Optimal Cost, Collaborative, and Distributed Response to Zero-Day Worms - A Control Theoretic Approach

Senthilkumar G. Cheetancheri[1,⋆], John-Mark Agosta[2], Karl N. Levitt[1], Felix Wu[1], and Jeff Rowe[1]

[1] Security Lab, Dept. of Computer Science, Univ. of California, One Shields Ave., Davis, CA - 95616, USA
[2] Intel Research. 2200, Mission College Blvd., Santa Clara, CA - 95052, USA
cheetanc@cs.ucdavis.edu

Abstract. Collaborative environments present a happy hunting ground for worms due to inherent trust present amongst the peers. We present a novel control-theoretic approach to respond to zero-day worms in a signature independent fashion in a collaborative environment. A federation of collaborating peers share information about anomalies to estimate the presence of a worm and each one of them independently chooses the most cost-optimal response from a given set of responses. This technique is designed to work when the presence of a worm is uncertain. It is unique in that the response is dynamic and self-regulating based on the current environment conditions. Distributed Sequential Hypothesis Testing is used to estimate the extent of worm infection in the environment. Response is formulated as a Dynamic Programming problem with imperfect state information. We present a solution and evaluate it in the presence of an Internet worm attack for various costs of infections and response.

A major contribution of this paper is analytically formalizing the problem of optimal and cost-effective response to worms. The second contribution is an adaptive response design that minimizes the variety of worms that can be successful. This drives the attacker towards kinds of worms that can be detected by other means; which in itself is a success. Counter-intutive results such as leaving oneself open to infections being the cheapest option in certain scenarios become apparent with our response model.

Keywords: Worms, Collaboration, Dynamic Programming, Control Theory.

1 Introduction

Computer worms are a serious problem. Particularly in a collaborative environment, where the perimeter is quite secure but there is some amount of trust and implicit security within the environment. Once a worm breaks the perimeter

⋆ Corresponding Author.

R. Lippmann, E. Kirda, and A. Trachtenberg (Eds.): RAID 2008, LNCS 5230, pp. 231–250, 2008.

defense, it essentially has a free run within the collaborative environment. An enterprise environment is a typical example of a network with this 'crunchy on the outside – chewy on the inside' characteristic. In this paper, we try to leverage the collaboration to collectively defend against such worm attacks. Dealing with known worms is a solved problem – signatures to be used by Intrusion Prevention Systems(IPSs) are developed to prevent further infections, and patches are developed to fix vulnerabilities exploited by these worms. Dealing with unknown worms – worms that exploit zero-day vulnerabilities or vulnerabilities for which patches have either not been generated or not applied yet – is still a research question. Several ingenious proposals to detect them automatically exist. Many sophisticated counter measures such as automatic signature generation and distribution [17,13,16,20] and automatic patch generation to fix vulnerabilities [18] have also been developed.

Often times, even if automated, there is not much time to either generate or distribute signatures or patches. Other times, system administrators are skeptical about applying patches. During instances when response based on the above mentioned techniques are not feasible, the only option left is to either completely shut-down the vulnerable service or run it risking infection. It is usually preferred to shut-down the service briefly until a mitigating response is engineered manually.

However, making a decision becomes hard when one is not certain if there is really a worm, and if the service being offered is vulnerable to it. It is not desirable to shut-down a service only to realize later that such an action was unwarranted because there is no worm. However, suspending the service in an attempt to prevent infection is not considered bad. Intuitively, it is desired to suspend the service briefly until it is clear whether there is an attack or not. Balancing the consequences of providing the service risking infection against that of not providing the service is of the essence.

This paper captures this intuition and devises an algorithm using Dynamic Programming(DP) techniques to minimize the overall cost of response to worms. Cost is defined as some mathematical expression of an undesirable outcome.

These algorithms use information about anomalous events that are potentially due to a worm from other co-operating peers to choose optimal response actions for local application. Such response can be later rolled-back in response to changes to the environment such as a curtailed worm. Since peers decide to implement response independently, the response is completely decentralized.

We surprisingly found that in certain scenarios, leaving oneself open to infection by the worm might be the least expensive option. We also show that these algorithms do not need large amounts of information to make decisions. One of the key achievements here is that we use weak Intrusion Detection Systems(IDSs) as sensors that have high false positive rates. By corroborating alerts raised by them with other collaborating sensors, we are able to minimize the false positives and achieve better fidelity in detecting worms.

