and security, not user focus groups and penetration testing analysis.

Key-words:

Antinomy, Authentication, Convenience, Decryption, Dichotomy, Email, Encryption, Key Escrow, FFIEC, GLBA, HIPAA, IBE, Identity-Based Encryption, ISO 17799, Key, Mail, Metric, Password, PGP, Phishing, PKI, Pretty Good Privacy, Voltage, Regulatory Compliance, Reliability, SB-138, Security, SOA, Spamming, Spoofing, SSL, Synergism, TLS, Trojan Horse, Two-factor Authentication, Usability, Virus, Worm, X.509, Zmail, ZSentry Mail.

1. Perspective

"The email message addressed to a Booz Allen Hamilton executive was mundane—a shopping list sent over by the Pentagon of weaponry India wanted to buy. But the missive turned out to be a brilliant fake. Lurking beneath the description of aircraft, engines, and radar equipment was an insidious piece of computer code known as "Poison Ivy" designed to suck

[a] Professor Vernon Neppe, Seattle, WA. www.PNI.org. psyche@PNI.org.

sensitive data out of the $4 billion consulting firm's computer network.

The Pentagon hadn't sent the email at all. Its origin is unknown, but the message traveled through Korea on its way to Booz Allen. Its authors knew enough about the "sender" and "recipient" to craft a message unlikely to arouse suspicion. Had the Booz Allen executive clicked on the attachment, his every keystroke would have been reported back to a mysterious master at the Internet address cybersyndrome.3322.org, which is registered through an obscure company headquartered on the banks of China's Yangtze River.

The U.S. government, and its sprawl of defense contractors have been the victims of an unprecedented rash of similar cyber attacks over the last two years, say current and former U.S. government officials. "It's espionage on a massive scale," says Paul B. Kurtz, a former high-ranking national security official. Government agencies reported 12,986 cyber security incidents to the U.S. Homeland Security Dept. last fiscal year, triple the number from two years earlier. Incursions on the military's networks were up 55% last year, says Lieutenant General Charles E. Croom, head of the Pentagon's Joint Task Force for Global Network Operations. Private targets like Booz Allen are just as vulnerable and pose just as much potential security risk. "They have our information on their networks. They're building our weapon systems. You wouldn't want that in enemy hands." Croom says. Cyber attackers "are not denying, disrupting, or destroying operations—yet. But that doesn't mean they don't have the capability."… [1]

On the Business Week blog responses for that article, we also read from a blogger called Michele:

I have been reading all I can about cyber attacks and warfare. The former Chief Strategist of Netscape, Kevin Coleman, has warned that we are at great risk in business, government and industry. Why is it we never listen to the experts before it is too late?[2]

Another entry in the same blog, confirmed to be from Ed Gerck[b], commented that these events should not actually be expected at that level of communication security:

It's surprising to know that a top US defense contractor even had to read such email. The reason that organizations do not better protect their email and authenticate each sender is not just cost (PKI) but because it has been difficult to do so.[3]

In the same comment, Gerck also cited a case where even in the absence of any attack, the diplomatic mail of some governments could be read by someone who legitimately operated a non-secure server, which server those countries nonetheless used to send purportedly secret communications, hoping perhaps to be hidden "in the forest" 3. The case refers to Dan Egerstad, a 21-year-old security researcher, who revealed that he was able to capture the information by setting up his own node in a peer-to-peer network used by the embassies to make their Internet traffic anonymous. Egerstad collected thousands of sensitive emails and passwords from the embassies of countries such as Russia and India and blamed systems administrators for not using encryption to shield their traffic from snooping.[4]

Were those just radical examples? Perhaps, but we see similar problems often, in our own inbox. Many problems remain innocuous because they are defused before they can cause harm. However, there are sometimes crushing problems, even when care is taken to avoid them. Such events often shock both those who did not know the extent of the problem and even experts.

---

[b] Dr. Ed Gerck, well-known for his work in Internet security and cryptography, is the developer of secure email Zmail discussed in this article.

2. REGULAR EMAIL

2.1 The hazards of regular email

Email is the most important single service running on the Internet. Regular email is also the number one source of security risks, making our messages vulnerable to inside and outside threats.

Email messages have no protection whatsoever. An email message is like an open postcard. Anyone can read an email you send or receive, and even change them, at any time while in transit on the Internet, at rest in a server, in a back-up file, or even in your computer.

Anyone receiving an email identifying you as the sender has no way of knowing if you really sent it or not. I frequently receive returned emails supposedly sent from my email address that I have never sent. Business customers may also receive emails using your name. That can cause you and them great harm, and yet many will not even be able to identify those emails as fraudulent. Emails may also be confused with spam or identity theft attempts, and deleted or delayed by overloaded systems and people.

Regular email is broken in many ways. Every email or attachment you send over a computer network is also copied (and perhaps even backed up) on many different computers, without your explicit knowledge and consent. That is the way computers pass data around -- they make copies.

But the lack of security in regular email is not limited to transmission exploits. Regular email is stored, sometimes even temporarily, in online servers and clients. This makes email vulnerable to a wide range of attacks, with external and internal sources that may also be exploited in combination.

Email messages are markedly at risk in our own facilities: Employees and contractors and sometimes clients need to access data and servers as part of their routine work. Over half of known IT security breaches occur from within organizations.[5] There is an economic consequence, too: Email security breaches significantly affect company stock prices.[6] Virus attacks were named the leading culprit of financial loss by U.S. companies in 2006.[7] Moreover the average annual loss in the Computer Crime and Security reported by U.S. companies more than doubled, from 2006 to 2007.[8]

Some email-specific hazards are listed in Table 1.

