# The Response Continuum

Sergio Caltagirone, Deborah Frincke

*Abstract*— **Active response is a sequence of actions performed specifically to mitigate a detected threat. Response decisions always follow detection: a decision to take 'no action' remains a response decision. However, active response is a complex subject that has received insufficient formal attention. To facilitate discussion, this paper provides a framework that proposes a common definition, describes the role of response and the major issues surrounding response choices, and finally, provides a model for the process of response. This provides a common starting point for discussion of the full response continuum as an integral part of contemporary computer security.**

## I. Introduction

Response is implicitly present in every security defense system, whether that response is to inform a system administrator, to close off access, to involve law enforcement, or even to ignore the misuse. A reasoned debate as to how to determine an acceptable level of response seems appropriate. However, only recently has "response"[1] been addressed with academic rigor by the mainstream research community. We propose a framework for evaluating response possibilities both qualitatively and quantitatively. This framework provides a common starting point for response discussions with an emphasis on active response as an integral part of contemporary computer security.

Many factors may have influenced the hesitation in addressing response. First, researchers have logically argued that it is of primary importance to reduce system design vulnerabilities and/or accurately and rapidly detect misuse, as a precursor to formal response. Second, experimenting with the more extreme forms of response is difficult to do safely within most university environments, and has associated costs: increased supervision of students, separation of equipment from the mainstream, potential for bad publicity, harm caused if experiments overflow the university testbed, to name a few. Too, some aspects of the discussion of response techniques are akin to those about whether students learn more from a defensive, or an offensive, philosophical approach [1]. For researchers teaching from a defensive posture, it would be somewhat foreign to study response in detail and correspondingly natural to incorporate primarily protective measures into their research and classrooms. Further, response has tended to be "folded in" as a sideline to detection, rather than designed in as a key characteristic of a protective system, and this may also have led to the lack of focused attention upon active response. Finally, some equate response research with *advocacy* of extreme forms of response. These are all valid concerns.

However, just because a discussion is awkward does not mean it should be avoided. There is a growing frustration among some security practitioners, many of whom are increasingly dissatisfied with the effectiveness of current remedies. This frustration has led a few to take more aggressive measures [2–4]. Therefore, we believe it is incumbent now more than ever for the academic community to seriously re-examine the question of response, and to lead the public debate regarding where/when/why various forms of response are appropriate.

To support this debate, we offer three things:

- a common definition of response,
- an ordered continuum of response actions, and
- a straw-man schema for evaluating choices among possible response actions.

## II. The Need for Response

Response has long been incorporated as part of most good defense strategies, and computer defense is no exception. We note some uses of "response" to computerized threats. In 1998, the US Department of Defense, while responding to an attempted denial-of-service (DoS), launched an applet that shut down the browsers of attackers preventing them from attacking further [5]. In 2001, a U.S. District court judge allowed the FBI to compromise a Russian hacker's computers and install a keylogger to gather evidence on his illegal activities regarding computers inside the U.S. [6].

In the context of decisions such as these, it is legitimate to ask: why were these particular response actions chosen? Were they at the appropriate level of force given the context of the perceived threat and the degree of certainty of the defenders about their circumstance? What was the external cost of these response decisions and did it match what was predicted? The problem is that there is no agreed-upon tool, technique, method, standard or policy, upon which one can rely for making good response choices when

Sergio Caltagirone (scaltagi@acm.org): University of Idaho, Moscow, ID.

Deborah Frincke (deb.frincke@pnl.gov): Pacific Northwest National Laboratory, Richland, WA.

[1] Alternately called "active defense"

a threat is identified, perceived, or predicted. Currently, those responsible for defending the front lines of most computing systems rely heavily on relatively ad hoc policies or instincts guided by experience and intuition, rather than a well-defined scientific model.

Any ad hoc method of making response decisions is most unsatisfactory, and is particularly unacceptable when system value is high and/or availability cannot be interrupted, as in life/safety/national security critical systems. Ideally, it would be replaced with a rigorous and scientific model that can support appropriate response selection. With the growing number of exploitable vulnerabilities in critical systems (e.g. air traffic control [7], nuclear power safety systems [8], etc.), and corresponding pressure to protect such systems, the immediate need for such a model is evident.

## III. Related Work and Previous Discussion

Much discussion of response has emphasized hack-back,[2] and so we begin by outlining the essence of that discussion. In 2002, Jayawal, Yurcik and Doss called for more effective ways of protecting networked systems from attack and examined the possibility of hack-back [9]. In the same year, Mullen presented justifications for strike-back at defcon [10], wrote a corresponding article at SecurityFocus [4] and published a whitepaper on strike-back [11].