2 Dynamic Programming

This section provides a brief introduction to the theory behind Dynamic Programming [6]. DP as applied to the current problem balances the low costs presently associated with operating a system against the undesirability of high future costs. The basic model of such a system is dynamic and discrete with an associated cost that is additive over time. The evolution of such a system can be described as:

$$x_{k+1} = f_k(x_k, u_k, w_k), \quad k = 0, 1, \ldots, N-1 \ , \tag{1}$$

where k indexes discrete time, x_k is the state of the system and summarizes past information that is relevant for future optimization, u_k is the control or decision variable to be selected at time k, w_k is a random parameter, also called disturbance or noise depending on the context, N is the horizon or the number of times control is applied and f_k is the mechanism by which the state is updated. The cost incurred at time k is denoted by $g_k(x_k, u_k, w_k)$, which is a random function because it depends on w_k. The goal is to minimize the total *expected cost*

$$J_\pi(x_0) = \underset{w_k}{E}\left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\} \ .$$

This is achieved by finding a sequence of functions called the *policy* or *control law*, $\pi = \{\mu_0, \ldots, \mu_{N-1}\}$, where each $\mu_k(x_k) \rightarrow u_k$ when applied to the system takes it from state x_k to x_{k+1} and minimizes the total *expected cost*. In general, for a given π, we use $J_k(x_k)$ to denote the *cost-to-go* from state x_k at time k to the final state at time N.

Dynamic Programming Algorithm: The optimal total cost is given by $J_0(x_0)$ in the last step of the following algorithm, which proceeds backwards in time from period $N-1$ to period 0:

$$J_N(x_N) = g_N(x_N), \tag{2}$$

$$J_k(x_k) = \min_{u_k} \underset{w_k}{E}\Big\{ g_k(x_k, u_k, w_k) + J_{k+1}(x_{k+1}) \Big\}, \quad k = 0, 1, \ldots, N-1 \ . \tag{3}$$

2.1 Imperfect Information Problems

DP problems as described above have perfect information about the state of the system, x_k. Often, x_k cannot be determined accurately; only an estimate,

$$z_k = h_k(x_k, v_k) \ , \tag{4}$$

can be made, where h_k is a sensor that maps x_k and a random disturbance v_k, into an observation, z_k. Such problems are solved by reformulating them into a perfect state information problem by introducing an augmented state variable I_k, which is a vector of the past observations and controls applied.

$$I_{k+1} = (I_k, z_{k+1}, u_k), \quad k = 0, 1, \ldots, N-2 \, ,$$
$$I_0 = z_0 \ . \tag{5}$$

3 Response Formulation with Imperfect State Information

In this section we formulate the computer worm response problem as a DP problem with imperfect state information. We assume that there could be only one worm and that the worm is a random scanning worm. We also assume that there is a sensor, such as an IDS albeit not very accurate. This DP formulation tells us which control should be applied to minimize the costs incurred until the worm detection process is complete.

3.1 Problem Statement

System Evolution: Consider a machine that provides some service. This machine needs to be operated for N steps or N time units. This machine can be in one of two states, P or $\overline{P}$, corresponding to the machine being in proper(desired state) or improper(infected by a worm) state respectively. During the course of operating the machine, it goes from state P to $\overline{P}$ with a certain probability λ and remains in state P with a probability $\overline{\lambda} = (1-\lambda)$. If the machine enters state $\overline{P}$, it remains there with probability 1. The infectious force λ, is an unknown quantity and depends on how much of the Internet is infected with the worm, if at all a worm is present.

Sensor: The machine also has a *sensor*, which inspects the machine for worm infections. However, it cannot determine the exact state of the machine. Rather, it can only determine the state of a machine with a certain probability. There are two possible observations; denoted by G (good, probably not infected) and B(bad, probably worm infected). Alternatively, instead of infections, we can imagine that the *sensor* looks for infection attempts and anomalies. The outcome would then indicate that there is probably a worm on the Internet (B) or not (G) as opposed to whether the host machine is infected or not. It is this latter interpretation we adopt for the rest of this paper. For the time being, let us assume that the inspections happen proactively at random intervals and also when alerts are received from peers. We also assume that the *sensor*'s integrity is not affected by the worm.