Table 1: Some Email Specific Hazards

Identity theft: Generic name for any fraudulent activities using someone else's identity such as unauthorized access, interception, interference, deletion, alteration, forgery or suppression of data.
Impersonation: A more accurate name for Identity Theft.
Phishing: Attempted fraudulent acquisition of sensitive information e.g. usernames, passwords and credit card details, by masquerading as a trustworthy entity.
Pharming: A hacker's attack aimed at redirecting website traffic to another, bogus website.
Spam: Unwanted, unsolicited email. However, unwanted, unsolicited email that seems legitimate and is either received once or has an "unsubscribe address" is not considered spam.
Spoofing: Fraudulently using the sender address and/or other parts of the email header to make the email look as if it originated from a different source.

Spam has become not only a major nuisance but also a very important factor in achieving success in attacks. Because the probability of each single success event is low, spam is used to increase the total probability by massively sending the message with the desired attack vector (e.g., phishing). The success of an attack is basically a numbers

game – send more spam, achieve more success.

The hazards listed in Table 1 also are often combined with attacks using worms, Trojan horses, and viruses, that use email to propagate directly and also indirectly (e.g., in a visited web page, clicked from an email link). Thus, non-email related vulnerabilities have also become major problems in disrupting email communications.

In addition to the email-specific hazards introduced in Table 1, other email-specific risks can be quite significant and have been exploited. Some of these additional hazards are listed in Table 2.

Table 2: Additional Regular Email Hazards

Eavesdropping
Message Modification
False Message Date
Message Replay
Unprotected Backup followed by disclosure
Sender Repudiation (sender denies that she ever sent it)
Recipient Repudiation (recipient denies that she ever read it)

Email "soft spots" (e.g., Table 1 and Table 2) and possible attack vectors allow for a large number of risk factor combinations, including hackers, automatic password crackers, virus, worms, buffer-overflow, software bugs, zero-day-exploits, delayed patching, patch conflicts, security gaps, collusion, conflicting business interests, lack of legal protection for data at rest, and other security breach factors. These risk factors for regular email are not going away anytime soon, while their number keeps growing larger and the attacks more subtle.

Email security problems plague Internet servers, users' desktops, laptops, even their cell phones and off-line machines. Email security problems can cause a direct *invasion of privacy* in an otherwise trusted situation. Private data in regular email can be compromised, inter alia, in improper file access by a service provider employee or by their contractors, in an overly broad legal discovery processes, and in government-mandated broad surveillance.

All these actions can be rather easy to perform unilaterally in some jurisdictions. For example, email can be sent abroad, sometimes just for processing purposes (e.g., for spam or virus detection), which situation may be unknown to the sender and recipient who are both local. While they may correctly trust their local protection laws, their email is nonetheless subject to compromise risks in a foreign jurisdiction.

Online email services such as Gmail™ up the ante on the already large—and growing—Internet risk by using a multi-tenant architecture where many users share the same servers—if one or more servers are compromised, many user accounts may fall in wrong hands.

The risks of regular email (e.g., disclosure, tampering, and spoofing) are not limited to email transmission risks, that could be protected by SSL, and increase as the number of customers sharing the same servers increase for a given provider.

Thus, even when you have nothing to hide, you may be the victim of email fraud, disclosure, spam, spoofing, phishing and pharming attacks. These attacks are increasing in number and reach, with major losses resulting from worms and viruses that use email to propagate.

The problem of making email information secure includes further levels: We may

want to establish proof of sending, proof of receipt (who received, date, time, where, what client machine, what operating system, what application). We may also want to electronically notarize (message was not tampered with, message was delivered as sent) and digitally sign (message is legally valid as a signed document) the message.

These are no mean tasks. Not even our regular postal mail can do all of this. Imagine sending a registered letter by USPS, or through special mail services such as FedEx. The receiver may sign for it but claim that there was no letter inside, or that the contents were not as described. This is where the need for document notarization and document delivery registration (as opposed to envelope delivery registration) comes in. Further, sending a postal letter is inherently not secure. It can be read by others in transit (e.g, using the "tea kettle attack" or re-enveloping after reading), or read when the recipient is out of the office.

In short, regular email communication is just not a secure method of communication for several reasons, some of which also plague postal mail but to a lesser extent.

What happens if one sends a fax instead? This is still not secure. It can be read by anyone in the office. Moreover, it can easily end up at the wrong office. And a fax to your computer (an efax) does not help for the same reasons. Further, the same security breach factors that plague email can also be used to breach fax security, to your phone or to your computer.

2.2 But I have nothing to hide…

Ordinary persons may ask legitimately why they cannot just use regular email. Of course, they can. And we all do—all the time. There is a convenience to it. But there are hazards, and we need to select a balance that suits our purpose and purse.

Clearly, most people today would not send their credit card information by regular email. And, even if we think we have nothing to hide, such determination may fail in the future if that information is misused (e.g., by a health insurance company reading the grocery bill of an unsuspecting customer and making health care policy decisions based on what the consumer apparently eats and drinks).

We realize some of the bad parts about email when we receive unauthenticated, unsolicited mail such as spam, where we do not know their real origin. When the email comes with attachments that we would like to open, even if they are purportedly from people we know, the risk is greatly increased.

Yes, we still use email. And we do it daily. But we use it at our own risk—risks that most have not even conceived of.

If one wants to use regular email in a professional setting, the situation becomes far more complex. Regulatory compliance requirements such as the Health Insurance Portability and Accountability Act (HIPAA), Federal Financial Institutions Examination Council (FFIEC), Gramm-Leach Bliley Act (GLBA), and the Sarbanes-Oxley Act (SOA), as well as assistive technology requirements, professional guild rules, and organizations' policies may be difficult to meet, or even forbid use, for regular email communications in many cases (see also Tables 7A and 7B).

