IMPOSING COSTS ON CYBER ATTACKERS

Stephen J. Lukasik

Hicks & Associates, Inc.
Prepared under contract to the Office of the Secretary of Defense

October 2005

TABLE OF CONTENTS

IMPOSING COSTS ON CYBER ATTACKERS

Stephen J. Lukasik

INTRODUCTION

This is the third in a series of short papers dealing with cyber attacks and cyber attackers. The first "mapped the landscape" and discussed the targeting principles attackers will employ, the targets sets they will attack, the origins of attackers, and the steps in the cyber attack process at which costs might be imposed on attackers.[1] The second paper examined the specific case of cyber burglary, the unauthorized copying of files, using Titan Rain as an examplar.[2] It looked into possible policy approaches to preventing such thefts. The paper offered ten policy proposals to minimize the effects of, or to prevent, such attacks. The proposals can be separated into those that can be implemented relatively quickly and those that can only be effective over a longer term.

One of the short term proposals was to directly impose costs on the attacker by removing the prospect of a "free" attack, i.e. one for which the attacker pays no penalty for attempting and thus which can be repeated until the attacker succeeds. This would require a counter-attack, preferrably immediate though a delayed counter-attack would also be possible depending on the dynamics of the situation. The downside of such a response, in view of the high degree of anonmity that exists in public networks, is the high probability of collateral damage.

Three other relatively short-term proposals were to undertake damage limitation while an attack is underway, blocking that attacker's access, sequestering files being accessed, or substituting erroneous files for copying. A second proposal, one probably requiring R&D, is to develop a detailed understanding, at the bit level, of what software is in a machine and what processes are being executed at each machine cycle. The third proposal is to manage trust on a much finer-grained time scale than to simply to vet a user at one point in time and to do little or nothing to re-establish a basis for trust of that person. In effect this says that users and operators are as suspect as actual attackers and should be viewed, if not as guilty, then as suspect at all times.

Of these three proposals, the first is not cost-imposing but rather "profit-reducing." It diminishes the extent of a theft for a given level of effort by the attacker. The second and third increase the cost to the attacker either by requiring greater effort to introduce malicious code surreptitiously or to maintain freedom of action within the attacked machine.

Five of the remaining longer-term proposals in footnote 2 similarly increases the cost to an attacker. Removing machines from open networks will require the attacker to either give up

[1] Stephen J. Lukasik, "Mapping the Landscape of Cyber Attacks." H&AI, Aug 05, FOUO.
[2] Stephen J. Lukasik, "Combating Cyber Burglary: the Case of Titan Rain," H&AI, Sep 05, FOUO.

on that machine or to devise other ways to attack it, such as through the insertion of agents into the secure facility. Increasing the number of "watchers," paid or unpaid, will force the attacker to adopt deeper and constantly-changing cover identities, with consequent reductions in productivity. Prompt installation of security patches will reduce the window of vulnerability within which an attacker must act. Defensive design will, in the long-term, result in harder systems and less naïve system operators. Making users aware of the power at their disposal when on a public network is likely to reduce their tendency to behave irresponsibly.

IDENTIFYING AND MONITORING ATTACK TEAMS

The above recounting of earlier security proposals relating to cyber burglary has omitted the one that forms the subject of this paper, that of identifying and monitoring organized attack groups. The viewpoint here is the same as for military intelligence generally. Identifying large teams of attackers is nothing more than *order of battle* intelligence, and monitoring their actions has as its goal *indications and warning*.

There are two types of organized attack teams to consider.[3] One type is state-supported groups, whose goal is the acquisition of immediately useful information or the preparation of the cyber battlefield for a later attack. The targets of such groups will be predictable for each sovereign state, and will involve that state's current and possible future adversaries and competitors. State-supported attack groups thus have a larger political and economic reason for existing and their actions support national objectives. This provides an important starting point for searching for them and for responding to their probes and attacks.

The second type of organized team is one that operates for private profit. They will do things such as selling stolen goods, or will sell their services using their penetrations and their track record as a marketing tool. Unlike state-supported groups, these groups will be global in their reach and their potential customer base.

