

## DEFENSIVE INFORMATION WARFARE WITH NON-LOCALIZABLE COMMAND AND CONTROL

Randy Browne

New Jersey Computer and Communications  
2001 U.S. Highway 46 East, Suite 310  
Parsippany, New Jersey 07054

### INTRODUCTION

I loosely define "hard" information defenses as those assurance mechanisms that tend more to prevent (rather than recover from) the ill effects of security breaches and which tend to be more static in nature. Similarly, I loosely define "soft" defenses as assurances that tend to rely more on recovery from security breaches and offer a more dynamic/adaptive kind of information protection. Information protection, especially for critical information infrastructure, should balance the use of "hard" and "soft" information defenses owing to their often-complementary characteristics.

Unfortunately, rather than being complementary, "soft" information defenses appear to be largely replacing "hard" defenses, which is a particular concern within critical information infrastructure. Information warfare is rapidly becoming an excuse for simply not bothering with "hard" assurances, and my opinion of the cause is that our society (including the information technology industry, itself) has failed to truly grasp the economics of critical reliance on information technology, which must consider assurance costs. I grant that a "cyberspace arms race" is unavoidable by the very nature of the information protection problem. My concern is that the "arms race" created by excessive reliance on "soft" defenses is unnecessarily fast-paced and costly to society, and exposes critical infrastructure to unnecessary risk. In this paper, I argue for a more balanced use of "hard" and "soft" information protections.

### A DEFENSIVE INFORMATION WARFARE GAME WITH NON-LOCALIZABLE DEFENSE

I'll start by giving a simple result in the theory of multi-player games. For this paper, a defensive information warfare (DIW) game is a 5-tuple  $\langle n, S, P, T, G \rangle$ , defined as follows. Parameter "n" is an integer ( $> 0$ ), that defines the number of players on an attacking Team A, and on a defending Team D, with "2n" players in all. Parameter "S" is just a set of integers encoding the decisions made by the attacking and defending teams. Parameter "P" is a set of protection decision functions of the form  $Y = D(X_1 \dots X_n)$ , where each  $X_j$  is a member of "S" and represents the decision of the "j-th" attacker on Team A, and where "Y" is a member of "S" and represents the decision of some yet-to-be-specified defender on Team D. Note that each function  $D(X_1 \dots X_n)$  can be informally regarded as the potential behavior of a defender on Team D. Similarly, parameter "T" is a set of threat decision functions, each element of which can be viewed as the potential behavior of an attacker on Team A. The functions in "T" have the form  $X = A(D_1 \dots D_n)$ , where "X" is a member of "S" and represents the decision of some yet-to-be-specified attacker, and each  $D_j$  is a member of "P". Finally, function "G" is the game valuation function having the form  $G(X_1 \dots X_n, Y_1 \dots Y_n)$  into  $\{0, 1\}$ . Given the "2n" attacking and defending decisions,  $G = 0$  iff Team A prevails, and  $G = 1$  iff Team D prevails.

For the first step in the game, the defending Team D chooses "n" protection functions  $D_1 \dots D_n$  to anticipate the attacking Team A's choices. Second, or in

parallel with the first step, Team A chooses "n" threat functions "A1...An". Third, an impartial referee determines the attack "X1...Xn" by computing "Xj" from the "j-th" threat function "Xj = Aj(D1...Dn)" for each "j" in "{1..n}". Fourth, the referee computes Team D's decisions "Y1...Yn" by evaluating the protection functions of the form "Yj = Dj(X1...Xn)". The winning team is finally determined using the game valuation function by computing "G(X1...Xn,Y1...Yn)". I cannot give a detailed justification for the game structure in a short paper of this sort, except to say that it parallels prior work [3] using decision functions in formal theories of information security.

I use the term "strategy" to refer to a team's choice of decision functions, so Team A's choice "A1...An" is the "attack strategy", and Team D's choice "D1...Dn" is the "defense strategy". Ignoring weaker kinds of optimality for this paper, a strategy is "categorically optimal" when the corresponding team wins regardless of the opposing strategy. Also, I'll say that a strategy is "localizable" if every decision function in the strategy is equivalent to a function with strictly fewer than "n" arguments. The "localizability" of a strategy captures the important idea that none of a team's members has to have "global awareness". In the case of defense, Team D's defensive command and control (C2) problem is simpler if each team member only needs to worry about a subset of the attack strategy. The localizability of defensive C2 (strategies) is essential if we want to implement information defenses that completely avoid global decision-making. Unfortunately, there are DIW games where localizability of the defense strategy is mathematically impossible, even though the defender can always win by using some global decision-making. More precisely:

**THEOREM.** There are DIW games such that (A) there exists some categorically optimal defense strategy, but (B) for every localizable defense strategy, there exists a categorically optimal attack strategy.