Mullen's work inspired a Reuter's news article about strike-back, especially the ethical and legal implications [12]. As well, researchers such as Schneier criticized Mullen's position, using an analogy to the Recording Industry Association of America's (RIAA) attempt to attack copyright infringer's computers [13]. Mullen responded in [14]. However, active response considers a greater range of actions than only hack-back.

Some researchers emphasize the response decision-making process. Loomis, in [15], while implying hack-back, presents an objective discussion of the ethical and legal aspects of response decisions. Grove, et al., distinguishes 'active defense' from passive defense and undertakes a thorough discussion of international law implications of active defense [16]. Bruschi and Rosti [17] discuss a response strategy for DoS attacks in which they limit the capabilities of the attacker rather than strengthening defenses. Additionally, they provide AngeL, an implementation of their strategy [18,19]. [20] provides an organizational model and structure for codifying how decisions about response could be made.

## IV. Active Response: Definitions and Models

### A. Characteristics of a Definition

We propose the following requirements for any definition of active response:

- Active response is *time-bound*, and takes place exclusively during a period when an attack is known[3] to be in progress.
- Active response is *purposeful.* Actions would only be considered 'response' when used to mitigate the threat for the purpose of returning to a more secure state and no further.
- Active response, being purposeful, has *limitations.* Threat mitigation need not mean threat elimination, but instead indicate that the actions seek to diminish or contain the threat with respect to the resource considered to be threatened. Neither punitive responses nor warning responses, which might be intended to have the general effect of reducing threats from a broad perspective, are not included in our definition of active response.
- The decision to apply specific active response actions is *controllable* and *deliberate*,[4] though an active response action sequence's consequences might be neither.
- Active response is comprised of a *sequence* of actions.
- Active response is *technologically independent* and can be executed by an operator or automatically by an intrusion detection system using a pre-defined ruleset.
- The timeline of active response is to be considered considered *subjectively*, from the perspective of the decision-making entity, and not from any other body.

We have been deliberately general in some of these requirements. For example, when precisely an attack should be considered "detected" or to have "finished" depends on several factors and in our opinion is again a matter for the academic, legal, and other stakeholder communities to discuss. We identify some of the relevant factors in [20].

### B. Definition

This leads us to the following definition of active response.

Definition 1. **Active Response**: *Any action sequence deliberately performed by an individual or organization between the time an attack is detected and the time it is determined to be finished, in an automated or non-automated fashion, in order to mitigate the identified threat's negative effects upon a particular asset set.*

How well does this match our requirements?

- **Time bound** and considered **subjectively**. Active response takes place only when the organization believes the attack to be in progress.

[2] Hack-back: retaliating against the attacker using techniques that share many attack characteristics (a.k.a. strike-back)

[3] Accuracy of detection clearly has a role to play, including the degree of certainty to which an attack is known to be in progress. We have chosen not to require that aspect of the issue in our initial identification of required characteristics of active response: it is an area for future work.

[4] The decision to apply certain responses when certain threats are identified could still be made in advance of the particular threat, as with an automated active response.

- **Purposeful.** The reason for the action sequence must be to address a specific problem, and the organization must choose to take these particular actions.
- **Limited.** It is sufficient for the response to be intended to improve the situation.

This definition disallows reactions to threats such as retaliation and retribution, even though these might be considered by some to have deterrent value and thus be "response." We believe that it would be better to consider policies about actions in those categories separately, as the issues involved are significantly different. Also note that the term 'attack' is used, but no implications as to the motivations behind the security event are made.

We end this section by noting the prevalence of alternate terminology to active response: active defense. While *active defense* may well be more descriptive of the rationale behind the *actions* being taken, the phrase active defense has a difficult history of prior use in the military [21–23]. Militaries use the phrase active defense as meaning limited offensive action to deny a contested resource or position, implying target destruction. Since we are allowing a continuum of potential responses in our definition, the phrase active response seems more suitable.

### *C. A Model of Response*

When a system administrator chooses to manually respond to a threat, normally this is not done in a single "response", but rather in an almost interactive collection of actions. More formally, response is usually realized as a temporally ordered sequence of actions executed in a manner that supports feedback regarding the actions' effectiveness through analysis and continued detection.

For example, an attack may be suspected and actions $a_1, a_2, \ldots, a_k$ are chosen for mitigation. If the confidence level in the detection is low, then the $a_1, a_2, \ldots, a_k$ would, we anticipate, be relatively benign. Suppose, though, that ongoing analysis of the suspected attack yields new information, which leads to a greater certainty. The defender might then choose to perform additional actions $a_{k+1}, \ldots, a_{k+n}$. If the results are unsatisfactory, and/or the risk increases, a stronger action $a_{k+n+1}$ may be taken next.