The defensive actions one takes are different for these two situations. Defenders against state-supported group can benefit from the international sharing of information and joint actions along existing political axes will be possible. Friends will cooperate and more remotely-related states are less likely to, or may not be trusted. On the other hand, the criminal consortium will face more global opposition and opprobrium since all states are potential victims or unwitting accomplices. Actions against state-supported groups are likely fall to intelligence agencies while those against criminal consortia will more likely be the mission of law enforcement agencies. Mixed defensive teams, domestically and internationally, against both types of adversaries, are possible and practices will vary among states.

[3] The adjective "organized" is intended to convey the idea of large groups in an operational mode for a serious purpose, as opposed to casual hackers engaging in sport and seeking bragging rights among peers.

Thus the Titan Rain thefts need not be viewed only as short-term efforts to obtain information. They may also be "graduation exercises" or "journeyman training" intended to develop attackers and attack techniques. In this view, a theft is simply proof of entry or OT&E of new attack alternatives. Thus the implication of Titan Rain types of theft may not simply be further losses of information. It is just as likely that of more attackers and more powerful attacks can be expected. The Titan Rain thefts can be viewed as attacker *failures*, where the intended covert nature of the entry was unsuccessful. The outlook in such cases may not only be more attacks, but for attacks that will be even more sophisticated.

The number of such attack teams may not be impossibly large to detect and monitor. Private groups, like organized crime everywhere, will be limited in number as the parent criminal enterprises divide up the market by geographical region or by target specialty and enforce "local" monopolies for the production and sale of stolen information, the operation of protection rackets, and for maintaining their primacy in various markets. The Titan Rain paper sized such groups as likely to consist of modules of perhaps fifty people, divided between production, management, and distribution. A rough estimate might be that 20–40 such groups could operate worldwide as adjuncts to major criminal consortia. The market for such stolen information requires study. For example, the large number of identity thefts means there must be a distribution system for their sale, just as there is for cloned cell phones and untraceable firearms. The distribution systems may be good starting points back to the cyber sources. They are likely already known to law enforcement agencies; some will have been penetrated; and as large numbers of people participate in the market, they will make mistakes and leave trails.

The number of state-supported groups will be similarly small. There are over 200 recognized sovereign entities, but many are small and are unlikely to have enemies for which cyber penetrations for state purposes is required. perhaps at most 40–50 nations will have adequate incentives to operate cyber attack groups, although in the case of Titan Rain it appears that China may have stood up at least three. On the other hand, corrupt states could well countenance private operations for the private benefit of government leaders in another example of public–private partnership.

Thus a key issue on which the type and size of defensive measures should be based is whether Titan Rain is a major part of such activities or is a very small indication of a much larger problem. Even should Titan Rain be unique any prudent state would plan for the eventual flowering of both private criminal and state-supported cyber attack groups.

Detecting and identifying cyber attack groups can be done by looking for evidence in three phases of their development. All such groups will start small, and they may not be visible when in embryonic form. But as they develop in size the indicators will increase in number and one may hope to catch them in their early stages. Indicators will arise during their three life cycle phases: recruitment and basic training; advanced training and journeyman development; and from their operations. Tables I and II suggest models for the development, and thus detection, of the two major typer of cyber attack groups. Table I applies to the development cycle of state-supported groups while Table II describes non-official groups.

Table I
Model of Officially-Sanctioned Cyber Attack Groups

RECRUITING/BASIC TRAINING	ADV TRAINING/JOURNEYMAN	OPERATIONS
Motivation of recruit - official assignment - perks	Official (closed hacker circles)	Integrated teams to handle HR, target value assessment, translation and analysis of collection, and line production staff
Source process - school performance - competitions	Supervised "field trips" by senior instructors	Organized structure - relatively fixed order of battle - TO&E - 7 x 24 operation - organized around tgt time zones
Work schedule - recruit young, as early as age 10 - daytime hours over time zones	Small group "mentoring"	Systematic view of targets - attacker currently need access - attacker is preparing battlespace
Locations - centralized facilities - urban	Small test penetrations at low security/low sensitivity organizations	
	Practice attacks against own realworld systems both to for quality assurance and to train own defenders	
	R&D in computer and network security, both to work with leading-edge people and as a recruiting tool	
	Establishment of start-up companies in R&D and cyber security	

Indicators or signatures can be constructed on the basis such a model. The recruiting process is apt to be visible since it by definition must interact with the attacker nation's populace. The fact that it will start with very young people may make it more easily visible. The time zone argument means that if there is a defensive effort to analyze cyber incidents and attacks, it will likely have a 24-hour periodicity that will provide a spectral signature. Similarly the training locations will be in urban areas having the largest populations.