There are additional concerns. For example, organizations often want to protect their brands, and people often want to protect their name, against spoofing. Users also want to prevent spam and be able to read attachments without worms or viruses infecting their computer. In addition, it is useful for both sender and recipient to make legitimate email stand out against spam and phishing emails that may pass through spam detection.

In such a scenario, while it is not possible to protect the recipient in regular email,

senders often protect themselves at least legally. The Missouri Bar Disciplinary Council requires all Missouri attorneys to warn recipients about the risks of using regular email. For example, and as an example only, the disclaimer in Figure 1 could be added to a regular email for this purpose:

Figure 1: A Typical Disclaimer For Using Regular Email Communication

DISCLAIMER: CONFIDENTIALITY NOTICE

- Regular email communication, such as this one, is not a secure method of communication. Anyone can send a regular email using my name and email address.
- Any regular email that is sent to you or by you may be copied, changed, and held by various computers it passes through as it goes from us to you or vice versa.
- Persons not participating in our regular email communications may intercept our communications by improperly accessing your computer or our computers or even some computer unconnected to either of us which the email passed through.

The above disclaimer graphically shows the security and privacy problems of regular email.

3. SECURE EMAIL

3.1. The History of Secure Email

To be usable, secure email technologies must work together (interoperate) with standard email servers and applications, such as browsers. A basic objective of secure email is to achieve seamless operation from the viewpoint of the user. Operation should also remain within acceptable performance levels, usually referring to availability, reliability, error rate, and error recovery specifications. Performance levels may vary for each class of use, for example, the needs and requirements differ in military, home, financial institution, small business, medical and legal contexts. Fulfilling these requirements has not been an easy task and has often presented compatibility difficulties. These have been gradually reduced due to the increased standardization promoted by the IETF, W3C, ISO, ITU and other organizations.

One example of successful interoperation is intermediated by S/MIME, which stands for Secure/Multipurpose Internet Mail Extensions, a security-oriented protocol that adds encryption and support for digital signatures to the widely used MIME email protocol. S/MIME is not a technology providing secure email but works as an interface between the cryptography and regular email functionality. Table 3 shows the major cryptographic technologies, and solution examples, that are available today for secure email.

The table includes Postini and Google Mail, which provides a form of secure email based on SSL but fail to meet email security requirements (see item 3.4). The other technologies are discussed next. The security and usability features and difficulties of each technology are discussed elsewhere. [9, 10, 11, 12]

X509 and PKI (Public Key Infrastructure):

About 1988, the technology X.509[c] was introduced to define strong authentication, where a digital certificate cryptographically binds a name to a public-key. X.509 has been

[c] The ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T Recommendation X.509 has been implemented as a de facto standard to define a

Table 3: Examples Of Secure Email Technologies.

TECHNOLOGY	PKI/ X.509	PGP	IBE	ZMAIL
	Public-Key Infrastructure	*Pretty Good Privacy*	*Identity-Based Encryption*	*ZSentry Mail*
Year of initial development	1988	1991	2000	2004
Reference URL	**PKI** http://rsa.com http://entrust.com **X.509** http://verisign.com http://thawte.com	http://pgp.com http://hushmail.com	http://voltage.com	http://zmailservice.com
EXAMPLES	Entrust RSA Microsoft Outlook Cryptzone Rpost Postini (SSL) Google Mail (SSL)	PGP HushMail	Voltage MessageGuard	Zmail

used to define what we call a PKI,[d] in principle enabling entities to solve the basic problems of secure email communication.

PGP (Pretty Good Privacy):

PGP was released in 1991 to simplify and reduce costs in solving the basic problems of secure email communication for small-scale groups, as in a group of friends.[e] PGP technology provides both the digital certificate technology and the encryption technology.

Problems with X.509/PKI and PGP:

Both X.509/PKI and PGP require knowing the digital certificates of the recipient before the email is sent, which is often impractical. This is especially true if the recipient has none. X.509/PKI and PGP have additional usability difficulties, most related to digital certificate management. When compared with PGP, X.509/PKI seems to have the additional difficulty of requiring a PKI. This difficulty, however, does not apply for the same case where PGP is applied, with a small group. See further discussion elsewhere.[9, 12]

IBE (Identity Based Encryption):

X.509/PKI and PGP usability difficulties led to the development of IBE in 2000, where the digital certificates are simply the email addresses of each user.[f] With IBE, the basic problems of secure email are trivially solved but this happens at the potential cost of privacy and security as all private keys are stored and at-risk in central servers.

Problems with IBE:

The lack of security of IBE-based solutions (e.g., MessageGuard, Voltage) are mainly due to the fact that IBE inherently mandates key-escrow, where a central location becomes a single point of failure for decryption of all communications for all parties. See further discussion elsewhere.[9, 12]

Zmail (ZSentry Mail):

The motivation behind the Zmail technology development is that existing technologies for secure email (X.509/PKI, PGP, and IBE) were either not usable enough or not secure enough. Zmail uses the NMA ZSentry[g] technology. Zmail uses standard cryptographic

protocols. However, if security is breached or even if an attacker physically walks away with any servers, no customer access data or customer data can be recognized or accessed. With Zmail, the basic problems of secure email are solved but there is no usability-related burden of certificate management (as in X.509/PKI and PGP) and no security-related burden of key-escrow (as in IBE).