**PROOF.** Consider some DIW game  $\langle n, S, P, T, G \rangle$ , with arbitrary " $n > 0$ ", with " $S = \{0,1\}$ ", and having function " $G(X1...Xn,Y1...Yn)$ " defined as follows. We let (a) " $G=0$ " if all of " $X1...Xn,Y1...Yn$ " are zero, or (b) " $G=0$ " if at least one of " $X1...Xn$ " is non-zero and at the same time, at least one of " $Y1...Yn$ " is non-zero, or (c) otherwise, " $G=1$ ". To prove part (A) of the theorem, define " $D1...Dn$ " as follows. Take " $Dn$ " to be the function into " $\{0,1\}$ " which equals one iff all of its arguments are zero. For " $j$ " other than " $n$ ", we take " $Dj$ " to be the zero function. The defense strategy " $D1...Dn$ " so defined is categorically optimal (i.e. " $G=1$ " always, but note the strategy is not localizable by my choice of " $Dn$ "). For part (B) of the theorem, let " $D1...Dn$ " be any given localizable defense strategy. I give a procedure for defining a categorically optimal attack strategy " $A1...An$ " simultaneously against " $D1...Dn$ ". If " $D1...Dn$ " are everywhere zero, then define all functions " $A1...An$ " to be everywhere zero. Otherwise, for some " $j$ ", and some choice of  $X1...Xn$ , we have " $Dj(X1...Xn)=1$ ". Since the defense strategy is localizable, there is some " $p$ " for which " $Dj$ " does not depend on " $Xp$ ". In this case, define " $Ap(D1...Dn)=1$ ", and for every " $k$ " other than " $p$ ", we let " $Xk=Ak(D1...Dn)$ ". Proceeding over every localizable strategy " $D1...Dn$ ", this defines functions " $A1...An$ " so that " $G=0$ " on every localizable defense strategy, thus constructing an optimal attack strategy for such cases. QED.

My notion of a DIW game is a bit abstract, so I'll offer an interpretation of the above results. Suppose we have a military real-time computer network of three identical hosts processing classified and unclassified information. Suppose also that real-time applications can tolerate no more than a 25% reduction in the raw, total computational performance of the three hosts. Also, the network contains what I'll call a "three part aggregation channel". This

aggregation channel is a security flaw (a "covert channel") in the host software that can result in leaking classified information, provided that all three hosts are monitored in parallel by invading viruses, worms, or "Trojan horses", etc.. Similar to the notion of key shadowing in cryptography, the existence, in theory, of "K part aggregation channels" has been proved for arbitrary "K" [1], where at least "K" information sources must be monitored, and monitoring fewer than "K" sources reveals no classified information.

Suppose the attacking Team A has penetrated each of the real-time hosts with an identical rogue program, where each rogue program (representing a member of Team A) can make one of two mutually exclusive decisions. Each rogue can either (1) cooperate with other rogue programs to monitor the three part aggregation channel, with negligible performance impact, or (2) impair real-time service by loading their local "victim" host resulting in 20% loss of performance per host. Also, the security defenses on each of the three hosts (each representing a member of Team D) can also make one of two decisions, either to (1) do nothing, or (2) do an expensive noise-injection operation to suppress the network aggregation channel. The noise-injection will prevent the three-host monitoring by the rogue programs, but results in a 66% performance loss on the affected host. Fortunately, the operation only needs to be done on one host to adequately suppress the covert channel monitoring on all three hosts. It turns out that the above three-host scenario offers an interpretation for the theorem. The attacker and defender both have two decisions per host as stated in this paragraph. If the two pairs of attacker and defender decisions are encoded as "{0,1}" in the order listed, then the game valuation function "G" (in the proof of the theorem) is zero, iff either (a) the network total/host average performance falls below 75% so that real-time service is denied, or (b) all three hosts are monitored without any defender-injected noise and classified information is thereby disclosed.