The advanced training process will also offer indicators having a 24-hour periodicity whose phase says something about its geographic location, but now the place to look will be low-sensitivity but nontheless real targets that can be instrumented by defenders. R&D centers and cyber-security start-ups will be good ponds in which to fish.

The indicators from the operational activities of a full-up team will be characterized by it rate of attacks, all coming from a few number of attacker locations. This is the case of Titan Rain. One can hade small levels of effort but eventually something akin the traffic analysis becomes possible. The attack organization will, for reasons of quality control and productivty will remain relatively fixed over time. Growth is likely to come from expansion of an

operational group followed by fissioning when two critical masses can operated successfully. Alternatively, a new top performer can be given a charter to establish a new group.

Now consider a model for non-official cyber attack groups.

Table II
Model of Non-Official Cyber Attack Groups

REC RUITING/ BASIC TRAINING	ADV TRAINING/JOURNEYMAN	OPERATIONS
Motivation - fun/thrill - money	Fun /thrill - self-organizing - moderately open to watching - follow screen names - recruitment of insiders - possibility of undercover penetr.	Fun/thrill - attacker will escalate - eventually makes fatal error - defender use for training - defender use for own recruiting - susceptible to gaming - defender should view as exploitable asset
Source process - self-organizing - internet chat rooms/downloads	Money - organized gangs - organized criminal groups - straight employees in trouble - disgruntled employees - advertsising of skills - careful auditing - use of accounting regulations - prisons as training ground	Money - the "casino model" of steady income - the "Brinks model" of a big kill ancillary other operations such as physical theft, money laundering, extortion, blackmail
Schedule - start young, age 10 - after school or work hours		
Location - urban - rural/bored kids		

The recruiting process is even more open here than in the case of the official groups since the individuals involved do not at that point see the need for operational security, nor is they anyone to assist with it. The phasing of basic training activity will differ from the business-day phasing for official groups. The geographical location of the early training may be outside urban centers. The advanced training process will differ between that organized for thrill and that for profit. The latter will be more visible, either because law enforcement agencies are already aware of criminal groups or because attacker activities are more directed to places that could offer profits. Non-official groups are likely to suffer a higher rate of infant mortality than those officially sanctioned, but those that do survive will be fully as capable as oficial groups. Operationally the criminal group is likely to go for a few large attacks while the official groups are in it for the long haul and will seek a steady stream of results.

DEFENDER STRATEGIES

Defender straegies for each type of attacker group are suggested in Table III

Table III
Defensive Strategies

TYPE OF ATTACKER	RECRUITING/ BASIC TRAINING	ADVANCED TRAINING/JOURNEYMAN	OPERATIONS
	Intelligence goals		Defense/Law Enforcement goals
Officially sanctioned	Establish order of battle	- Monitor for threat assessment - Study methods for countermeasures	- Strengthen defense of own assets - create plans for counter-attack - expose for political purposes -assess strategic balance of offense and defense
Non-official group	- Recruit for own defense - turn in place	- use as observables to understand the larger criminal purpose - turn in place	Follow paths to buyers

The defensive view suggested here is that for both official and non-official groups, the recruiting/basic training and advanced training/journeyman development processes should be treated as intelligence collection oppertunities, where the intent is to obtail information that will be useful later however the defender chooses to use it. When operations are in place the strategy should shift to either (1) active defense, (2) rolling up the attacker groups when that can be done with international support to pursue attackers, collect legally admissable evidence, and apprehend perpetrators, ore (3) to escalate the costs the attacker in other ways.