How actions $a_1, a_2, \ldots, a_k$ are chosen is dependent on the decision model. One possible decision model is our prior work on ADAM [20] summarized here. ADAM supplies a framework an organization can use to inform their decisions as to the active response actions to execute. ADAM weighs multiple categories of organizational 'cost'[5] with the relative probability of success of the action to mitigate the threat. This model allows organizations to balance cost with success to obtain a sequence of graduated responses — and to indicate reasonable cutoff points for when those responses are useful. These principles of cost and success should be a guiding factor in action choice, although their balance is arguable (see [24] for a thorough discussion of military theories to guide offensive action, for example).

In the following subsections, we provide both a graphical and a vector representation to illustrate active response costs. We believe that this representation can be helpful to clarify the relationship between a response policy and the active response measures taken to mitigate a threat.

#### C.1 Vector Representation

We propose a vector representation of the cost categories, which can be mapped onto our model, to permit direct comparison of the costs. Let $\overrightarrow{C} = < C_{ethical}, C_{legal}, C_{risk}, C_{technical}, C_{unintendedconsequences} >$. Then $\overrightarrow{C}$ is the cost vector of all of the components making up the "costs" of the risks of a particular active response action.

We can use this in several useful ways, which we will only enumerate here:
1. For a given entity wishing to make a judgement about a particular action, we can establish a $\overrightarrow{CMAX}$, where $\overrightarrow{CMAX}$ is the vector of the maximum cost we're willing to accept in each category: $< C_{ethical}, \ldots >$.
2. We can define an ordering function $\leq$, where $\overrightarrow{C} \leq \overrightarrow{T}$ (where both are cost vectors for active defense actions) iff each of $C_{ethical} \leq T_{ethical}$, $C_{legal} \leq T_{legal}$, $\ldots$
3. We can establish a time-sequence or environment-dependent version of the $C_{MAX}$, so that at a given time or in a given environment we change the thresholds allowed for the costs. This allows any changes in tolerance for active defense results to be recorded over time, or over environmental change.
4. We can establish a weighting vector $W$, where $W_{ethical}$ has the weight associated with the value placed on the ethical components, likewise for legal, financial, etc. This allows different organizations to weight costs based on their own particular goals and mission. The costs which an organization wishes to include are dependent on the organization's own situation and environment.

#### C.2 Graph Representation

On the $Y$ axis we represent the cost of an active response action; where as $Y$ increases, the cost for the responder increases. The $X$ axis represents the cost of the attacker's action to the asset; where $X < 0$ are attacker activities before the initial compromise (e.g. port scanning, vulnerability scanning, and other intelligence gathering techniques), $X = 0$ is the compromise event, and $X > 0$ are attacker activities after the initial compromise (e.g. adding users, copying data, installing back-doors, etc). The line that connects response actions is determined by the policy (as described in Section VI) and is referred to as the ***policy***

[5] Cost here has to do with predicted loss to the organization for employing active defense — whether financial, legal sanction, or violation of corporate ethics.

***determinate*** because the policy is determining the actions, which by proxy determines the cost to whomever is responding.

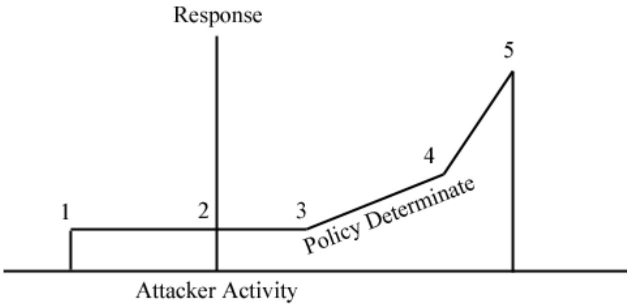


Fig. 1. An Example Response Continuum: (1) Attacker found portscanning, IDS utilizing more resources to watch for threat (2) attacker compromises system, (3) IDS detects intrusion, alerts internal and external authorities (4) files being copied, network tools used to determine source of threat (5) firewall rule changed, counter measures stop intrusion

### *D. Towards a Response Taxonomy*

This section provides a top-level view of a taxonomy for organizing active response actions, loosely based on the degree of control and the scope of the possible effects of what the responder does. This taxonomy has been adapted from [20] and [25]. We anticipate future work that will further subdivide this taxonomy or ordering.