Controller: The machine also includes a *controller* that can continue(C) or stop(S) operating the machine. The machine cannot change states by itself if it is stopped. Thus the *controller* can stop the machine to prevent a worm infection and start it when it deems it safe to operate the machine. There are certain costs involved with each of these actions under different conditions as described in the next paragraph. The controller takes each action so that the overall cost of operating the machine for N steps is minimized.

Costs: Continuing(C) to operate the machine when it is in state P costs nothing. It is the nominal. We incur a cost of τ_1 for each time step the machine is stopped(S) irrespective of whether it is infected or not, and a cost τ_2 for each step an infected machine is operated. One might argue that τ_1 and τ_2 should be

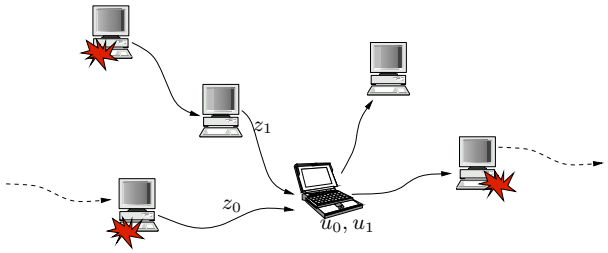

Fig. 1. Alert Sharing Protocol. The laptop is our machine of interest. It uses information, z_0 and z_1, from different chains to choose, actions, u_0 and u_1. It may or may not have seen an anomaly while the machines shown with a blast have seen an anomaly.

the same because an infected machine is as bad as a stopped machine. If that argument is true, the problem becomes trivial and it can be stated right away that the most cost effective strategy is to operate the machine uninterrupted until it is infected. On the contrary, we argue that operating an infected machine costs more as it can infect other machines also. Hence, $\tau_2 > \tau_1$.

Alert Sharing Protocol: Since a computer worm is a distributed phenomenon, inspection outcomes at one machine is a valid forecast of the outcome from a later inspection at another identical machine. (This is an assumption we make to develop the formulation and will be relaxed later on when we discuss a practical application.) Hence, a collection of such machines with identical properties seek to co-operate and share the inspection outcomes. Under this scheme, an inspection outcome at one machine is transmitted to another co-operating peer chosen randomly. The *controller* on the randomly chosen machine uses such received messages to select the optimal control to apply locally. This has the effect of a machine randomly polling several neighbors to know the state of the environment. This gives the uninfected machines an opportunity to take actions that prevent infection. Refer to Fig. 1. In addition to the inspection outcome, peers share information about the anomaly observed in what we call an *anomaly vector* – the structure, form and generation of which we leave undefined. Any two peers observing the same anomaly generate identical *anomaly vector*s.

Goal: Now, the problem is to determine the policy that minimizes the total expected cost of operating the machine for N time periods in an environment that could possibly be infected with a worm. DP problems are generally plagued with state space explosion with increasing number of stages to the horizon. However, since we solve the DP formulation of our problem offline, the value of N does not have any impact on the operational efficiency of the model. Moreover, DP problems can be solved approximately, or analytically for larger Ns significantly reducing the computational needs of the original formulation. The rest of this section develops the formulation for the current problem and provides a solution for $N = 3$. Computer generated results for larger Ns are presented and discussed in later sections.

3.2 Problem Formulation

The above description of the problem fits the general framework of Sect. 2.1, "Problems with imperfect state information." The state, control and observation variables take values as follows:

$$x_k \in \{P, \overline{P}\}, \quad u_k \in \{C, S\}, \quad z_k \in \{G, B\} \ .$$

The machine by itself does not transit from one state to another. Left to itself, it remains put. It is transferred from P to $\overline{P}$ only by a worm infection, a random process – an already infected victim chooses this machine randomly. The evolution of this system follows (1), and is shown in Fig. 2. The function f_k of (1) can be derived from Fig. 2 as follows:

$$\begin{array}{l} P(x_{k+1} = P \mid x_k = P,\, u_k = C) = \overline{\lambda}\,, \\ P(x_{k+1} = \overline{P} \mid x_k = P,\, u_k = C) = \lambda\,, \\ \vdots \\ P(x_{k+1} = \overline{P} \mid x_k = \overline{P},\, u_k = S) = 1 \ . \end{array} \tag{6}$$