The intelligence-oriented approach has features in common with some well-known WW II actions. They would allow basic and advanced training of adversaries to occur in order to draw them out into the open so they can be identified, their operations characterized, and their attacks analyzed for their future potential. But like the breaking of the Enigma machine, it means that some losses will have to be sustained for the larger intelligence purpose. One tactic such watch-and-wait strategy would allow is to turn penetration agents and use them as a, possibly unwitting, path to feed tailored information to their masters. The WW II XX operation, where all German agents in the U.K. were captured on entry and turned, is a comparable case.
In the case where the penetration is not for immediate action but one intended to prepare a future battlefield, the appropriate response, even in the case of attacks by operational groups, is to search for leave-behind code. This would be critically important in the case where a future pure cyber, or mixed military–cyber attack is planned, and where disrupting critical national infrastructures is the intent. Finding, analyzing, and walling off such code would then be a high priority. It would be the equivalent having the opponents strategic war plan,

and would provide impotant I&W capabilities. The defender could also simulate the activation of such infrastructure attacks and could, thereby, learn about vulnerabilities otherwise not considered.

There is a great deal of game-playing here, as the attacker checks on continued access to the desired systems and functions, all of which will point to attacker intentions. Such "cyber-mines-in reverse" could add to the nation's suite of predictive capabilities.

SETTING PRIORITIES

From the discussion thus far is is clear one can impose costs on cyber attackers. The issue is whether this can be done without disproportionate costs to the defender. As with nuclear smuggling, the situation is highly asymmetric in a way that favors the attacker.[4] Some of the factors operating are:

(a) attack costs are modest, involving relatively few people and trivial operating costs;

(b) the bigger defender has much more to lose than the smaller attacker;

(c) the defender has many targets to protect and thus has many entry points that can be breached or inadvertantly left vulnerable;

(d) the defenders' assets are distributed over a large number of owners, each of whom can do as little as they want, despite the fact that all are connected by a public network;

(e) private sector owners have little financial incentive to invest in defense, and if they did they would find themselves disadvantaged vis a vis competitors who did not;

(f) government regulation is not seen as a desirable approach;

(g) massive cyber attacks have not occured so denial is a typical reaction.

As noted in the referenced analysis of nuclear smuggling, the relevant cost-exchange ratio involves *relative* costs, not absolute costs. The richer defender can and, however grudging, will spend more than the poorer attacker.

The starting point for the analysis of steps with the most attractive cost-exchange ratios is to outline the process the attacker must go through to mount an attack; to examine the interventions the defender can make at each step in that process; to estimate the regret to the

[4] Stephen J. Lukasik, "Imposing Costs on Nuclear Smuggling by Terrorist Groups," Center for Adaptive Strategies and Threats, H&AI, 8 Aug 05, FOUO

attacker should the defender succeed with such an intervention; and to sort possible defender interventions by lowest relative cost to the defender and then by decreasing regret to the attacker. Thus low cost interventions that hurt the attacker most rise to the top, with increasingly costly defender interventions having decreasing impact on the attacker following. For more detail the reader is referred to the report indicated in footnote 4.

The steps for the cyber attacker are taken from the report indicated in footnote 1:

(a) decision made to create a cyber attack capability;

(b) acquire personnel and develop as indicated above;

(c) acquire target information to plan offensive action and choose among alternatives;

(d) set up a (distributed) facility from which to stage training, exercises, probes, and operations;

(e) formulate a campaign plan against an adversary nation or target industrial, government, or commercial sectors worldwide;

(f) deploy forces into operational staging areas;

(g) execute attack campaign over time.

The analysis leading to establishing priorities is shown in Table IV. The first step is to mine this and the previous two cyber papers for possible interventions in the cyber attack process. They are listed in column 2 and numbered for future reference. There are 27 such interventions shown, but there is no implication that this is a complete set. The reader is invited to add others as may suggest themselves to the list and then carry through with the analysis as presented below.

Next one identifies each listed intervention with one or more of the cyber attack process steps. These identifications are shown by a “1” in the seven columns labled “step A” through “step G.” Following each of these seven columns are two more columns labeled “defender cost” and “attacker regret.” The methodology is now the same as that used for nuclear smuggling referenced in footnote 4.

