Spearphishing Defense Through Improved Usability

Robert Zager and John Zager

Abstract

The information environment consists of three interrelated dimensions which are the physical, the informational and the cognitive. The most important of these is the cognitive dimension which encompasses the minds of people who interact with information and make decisions using information. Employing cyberspace operations, attackers use spearphishing to engage in espionage, sabotage and other intelligence-related activities. We discuss the use of an enhanced user interface as a means of counterintelligence to disrupt spearphishing attacks. We describe how improved usability can help the targets of spearphishing act as a sensor network which can provide information superiority over the spearphishing attackers.

Keywords: intelligence, cyberspace operations, counterintelligence, information operations, operational security, military deception, phishing, spearphishing, APT, computer network defense, human sensors, cyber counterattack, deception countermeasures, usability, Compliance Budget

The Challenge of Cyberspace

We have known for a long time that there are significant vulnerabilities, and that these vulnerabilities are going to accelerate as time goes by, both in systems within government and within the private sector.

- President Barack Obama[1]

On 05 February 2015, Anthem, a Blue Cross Blue Shield insurance company, reported that "hackers" had compromised information regarding about 80 million people in a "very sophisticated" cyberattack.[2] The report of this attack was followed in short order by similar reports from health insurance companies Premera, CareFirst, and EmpireBlue.[3] On 04 June 2015, the Office of Personnel Management reported the compromise of records regarding about four million current and former federal employees.[4] Within hours of the President's remarks quoted above, the U.S. Army reported that its official website, army.mil, had been compromised.

Hackers engaging in very sophisticated cyberattacks raise the specter of adversaries commanding technologies far more advanced than our own. What tools are being used to compromise millions of insurance records, the defense secrets of the United States, major news organizations and the President's own calendar?[5] The answer is targeted deceptive email, otherwise known as spearphishing.

It is estimated that 91% of Advanced Persistent Threat ("APT") infiltration involves spearphishing and of that 95% uses email attachments.[6] Reports of state affiliated APT often find that 100% of the attacks use spearphishing.[7] This paper discusses the application of the principles of intelligence and information operations to defend against spearphishing.

What is Cyberspace?

Cyberspace consists of many different and often overlapping networks, as well as the nodes (any device or logical location with an internet protocol [IP] address or other analogous identifier) on those networks, and the system data (such as routing tables) that support them.

- Joint Publication 3-12(R), Cyberspace Operations[8]

Joint Publication 3-12(R), *Cyberspace Operations*, sets forth a three-tiered architecture to describe Cyberspace:

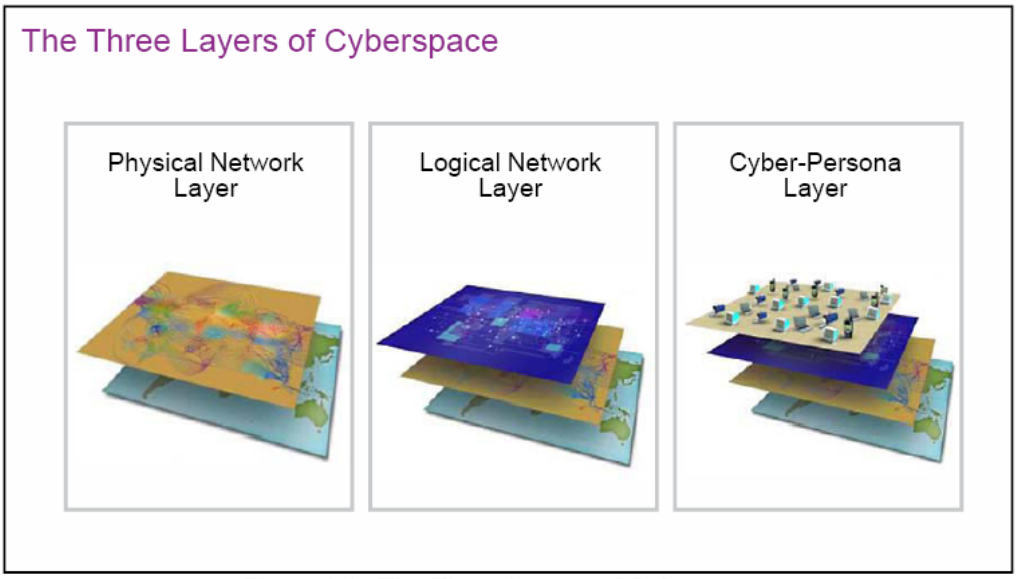


Figure I-1. The Three Layers of Cyberspace

Illustration 1

The physical layer consists of the physical components of the network that are geographically fixed. The logical network layer consists of those elements of the network that are related to one another but are not geographically fixed, for example, a website can be allocated over several physical servers, yet accessed by a single URL. The cyber-persona layer consists of the entities that are using the physical network layer and the logical network layer. The logical network layer contains digital representations of the cyber-persona layer. Rights of entities in the cyber-persona layer to access and use the physical network layer and the logical network layer are controlled by these digital representations. These rights are commonly called “credentials.”

What are attackers doing? Attackers are using cyberspace operations to seize control of computers in order to instruct the computers to perform tasks for the attacker. Malware, bugs and misconfigurations are merely methods used to seize control, substituting the attacker for the authorized user. For example, if equipment is configured to use the default factory passwords or weak passwords, the attacker can use this configuration error to gain control of the equipment. Likewise, attackers can gain control of systems by exploiting programming errors such as stack overloads, cross-site scripting errors or SQL injection.

Attackers frequently exploit a non-technical attack surface. In the case of APT, and particularly cyber-espionage, the attack surface is not the data processing system, but the system controls that are external to the computers - the user. In these cases, the attackers use social engineering

techniques to convert the authorized user into an Unintentional Insider Threat ("UIT"). In research funded by the Department of Homeland Security, CERT Division of Carnegie Mellon University provides this definition of UIT:

> An unintentional insider threat is (1) a current or former employee, contractor, or business partner (2) who has or had authorized access to an organization's network, system, or data and who, (3) through action or inaction without malicious intent, (4) unwittingly causes harm or substantially increases the probability of future serious harm to the confidentiality, integrity, or availability of the organization's resources or assets, including information, information systems, or financial systems.[9]

This research replaces the previous term, UIT-HACK, with two new terms.

> PHISHING/SOCIAL, Malware:
>
> *An outsider's electronic entry acquired through social engineering (e.g., phishing email incident, planted or unauthorized USB drive) and carried out via software, such as malware and spyware.*[10]
>
> PHISHING/SOCIAL, Credentials:
>
> *An outsider's electronic entry acquired through social engineering (e.g., phishing email incident) and carried out through compromised credentials, including passwords and other identifying information.*[11]

As CERT discusses, these two definitions are important because they identify a common technique, social engineering, and two different implantations, malware and credentials. The similarities and differences help identify appropriate defense and mitigation strategies.

Information Operations

The nation's state and non-state adversaries are ... aware of the significance of [the information environment], and will use information-related capabilities (IRCs) to gain advantages in the information environment, just as they would use more traditional military technologies to gain advantages in other operational environments.

- *JP 3-13 Information Operations*[12]

UIT adds two important concepts, information and people, to the definition of Cyberspace. With the addition of information and people to the operational environment, a more expansive view of the cyber domain is revealed. Rather than a universe comprised of data processing systems, Cyberspace is seen as a subset of the physical dimension of the Information Environment. Similarly, cyberspace operations (CO) is seen as a subset of information operations.[13]

JP 3-13 *Information Operations* states the information environment is made up of individuals, organizations, and systems that collect, process, disseminate, or act on information. This environment is comprised of three interrelated dimensions which continuously interact with individuals, organizations, and systems. These three dimensions, the physical, informational, and

cognitive, are illustrated in this figure from JP3-13:[14]