The connection of the theorem (and its interpretation) with the theme of this paper is that it might be mathematically impossible to fully localize defensive C2 (strategies) near the point of attack within distributed DIW systems. Using "soft" defenses may be more complex than it appears, since an "adaptive, soft" defense may require solving a global defensive C2 problem. Given a choice, it may be easier to just "harden" the host security (e.g. to eliminate the covert channel) and dispense with the "soft" defense.

#### OTHER NOTIONS OF NON-LOCALIZABILITY AND WEAKNESSES WITH "SOFT" DEFENSES

Now consider two processors P1 and P2, which are identical except that P2 contains a "firmware time bomb". The instruction set microcode in P2 has been sabotaged so that a counter is updated in parallel with the start of each instruction fetch cycle. After the counter expires, P2 will no longer operate. However, before the counter expires, P1 and P2 have identical observable behavior down to the smallest timing detail, except possibly for electromagnetic characteristics. Basic information theory tells us that it is impossible to distinguish P1 and P2 before-the-fact by monitoring the processor, alone. If clandestine substitution of P2 for P1 is part of a larger attack strategy, then "soft, adaptive" DIW might detect higher level clandestine activity (e.g. hostile agent migration patterns) intended to exploit the substitution. However, there is no "intelligent, soft" defense that can adaptively predict the presence of processor P2 versus P1 specifically via monitoring. By contrast, "hard" examination of the microcode before deployment would go a long way to avoiding the problem outright.

While a firmware time bomb may not pose a practical risk, it illustrates a theoretical weakness with "soft" defenses, when the attack actually occurs prior to system operation. There are analogous situations with operating systems that are almost as difficult to defend against, and are much more dangerous. An operating system might accept a covert shutdown command triggered by a long sequence of network messages. The trigger pattern could be encoded by some sequential combination of header options, source address, packet length, etc.. Unlike the firmware time bomb, it may be theoretically possible to detect such behavior at run time, if the trap door affects system timing. In any case, it should be clear that many trap doors would be virtually impossible to detect by either local or remote monitoring. Traditional "hardening" of the operating system itself seems the only effective solution. Thus, denial of service by time bombs and trap doors also pose difficulties in trying to localize defensive C2 near the point of attack. For certain such problems, any kind of "soft" run-time defenses would be ineffective, yet solutions can be found with traditional "hard, off-line" assurances.

#### CONCLUSION

Defensive information warfare, and "soft" defenses generally, are difficult to implement when defensive C2 resists localization. In some cases, such as with the theorem proved in this paper, non-localizability of defensive C2 might simply frustrate any distributed approach to DIW. In other cases, as with many time bombs and trap doors, non-localizability of defensive C2 makes "soft" defenses virtually impossible. What's more, information warfare is not the only, nor necessarily even the central issue in information survivability [2]. Our society is growing too dependent on "soft" defenses and we need a more thoughtful balance between "hard" and "soft" information protections, especially for critical infrastructure. Using "soft" defenses is essential to cover unforeseen threats that "hard" defenses cannot cope with. However, "soft" defenses should not be used to "patch up" a system lacking "hard" defenses against easily foreseeable threats. Furthermore, "soft, adaptive" defenses are inappropriate where the opportunities to learn from mistakes are limited, such as where information attacks can have catastrophic effects. We are facing a "cyberspace arms race" which can never be eliminated, but which has a rapid pace that can be, and must be greatly reduced by broader use of "hard" defenses. Moreover, this paper has not addressed how "soft" defenses create very special problems for (1) Internet information protection for small businesses and individuals, (2) mobile computing devices, and (3) military systems where information survivability issues are subordinate to broader survivability problems [2]. Such problems, among others, are part of the motivation for seeking a better balance between the use of "hard" and "soft" information defenses.

#### REFERENCES

- [1] Browne, R., "An Entropy Conservation Law for Testing the Completeness of Covert Channel Analysis", In Proceedings of the 2nd ACM Conference on Computer and Communication Security, Fairfax, Virginia, November 1994.
- [2] Browne, R., "C4I Defensive Infrastructure for Survivability against Multi-Mode Attacks", To appear, In Proceedings of MILCOM 2000.
- [3] Wittbold, T., and Johnson, D., "Non-Deducibility on Strategies", In Proceedings of the 1990 IEEE Symposium on Security and Privacy.