1. *No Action*: A threat is detected, but no action is taken.
2. *Internal Notification*: Using the organizational structure to notify the designated responder(s) of an active response situation.
3. *Internal Response*: Applying active response actions within the domain over which the responder has authority (e.g. close a threat vector's associated port).
4. *External Cooperative Response*: Employing entities external to the responding organization to mitigate a threat.
5. *Non-cooperative Intelligence Gathering*: Using external services (e.g. finger, nmap, netstat, etc.) to gather intelligence on the threat. Sometimes referred to as "look but don't touch."
6. *Non-cooperative 'Cease and Desist'*: Stopping harmful and unauthorized services (e.g. zombie control processes) without compromising legitimate usability.
7. *Counter-strike*: An external action to reduce or deny the capabilities of an attacker to continue the attack.
8. *Preemptive Defense*: With knowledge of a forthcoming attack, execute active defense actions to preempt (and disable) the upcoming attack.

#### D.1 Examples of the Response Continuum

There are several publicly available systems which implement response all along the response continuum outlined in the previous section, ranging from rudimentary support for notification through relatively sophisticated and/or powerful responses. To illustrate our response continuum, we outline a few of them here.

D.1.a Internal Notification. Snort,[6] the open source intrusion detection system (IDS), is a very popular notification-based system. Most other IDSs include a notification feature.

D.1.b Internal Response. Internal response can come in the form of firewall rules, detaching a machine from the network, or simply destroying TCP connections, among other actions. An extension of Snort is flexresp2, which is an active response tool that terminates TCP connection attempts [26]. Snort can also be run in 'inline' mode allowing it to drop or modify packets that are flagged as malicious.

D.1.c Automatic or Manual External Cooperative Response. Relatively few systems have automated external cooperative response. However, they do exist. DShield,[7] a distributed IDS which collects and analyzes firewall logs from several commercial products, has implemented a 'Fightback' feature that allows an ISP to be notified if firewall logs show attacks emanating from their network.

D.1.d Non-cooperative 'Cease and Desist'. These methods are difficult to implement properly because they must stop harmful services without impinging upon the usability of a network or host. However, there are tools which are reasonably successful within limited domains. Bindview devised ZombieZapper, which acts as a zombie master in to seek shutdown all zombie bots in a network [27].

D.1.e Counter Strike. Use of counter strike is highly controversial. Lycos recently released, then retracted, the 'makelovenotspam' screensaver with some counter strike characteristics. The Lycos screensaver would continually request information from a list of websites known to send spam — effectively creating a pseudo-zombie army to launch a DDoS attack against spam sites [28]. However, ISPs began blocking this traffic as they would any DDoS attack, and public pressure mounted until Lycos ultimately removed the product from service. Additionally, Symbiot Inc. created iSMS, which can run the gamut of responses, from blocking traffic, to denial of service attacks, to gaining administrator access on an attacker's machine [29].

D.1.f Preemptive Defense. One example of preemptive defense was designed cooperatively by Mazu Networks and Asta Networks. Their tool, from an ISP's edge router, can detect and block a denial of service (DoS) attack [30] before it leaves their network. Although the attack has already been launched, from the perspective of the potential victim the attack was preempted.

[6] http://www.snort.org
[7] http://www.dshield.org

## V. Evaluating a Response Sequence

We define response evaluation using five components: ethical, legal, risk analysis, technical, and unintended consequences. Each of these can be assessed as a "cost", though "cost" in the case of ethics, etc, should be read as "relative value" rather than "money." This section briefly outlines these evaluation areas.

### A. Ethical

One hotly contested issue surrounding active response decisions is whether certain response actions are ethical, particularly those with the potential to overflow the boundaries of the responder's domain of responsibility. As an example of the distinctions drawn, many who would consider launching a DoS attack against an attacker's firewall unethical would consider modifying one's own firewall rules ethical. The key is that the calculus of response decisions should include components that allow such distinctions to be drawn clearly and rationally, in a way that can be supported within a logical framework. This is particularly important because the speed with which most responses would have to be launched will necessarily lead to automation, so codification of the ethics involved should be done in advance.

There are two frameworks we might utilize in the ethical debate: the teleological (only consequences matter) and deontological (only the actions and types of actions matter). Rawls in [31], and Davis in [32], both argue that teleological and deontological theories "exhaust the possibilities regarding theories of right action." There are strong arguments made using both approaches. Spafford [33], argues from deontology that offensive action would be unethical. Welch et al [24] argue from the teleological that if life, safety, or national security critical systems were significantly threatened, offensive action could be supported.

The key, however, is to have these discussions and make these determinations *a priori*, and integrate them within the decision making process.