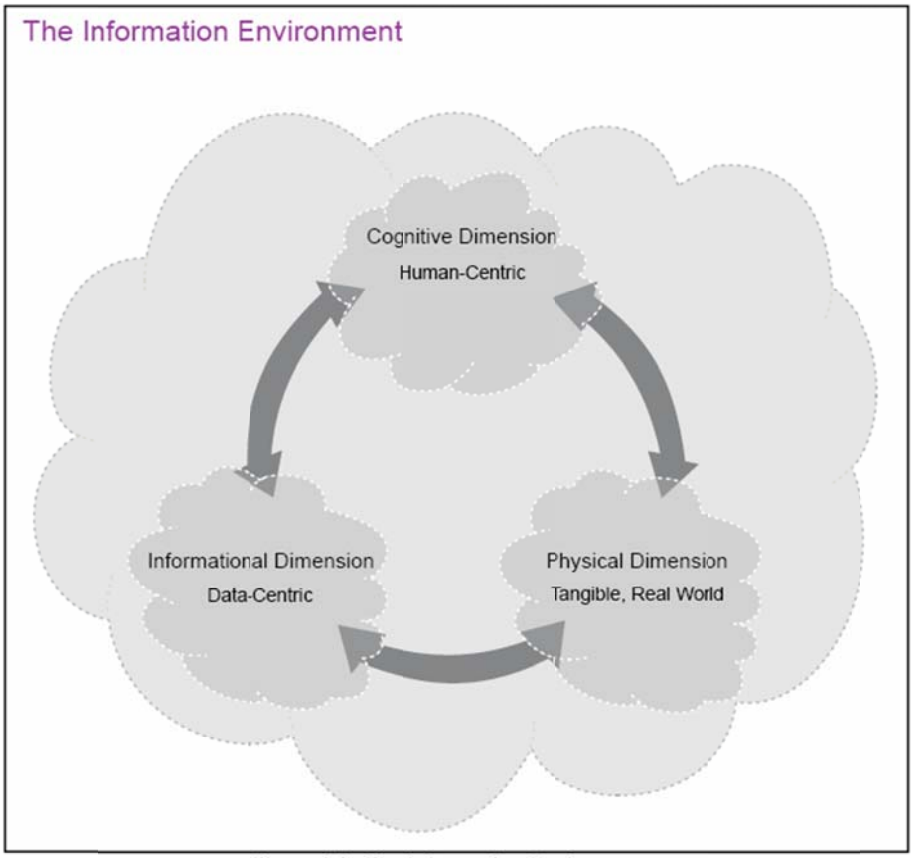


Figure I-1. The Information Environment

Illustration 2

The physical dimension consists of command and control (C2) systems, decision makers, and supporting infrastructure that enable individuals and organizations to create effects. The physical dimension contains the physical platforms and the communications networks that connect them. This dimension includes, but is not limited to, human beings, C2 facilities, newspapers, books, computer processing units, laptops, smart phones, tablet computers, or any other objects that are subject to empirical measurement. The physical dimension is a worldwide network connected across national, economic, and geographical boundaries. Cyberspace is part of the physical dimension.

The informational dimension consists of where and how information is collected, processed, stored, disseminated, and protected. This is the dimension in which the C2 is exercised and where an actor's intent is conveyed. Actions in this dimension operate on the content and flow of information.

The cognitive dimension encompasses the minds of the people who transmit, receive, and respond to or act on information. It refers to individuals' or groups' information processing, perception, judgment, and decision making. Human decision-making is not a computer process; it is an incredibly complex process of human psychology.[15] Factors influencing decisions include the environment in which the decisions are being made and incorporate individual and cultural beliefs, norms, vulnerabilities, motivations, emotions, experiences, morals, education, mental health, identities and ideologies. Decisions can be rational or irrational. Understanding these factors creates opportunities to influence the decision-making to create the desired effects. "As

such, **[the cognitive] dimension constitutes the most important component of the information environment**. [emphasis added]"[16]

JP 3-13 observes, "Information is a powerful tool to influence, disrupt, corrupt, or usurp an adversary's ability to make and share decisions."[17] The tools of information operations are information related capabilities, which include any means of communicating information to the target audience, including, but not limited to, public affairs, cyberspace operations, civil-military affairs and military deception.[18] The purpose of information operations is to apply information related capabilities to affect the ability of the target audience to collect, process, or disseminate information before and after decisions are made in order to achieve the desired ends.[19] Effective information operations can undermine the adversary's intelligence.

Intelligence

Ultimately, intelligence has two critical features that distinguish it from information. Intelligence allows anticipation or prediction of future situations and circumstances, and it informs decisions by illuminating the differences in available courses of action (COAs).

- *JP 2-0 Intelligence*[20]

The purpose of intelligence is to give decision-makers an understanding of the operational environment. The most useful information provides insight into the future of the operational environment. When information is related to other information and evaluated in the context of past experience, a more complete understanding of the information emerges – that understanding is intelligence.

JP 2-0 sets forth ten principles of Joint Intelligence which are summarized in this illustration:[21]

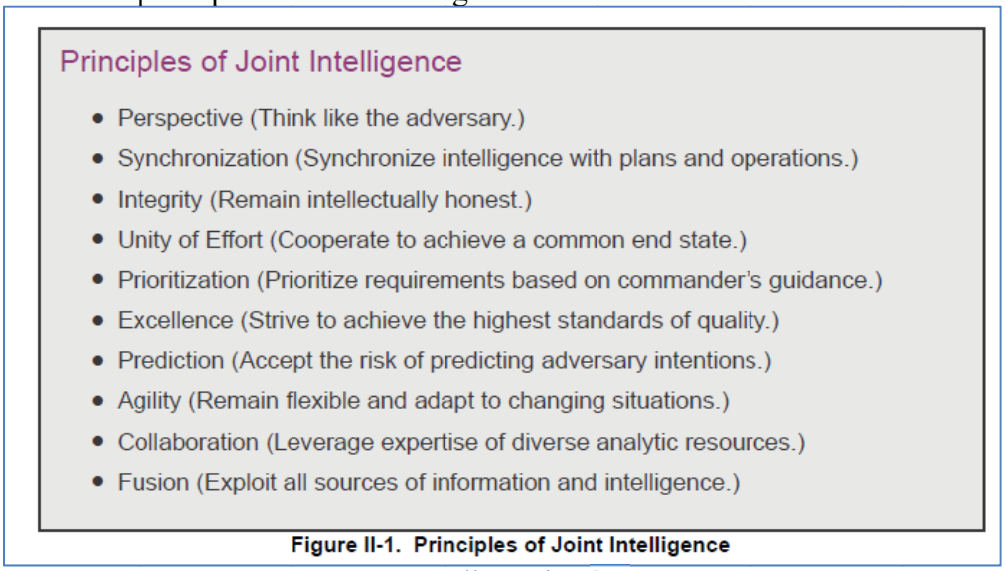


Figure II-1. Principles of Joint Intelligence

Illustration 3

While all of the Principles of Joint Intelligence are crucially important, it is worth noting the

eight attributes of Excellence.[22] These attributes are:

a) Anticipatory. Intelligence must anticipate the needs of the user in order to support planning and decision-making. The intelligence should provide understanding of the current situation including the courses of action (COAs) available to the friends and foes.
b) Timely. The intelligence must be provided in time to be useful. Timely intelligence avoids surprises.
c) Accurate. The intelligence must be factually accurate to the extent possible. The intelligence must be based on information which is reliable, with more emphasis placed on reliable sources. The intelligence must consider the possibility that the source information is the result of the adversary's deception efforts.
d) Usable. The intelligence must be crafted to meet the user's needs and should be in a form that is readily applicable to the user's tasks.
e) Complete. The intelligence should identify factors that impact the mission, courses of action open to the adversary, and gaps in the intelligence.
f) Relevant. The intelligence must be relevant to the planning and execution of the mission.
g) Objective. The intelligence should be free of bias to favor the user's preferences.
h) Available. Not only must intelligence by timely and accessible to the decision-maker, it must be classified at the appropriate level to facilitate any required sharing.

Intelligence influences the chosen course of action. Adversaries are thinking and adaptive opponents. Upon observing the chosen course of action, that course of action will influence the actions of the thinking adversary. For example, if intelligence determines that the enemy will attack at Point A and Point A is reinforced, the enemy will observe the reinforcement and will modify its attack in response to the reinforcements. This establishes a process of action and reaction known as the Paradox of Warning, which is illustrated by this diagram:[23]

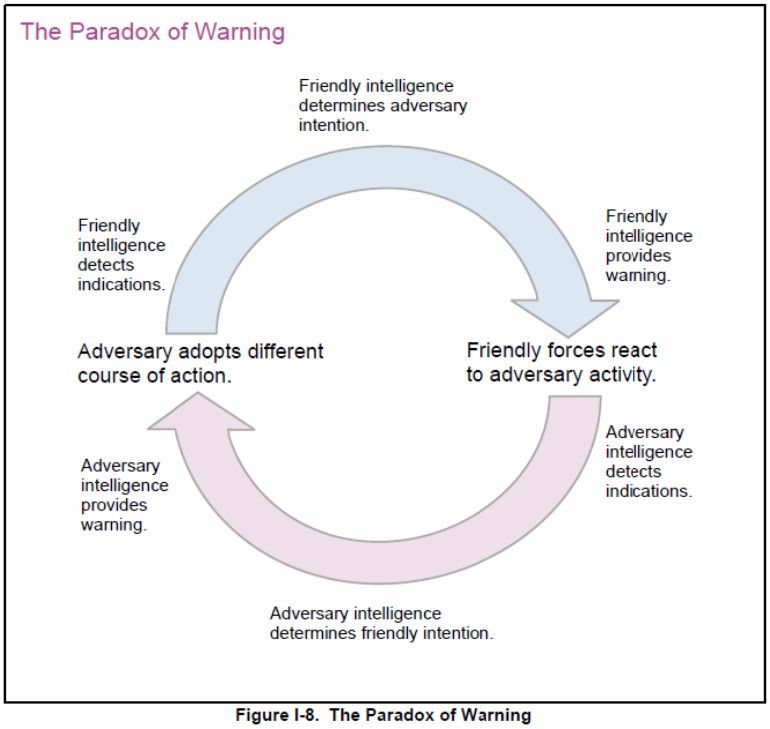


Figure I-8. The Paradox of Warning

Illustration 4

The Paradox of Warning describes an important feedback loop in which a chosen course of

action may be negated by enemy adaptation. Thus, JP 2-0 observes:

> **To prevail, [our] decision and execution cycles must be consistently faster than the adversary's and be based on better information.**[24]

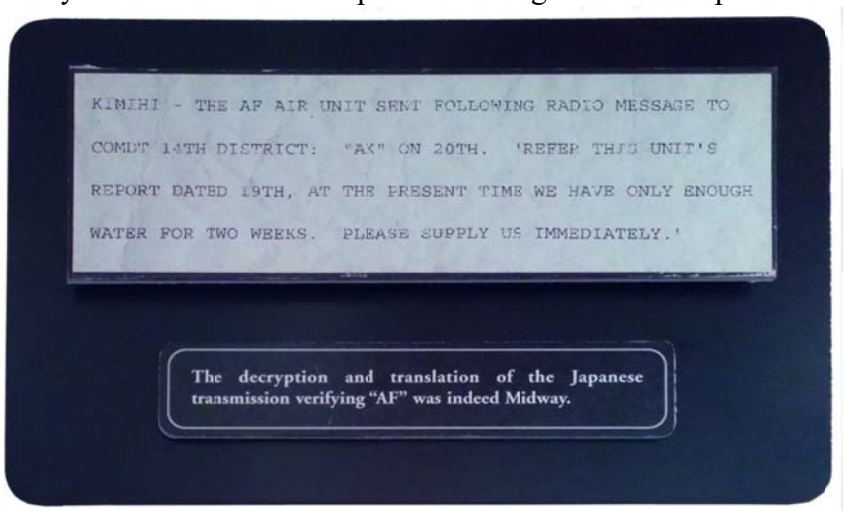
Information Operations and Intelligence Integration (IOII)

information superiority. *The operational advantage derived from the ability to collect, process, and disseminate an uninterrupted flow of information while exploiting or denying an adversary's ability to do the same.*

- *JP 2-0 Information Operations*[25]

Intelligence can provide vital support information operations. Intelligence can provide insights into the information environment and the relationships between the physical, informational, and cognitive dimensions. Intelligence can provide information about the psychological factors of the cognitive space that drive the target audience's decision-making processes. Intelligence can characterize the physical dimension and the information that flows over it. Information operations can disrupt the target audience's intelligence processes. By integrating intelligence and information operations, information superiority can be obtained. As the following example demonstrates, information superiority can have decisive results.

During World War II, the Axis Powers believed that their communications encoded by the German Enigma machine and its Japanese counterpart, Purple, were unreadable. This meant that intercepted encoded information was highly reliable. In 1942, intercepted Japanese naval communications indicated that an attack was being planned for the Central Pacific. However, the target of the attack was referred to only as "AF." Despite observing increased Japanese reconnaissance of Hawaii, Admiral Nimitz believed that AF was Midway. In order to confirm that AF was Midway, an information operation of military deception was devised. Midway reported by its wireless that its fresh water plant had failed and it had only two weeks of water remaining. A few days later this encoded Japanese message was intercepted:



Source: U.S. National Cryptologic Museum
Illustration 5

On 27 May 1942, the main Japanese fleet sailed toward Midway while a diversionary force attacked the Aleutian Islands. On 06 June 1942, the Japanese attack fleet anticipated a repeat of

their successful surprise attack on Pearl Harbor six months earlier. Instead, the Japanese were met by the U.S. fleet ready for battle. During the Battle of Midway, the Japanese Navy lost four aircraft carriers and about 300 of its finest pilots.

Spearphishing –Military Deception Exploits "Compliance Budget"

Because APTs are not traditional threats, they cannot be treated as traditional threats, and the traditional security mechanisms are ineffective in detecting and defending against them. For example, with all of the known APTs that have been uncovered, none has ever triggered an [intrusion detection] system.

- Li & Clark, *Security Intelligence*[26]

A successful spearphishing attack requires the attacker to complete a series of steps known as a "kill chain." The kill chain integrates Intelligence and Information Operations using Cyberspace Operations to accomplish its mission. This kill chain from CERT's *Unintentional Insider Threats: A Review of Phishing and Malware Incidents* is representative of an APT spearphishing kill chain:[27]

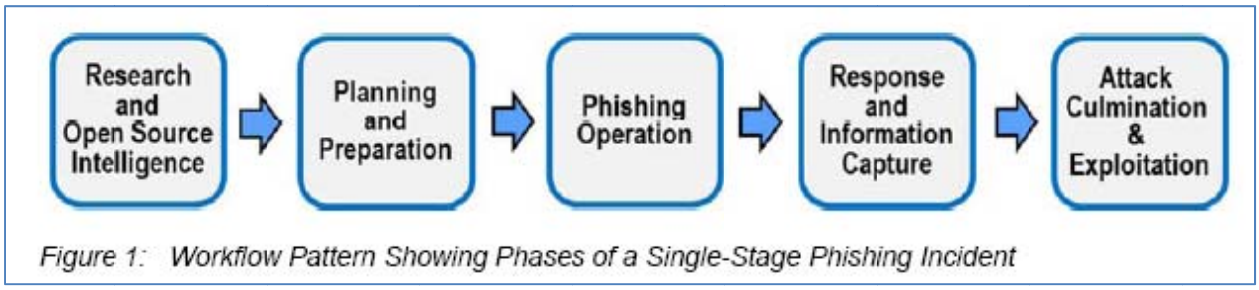


Figure 1: Workflow Pattern Showing Phases of a Single-Stage Phishing Incident

Illustration 6

A single stage phishing incident is the simplest case and is the basis for multi-stage incident.[28] In the Research and Open Source Intelligence phase, the attacker researches its targets. Based on the results of Research and Open Source Intelligence phase, the attacker plans the attack and prepares the attack tools, such as attack domains, email text, malware, fake websites, fake documents, and data collection tools. The Phishing Operation is the transmission of the attack emails. The Response and Information Capture phase occurs when targeted individuals fall victim to the attack email by taking its call to action. In this phase, user credentials are captured. In the final phase, Attack Culmination & Exploitation, the attacker uses the seized credentials to execute system commands in furtherance of the attack objectives.

The 2015 compromise of the White House demonstrates a multi-stage incident in which the breach of one entity can be leveraged to attack other entities.[29] After compromising the State Department using a spearphishing attack, the attackers used real State Department email accounts to send spearphishing emails to the White House. Although the email account was real, the actual author of the attack email was not the authorized State Department employee.

In both attack methods (PHISHING/SOCIAL, Malware and PHISHING/SOCIAL, Credentials), the attacker uses intelligence to create the messaging to be delivered to the target audience for the purpose of inducing the target audience to take the call to action. This social engineering

aspect is always the information-related capability (IRC) military deception.[30] This diagram from JP 3-13.4, *Military Deception,* summarizes the three-tiered cognitive process of deception which culminates in the target audience taking the desired action:[31]

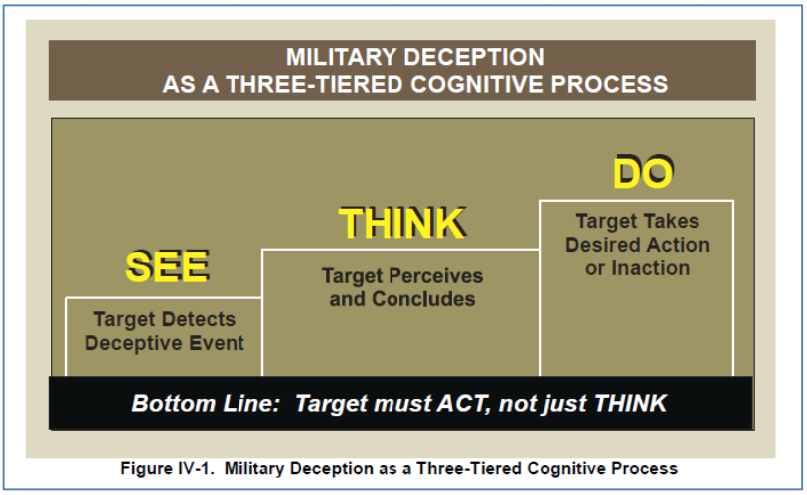


Illustration 7

The three-tiered cognitive process is demonstrated by this spearphishing email that compromised the Israeli security forces:[32]