Problems with Zmail:

Some of the promised usability and security features such as the possibility of replying securely to the sender without registration, secure web form processing, and the choice of marking a message to be read only are either still in tests or not implemented. See further discussion elsewhere. [9, 12]

3.2. What is needed in secure email?

Email is facing a dilemma today: extinction or change. Either option is unfavorable. Extinction of email is a real possibility. Other Internet protocols, widely used less than 15 years ago, have become extinct at least to the general public, such as FTP (File Transfer Protocol), and Telnet (Terminal Emulation Program for TCP/IP). Email extinction, killed by spam, spoofing, phishing, virus, eavesdropping, and difficulty of integration with different organizations' policies, may likewise come quite quickly. On the other

---

framework for the provision of authentication services, under a central control paradigm represented by a "Directory". It describes two levels of authentication: simple authentication, using a password as a verification of claimed identity; and strong authentication, involving credentials formed by using cryptographic techniques. It is this second level that is of interest for secure email, with X.509 digital certificates binding the public keys of a user, together with some other information, with the private key of the certification authority which issued it. The binding is rendered unforgeable by encipherment. The X.509 standard is available free of charge for a previous release, at the ITU web site http://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-X.509-199708-S!!PDF-E&type=items

[d] A public key infrastructure (PKI) binds public keys to entities, enables other entities to verify public key bindings, and provides the services needed for ongoing management of keys in a distributed system. See http://csrc.nist.gov/publications/nistpubs/800-32/sp800-32.pdf

[e] Philip R. Zimmermann is the creator of Pretty Good Privacy, as an email encryption software package. Originally designed as a human rights tool, PGP was published for free on the Internet in 1991. See PGP Inc, http://www.pgp.com

[f] An Identity Base Encryption (IBE) scheme is a public-key cryptosystem where any string is a valid public key. In particular, email addresses and dates can be public keys. See http://crypto.stanford.edu/ibe/

[g] Ed Gerck is the creator of ZSentry and ZSentryID, and Zmail. ZSentry solutions provide two-factor authentication and authorization, spoof and phishing prevention, password hardening, access control, key and certificate management, and other security services. ZSentry shifts the information security solution space from the yet-unsolved security problem of protecting servers and clients against penetration attacks to a connection reliability problem that is easily solved today. ZSentry does not have or access copies of customer login data and keys—there are no customer targets to be attacked and cracked. The best defense against data theft is to not have the data in the first place. See NMA Inc., http://nma.com/zsentry/

hand, email changes could be as impractical to implement as the change from IPv4 to IPv6, the basic address system on the Internet, which never took off even after 15 years of an all-out effort by the IETF and other organizations, and is still experimental and "next generation". Conventional email is so ingrained in users' habits and the make-up of basic Internet operations that it has a large inertia to change.

But dilemmas do not exist for long in a free market. The market solutions in the past 20 years have been to complement email with add-ons that provide what is missing, security, while keeping email unchanged at least on the surface.

These solutions, represented by the technologies X.509/PKI, PGP, IBE, and Zmail, as reviewed in item 3.1, provide different approximations to the desirable email solution that must be at the same time secure, usable, and compatible with everything else.

Usability and security are thus the drivers of an email soft-evolution that must not change email but complement it.

Are these contradictory—is there a necessary antinomy? Or can security and usability be both improved to the levels that we need to have?

Indeed, there appears to be a necessary antinomy between the two conditions. The most obvious example would be regular email which is very usable, but not at all secure. The early major technologies for secure email have not progressed to the level that they are both usable and secure.

To attain appropriate practical use, the technology must be as usable as regular email and as secure as our changing needs require.

First, we need to develop a technical perspective of the basic requirements of secure email. This is reflected in Table 4.

Table 4: Basic Requirements Of Secure Email[9]

message confidentiality: only the dialogue parties are privy to the message
message integrity: the message was not tampered with
authentication: the dialogue parties have verified identities and/or credentials

It is also important to bear in mind that fallacious ostensible strengths, often claimed in secure email technologies, such as claims of an added special security attribute because the text cannot be forwarded, or that the contents cannot be printed, or copied, do not make the message more secure because no technology can protect against taking a picture of the screen, either via computer or failing that by using an outside camera.

3.3. Development of quantifiable metrics for secure email evaluation

The basic requirements given in Table 3 do not fully address the issue of trust in Internet communication, are not sufficient for secure email, and do not address any usability issue, as noted by Gerck in 2005.[12]

Some may wonder about the role that trust should play, if any, in terms of secure email. Trust has been used with a variety of meanings, but in Internet communications it usually has to do with our reliance on someone or something in regard to "matters of x", where "x" could be visiting a site, opening an email, or replying to an email. In 1997, Gerck led an Internet discussion on the definition of trust in terms of Internet security, published in 1998.[10] Gerck's answer to the question "How can I trust a set of bytes?" provides a framework for understanding human trust (as expected fulfillment of behavior) and for bridging trust between humans and machines (as qualified information based on factors independent of that information).

Such was the starting point to apply the same concepts to understand the needs for

trust in secure email by Gerck in 2003, first informally in discussion lists and then more formally in 2005.[12] A (+) (-) metric was developed 12 with secure email performance requirements for both security and usability in terms of desirable features (+) and shortcomings (-), in order to rate existing systems and also guide development of new email systems. Table 5 shows the (+) and (-) criteria involving essential security issues, where those in italics also imply usability issues.[h] Based on the criteria suggested for current secure email systems, Gerck also began to develop Zmail, by attempting to apply these requirements.

Table 5: Criteria Of Desirable Features And Shortcomings In Secure Email Technologies.