The random disturbance, w_k is provided by λ and is rolled in x_k. λ is the infectious force, a function of the number of the machines infected on the Internet. Assuming the machine initially starts in state P, the probability distribution of x_0 is

$$P(x_0 = P) = \overline{\lambda}, \qquad P(x_0 = \overline{P}) = \lambda \ . \tag{7}$$

(This assumption is for exposition only. In practice, we do not have to know the initial state the machine starts in.) Recollect that the outcome of each inspection of the machine is an imperfect observation of the state of the system. Thus,

$$\begin{array}{l} P(z_k = G \mid x_k = \overline{P}) = \mathrm{fn}\,, \\ P(z_k = B \mid x_k = \overline{P}) = (1 - \mathrm{fn})\,, \\ P(z_k = G \mid x_k = P) = (1 - \mathrm{fp})\,, \\ P(z_k = B \mid x_k = P) = \mathrm{fp} \ , \end{array} \tag{8}$$

where fp and fn are properties of the *sensor*s denoting the false positive and false negative (miss) rates.

Assuming the cost function remains the same regardless of time, the sub-script k can be dropped from g_k. We define the cost function as follows:

$$\begin{array}{l} g(P, C) = 0, \qquad g(\overline{P}, C) = \tau_2, \\ g(P, S) = g(\overline{P}, S) = \tau_1, \\ g(x_N) = 0. \end{array} \tag{9}$$

$g(x_N) = 0$ because u_N is chosen with accurate knowledge of the environment, (i.e) whether there is a worm or not. If there is a worm, $u_N = S$, else $u_N = C$.

Our problem now is to find functions $\mu_k(I_k)$ that minimize the total expected cost

$$\underset{x_k, z_k}{E} \left\{ g(x_N) + \sum_{k=0}^{N-1} g\big(x_k, \mu_k(I_k)\big) \right\} \ .$$

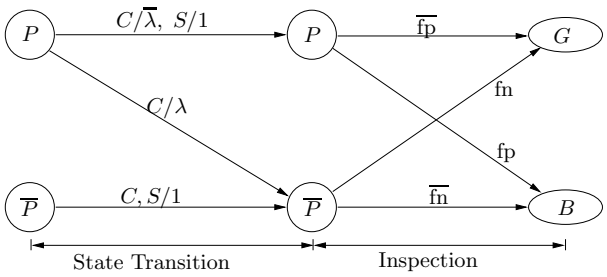

Fig. 2. The left half of the figure shows the state transition probabilities for each action. For example, the system goes from state P to P with a probability of $\overline{\lambda}$ when action C is applied. The right half of the figure shows the observation probabilities for each state. For example, when the system is in state P, the sensors output a G with a probability of $\overline{\text{fp}}$.

We now apply the DP algorithm to the augmented system (refer Sect. 2.1). It involves finding the minimum cost over the two possible actions, C and S, and has the form:

$$J_k(I_k) = \min_{\{C,S\}} \Bigg[\Big(P(x_k = P \,|\, I_k, C) \cdot g(P, C) + P(x_k = \overline{P} \,|\, I_k, C) \cdot g(\overline{P}, C) \Big) + \underset{z_{k+1}}{E} \Big\{ J_{k+1}(I_k, C, z_{k+1}) \,|\, I_k, C \Big\} \;, \Big(P(x_k = P \,|\, I_k, S) \cdot g(P, S) + P(x_k = \overline{P} \,|\, I_k, S) \cdot g(\overline{P}, S) \Big) + \underset{z_{k+1}}{E} \Big\{ J_{k+1}(I_k, S, z_{k+1}) \,|\, I_k, S \Big\} \Bigg] \quad (10)$$

where $k = 0, 1, \ldots N-1$ and the terminal condition is $J_N(I_N) = 0$. Applying the costs (9), and noticing that $P(x_k = P \,|\, I_k, S) + P(x_k = \overline{P} \,|\, I_k, S)$ is the sum of probabilities of all elements in a set of exhaustive events, which is 1, we get

$$J_k(I_k) = \min_{\{C,S\}} \Bigg[\tau_2 \cdot P(x_k = \overline{P} \,|\, I_k, C) \; + \underset{z_{k+1}}{E} \Big\{ J_{k+1}(I_k, C, z_{k+1}) \,|\, I_k, C \Big\}, \; \tau_1 \; + \underset{z_{k+1}}{E} \Big\{ J_{k+1}(I_k, S, z_{k+1}) \,|\, I_k, S \Big\} \Bigg] \;. \quad (11)$$