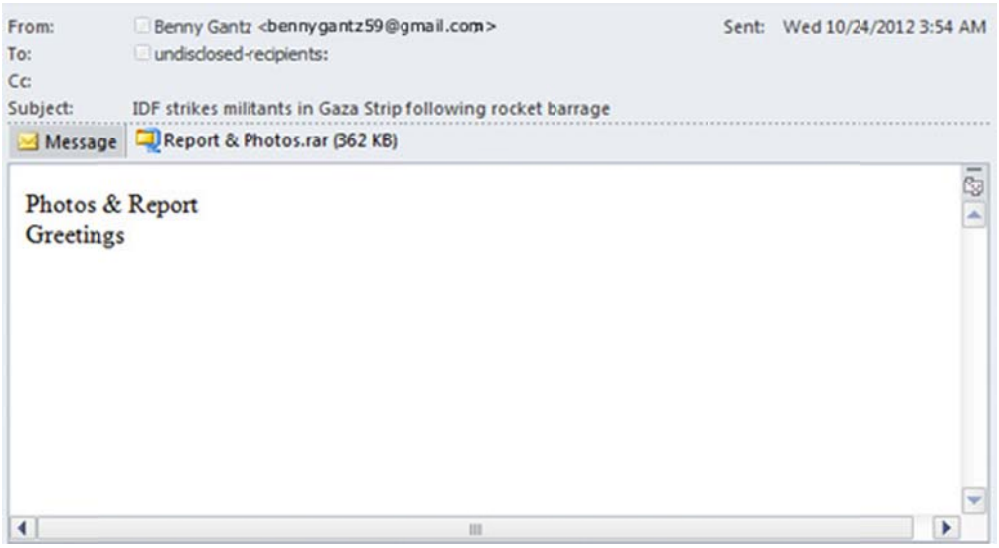


Illustration 8

The target audience (an Israeli police officer) saw an email from the Lt. Gen. Benny Gantz, trusted head of Israeli defense forces, delivering a report about enemy rocket attacks ***(See)***. The police officer drew the conclusion that he should open the attachment ***(Think)*** and took the action of opening the attachment ***(Do)***. The consistent constellation of content comprising the perceived trusted sender, the text and the call to action caused the target audience to conclude this email was real, despite the fact that the purported sender, an Israeli general, probably does not send critical security email to police from a Gmail account.

Spearphishing is a particularly pernicious form of military deception. After thousands of email interactions, email users shift from a decision-making framework of cognitive judgment to a decision-making framework of intuitive judgment.[33] Instead of a deliberative cognitive process, intuitive email decision-making is predominated by habitual responses to simple urgency clues

and perceived relevance.[34] The intuitive decision-making process implies that in the case of spearphishing, the "See, Think, Do" cognitive process is more accurately described as "See, Do."

Where PHISHING/SOCIAL, Malware and PHISHING/SOCIAL, Credentials diverge is that the Malware form uses malware to accomplish the theft of credentials while the Credentials form does not use malware. In both cases, the Paradox of Warning is at work as the defenders modify their defenses in response to the attackers and the attackers modify their attacks in response to the defenders. This process is summarized in this diagram from RAND's *The Defender's Dilemma: Charting a Course Toward Cybersecurity*:[35]

Diagram Depicting the Measure-Countermeasure Dance Between Defender and Attacker

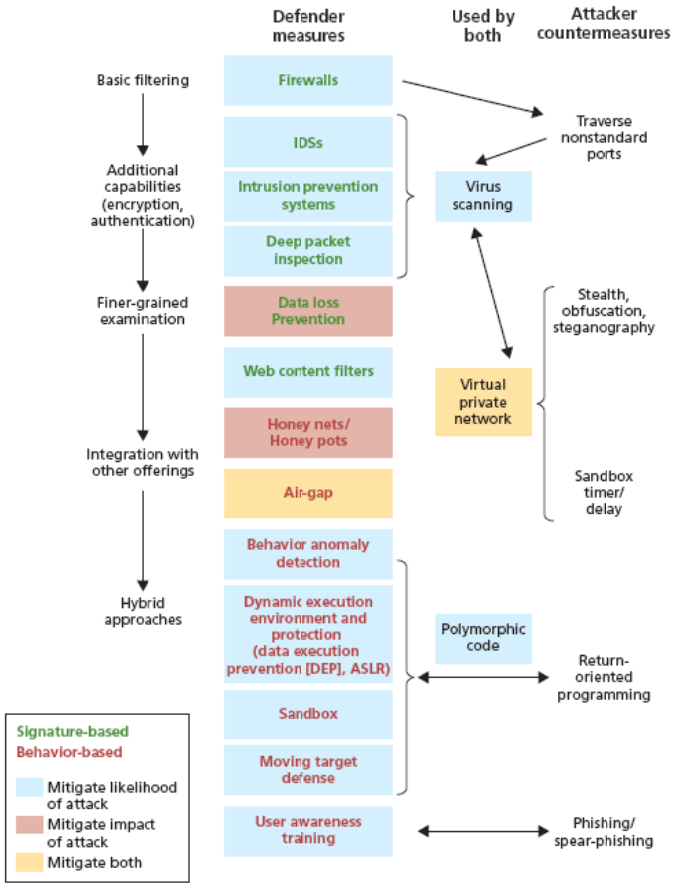


Illustration 9

Because APT evades intrusion detection systems, defenders are forced to use retrospective analysis to find the intrusions.[36] The adaptability of human attackers helps explain why infiltration is so successful and the median time on target is 205 days. [37] In 63% of APT events, the attack is found by third parties, such as law enforcement and financial institutions, evading the retrospective analysis by the victims.[38]

The 2014 compromise of Sony is a concrete example of the use of intelligence and information operations by attackers. Cylance Senior Security Researcher Brian Wallace and Cylance CEO Stuart McClure observe that Sony was compromised using a phishing attack against Apple IDs.[39] Wallace and McClure report that the spearphishing email looked like an Apple ID verification email which contained a link to a fake Apple website. When victims entered their Apple

credentials into the fake website, an error message was displayed and their credentials were recorded. The attackers correlated the victims' LinkedIn profiles with their Apple credentials. The attackers were also able to determine the Sony user names of the victims. In cases where the Sony employee reused his Apple ID password as his Sony password, the attackers were able to access Sony's systems. Once inside Sony, the hackers were able to distribute their own malware throughout Sony.

With stolen credentials in hand, the attackers continue to apply the information related capacity military deception in order to avoid detection. In this phase of the APT, the target audience shifts from the email user to the cybersecurity analysts. For the attacker, intelligence is crucial in the effective post-infiltration use of military deception. For example, in the Neiman Marcus breach, the attackers demonstrated a thorough understanding of the operational environment in which the victim's network analysts worked.[40] The custom malware that was created for this attack mimicked a normal function which triggered false alarms. When the cyber intelligence analysts saw 60,000 alarms over the course of less than four months, these alarms were ignored because the analysts mistook the malware alerts for normal false alarm traffic. The attackers employed four of the ten Maxims of Deception:[41]

1. "Magruder's Principal" which states that it is easier to exploit a preconceived notion than deceive the target into new beliefs. In this case, the notion being exploited was that alerts that appeared to be generated by the proprietary payment system were false alarms.
2. "Limitations to Human Information Processing" which consists of two concepts. The first is the "law of small numbers" which suggests that conclusions should not be made on small samples that are not statistically significant. The second is the inability of targets to detect small changes in normal indicators – even if the small changes result in a large cumulative change over time. In this case, although the analysts observed 60,000 alerts, these alerts were less than 1% of the background alert traffic.
3. "Choice of Types of Deception" in which the attacker decides if the deception should increase the target's confidence or induce confusion. In this case, by creating alerts that looked innocuous, the attackers sought to increase the victim's confidence in a wrong conclusion.
4. "Importance of Feedback" states that an intelligence, surveillance or reconnaissance plan should be adopted to determine if the deception is effective. In this case, the feedback loop consisted of the steady flow of customer credit card data for over three months.