Desirable features (33 criteria)

encryption (message confidentiality), decryption, message integrity, legal digital signature, key expiration, key revocation, identity certificate, private key not at server, *anti-spoofing*, sender two-factor authentication (end-point authentication), recipient two-factor authentication (end-point authentication), *message expiration, message release, message recall, verified timestamp, verifiable message notarization (fingerprint), send receipt, return receipt, mobile use encoding, compact encoding, attachment encoding (easy decryption), html encoding (very easy decryption), secure web form processing, decryption key self-revocation & reset, signature key self-revocation & reset, decryption key self-recovery, signature key self-recovery, send unique, protect Cc, protect subject, read once—read only— no login, read once—read only—no registration, read and reply once—no login.*

Shortcomings (17 criteria)

private key escrowed at server, break private key protection at server, break policy protection at server, weak authentication accepted, server spoofing, unverified sender's email address, phishing (email fraud), "lunchtime" attack, *key management, key revocation delay, lack of centralized key revocation, open message headers, must pre-enroll recipients, must register to read, must register to reply, must send own certificate, requires common root of trust.*

To reduce bias, Gerck decided to include a large number of criteria, and also to introduce a weight system to better reflect different use environments. To provide the needed outside validation of his (+) (-) criteria, of the weight system, and of the criteria allocation for the different technologies, Gerck created a blog which experts in Internet Cryptography could examine. This became a dynamic interchange.[11] The most recent version9 of the metrics was published in 2007.

The many public comments by both users and leaders in the field validated this approach and helped focus the development. Once these were established, a metric was developed giving a weight loading ("scorecard") for the 33 positive and 17 negative criteria to better reflect different use environments.[9] The scorecard generally ranged from 1 to 3 but two criteria were so important, they were loaded at 10, namely private key escrowed at server and key management. Seven others scored 5 namely, identity certificate, key expiration, private key not at server, break private key protection, break policy protection at server, unverified sender's email address and phishing (email fraud).

The full listings separated the criteria into the two groups —desirable (positive values) and shortcoming (negative values), as shown in Table 5. At that point, the three earliest technologies, PGP, PKI/X.509 and IBE, were scored.[12] The online blog continued after the later results were published. Consequently, even after the 2005 blog discussion, further updates were continuously occurring.

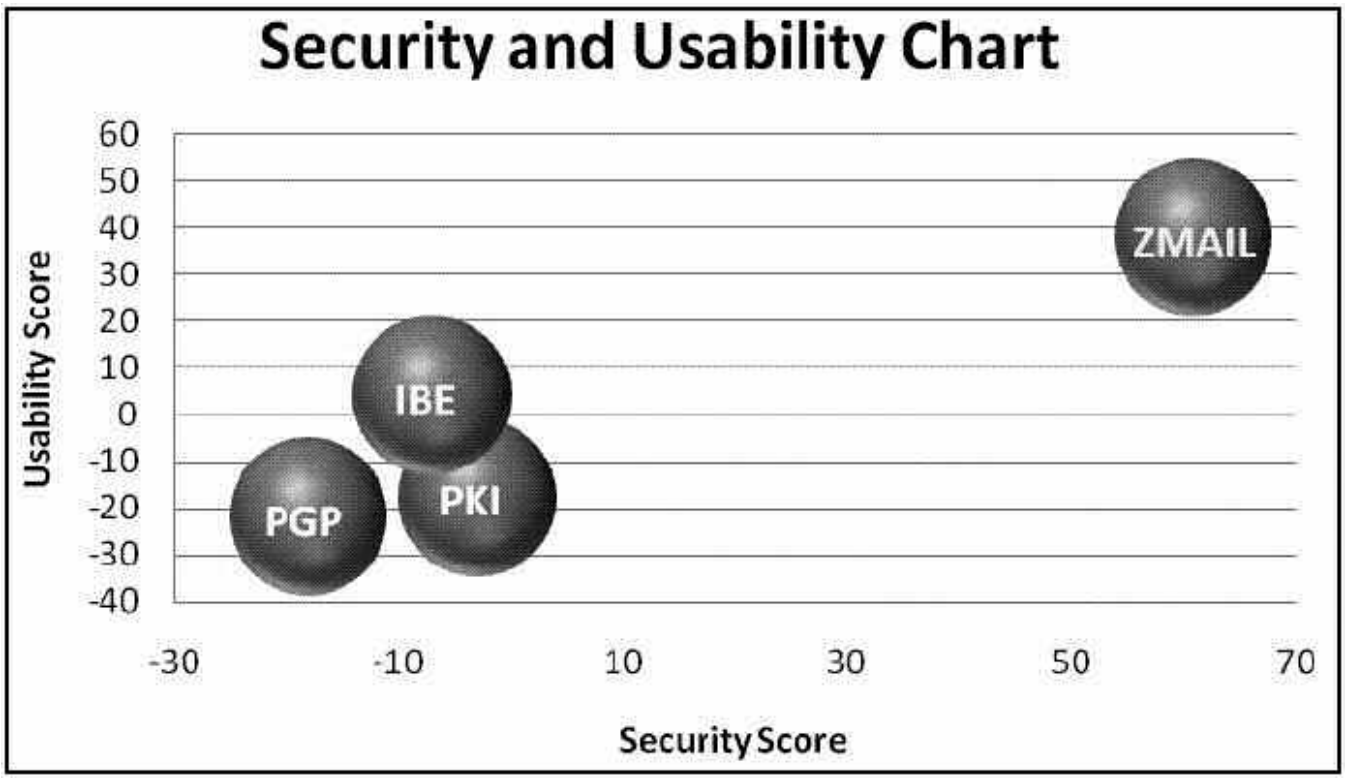
[h] All of these features by definition impact security. Entries in italics can considerably impact usability, consequently these criteria are used in both security and usability models.