This is the required DP formulation of response to worms. Next, we demonstrate a solution derivation to this formulation for $N = 3$.

3.3 Solution

Here we show a solution assuming that we expect to know with certainty about the presence of a worm at the receipt of the third message, that is, $N = 3$. The same procedure can be followed for larger Ns.

With that assumption, control u_2 can be determined without ambiguity. If the third message says there is a worm, we set $u_2 = S$, else we set it to C. This also means that the cost to go at that stage is

$$J_2(I_2) = 0 \ . \qquad \text{(Terminal Condition)}$$

Penultimate Stage: In this stage we determine the cost $J_1(I_1)$. Applying the terminal condition to the DP formulation (11), we get

$$J_1(I_1) = \min \Big[\tau_2 \cdot P(x_1 = \overline{P} \mid I_1, C) \ , \ \tau_1 \Big] \ . \tag{12}$$

The probabilities $P(x_1 = \overline{P} | I_1, C)$ can be computed using Bayes' rule and (6–8), assuming the machine starts in state P. (See Sect. B for exposition.) The cost for each of the eight possible values of $I_1 = (z_0, z_1, u_0)$ under each possible control, $u_1 \in \{C, S\}$ is computed using (11). Then, the control with the smallest cost is chosen as the optimal one to apply for each z_1 observed. The *cost-to-go*, $J_1(I_1)$, thus calculated are used for the zeroth stage.

Stage 0*:* In this stage we determine the cost $J_0(I_0)$. We use (11) and values of $J_1(I_1)$ calculated during the previous stage to compute this cost. As before this cost is computed for each of the two possible values of $I_0 = (z_0) = \{G, B\}$, under each possible control, $u_1 = \{C, S\}$. Then, the control with the smallest cost is chosen as the optimal one to apply for the observed state of the machine. Thus we have,

$$J_0(I_0) = \min \Bigg[\tau_2 \cdot P(x_0 = \overline{P} \mid I_0, C) \ + \underset{z_1}{E} \Big\{ J_1(I_1) \mid I_0, C \Big\} \ , \\ \tau_1 + \underset{z_1}{E} \Big\{ J_1(I_1) \mid I_0, S \Big\} \Bigg] \ . \tag{13}$$

The optimal cost for the entire operation is finally given by

$$J^* = P(G) J_0(G) + P(B) J_0(B) \ .$$

We implemented a program that can solve the above formulation for various values of $\lambda, \text{fp}, \text{and fn}$. A sample rule-set generated by that program is given in Table 1. Armed with this solution, we now show a practical application.

4 A Practical Application

4.1 Optimal Policy

Table 1 shows the optimal policies for a given set of operational parameters. The table is read bottom up. At start, assuming the machine is in state P, the optimal action is to continue, C. In the next time step, stage 0, if the observation is B, the optimal action is to stop, S. If $z_0 = B$ is followed by $z_1 = G$, the optimal action is to operate the machine, C. This is denoted by the second line in

Table 1. An optimal policy table

$\lambda = 0.50$, fp $= 0.20$, fn $= 0.10$ $\tau_1 = 1$, $\tau_2 = 2$			
	I_k	J_k	u_k
Stage 1	(G, G, S)	0.031	C
	(B, G, S)	0.720	C
	(G, B, S)	0.720	C
	(B, B, S)	1.000	S
	(G, G, C)	0.270	C
	(B, G, C)	1.000	S
	(G, B, C)	1.000	S
	(B, B, C)	1.000	S
Stage 0	(G)	0.922	C
	(B)	1.936	S
Start		1.480	C

stage 1. This shows that an undesirable response is rolled back when the environment is deemed not dangerous. In a practical application, such a table will be looked up for a given λ and observation to choose the optimal action. Note that the first, third, sixth and eighth states are unreachable because, for the given z_0, the control u_0 mentioned in the vector is never applied if the system operates in good faith.