Mandiant is currently reporting that 100% of breaches involved stolen credentials.[42] One must look at compromises lasting over 200 days and the loss of tens of millions of records and ask how current cybersecurity intelligence products stack up against the Excellence standard for intelligence.[43] Clearly, the attackers hold information superiority. It is little wonder that Chief Information Security Officers feel that attackers have the upper hand, and will continue to have it.[44]

Spearphishers Leverage the Compliance Budget

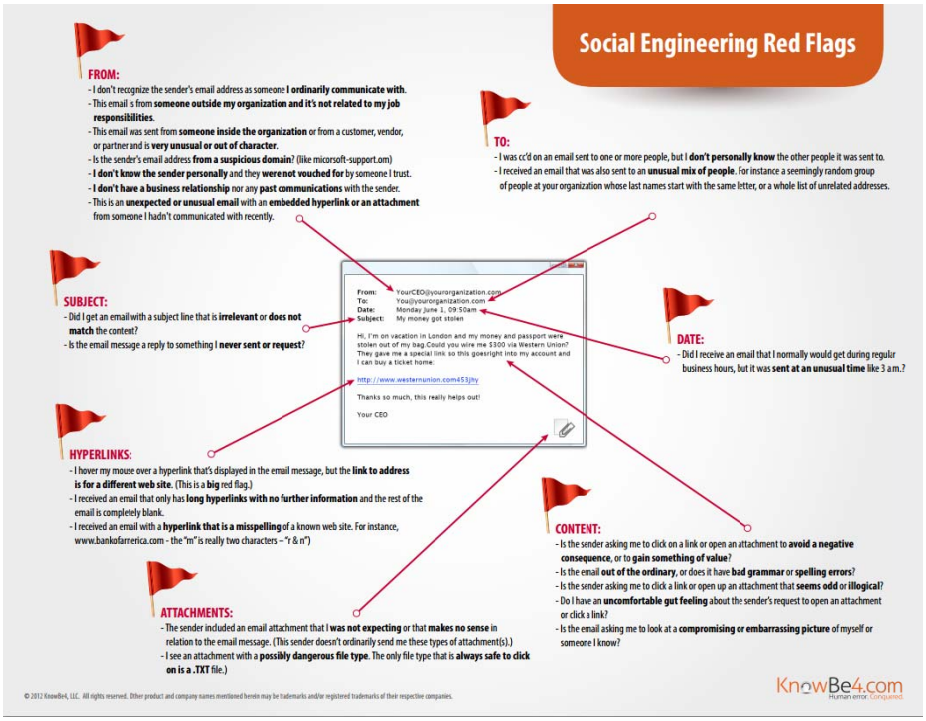
Employees focus on completing their primary (production) tasks, and the behaviour required by the security (enabling) tasks often presents an obstacle on the shortest path to primary goal. This misalignment introduces friction between security and business processes into the organizational system, and it is this friction that is at the heart of individual compliance issues.

- Adam Beautement. et al., *The Compliance Budget*[45]

The crucial element in the UIT attack is military deception based on tricking the target audience, the email recipient, into believing that the attacker is a trusted email sender. In every case, the attacker is trying to be something it is not. The attacker can masquerade as anyone that the attacker concludes is trusted by the target audience – be that a co-worker, the IRS, FedEx, or any other sender. That trust relationship is perceived by the human target audience in the cognitive dimension, not by computers.[46]

The solution proposed to address military deception in email is training.[47] Training is mandated by FISMA and is pervasive in the NIST guidance issued pursuant to FISMA.[48] Standing in opposition to the law and the common sense appeal of training is the absence of empirical evidence that training is effective against military deception when practiced in email.[49] Study after study finds that training is ineffective.[50]

Why does training fail to solve the spearphishing problem? In every spearphishing email there is a divergence between the actual sender and the perceived sender, a divergence which we term the Spearphishing Incongruity. It is straightforward to provide users with a set of rules that can be used to spot evidence of the Spearphishing Incongruity. This is a typical set of rules, in this case published by KnowBe4, a leading spearphishing training company:[51]



Source: KnowBe4

Illustration 10

By deconstructing every email using these twenty-two things to remember and applying these rules against a list of trusted URLs and file extensions, a user has a methodology to discover the Spearphishing Incongruity, thereby substantially reducing the likelihood of being deceived by a socially engineered email. These rules are effective in detecting a single phase spearphishing attack, thereby preventing a multistage attack. Of course few users (even technically advanced users) follow complex email processing rules. In fact, most users do not have access to a list of trusted URL's and even if they did, most people cannot decode URLs.[52] Instead users apply a simple triad of perceived relevance, urgency clues and habit to process email.[53] These three factors are easily manipulated by spearphishers applying the principles of military deception.

Why do users adopt simple rules of thumb over valid email security practices? This user behavior is explained by the paradigm of the Compliance Budget, a model of human security behavior developed by Beautement, Sasse, and Wonham.[54] In spearphishing attacks, the choices made by users, not technical defects, cause security compromises. Users are choosing to interact with the malicious emails. Users make these security choices in the context of their jobs, a context in which they seek to optimize employment performance, not security. This creates a conflict of interest in which the users focus on completing their primary (production) tasks and the requirements of security (enabling) tasks can present impediments to achieving the primary tasks. There is a limit to how much effort users will put into non-productive security tasks, which they named the Compliance Budget. They summarize the Compliance Budget:

> The decision to comply or not comply can be addressed by the following simple formula:
>
> Task cost = individual benefits – individual costs.
> Compliance if task cost < compliance threshold.

They provide examples of users' perceived costs and benefits:

> Costs:
>
> 1. Increased physical load. These are the extra steps and delays introduced into the production task by the security measures.
> 2. Increased cognitive load. This is the extra information that needs to be stored and recalled on demand.
> 3. Embarrassment. Embarrassment occurs when the compliance measures interfere with a production task causing the user embarrassment. For example, if a user would be embarrassed if the user was unable to complete a presentation due to security uncertainties about an email attachment.
> 4. Missed opportunities. These are business opportunities that are foregone because security measures make production tasks more burdensome or impossible. For example, by adopting a "better safe than sorry" approach, an important email may be ignored because is easier to skip the email than go through all the security steps required to safely process the email.
> 5. The "Hassle Factor." This is the perceived cost of compliance, which may be greater than the actual cost of compliance. For example, undertaking the same task repeatedly; frequent repetition increases the perceived costs or waiting for a patch to install when facing a deadline.

Benefits:

1. Avoiding the consequences of a security breach. A security breach can impact the individual, such as losing data.
2. Protection from sanctions. The danger of being caught and punished for breaching security policies.

Applying these cost/benefit factors to email it is obvious that the costs of applying twenty-two criteria across hundreds of email decisions per day triggers all five of the costs.[55]

Spearphishing awareness training is an example of human failure caused by security mechanisms that are too difficult for ordinary users to implement.[56] When spearphishing training is analyzed using the Compliance Budget framework, it becomes clear that training merely reinforces users' perceptions of the cost of elaborate email processing rules. Users discount the benefits of compliance, placing a false sense of security on security systems[57] and perceive little personal risk of sanctions for non-compliance.[58] In the operational environment of completing production tasks, users must make rapid assessments of large numbers of emails. Spearphishing training such as this evidences an organizational attitude that is unaware of, or ignores, the impact of security mechanisms on users and operates on the assumption that users can and will perform every task associated with security compliance.[59]

In reality, the priority to accomplish production tasks – processing email – overwhelms the security tasks that are required to protect the organization. As a result, when processing emails, users apply simple heuristics to get through their inbox rather than applying costly email analysis rules. As the fake Benny Gantz email (Illustration 8) demonstrates, this low effort approach to email processing creates the ideal opening for military deception. The target audience does not apply twenty-two red flags. When processing email, the Compliance Budget drives cursory email processing. The attackers craft deceptive messages that align with the realities of the email Compliance Budget, thereby inducing the target audience to act. Intelligence permits attackers to understand the workloads and cognitive dimension of the target audience. Attackers can use this intelligence to focus their attacks on users operating under the most Compliance Budget stress, making these users more likely to forego email security protocols.[60]

Beautement, et al. identify five external factors created by the production task that an organization can use to impact the Compliance Budget. These factors are:[61]

1. Design. The most direct way to influence the Compliance Budget is to design the production function so that compliance is aligned with the production task, thereby reducing the gap between the production task and compliance.
2. Awareness, Training, Education. Training can reduce the costs of compliance if it improves the users' ability to perform the compliance task. Increased awareness can increase the perceived benefits of compliance.
3. Culture. A culture of compliance induces users to "fit in" with their colleagues, reducing the perceived costs of compliance.
4. Monitoring. Compliance monitoring increases a user's perception that non-compliant behavior will be discovered, thereby increasing the perceived costs of non-compliance.
5. Sanctions. Avoiding sanctions is a perceived benefit. In order to be effective, sanctions must be enforced.

They observe that the most effective compliance strategy is to focus on design and improve the integration of production tasks and security, thereby decreasing the costs of compliance.

Improved Usability In Support of Counterintelligence

Usability is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"

- ISO 9241-11

What can be done to counter the UIT? When attackers challenge the Compliance Budget with military deception, the target audience overlooks the Spearphishing Incongruity. Applying the principles of Information Operations and Intelligence Integration (IOII) to understand the Compliance Budget and create appropriate military deception, attackers are able to infiltrate systems using UIT.