Table IV
Most Attractive Relative Cost-Exchange Ratios for Defensive Intervention

		Decision			Personnel			Tgt info			Facilities			Plan			Deploy			Attack		
Int no.	possible interventions	step A	def cost	att reg	step B	def cost	att reg	step C	def cost	att reg	step D	def cost	att reg	step E	def cost	att reg	step F	def cost	att reg	step G	def cost	att reg
1	identify and monitor attack teams	1	2	1	1	1	1	1	3	3	1	2	1							1	2	1
2	limit theft transattack																			1	2	1
3	limit networking																			1	2	1
4	launch direct counterattack																			1	2	1
5	increase number of watchers				1	3	2	1	3	2	1	3	2							1	3	2
6	force installation of security patches																			1	3	2
7	maintain software/process awareness																			1	2	2
8	manage trust/distrust continuously							1	3	3										1	3	2
9	practice defensive design of systems							1	3	2										1	3	2
10	elevate user awareness of IT power																			1	2	2
11	block attacker access transattack																			1	3	1
12	sequester files transattack																			1	2	1
13	substitute erroneous files transattack																			1	1	1
14	establish adversarial pairs of entities	1	3	3																		
15	report all incidents centrally in RT				1	3	3	1	3	2	1	3	2				1	2	2	1	3	2
16	focus monitoring at high-level sites										1	3	2				1	2	2	1	3	2
17	search for leave-behind code							1	2	1	1	2	1				1	2	1	1	2	1
18	exploit leave-behind code							1	2	1	1	2	1				1	2	1	1	2	1
19	broadcast alert levels in RT																			1	3	2
20	track recruitment of young hackers				1	3	1															
21	track disappearance of known hackers				1	2	1															
22	monitor prisons and ex-prison inmates				1	3	2															
23	auditing users for change in behavior				1	3	2	1	3	2										1	3	2
24	do time-zone filtering of traffic statistics										1	2	3				1	2	2	1	3	3
25	analyze attacks and plan active CM										1	3	2							1	2	1
26	turn attackers				1	1	1	1	1	1	1	1	1									
27	prosecute selected attackers													1	3	2	1	3	1	1	3	2
	TOTAL	2			8			9			9			1			6			22		

Costs to the defender for a particular intervention and the regret level of the attacker should that intervention be successful are shown as one of three levels. The lowest level of cost and regret is shown as "3." For the defender this means that the cost to succeed in such an intervention is comparable to what is currently being expended for cyber security and that the technology to do so is in-hand and deployable. A further implication of this lowest level of cost is that there is no unbearable political cost that would be triggered should such an intervention be mandated. A level "3" regret to the attacker is that he will not be critically impacted should that intervention be implemented and succeed. The attacker may be inconvenienced but can find a work-around.

Level "2" for each is one qualitative notch higher. For the defender the technology may not be available and may require development and/or testing, which implies a time delay, R&D costs, and technical risk. There may also be a higher political cost involved, such as legal isues, economic downsides, etc. For the attacker a level "2" regret is something that constitutes a serious setback, probably requiring going back to an earlier step of probing the target or impacting an attack milestone.

Level "1" is another qualitative notch up for each. For the attacker a "1" means that there may be no way to accomplish the intervention, or the deployment costs may be extremely high, or severe political pain will follow from mandating the intervention. For the attacker a "1" means complete failure, a show-stopper, total compromise of the operation, the attack team, the attack plan, and the ultimate sponsor. Note that this scheme reflects the asymmetric nature of a cyber attack. The issue is not absolute dollars but the relative costs to both sides.

The analysis proceeds by deciding on the level of cost and regret for each intervention, keeping in mind the action the attac ker is undertaking at each step. For example, in Table IV identifying and monitoring an attack team is a level "2" in terms of cost of detection (step A) but a level "1" in terms of cost to identify the individuals involved. Since each intervention can impact more than one attacker step, there are a total of 57 sets of cost and regret judgments to make.

The greatest number of interventions, 22 (38%) impact the attack phase. This is a good news–bad news result. While it is comforting to see that so much can be done to defeat an attack, a "dodging the bullet" strategy leaves one more than a little uncomfortable. As a general rule one prefers to defeat an adversary as early in the process as possible. Another factor, not supported by this analysis but by other studies of attacks is that they pick up speed as they unfold. The early stages of planning, recruiting, and training take time but once the team is in place, the targets selected, and the attack tools mastered, things can move too quickly for comfort.