4.2 Choosing λ

The value of λ varies with the extent of infection in the Internet. Given we are uncertain that there is a worm in the Internet, λ cannot be determined with any accuracy. Rather, only estimates can be made. Hence the distributed Sequential Hypothesis Testing developed earlier is used to estimate λ [9].

Given a sequence of observations $\boldsymbol{y} = \{y_0, y_1, \ldots, y_n\}$, made by a sequence of other participating nodes, and two contradicting hypotheses that there is a worm on the Internet(H_1) and not(H_0), the former is chosen when the likelihood ratio $L(\boldsymbol{y})$ of these hypotheses is greater than a certain threshold η [9]. This threshold η is determined by the performance conditions required of the algorithm. Assuming the observations are independent, $L(\boldsymbol{y})$ and η are defined as follows:

$$L(\boldsymbol{y}) = \prod_{i=1}^{n} \frac{P(y_i|H_1)}{P(y_i|H_0)}, \quad \eta = \frac{DD}{DF} \ , \tag{14}$$

where DD is the minimum desired detection rate and DF is the maximum tolerable false positive rate of the distributed Sequential Hypothesis Testing(dSHT) algorithm. We define each of the above probabilities as follows:

$$P(y_k = B \mid H_1) = [\lambda\,(1 - \text{fn}) + (1 - \lambda)\,\text{fp}]\,,$$
$$P(y_k = G \mid H_1) = [(\lambda\,\text{fn}) + (1 - \lambda)(1 - \text{fp})]\,,$$

$$P(y_k = B \mid H_0) = \text{fp}\,, \qquad (15)$$
$$P(y_k = G \mid H_0) = (1 - \text{fp})\ .$$

The first equation in the above set is the probability of observing a B given hypothesis H_1 is true. It is the sum of probability of getting infected (λ) times the probability of detection, and the probability of not getting infected($1 - \lambda$) times the probability of false positives. The others in (15) are defined similarly.

For a received sequence of observations, a node calculates $L(\boldsymbol{y})$ for several values of λ – say for ten different values in steps of 0.1 starting at 0.1. The lowest λ for which the $L(\boldsymbol{y})$ exceeds η is then taken as the current levels of infection and used in determining the optimal response. The reason for choosing discrete values of λ will be apparent shortly.

An observation at a node can be conveyed to another by transmitting the observation vector, $\boldsymbol{y} = \{y_0\}$. The recepient can add its own observation to this vector making it $\boldsymbol{y} = \{y_0, y_1\}$. Such a sequence accumulates information leading to larger vectors with each hop. Given $L(\boldsymbol{y})$ in (14) is essentially a digest of such vectors, no node has to transmit a whole vector. Instead, it suffices to transmit just one number, $L(\boldsymbol{y})$. A recepient can update the received $L(\boldsymbol{y})$ using (14), (15), and its own observations. It is indeed a conundrum to estimate λ using (15), which is a function of λ itself. This problem is solved as described in the previous paragraph – the lowest λ for which $L(\boldsymbol{y})$ exceeds η is taken as the current operating λ.

In operational practice, a policy in the form of a table is calculated offline for several candidate values of λ. Each row in these tables gives a u_k for a given I_k. For each new λ estimated, the corresponding table is consulted to choose u_k given I_k, where I_k is the node's own past observations and corresponding actions. Thus, each node only receives a likelihood ratio of the worm's presence from its peers and also has to remember only its own I_k. Limiting the number of such tables is the reason for choosing discrete λs in the preceeding paragraphs.