In a single stage UIT, the Spearphishing Incongruity is readily detectible. Referring back to the fake Benny Gantz email in Illustration 8:

The Spearphishing Incongruity

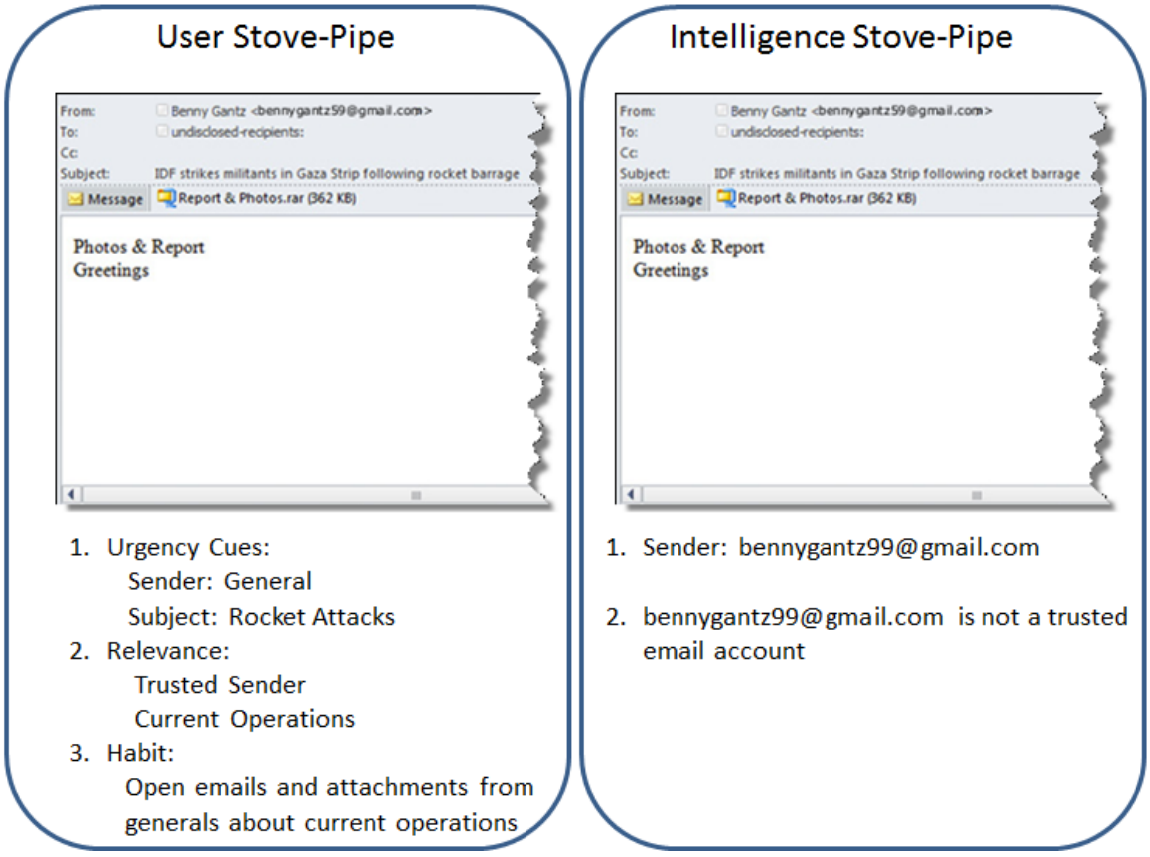


Illustration 11

While the details of the divergence between the two stove-pipes varies from case to case, in every single stage UIT, the User Stove-Pipe is inconsistent with the Intelligence Stove-Pipe.[62] When the two stove-pipes are viewed simultaneously, the observer can quickly see that the trust

which is assumed in the first stove-pipe is unwarranted – the email is suspicious. **However, the Spearphishing Incongruity is not fused**. In the User Stove-Pipe, the fusion of the stove-pipes fails due to the exploitation of the Compliance Budget by military deception. One can only imagine the embarrassment and dismay suffered by the victim of this attack if, after the attack, the victim's attention was focused on the sending email address. In the Intelligence Stove-Pipe, the cognitive dimension is not accessible. Intelligence does not know the psychological attributes (i.e., the available Compliance Budget together with the urgency clues, perceived relevance and habits) which comprise the cognitive dimension of the target audience at the moment of the attack. Intelligence lacks the user context required to operate in the User Stove-Pipe.

It is possible to seize the Spearphishing Incongruity and turn it against the attacker by fusing the two incongruous pieces of information at the time of the attack. This fusion is accomplished by providing the intelligence from the Intelligence Stove-Pipe to the target audience, the user, in real time.[63] When fused information is delivered such that the security task becomes supportive of, instead of in conflict with, the production task of processing email, usability is improved. When the security task is consistent with the production task, less Compliance Budget is consumed and users can more quickly, easily, and securely process email.

This improved usability can be accomplished by providing the targeted person with trust indications from Intelligence. These trust indications allow the intended victim to fuse the perceived sender (the User Stove-Pipe) with the actual sender (the Intelligence Stove-Pipe), allowing the user to make a better assessment of the risk of the email interaction. This simulated inbox demonstrates the fusion of the real sender and the perceived sender:

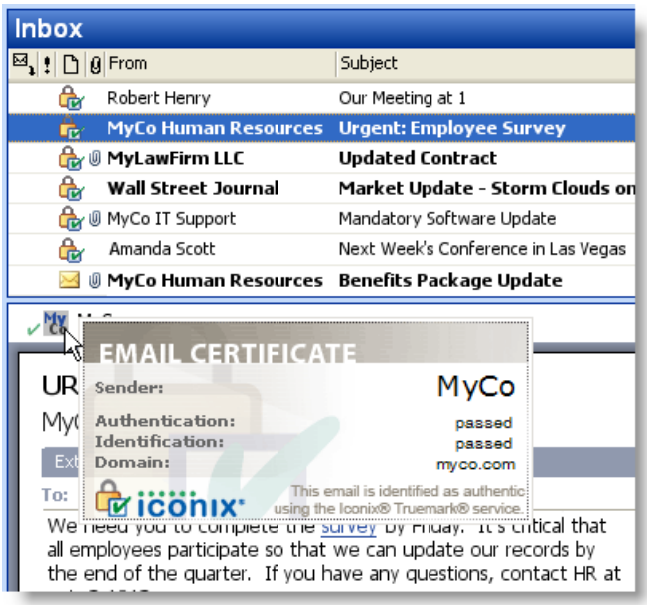


Illustration 12

In this example, all of the senders that Intelligence deems trustworthy have a trust indicator in the inbox. This trust indicator is reinforced in the message with additional iconography and a hover-over information window. The last message in the inbox is a spearphishing attack in which the attacker is attempting to be perceived as the HR Department. The display now integrates the knowledge of Intelligence (the attacker is not the trusted HR Department) with the perception of the victim (the sender is the trusted HR Department), thereby revealing the Spearphishing Incongruity and uncovering the attack. The security information is presented without a large

Compliance Budget task cost. Instead of a vague admonition to avoid suspicious emails, the user is now provided with actionable warning intelligence that permits the intended victim to identify and report the attack rather than being tricked into compromising the system. Rather than the kill chain advancing to the exploitation phase, the attack is interdicted during infiltration.

The enhanced user interface becomes an important tool of counterintelligence by identifying and disrupting the attempted infiltration.[64] Moreover, when the user provides early warning back to Intelligence, Intelligence can engage in operations against the attacker.[65]

The impact of this improved user interface can be seen by applying the eight attributes of Intelligence Excellence to the standard interface and the enhanced interface:

Intelligence Excellence
Comparison of Standard and Enhanced Email Interfaces

	Standard Email Interface	Enhanced Email Interface
Anticipatory	No	Yes
Timely	No	Yes
Accurate	No	Yes
Usable	No	Yes
Complete	No	Yes
Relevant	No	Yes
Objective	Yes	Yes
Available	No	Yes

Illustration 13

Similarly, applying the cost and benefit factors of the Compliance Budget shows the benefits of the improved usability of the enhanced email interface over the standard interface supplemented with training.

Compliance Budget Costs/Benefits
Comparison of Standard and Enhanced Email Interfaces

		Standard Email Interface with Training	Enhanced Email Interface
Costs			
	Increased Physical Load	Substantial non-task load	Display supports email task
	Increased Cognitive Load	Substantial non-task load	Display supports email task
	Embarrassment	"Chilling effect" of hard to discern risk interferes with email task	Display identifies low risk senders
	Missed Opportunities	"Chilling effect" of hard to discern risk interferes with email task	Display identifies low risk senders
	Hassle Factor	Highly repetitive, burdensome	Display supports email task
Benefits			
	Avoiding Breaches	Guessing game	Intelligence
	Sanctions Protection[66]	No credible sanctions	Present

Illustration 14

The enhanced user interface is resistant to the Paradox of Warning because it is built on the robust and widely deployed system of standards based email authentication. In order to

undermine the Intelligence Stove-Pipe, a single-stage spearphishing attack requires the compromise of the trusted sender's DNS records, a technical task far more challenging than sending a well-crafted work of military deception. A multi-stage attack requires a first stage compromise in order to take control of an email account for multi-stage operations. Of course, users are still able to make compromising decisions in the enhanced interface. However, incorporating security information into the usability factors will help users avoid the allure of military deception.