### B. Legal

The legal risks to active response are significant. Not only are there questions of the applicability of existing law to a cyber domain, but also the question of which actions are permitted under the environment governing the decision-maker. This component of the calculus incorporates assessment of all facets of law, criminal, civil, domestic, international and foreign domestic. The international question is particularly difficult because the nature of networks do not (generally) limit themselves to national boundaries; and if an action is taken in another nation, then the question is if the action constitutes a 'use of force' and what diplomatic repercussions there may be.

Information warfare best informs the international issue. Grove et al [16], Barkham [34], and Yurcik [35,36] have all contemplated and analyzed the international question. Additionally, Schmitt proposed the 'Schmitt Analysis,' which is a useful tool in determining whether a cyber attack is a 'use of force' in international law [37].

With regards to domestic law in the United States, the Computer Fraud and Misuse Act (§18 USC 1030), the Wiretap Act (§18 USC 2511), and the Electronic Communications Privacy Act (§18 USC 2510), with corresponding case law, are the primary guides. There are also state statutes regarding computer trespass (see Rhode Island §11-52-3, Virginia §18.2-152.5, and the University of Dayton School of Law Model State Computer Crime Code §4.01.1 [38] for examples).

Some additional areas for consideration include the 'necessity defense' (a.k.a. choice of evils justification) (see Model Penal Code §3.02 [39]) and the 'use of force in defense of property' state statutes (see Utah §76-2-406 and North Dakota §12.1-05-06 for examples). Extending the use of force in self defense where 'self' encompasses electronic assets, is difficult. However, there are certain legal theories that may be helpful: minimal force, proportional force, and immediate (or immanent) threat. These theories are supported in both United States domestic law and International Law (Article 51 of the UN Charter [16] and the Model Penal Code §3.02 [39]). Further identification of risks and identification of the appropriateness of categories of active response can be expected from the legal community.

### C. Risk Analysis

Risk analysis must be performed properly if it is going to serve as the basis of decisions about response. However, can a satisfactory risk analysis be accomplished? Can the ethical and legal risks be realistically evaluated in the face of enormous unknowns? This question should frame ongoing research — much information would be gained if a few organizations attempt the task and report on the outcome via case study, or researchers propose models to examine in abstract, or a combination.

Given the challenges of risk analysis we propose two limiting factors. One, that the extent of risk analysis performed be treated as a due diligence requirement. Second, the risk analysis may not include ALL risks. Accuracy of prediction normally is reduced the further forward one looks, and the ripple effect of an event may take it well out of the range of what an organization is capable of assessing. In practice, risk analysis will be limited to a finite time-frame. This pragmatic consideration has implications on the ethical side (is it ethical to take external actions for which risk analysis cannot be performed properly?) and the legal side (what is due diligence in this context?).

### *D. Technical*

There are a number of technical issues when it comes to active response. Some of the more significant questions are:

1. Do contemporary response-triggering systems provide enough confidence in their alerts to base particular responses upon? Is there enough information incorporated in the alert to support response?
2. Can response be quick enough to be effective?
3. Is identification and authentication of the attacker/attacker's resource via trace-back[8] and other means viable in this environment, particularly given the prevalence of anonymity techniques and utilization of "innocent bystander" resources?

In essence, technical questions of active response can be summarized by asking whether our technology is reliable and accurate enough to execute response, particularly the more extreme forms of response on our continuum, or if the modern network environment precludes use of certain forms of response. We note that changes in technology make evaluation of this component of the response decision calculus one that changes rapidly; there are advances in intrusion detection systems (see [40, 41]) and network tracking (see [42–44]) and evaluation of confidence level that may cause evaluations in this category to change significantly.

### *E. Unintended Consequences*

Unintended consequences are placed in a separate category in our calculus of response, primarily because concern about these is a key issue for response. Discussions of unintended consequences could be placed under legal, ethical, or technological - but because of the magnitude of the issue, we prefer to treat it separately as well, to emphasize the importance of analysis in this area.

The costs of unintended consequences derive from:

1. An unintended (counter) response elicited from the attacker (i.e. you want them to stop, but an unexpected result occurs — perhaps behavior escalates or is diverted)
2. Damage to the perceived source of the threat excluding the attacker (i.e. damage to a zombie or co-opted system)
3. Unplanned damage to the responder's domain.

These sources will now be discussed.

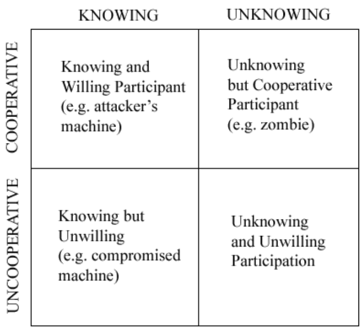
#### E.1 Unanticipated Attacker Response

It is possible that actions executed to mitigate an attack will alter the attacker's behavior. As with any form of self defense, there is always the possibility that resistance will lead to escalation (though it may also lead to cessation of the attack). Some examples of attacker tactic change in the cyber domain: they divert their attack to another resource, they become angry and launch a more vicious attack, or even alert other attackers to join the assault.[6] Without sufficient knowledge of the attacker, such as their abilities, their contacts, and resources at their disposal, this variable is outside the responder's control, and is an area where additional external information would be useful in forming a decision.