One reads Table IV in terms of cost–regret pairs. The most attractive interventions for the defender are the 3:1 combinations, lowest in cost to the defender and highest in impact on the attacker. There are, unfortunately, only 3, and only one is an early-stage intervention, the identification of recruits. The other two, selective prosecution of attackers identified and blocking attacker access transattack are late-stage efforts.

Table IV shows 41 (72%) of the entries highlighted. These are what are cost-attractive to the defender and thus *relative cost*-imposing on the attacker. So there is a lot that can be done with even this admittedly incomplets set of possible interventions. The criterion for highlighting an entry is that the cost to the defender is less than the regret to the attacker. This leaves out even-trade actions: 1:1 (nice if you could do it), 2:2, and 3:3 (not worth the effort.) Interestingly, however, there is only one where the trade is unfavorable to the defender, a 2:3 intervention about using time-zone filtering on attack traffic analysis.

Of the 38 highlighted interventions that are not 3:1s, they divide into two sets. The next most attractive, and which should form the basis for a rational cyber security program, are the 2:1s. Their cost is manageable and they exact a level "1" regret on the attacker. There are 17 (30% of the total) of them, although there is likely to be an optimum set of a smaller number that constitute lowest cost/highest effectiveness interventions. Establishing this minimum cost optimum effectiveness interventions will require a deeper level of analysis and design than is possible here.

This leaves 20 entries (35% of the total) that are 3:2s. One is likely to find there is utility in some of these because they are still level "3" costs to the defender. As a more fully articulated program of cyber defense is formulated they find a place, either from providing back-up to other actions or as incremental cost-free by-products.

CONCLUSIONS

This analysis suggests that there is a lot that can be done on the cyber defense problem. The task is to select a set of interventions that will provide the maximum protection for the least expenditure, public and private. The PCCIP first drew attention to the issue.[5] Since that time the national response seems to be that everyone is doing everything they can think of they can find funding for and those responsible for securing systems often see only the costs and not the benefits of defensive interventions. The point is not simply to spend more money, or mandate new rules, but to apply defensive resources to the most cost-effective actions.

The "rejects" from this analysis, the 14 n:n pairs, plus the one unfavorable cost-exchange intervention, constitute only 24% of the interventions examined. This suggests, if the sample of possible interventions considered here is in some sense statistically representative, that there is a 3:1 chance that even a random intervention will have merit. The management issue is not whether random choices are ineffective, but whether the most cost-effective actions are being pursued.

Returning to Table III, one can ask if the defensive strategies indicated there are supported by the subsequent analysis. With respect to officially sanctioned attack teams, the answer is yes. There are a number of interventions that offer promise in identifying and monitoring attack

[5] "Critical Foundations: Protecting America's Infrastructures," The Report of the President's Commission on Critical Infrastructure Protection, October 1997.

teams in the pre-attack period yet still provide ways of meeting the essential requirement for adequate transattack responses. The same assessment can be made for non-oficial attack teams also. But there is oine critical caveat to both. This is that the privately-owned targets of such attacks will be resistent for the reasons indicated. The government's reach in mandating cyber defense is quite limited and absent a convincing disaster private-owned targets are unlikely to cooperate.

While a real attacker will pull no punches, it might be useful to contemplate serious demonstrations of the magnitude of the threat. Controlled demonstrations of vulnerability might be considered. The DoD, for example, might secure the cooperation of one or more of its contractors to allow it to bring down an entire facility in a remote region (to minimize collateral damage) through a combination of electric power, communications, and local transportation. Since facilities supposedly practice disaster responses, there would seem to be a basis for such demonstrations. It is important, however, that they be *real* attacks and not the usual scripted and simulated exercises. The Department could perhaps line up five locations and even warn them that one of them would be taken down for one day during a specified one-month period. FEMA and other agencies might also participate. Efforts to secure information concerning the attack path would be required since the vulnerabilities exploited are likely to be replicated elsewhere.