4.3 Larger *N*s

As N increases, the dimensions of I_k increases, which in turn increases the number of the calcuations involved exponentially. This problem can be overcome by reducing I_k to smaller dimensions containing only the *Sufficient Statistics* yet summarizing all essential contents of I_k as far as control is concerned. There are many different functions that can serve as *sufficient statitics*. The one we use here is the conditional probability distribution $P_{x_k|I_k}$ of the state x_k, given the information vector I_k [6]. Discrete-time stochastic systems can be described by the evolution

$$P_{x_{k+1}|I_{k+1}} = \Phi_k(P_{x_k|I_k}, u_k, z_{k+1})\,, \qquad (16)$$

where Φ_k is a function that estimates the probalistic state of the system $P_{x_k|I_k}$ based on $P_{x_{k-1}|I_{k-1}}$, z_k and u_{k-1} , and can be determined from the data of the problem [5]. Figure 3 explains this concept. The actuator $\overline{\mu}_k$ then selects the optimal response based on $P_{x_k|I_k}$.

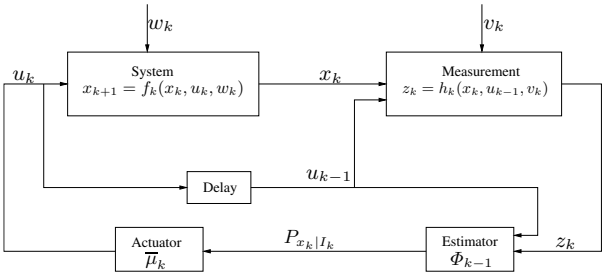

Fig. 3. The controller split into an *Estimator* and an *Actuator*. The *Estimator* Φ_{k-1} estimates the probabilistic state of the system $P_{x_k|I_k}$ while the *Actuator* $\overline{\mu}_k$ picks the appropriate control u_k.

This re-formulation makes it easy to apply the response model for larger Ns. We implement this model and evaluate it in a simulation. The evaluation and the results are discussed in the next section.

5 Evaluation

The sufficient statistics formulation discussed in the previous section was implemented and evaluated with a discrete event simulation. The simulation consisted of a world of 1000 participants with 10% of the machines being vulnerable. We set the number of stages to operate the machine, $N = 4$ to calculate the rule-sets. Note that $N = 4$ is used only to calculate the rule-sets but the machines can be operated for any number of steps. N is essentially the number of past observations and actions that each machine remembers. The local IDSes were set to have a false positive and false negative rates of 0.1. These characteristics of the local IDS is used to calculate the probability of infection, λ with a desired worm detection rate of 0.9 and failure rate of 0.1. In all the following experiments, we used a random scanning worm that scans for vulnerable machines once every unit-time.

5.1 Experiments

Parameters of Evaluation: A set of experiments was designed to understand the effect of various parameters on the effectiveness of the model in controlling the spread of the worm. The only free variable we have here is the ratio τ_2/τ_1. There is no one particular parameter that can measure or describe the effectiveness of the response model. Rather, the effectiveness is described by the number of vulnerable machines that are not infected and of those the number that provide service, i. e. in state C.

Algorithm: The algorithm for the discrete-event simulation is as follows. At each time cycle.

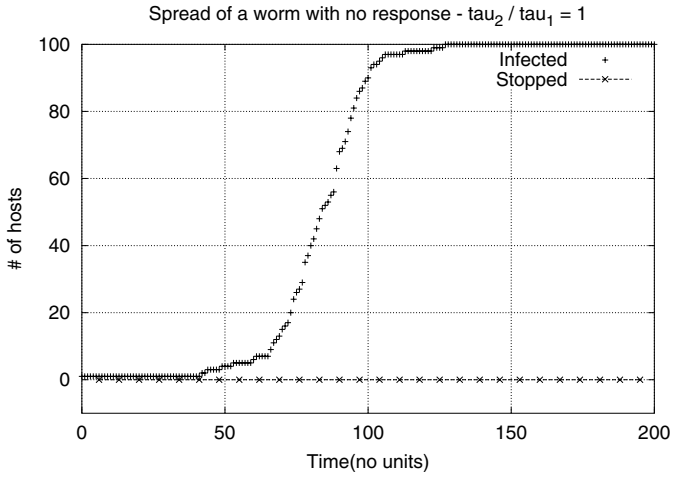

Fig. 4. No machines are stopped when the cost of being infected is the same as cost of shutting down the machine. fp = fn = 0.1, $DD = 0.9$, $DF = 0.1$

- all infected machines attempt one infection,
- all machines that had an alert to share, share the likelihood ratio that there is a worm on the Internet with another randomly chosen node,
- and all vulnerable machines that received an alert earlier take a response action based on the information received and the current local observations.