However, deception, diversion or tactical change may be useful also. The attacker could be diverted to a honeypot/honeynet or other disposable resources — which then may provide additional data as to the source of the threat and better inform any future actions against this attacker.

#### E.2 Damage to Non-Attackers

Accurate tracking, an identification/authentication aspect mentioned in the technology section is difficult to ensure. Attackers may utilize Internet Protocol (IP) spoofing, Media Access Control (MAC) spoofing, and the use of zombies, handlers, or other proxies, for instance. Any action taken outside of resources controlled by an organization or cooperating entities necessarily involve uncertainty with regards to whether the attacker has been correctly identified, and also which of the resources associated with the attack belong to the attacker in the sense of ownership and not simply controlled.



| | KNOWING | UNKNOWING |
|---|---|---|
| COOPERATIVE | Knowing and Willing Participant (e.g. attacker's machine) | Unknowing but Cooperative Participant (e.g. zombie) |
| UNCOOPERATIVE | Knowing but Unwilling (e.g. compromised machine) | Unknowing and Unwilling Participation |

Fig. 2. Attack Participation Taxonomy

For example, it is possible that the resource targeted for response was either incorrectly identified, or was engaged without the knowledge/support of the true owner of the resource. There are real risks that the target of response might be a life, safety, or national security critical system, possibly of more value than the one under attack. A byproduct of active response could be an increase of threat to critical systems, to be used as shields from active responses. This unintended consequence has both ethical and technological effects, with the potential for legal as well.

#### E.3 Damage to Own Resources

The third source of unintended consequence costs come from effects upon one's own resources. This risk is com-

[8] Trace-back is the attempt to find the attacker by tracing their network traffic through the network signaling equipment that composes contemporary networks.

mon whenever an organization makes a change in policy or design. The risk is that by changing a policy or design, the responder may unintentionally block legitimate users from a resource or harm internal assets, causing even more damage than the original attack.

## VI. The Eight Stages of Response

We propose study of the response process in eight stages: planning, detection, evaluation, decision, action, analysis, escalation, maintenance. These stages also meet our goals for the response definition in Section IV-A.

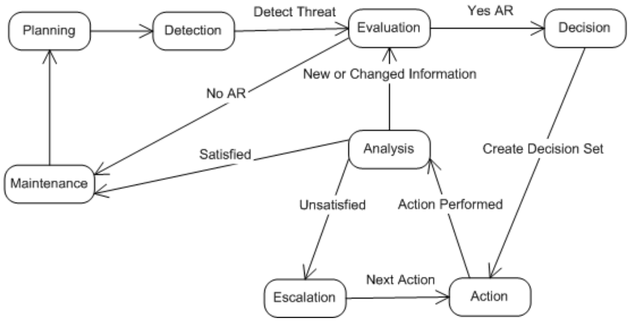


Fig. 3. The Eight Stage Response Cycle

### A. Planning

Every stage of the response process is informed by the policy developed in this first stage, giving a planning-centric model which we believe best serves response. Because of the risks and unknowns involved in assuming an active response position, this stage requires the greatest investment. However, with proper planning as incorporated with risk analysis, the risks can be reduced and active response more likely to be a viable option, or else a reasoned decision to limit response to low risk categories can be made.

In the planning stage, an ***active response policy*** is developed. The policy is an unambiguous and complete analysis of the risks and costs (in every category) of each threat and potential mitigating active response action allowing the risks and costs to be compared. The policy takes as input: the assets to be protected, the threats to those assets, the risks/costs if the threats were successful (or partially successful), and the potential mitigation active response action and corresponding risks/costs. After the inputs have been provided, then, using a scaling method, such as ADAM [20], the actions are ordered based on their relative probability of success and risk. An important step is for each threat to be assigned an unambiguous goal that defines the state at which a threat is successfully mitigated.

To develop the policy, we would anticipate involvement of an array of stakeholders, internal and external. The output of this process should be a clear analysis of all of the risks and costs involved with each asset and potential mitigating action allowing for the most informed decision.