After the development of Zmail, guided by such studies, the information was updated.9 The development of Zmail showed that it was possible to satisfy the criteria developed during that blog for security and usability at the same time. 11 Some of the missing positive points are motivation for future work. Meanwhile, the Zmail technology continued to develop given the ongoing feedback.

Figure 2 shows a graph with the respective scores for usability and security for each solution technology X.509/PKI, PGP, IBE, and Zmail, as evaluated in 2007.[9]

Positive scores are better: More usability and more security mean respectively higher scores. The chart shows that IBE, PGP, and PKI are not very far apart from each other. IBE is technically more usable than PGP and PKI and this conclusion fits the usability of respective solutions offered in the market today. However, IBE and PKI and especially PGP all have significant security issues. When these three earlier technologies are seen in comparison with the Zmail technology, it is apparent that Zmail represents a significant qualitative improvement in both security and usability, reaching near the maximum range of each score.

Figure 2: Comparison of main secure email technologies[13]. The Usability Score range is from –37 to 41. The Security Score range is from –70 to 64.



The scores above are also presented in a table combining comparative strengths and shortcomings of the four technologies.[13]

Additional experience with Zmail will become an important part of later analysis. Further, controlled comparisons can still be done in the field with user focus groups and penetration testing analysis.

Table 6: Comparative Scores of Secure Email Technologies[9]

Scores	PKI/ X.509	PGP	IBE	ZMAIL
Security	-41 38 **-3**	-51 33 **-18**	-34 27 **-7**	0 61 **+61**
Usability	-33 15 **-18**	-36 13 **-23**	-11 15 **+4**	0 37 **+37**

Table 6 lists 3 scores (shortcomings, desirable features, total) for each technology.

3.4. Technologies That Fail To Meet Security Requirements

There are several other technologies or methods that many may be familiar with on the Internet. The two most common are SSL/TLS— almost every reader has probably bought items over the Internet using this technology; the second is the password authentication techniques that are so prevalent.

SSL/TLS: This technology has been used to announce secure email services, for example, by Postini, Google Mail, and other companies. The main attraction is simplicity. TLS (Transport Layer Security) was previously called SSL.[i] TLS was developed about 1996 and has been very successful in ecommerce. These cryptographic protocols usually employ server certificates based on the X.509 standard. There is no need for a PKI. However, TLS/SSL cannot solve the basic problems of secure email. For example, because SSL/TLS messages are only encrypted in-between end-points, third parties can compromise message security and integrity at the security gaps created at each SSL/TLS endpoint (i.e., not only at Google or Postini but also at the recipient's ISP), and at the recipient's machine.

Password Authentication: Another example is password authentication and encryption. This is cumbersome to use, has no first-contact capability, and is trivially open to exploits by spoofing and phishing attacks. In addition, because users are likely to choose a weak password (even though it may look strong) and not periodically expire them, password-encrypted email may be rather easy to crack by the same automatic dictionary attack tools already in use to crack password files effectively.

4. PRIVACY REGULATIONS

Let us now examine a key issue: Regulations pertaining to privacy.

4.1. The politics of regulatory compliance

"Beyond hackers and beyond members forwarding material is the US government, which routinely acquires all Internet traffic. Data storage is cheap, decryption is easy for the government (by federal law), and ongoing massive search for keywords is probable, with "security" justification.

From a practical viewpoint, in the short-term this may not matter much. From an ethical point of view, electronic communication is never ever absolutely confidential. Be it for their own good or not, it is misleading patients for us to pretend that anything might be confidential other than unrecorded and unreported 1-1 private conversation in psychotherapy with a psychiatrist who doesn't tell tales in hospital elevators or at dinner parties. In the long run, it is easy to imagine scenarios where evolution of the government into idealistic totalitarianism will lead to dredging-up of long-stored evidence of insufficient patriotism, anti-religious leanings, internationalism or other seditious and destabilizing tendencies to the disadvantage of one or more medical personnel, others of whom will crow, "I told you so!" [j]

These comments appear very serious with regard to possible confidentiality and privacy breaches, individually, within professions and governmentally.

We need to recognize that the security of medical documents such as provided in

---

[i] The primary goal of the TLS Protocol is to provide privacy and data integrity between two communicating applications. See http://www.ietf.org/html.charters/tls-charter.html

[j] A psychiatrist sending an email to me.

HIPAA also depends on areas that are not presently covered by HIPAA (e.g., companies such as Microsoft and Google providing online storage services for medical documents by users, while legally disclaiming HIPAA assurances), in addition to built-in mandates within HIPAA regulation for patient file confidentiality exemptions.

Not surprisingly, several different regulatory bodies and laws have tried to ensure appropriate privacy by setting up their own requirements (Table 7A). Additionally, one needs to take into account the recommendations and laws passed to assure that everyone is potentially able to access the technologies (Table 7B).

Table 7A – Regulatory Compliance Laws Relevant To Secure Email
Health Insurance Portability and Accountability Act (HIPAA)
Federal Financial Institutions Examination Council (FFIEC)
International Standards Organization (ISO) 17799,
Gramm-Leach Bliley Act (GLBA)
Sarbanes-Oxley Act (SOA)
California SB-1386

Table 7B – Assistive Technologies Relevant To Secure Email
Section 508 of the U.S. Rehabilitation Act,
U.S. Americans for Disabilities Act,
W3C Web Content Accessibility Guidelines.

4.2 Technologies and Compliance.

The technologies X.509/PKI, PGP, IBE, and Zmail can satisfy the above requirements within varying degrees of security and usability.