Systems are under attack from clever and determined opponents employing deceptive spearphishing emails as a means of infiltration. As has been demonstrated in numerous cases, the opponent is persistent – eventually the deception target will respond to a carefully crafted email and that response will initiate a series of damaging events. Deceptive email is so effective in eliciting the ***See, Think, Do*** deception process that Proofpoint estimates an attack of as few as ten emails is guaranteed to infiltrate the target.[67] The attacks are crafted to avoid technical defenses (small email volumes that are "under the radar", zero day exploits and other APT countermeasures).[68] The initial point of vulnerability is the person interacting with a compelling email.

Because of Compliance Budget constraints, training is not effective in defending at the point of vulnerability – military deception.[69] At that point of vulnerability, providing the user with an interface that provides a countermeasure[70] against the spearphishing attack decreases the likelihood of infiltration by email and increases the chances of interdicting the infiltration and reporting it to Intelligence. Additionally, by reducing the deceptive power of spearphishing emails, the attacker must send more emails, further increasing the likelihood of discovery. Fewer successful infiltrations will result in fewer alerts for IT to process. Finally, identifying the attack at its earliest stages permits the intended target to mount an effective counterattack.[71]

The enhanced user interface improves usability and supports information superiority, permitting the defender's decision and execution cycles to be consistently faster than the adversary's and be based on better information.

About the Authors

Robert Zager

Robert Zager is an inventor and entrepreneur. He has been granted eleven United States patents in the areas of computer networking and email. He holds a BA degree from the University of California, Berkeley and a JD degree from Santa Clara University. He is currently a security researcher at Iconix, Inc. in San Jose, California

John Zager

John Zager is a graduate student at Hofstra University, Hempstead, N.Y. where he is studying Industrial Organizational Psychology with an emphasis on decision-making. He holds a BA degree in psychology from Hofstra University.

The views expressed herein are the views of the authors and do not reflect the views of Iconix, Inc. or Hofstra University.

References and Notes

[1] Remarks by President Obama in Press Conference after G7 Summit, 08, June 2015, Web. 09 June 2015, https://www.whitehouse.gov/the-press-office/2015/06/08/remarks-president-obama-press-conference-after-g7-summit

[2] Anon., "Hackers Access Records for Millions of Anthem Customers," *The New York Times*, 05 February 2015, Web. 03 June 2015, http://www.nytimes.com/aponline/2015/02/04/us/ap-us-anthem-hack.html?_r=0

[3] Krebs, Brian, *Carefirst Blue Cross Breach Hits 1.1M*, 15 May 2015, Web. 01 June 2015, http://krebsonsecurity.com/2015/05/carefirst-blue-cross-breach-hits-1-1m/

Attackers commenced their attack using clever domain names that were different than, but to a human being, confusingly similar to the real domains.

Victim	Real Domain	Fake Domain	Deception
Anthem	wellpoint.com	we11point.com	Substituted numbers for letters
Premera	premara.com	prennara.com	Substituted "nn" for "m"
Empire Blue Cross	empireblue.com	empireb1ue.com	Substituted number for letter
Carefirst	carefirst.com	caref1rst.com	Substituted number for letter

[4] Dinan, Stephen, Congress ponders: OPM data breach could total 32 million Americans, *The Washington Times*, 24 June 2015,Web. 25 June 2015, http://www.washingtontimes.com/news/2015/jun/24/congress-ponders-opm-data-breach-could-total-32-mi/
As if to underscore the theme of this paper – that the adversaries of the United States are adept at applying Information Operations and Intelligence Integration (IOII) to infiltrate and then hide their subsequent activities, enabling them to compromise vast amounts of information.

[5] Perez, Evan and Prokupecz, Shimon, How the U.S. thinks Russians hacked the White House, *CNN*, 08 April 2015, Web. 09 April 2015, http://www.cnn.com/2015/04/07/politics/how-russians-hacked-the-wh/index.html

[6] TrendLabs APT Research Team, *Spear-Phishing Email: Most Favored APT Attack Bait*, 2013, Web. 06 February 2014, http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-spear-phishing-email-most-favored-apt-attack-bait.pdf

[7] Mandiant website, Web. 10 July 2014, https://www.mandiant.com/threat-landscape/

[8] President Obama (n 1).

[9] CERT Insider Threat Center. *Unintentional Insider Threats: A Review of Phishing and Malware Incidents by Economic Sector* (CMU/SEI-2014-TN-007). Software Engineering Institute, Carnegie Mellon University, 2014. P. ix, Web. 08 June 2015, http://resources.sei.cmu.edu/asset_files/TechnicalNote/2014_004_001_297777.pdf

[10] CERT (n 9), page ix.

[11] CERT, (n 9), page x.

[12] JP 3-13, *Information Operations*, page ix.

[13] **Cyberspace Operations**

(a) Cyberspace is a global domain within the information environment consisting of the interdependent network of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers. CO are the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace. Cyberspace capabilities, when in support of IO, deny or manipulate adversary or potential adversary decision making, through targeting an information medium (such as a wireless access point in the physical dimension), the message itself (an encrypted message in the information dimension), or a cyber-persona (an online identity that facilitates communication, decision making, and the influencing of audiences in the cognitive dimension). When employed in support of IO, CO generally focus on the integration of offensive and defensive capabilities exercised in and through cyberspace, in concert with other IRCs, and coordination across multiple lines of operation and lines of effort.

(b) As a process that integrates the employment of IRCs across multiple lines of effort and lines of operation to affect an adversary or potential adversary decision maker, IO can target either the medium (a component

within the physical dimension such as a microwave tower) or the message itself (e.g., an encrypted message in the informational dimension). CO is one of several IRCs available to the commander.
JP 3-13 Page II-9

[14] JP 3-13 Page I-1.
[15] Kahneman, Daniel and Tversky, Amos, *Choice, Values, Frames*, Cambridge University Press, 2000.
[16] JP 3-13, Page I-3.
[17] JP 3-13 Page I-1.
[18] JP 3-13, Page xi.
[19] See, JP 3-13, Chapter I
[20] JP 2-0, *Intelligence*, Page I-1.
[21] JP 2-0 Page II-1.
[22] JP 2-0, Pages II-6 – II-8.
[23] JP 2-0, Page I-29.
[24] JP 2-0, Page V-1.
[25] JP 3-13, Page GL-3.
[26] Li, Qing and Clark, *Gregory, Security Intelligence*, Wiley, 2015, p. 251.
[27] CERT (n 9) Page 4.
[28] CERT (n 9) Page 4.
[29] Perez, et al, (n 5).
[30] Zager, Robert and Zager, John, *Deploying Deception Countermeasures in Spearphishing Defense*, Science of Security, 10 March 2015, Web. 01 June 2015, http://cps-vo.org/node/18422
[31] JP 3-13.4, Page IV-2.
[32] Macalintal, Ivan, *Xtreme RAT Targets Israeli Government*, 29 October 2012, Web. 01 August 2014, http://blog.trendmicro.com/trendlabs-security-intelligence/xtreme-rat-targets-israeli-government/
[33] See, Gilovich, Thomas; Griffin, Dale; and Kahneman, Daniel, (Editors), *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press, 2002, for a discussion of intuitive judgement.
[34] Vishwanath, Arun "Vish", Ph.D.; and Rao, Raghav, H., Ph.D., University at Buffalo; Herath, Tejaswini, Ph.D., Brock University (Ont., CA); Chen, Rui, Ph.D., Ball State University; and Wang, Jingguo, Ph.D., University of Texas, Arlington, *Why Do People Get Phished? Testing Individual Differences in Phishing Susceptibility Within an Integrated Information-Processing Model*, 22 June 2010, Web. 5 August 2014, http://citation.allacademic.com/meta/p_mla_apa_research_citation/4/0/0/3/5/p400353_index.html
[35] Libicky, Martin; Ablon, Lillian, Webb, Tim, *The Defender's Dilemma: Charting a Course Toward Cybersecurity*, RAND, 2015, Page 32.
[36] Li, Qing and Clark, *Gregory, Security Intelligence*, Wiley, 2015, p. 251.
[37] *M-Trends 2015: A View from the Front Lines*, Mandiant, 2015.
[38] Mandiant website, Web 15 June 2015, https://www.mandiant.com/threat-landscape/
[39] Kerner, Sean, *Sony Hackers Used Apple ID Phishing Scheme, Researchers Claim at RSA*, eWeek, 21 April 2015, Web. 10 May 2015. http://www.eweek.com/security/sony-hackers-used-apple-id-phishing-scheme-researchers-claim-at-rsa.html
[40] Elgin, Ben; Lawrence, Dune; and Riley, Michael, *Neiman Marcus Hackers Set Off 60,000 Alerts While Bagging Credit Card Data*, Bloomberg Business, 21 February 2014, Web. 21 May 2015, http://www.bloomberg.com/bw/articles/2014-02-21/neiman-marcus-hackers-set-off-60-000-alerts-while-bagging-credit-card-data
[41] JP 3-1.4 Pages A1-A2.
[42] Mandiant website, Web. 29 January 2015, https://www.mandiant.com/threat-landscape/
[43] It is profoundly difficult to deal with threats that have not been seen before. In testimony before the House of Representatives, an official from the Department of Homeland Security observed this about the highly touted EINSTEIN system, "the trick with EINSTEIN is, as it currently is built, it has to know about a threat before it can detect or block it . . . One layer of depth we need to provide is a layer that will help us detect and block intrusions we have not previously seen." Ravindranath, Mohan, *Committee Grills DHS Official Over EINSTEIN'S Failure to Prevent OPM Attack*, Nextgov, 24 June 2015, Web. 25 June 2015, http://www.nextgov.com/cybersecurity/2015/06/house-committee-grills-dhs-official-failure-einstein-cdm-prevent-opm-attack/116242/
[44] Libicky, et al., (n 35), Page xiv.
[45] JP 2-0, Page B-9.
[46] Zager, Robert and Zager, John, "Combat Identification in Cyberspace," *Small Wars Journal*, 25 August 2013, Web. 1 August 2014, http://50.56.4.43/jrnl/art/combat-identification-in-cyberspace.