### B. Detection

Detection is the discovery of a threat, whether automated or non-automated. In most instances it is preferable that detection occur early in the attack (e.g., during portscanning) or at least prior to damage (e.g., when an attacker moves past target identification phase). Detection ideally will provide sufficient high-confidence data with regards to the origin, method, and target of the threat to enable response — degree of confidence required for individual response options would be a matter of policy, and considered during the next phase, Evaluation.

### C. Evaluation

Evaluation takes the threat data provided by the detection stage, and compares that to the policy developed in the planning stage. The calculus of response decisions is used to identify whether active response is appropriate, and then the next step is either the decision stage (choose a response), or else the data is passed elsewhere and the cycle moves to the maintenance stage.

### D. Decision

The decision stage determines exactly which actions, identified in the active response policy as potential mitigation techniques, are selected for execution. The actions selected are placed into the ***decision set*** $(a_1, a_2, \ldots, a_k)$, an ordered set of actions that will execute in sequence until the threat is mitigated to satisfaction.

### E. Action

It is this stage that the active response action determined by the decision set is executed. After execution of an action, response moves to the analysis stage.

### F. Analysis

The purpose of the analysis is to determine whether the threat was successfully mitigated (which will invoke maintenance), or the action was unsuccessful (which will invoke escalation) according to the policy set in stage one. Analysis includes considering whether environmental changes require (re)evaluation.

### G. Escalation

Escalation refers not necessarily to increased force,[9] but rather to executing the next action described in the decision set. Escalation comes when action is unsuccessful in providing sufficient mitigation. In the escalation stage, the next action in the decision set is selected and fed into the action stage for execution.

[9] Whether a gradual increase in force is utilized, is determined by the decision model chosen.

### H. Maintenance

Maintenance is the final stage of the response cycle. Maintaining an effective active response policy is essential when implementing policy-based response. The primary goal of the maintenance stage is to take as input the results of the evaluation or analysis stage and review the policy in view of any forensic or post mortem analysis that occurs after an active response (or failed evaluation) and update the policy accordingly.

## VII. Conclusions and Future Research

There are several opportunities for future work indicated by this work. The authors would like to first identify the need for greater discussion forums for this topic. Each of the risk categories identified in Section V needs more attention and potential solutions identified. Intrusion detection systems need to support response through increased information required for response decisions, including confidence levels associated with alerts. Lastly, the authors acknowledge that the taxonomy and model presented here are only a starting point and more discussion is needed regarding the sufficiency or deficiency of these elements.

## References

[1] M. Bishop and D. Frincke, "Who watches the security educators?," *IEEE Security and Privacy*, vol. 1, no. 3, pp. 56–58, 2003.

[2] W. Schwartau, "Striking back," *Network World*, 1999.

[3] N. R. Wyler, ed., *Aggressive Network Self-Defense.* Rockland, Maryland, USA: Syngress, 2005.

[4] T. Mullen, "The right to defend," *Security Focus*, 2002.

[5] W. Schwartau, "Can you counter-attack hackers?," *CNN.com*, 2000.

[6] M. Delio, "'stung' russian hacker guilty," *Wired News*, 2001.

[7] CNN, "Teen hacker faces federal charges," *CNN.com*, 1998.

[8] K. Poulsen, "Slammer worm crashed ohio nuke plant network," *Security Focus*, 2003.

[9] V. Jayawal, W. Yurcik, and D. Doss, "Internet hack back: Counter attacks as self-defense or vigilantism?," in *International Symposium on Technology and Society*, (Raleigh, North Carolina), pp. 380–386, 2002.

[10] T. Mullen, *Neutralizing Nimda: Technical, Moral and Legal Discussions of an Automated Strike-back.* Defcon, http://www.blackhat.com/presentations/bh-usa-02/bh-us-02-mullen-nimda.ppt, 2002.

[11] T. Mullen, *Defending Your Right to Defend: Considerations of an Automated Strike-back Technology.* http://www.hammerofgod.com/strikeback.txt, 2002.

[12] Reuters, "Computer under attack can hack back, expert says," *SiliconValley.com*, 2002.

[13] B. Schneier, *Counterattack.* Cryto-Gram Newsletter: Dec 15, 2002, http://www.schneier.com/crypto-gram-0212.html, 2002.

[14] T. Mullen, "Strikeback, part deux," *Security Focus*, 2003.

[15] C. Loomis, "Appropriate response: More questions than answers," November 28, 2001 2001.

[16] G. D. Grove, S. E. Goodman, and S. J. Lukasik, "Cyber-attacks and international law," *Survival*, vol. 42, no. 3, pp. 89–104, 2000.

[17] D. Bruschi and E. Rosti, "Disarming offense to facilitate defense," in *2000 Workshop on New Security Paradigms*, (Ballycotton, County Cork, Ireland), pp. 69–75, ACM Press, 2000.