Regarding privacy, Zmail further protects the user's privacy by adding the following properties:[14]

- Protects both the privacy of customer data and the keys that protect the privacy.
- Assures customer access audit trails and customer data storage are maintained with encrypted, de-identified numbers, satisfying regulatory compliance standards such as HIPAA without requiring a business associate contract.
- Enables user accessibility to the protected records directly, including behind company firewalls and also in the mobile web, so that everyone can be protected everywhere, any time, without requiring intermediaries such as customer service.

5. Necessary antinomy or potential synergism of usability and security?

In addition to model findings, one can also verify the evaluation metrics and examine the utility and ease of use of these technologies by empirically testing them out. The respective URLs for each solution provider are given in Table 4. One way would be to set up a series of practical day-to-day tasks for the ordinary person and test how easily each technology handles them. Another way would be to use an expert consultant to evaluate what security claims are valid in practice, for a likely use scenario. We can certainly apply our theoretical knowledge from the model but, as an empiricist, I also like to see practical data. I give my feedback in this regard in italics, based on my experience with Zmail (the "it" below).

- How long does it take you to send the secure email? *It takes only seconds longer than regular email.*
- How quickly is your copy received? *About the same time as a regular email.*
- How easy would it be for the barely computer literate secretary in a corporation? *15,000 plus users and over a million Zmails sent suggest it's not difficult. Shortly, it*

may become as convenient to use as a regular email with the development of a new non-plug using secure SMTP based technology that works with all email systems—called ZSEND.

- How truly secure is the specific technology being tested? *Based on comparative data above, it satisfies all the security requirements. There is added protection too from spamming, phishing, spoofing and viruses. If human error is to be considered in the threat mode, such as carelessly exposing passwords or user-codes, two-factor authentication can be used with the ZSentryID technology for Zmail.*
- How cross-platform are the technologies? *It is fully cross-platform: PC, Mac, Linux, and mobile web cell phones.*
- Are special installations required that would require time and some expertise, and versatility with different programs? *No installations are required; the technology is not an application requiring installing, de-installing, or updating a program; you click and read; you register, receive your data, and send.*
- How much may one rule out compromises of security because others can know the information? Can individuals from the company itself or its related groups acquire your secure information or read your messages? *If security were breached where someone physically walks away with any servers, no customer access data or customer data can be recognized or accessed because the usercode, user password, user keys, and email address are all required and not stored anywhere; additionally data can easily be set to expire. Some forms of human error can be mitigated with ZSentryID for Zmail.*
- Can we prove the person sending the message is truly that person? *No. Whereas higher level control can be used occasionally, the normal operation is that technically anyone can register any name. However, email addresses are authenticated by cryptographic challenge-response and one can trace the origin of the corresponding IP address. The email address is thus positively verified, fully providing for third-party sender authentication assurances.*
- Does the technology interfere with the way email works? *It does not. It is entirely compatible with all the commonly used email protocols. It integrates seamlessly with all email solutions, including desktop and webmail.*
- Can I use secure email immediately with anyone in the world? *Yes. It does not require any pre-registration, installation or pre-determined codes to be sent to sender and user.*
- When would I use the technology? *I use it for medical letters and information. Effectively the attachments are also electronically notarized with a timestamp and digital signature, so the immediate delivery is proven, unlike even FedEx where you can prove it was received but cannot demonstrate content. I also use it to send sensitive information (credit cards) and to communicate with attorneys in my medico-legal expert work for many reasons, for example, I can easily expire documents and also authenticate receipt if required. I use it, too, for financial transactions.*
- Surely any secure method of signing webmail requires a browser to execute code to produce the signature? *This may have been so in the past. Because Zmail does not have any user's keys stored in the servers or anywhere, the user can securely sign outside the browser from the user's own secure area.*

6. CONCLUSION

Regular email lacks confidentiality, is insecure and easily interceptable. Additionally, there are problems such as identity theft (made worse by phishing and spoofing), spam-

ming, and easy transmission of viruses. Indeed, there is, unfortunately, little that is truly compliant when privacy regulations (e.g., HIPAA) are considered. Faxes can be read by staff-members, often go to open areas in an office, or may be even misdialed and end up elsewhere. Limited confidentiality also applies for letters, voice mails and phone calls.

There is certainly a need for usability as secure email technologies become more and more needed and as new people enter the work force. Desirable security features include usability (as the first security feature), notarized document delivery, lack of data and key targets vulnerable in the servers and client machines. This is particularly important to militate against phishing, spoofing, spamming and virus transmission; to diminish liability; to ensure compliance with confidentiality assurances (e.g. HIPAA laws in the USA); and to permit the transmittal of electronic signatures with legal validity at both ends.

To answer the initial question posed: Is the dichotomy of email security and usability necessarily an antinomy? Are the two contradictory such that usability goes down as security increases and vice versa? Or can there be a potential synergism?

Based on the data presented, at least Zmail demonstrates extremely high grades for both usability and security at the same time. At a practical level, it is simple and convenient to use. As usability options have increased, security is better provided. Increased usability improves security. This is a synergic effect and it cuts both ways. *With more usability in a secure system, security increases. With less usability in a secure system, security decreases. A secure system that is not usable will be left aside by users.*

What about the other technologies? IBE, PGP, and X.509/PKI have their strong points as well. But limitations at the basic security and usability levels preclude solutions based on these technologies to foreseeably, if at all, reach the desirable simultaneous improvements.

There are, however, limitations of the above conclusions. These findings are based on heuristic models for usability and security, which, even though empirically justified, still have to be more widely tested. Further, these results are not from user focus groups and penetration testing analysis, which will also be required for further validation versus the model findings.