[47] Libicky, et al., (n 35), makes over 100 references to the value of training, leveraging the assumption that training can overcome the Compliance Budget.
[48] The Federal Information Security Management Act of 2002 ("FISMA", 44 U.S.C. § 3541, et seq.)
NIST, Managing Information Security Risk, Special Publication 800-39, March 2011.
[49] Vishwanath, et al., (n 34).
[50] Caputo, Deanne, MITRE; Pleeger, Shari, Dartmouth College; Freeman, Jesse, MITRE and Johnson, M. Eric, Vanderbilt University, *Going Spear Phishing: Exploring Embedded Training and Awareness*, Jan-Feb 2014, Web. 03 August 2014, http://www.computer.org/csdl/mags/sp/2014/01/msp2014010028-abs.html

Ferguson, Aaron, Ph.D., United States Military Academy, "Fostering E-mail Security Awareness: The West Point Carronade," *Educause Quarterly*, Number 1 2005, Page 54, Web. 03 August 2014, http://net.educause.edu"/ir/library/pdf/eqm0517.pdf

Jackson, James W.; Fergusson, Aaron J., Ph.D.; and Cobb, Michael J.; United States Military Academy, *Building a University-wide Automated Information Assurance Awareness Exercise*, October 2005, Web. 3 August 2014, http://fie-conference.org/fie2005/papers/1694.pdf

Dodge, Ronald C., Ph.D.; Carver, Curtis, Ph.D.; Ferguson, Aaron J., Ph.D., United States Military Academy, Phishing for user security awareness, *Computers & Security* 26 (2007) Page 73, Web. 7 August 2014, http%3A%2F%2Fwww.researchgate.net%2Fpublication%2F222730836_Phishing_for_user_security_awareness%2Ffile%2F3deec5161aaaf78130.pdf&ei=genjU9KbD4XaoASnhIKoAQ&usg=AFQjCNHBB6LUHNG3Rgq21OR8Bw8sQ0cm2A&sig2=v0y7aBNPy9_NCwzk_gXn-g&bvm=bv.72676100,d.cGU
[51] KnowBe4, Web. 29 June 2015, http://cdn2.hubspot.net/hubfs/241394/Knowbe4-May2015-PDF/SocialEngineeringRedFlags.pdf?t=1435267515481
[52] Rasmussen, Rod, *Global Phishing Survey: Trends and Domain Name Use in 1H2011*, Anti-Phishing Working Group, November 2011, Web. 20 June 2015, http://docs.apwg.org/reports/APWG_GlobalPhishingSurvey_1H2011.pdf, p. 15
[53] Vishwanath, et al., (n 34).
[54] Beautement, Adam; Sasse, M. Angela; and Wonham, Mike, *The compliance budget: managing security behaviour in organisation*s, Proceedings of the 2008 workshop on new security paradigms. ACM, 2009.
[55] Hyken, Shep, *Phishers target middle management*, 06 July 2015, Web. 06 July 2015, https://powermore.dell.com/technology/phishers-target-middle-management/. Middle managers get 100 to 200 emails per day.
[56] Beautement, et al., (n 53).
[57] Kothari, Vijay; Blythe, Jim; Smith, Sean; and Koppel, Ross, *Agent-based Modeling of User Circumvention of Security*, Proceedings of the 1st International Workshop on Agents and CyberSecurity, ACM, 2014.
[58] Kirlappos, I.; Beautement, A.; and Sasse, M. A., *"Comply or Die" Is Dead: Long Live Security-Aware Principal Agents*, Financial Cryptography and Data Security, 2013.
[59] Kirlappos, et al., (n 57).
[60] Hyken, (n 54).
[61] Beautement, et al., (n 53).
[62] https://otalliance.org/resources/email-security discusses the use of standards-based email authentication to identify email senders.
[63] These innovations are covered by one or more of the following patents held by Iconix: U.S. Patents 7,413,085; 7,422,115; 7,457,955; 7,487,213; 7,801,961; 8,073,910; 8,621,217; 8,903,742.
[64] JP 2-0, Page B-9.
[65] Yuill, James, "Defensive Computer-Security Deception Operations: Processes, Principles and Techniques." *Ph.D. Thesis, North Carolina State University*, 2006, Web. 10 July 2014, http://repository.lib.ncsu.edu/ir/bitstream/1840.16/5648/1/etd.pdf
[66] In order to have effective sanctions, two conditions must be satisfied. First, the enforcers must have a system that is enforceable. Kirlappos, et al., (n 58) observe that when the Compliance Budget considerations create an environment of rampant non-compliance, discipline for non-compliance is not possible. Second, there must be a method of compliance monitoring.
In the proposed interface, the first condition is satisfied by the ease of compliance which respects Compliance Budget considerations; high risk behavior is easily avoided by not clicking links and opening attachments in emails that are not deemed trustworthy by Intelligence. The second condition is satisfied by recording interactions with unmarked emails.

[67] Dave, Paresh, "Who's susceptible to email phishing attacks? Study says everyone," *Los Angeles Times*, 26 July 2013, Web. 04 August 2014, http://www.latimes.com/business/technology/la-fi-tn-whos-most-susceptible-to-phishing-attacks-20130726-story.html
[68] See, for example, Cisco Systems, *Email Attacks, This Time It's Personal*, June 2011, Web. 03 May 2014, http://www.cisco.com/c/dam/en/us/products/collateral/security/email-security-appliance/targeted_attacks.pdf
[69] Vishwanath, et al., (n 34) and the references at Note 50.
[70] **countermeasures** — That form of military science that, by the employment of devices and/or techniques, has as its objective the impairment of the operational effectiveness of enemy activity. Joint Publication 1-02, *Department of Defense Dictionary of Military and Associated Terms*.
[71] Much like small raiding parties which operate from a common secret base of operations, APT attackers employ unique elements in the attacks (e.g., malware, domains, messaging) which often use a common obscured infrastructure from which multiple apparently independent attacks can be mounted. By obscuring these infrastructures, attackers use attack infrastructures for long periods of time. Examples of infrastructures that remained undetected for years:

A cyber attack infrastructure that went undiscovered for more than 5 years. GReAT, *Red October Diplomatic Cyber Attacks Investigation*, 14 January 2013, Web. 06 July 2014, http://securelist.com/analysis/publications/36740/red-october-diplomatic-cyber-attacks-investigation/

A cyber attack infrastructure that went undiscovered for approximately 3 years, Mandiant website, Web. 10 July 2014, https://www.mandiant.com/threat-landscape/ GReAT, *The "Icefog" APT: A Tale of Cloak and Daggers*, 2013, Web. 07 June 2014, http://www.securelist.com/en/downloads/vlpdfs/icefog.pdf

A cyber attack infrastructure used by attackers in India for at least 3 years was discovered in Norway. https://www.bluecoat.com/company/press-releases/norman-shark-cyber-research-report-uncovers-first-large-cyber-espionage, Web. 10 August 2014.

Although attackers cannot be stopped from creating new attack infrastructures (See, for example, the deployment of a new attack infrastructure after the publication of Mandiant's *APT1* report, Danger, David and Perlroth, Nicole, "Hackers From China Resume Attacks on U.S. Targets," *The New York Times*, 19 May 2013, Web. 31 July 2014, http://www.nytimes.com/2013/05/20/world/asia/chinese-hackers-resume-attacks-on-us-targets.html?pagewanted=1&_r=1&hp), forcing attackers to abandon infrastructures and create new ones introduces delays into attacks and increases the cost of operation for APT attackers.