[18] D. Bruschi and E. Rosti, "Angel: A tool to disarm computer systems," in *2001 Workshop on New Security Paradigms*, (Cloudcroft, New Mexico), pp. 63–69, ACM Press, 2001.

[19] D. Bruschi, C. L., and E. Rosti, "Less harm, less worry or how to improve network security by bounding system offensiveness," in *16th Annual Computer Security Applications Conference*, (New Orleans, Louisiana), pp. 188–195, IEEE Computer Society, 2000.

[20] S. Caltagirone and D. Frincke, "Adam: Active defense algorithm and model," in *Aggressive Network Self-Defense* (N. R. Wyler, ed.), pp. 287–311, Rockland, MD, USA: Syngress Publishing, 2005.

[21] V. D. Sokolovskii, *Soviet Military Strategy.* Englewood Cliffs, New Jersey: Prentice-Hall, 1963.

[22] J. P. 1-02., "Department of defense dictionary of military and associated terms," 2001.

[23] J. P. 3-40., "Joint doctrine for combating weapons of mass destruction," 2004.

[24] D. J. Welch, N. Buchheit, and A. Ruocco, "Strike back: Offensive actions in information warfare," in *1999 Workshop on New Security Paradigms*, (Caledon Hills, Ontario, Canada), pp. 47–52, ACM Press, 1999.

[25] D. Dittrich, *Active Defenses To Cyber Attacks.* University of Washington Information School: Agora Workshop, September 12, 2003.

[26] J. Nathan, *Snort flexresp2 README.* http://cerberus.sourcefire.com/ jeff/archives/snort/, 2004.

[27] D. Radcliff, "Hack back," *Network World*, May 29 2000.

[28] P. Roberts, "Lycos, spammers trade blows," *PCWorld.com*, 2004.

[29] S. Gaudin, "Plan to counterattack hackers draws more fire," *Internet News*, 2004.

[30] R. Tadjer, "Detect, deflect, destroy," *Internet Week*, 2000.

[31] J. Rawls, *A Theory of Justice.* Cambridge, Massachusetts: Harvard University Press, 1999.

[32] N. Davis, "Contemporary deontology," in *A Companion to Ethics* (P. Singer, ed.), pp. 205–218, Cambridge, Massachusetts: Basil Blackwell, 1991.

[33] E. Spafford, "Are computer hacker break-ins ethical?," in *Computers and Ethics in the Cyberage* (D. M. Hester and P. J. Ford, eds.), pp. 332–344, Upper Saddle River, New Jersey: Prentice Hall, 2001.

[34] J. Barkham, "Information warfare and international law on the use of force," *New York University Journal of International Law and Politics*, vol. 34, pp. 57–113, 2001.

[35] W. Yurcik, "Information warfare survivability: Is the best defense a good offense?," in *5th Annual Ethics and Technology Conference*, (Loyola University, Chicago, IL), 2000.

[36] W. Yurcik and D. Doss, "Internet attacks: A policy framework for rules of engagement," in *29th Research Conference on Communcation, Information, and Internet Policy* (L. F. Cranor, ed.), (Alexandria, VA), MIT Press, 2001.

[37] M. N. Schmitt, "Bellum americanum: The us view of twenty-first century war and its possible implications for the law of armed conflict," *Michigan Journal of International Law*, vol. 19, no. 4, pp. 1050–1090, 1998.

[38] S. Brenner, "Model state computer crimes code," 2001 2001.

[39] A. L. Institute., *Model penal code: official draft and explanatory notes: complete text of model penal code as adopted at the 1962 annual meeting of the American Law Institute at Washington, D.C., May 24, 1962.* Philadelphia, Pa.: American Law Institute, 1985.

[40] X. Wang, D. S. Reeves, and S. F. Wu, "Tracing based active intrusion response," *Journal of Information Warfare*, vol. 1, no. 1, 2001.

[41] D. Yu and D. Frincke, "Alert confidence fusion in intrusion detection systems with extended dempster-shafer theory," in *43rd ACM Annual Southeast Conference*, (Kennesaw, GA), 2005.

[42] T. Kohno, A. Broido, and K. Claffy, *Remote Physical Device Fingerprinting.* to appear in IEEE Symposium on Security and Privacy, 2005.

[43] T. W. Doeppner, P. N. Klein, and A. Koyfman, "Using router stamping to identify the source of ip packets," in *7th ACM Conference on Computer and Communications Security*, (Athens, Greece), pp. 184–189, ACM Press, 2000.

[44] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Practical network support for ip traceback," *ACM SIGCOMM*, vol. 30, no. 4, pp. 295–306, 2000.