However, the method provided is objectively defined and should help our own empirical experience, and in defining our control variables.

A final word of caution is always useful. And this one comes from the past. The following statement was made in the past about PKI, one of the other four technologies studied here, in a customer report by Andress:[15]

Each of the past few years was supposed to herald the onset of PKI (public key infrastructure), the digital certificate-based technology that allows users to safely exchange data on insecure networks. But PKI has yet to gain a foothold. The technology's complexity and general lack of interoperability have stunted its growth in many enterprises. Yet the PKI dream refuses to die, and vendors are working to improve their offerings.

All that effort is finally starting to pay off, judging from our experience with a prerelease version of Entrust Technologies Inc.'s Entrust/PKI 6.0, slated for release this summer. The Entrust/PKI 6.0 beta impressed us with improved deployment options and enhanced integration with a wide range of applications, meaning that enterprises should gain a higher return on their PKI investments—and reap those benefits faster—if the final version of this product follows suit.

These are remarkable and impressive words. They were followed by an even more impressive summary:

Pros:
+ Integrates with Microsoft's Active Directory and CryptoAPI
+ Exports PKCS #12 certificates
+ Supports application signing
+ Requires minimal configuration
Cons: None significant

Andress wrote these words in 2001. How much has technology changed? Today we are able to easily understand the cons of PKI (as in Section 3.1), where the pros have not significantly materialized. But one part is correct in that 2001 statement, seven years later: PKI has yet to gain a foothold.

What will we be quoting in 2015 about the most promising technology of 2008?

Disclaimer

This work does not intend to cover all the nuances of the technologies and solutions reported. I have been actively working with Zmail for almost four years, first as an investor and then as a frequent user. In those capacities I have closely watched its step by step development and the gradual increase in convenience and ease of use. As an unpaid consultant, Advisory Board Member and also, later, as an User Advocate, I noted these improvements and continued to invest substantially in it. It is my opinion that Zmail will become a major means of our written future communication with colleagues, patients, attorneys and financial institutions. Nevertheless, I have tried to convey each pro and con of all technologies described in this work in the most neutral manner possible.

Acknowledgments:

I wish to acknowledge the mentoring over the past four years that I've received in the area of secure email from Dr. Ed Gerck. I also want to sincerely thank Dr Gerck and ISPE-er, Dr Andrew Mackie, along with several other reviewers, who have provided comments and evaluations for this peer reviewed article.

References

1. Grow, B, Epstein, K, Tschang, C-C. The New E-spionage Threat: A Business Week probe of rising attacks on America's most sensitive computer networks uncovers startling security gaps. (Also) http://www.businessweek.com/magazine/content/08_16/b4080032218430.htm *Business Week*, 2008, April 10.,
2. Michelle. The new Espionage threat: comments. http://app.businessweek.com/UserComments/combo_review?action=getComment&productId=29875&reviewId=258468#258468. 2008, 29 April.
3. Gerck, E. The new Espionage threat: comments. http://app.businessweek.com/UserComments/combo_review?action=getComment&productId=29875&reviewId=258468#258468. 2008, 29 April.
4. Kirk, J. Security researcher intercepts embassy passwords: He says systems administrators failed to use encryption. http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9035238&source=NLT_SEC&nlid=38, in *Computer World* (Security(2007, 10 September.,
5. (CSI), CSI. 2003 CSI/FBI Computer Crime & Security Survey. http://www.sunbelt-software.com/documents/csi_fbi_survey.pdf. 2003.

6. (CSI), CSI. 2005 CSI/FBI Computer Crime & Security Survey.www.asiaing.com/csi-fbi-computer-crime-and-security-survey-2005.html. 2005.
7. (CSI), CSI. 2006 CSI/FBI Computer Crime & Security Survey. www.gocsi.com/press/20060712.jhtml. 2006.
8. (CSI), CSI. 2007 CSI/FBI Computer Crime & Security Survey. http://www.alliedacademies.org/Publications/Papers/EE%20Vol%2012%202007%20p%2023-36.pdf. 2007.
9. Gerck, E. Secure email technologies X.509/PKI, PGP, IBE and Zmail, in *Corporate email management*, Chapter 12 Edited by Krishna SJ, Raju E. Hyderabad, India, ICFAI University Press, 2007, 171-196.
10. Gerck, E. Trust points (cited section), in *Digital Certificates: Applied Internet Security* authored by Feghhi J, Feghhi J, Williams P. New York, Addison-Wesley, 1998, 194-195.
11. Gerck, E. (moderator) Blog discussion from 2003-2006 regarding secure email technologies X.509 / PKI, PGP, and IBE. http://email-security.blogspot.com . 2006, 31 December.
12. Gerck, E. Comparison Of Secure Email Technologies X.509 / PKI, PGP, and IBE. http://email-security.net/papers/pki-pgp-ibe.htm/. 2005, 31 December.
13. NMA. Secure EMail Zmail™. http://www.nma.com/papers/zmail.pdf. 2008, 25 March.
14. ZSentry.com. Compliance Statement http://zsentry.com/privacy_security_compliance_zmail.htm. 2008, 1st May.
15. Andress, M. Improving ROI for PKI. *InfoWorld*. http://security.itworld.com/4360/IWD010625pki/page_1.html. 2001, 25 June.

Any time something is written against me, I not only share the sentiment but feel I could do the job far better myself. Perhaps I should advise would-be enemies to send me their grievances beforehand, with full assurance that they will receive my every aid and support. I have even secretly longed to write, under a pen name, a merciless tirade against myself.

~Jorge Luis Borges