Results: In the first experiment, we want to make sure that we have a worm that behaves as normal random scanning worm and validate the response model for

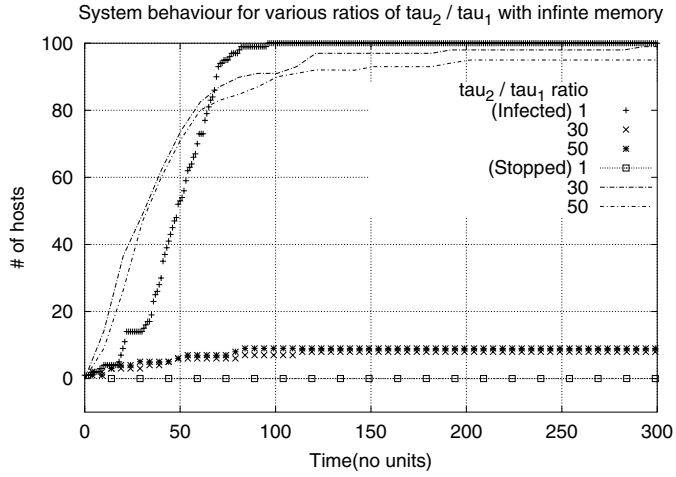

Fig. 5. When nodes are set to remember infection attempts forever, they never back-off their defensive posture. Once entered the S state, a machine stays there. fp = fn = 0.1, $DD = 0.9$, $DF = 0.1$.

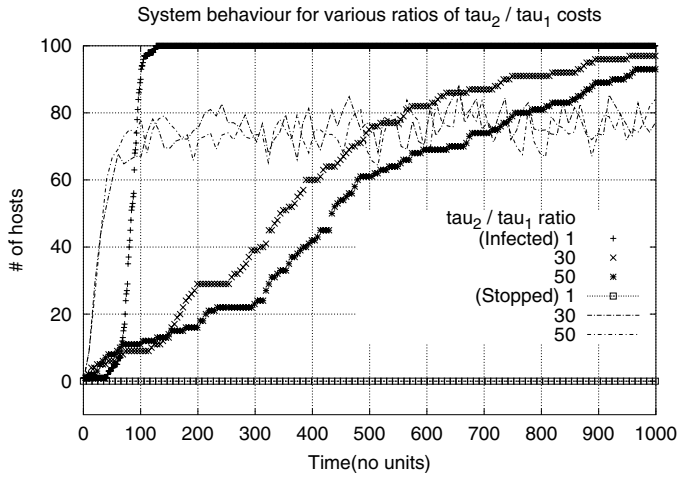

Fig. 6. Higher costs of being infected invoke stricter responses. fp = fn = 0.1, $DD = 0.9, DF = 0.1$.

the degenerate cases. We verify this by providing no response. This response can be achieved by setting the cost ratio to 1 – the cost of stopping the service is the same as getting infected. In this scenario, we expect the response model not to take any defensive measures against suspected infection attempts. As expected, we see in Fig. 4, that none of the machines are stopped (S state). The worm spreads as it would spread when there is no response in place. This validates our worm and also our response model.

As another sanity check we set the machines to remember infection attempts forever. Under this policy, once a machine enters the S state, it remains in that state forever. We see that in this case (Fig. 5) the number of machines infected are very low except when $\tau_2/\tau_1 = 1$.

In the next experiment, we try to understand the behavior of our response model in various situations. Since the only free variable is the ratio τ_2/τ_1, we repeat the previous experiment with various values for that ratio. The results for this set of experiments is shown in Fig. 6. This graph shows behavior of our response model in three different tests. There are two different curves for each test indicating the number of vulnerable machines being infected and the number of machines that are stopped. We can see that when the ratio is 1, the number of machines that are in S state is 0. As the ratio τ_2/τ_1 rises, the response becomes stricter. We see that the number of machines in the stopped(S) state is higher when the cost of being infected is higher. Also the worms spreads significantly slower than without any response in place or with a lower τ_2/τ_1 ratio.

5.2 Effects of Increasing N

The experiments shown earlier in this section were all conducted with $N = 4$. An interesting question to ask here, "What happens if we increase the value of

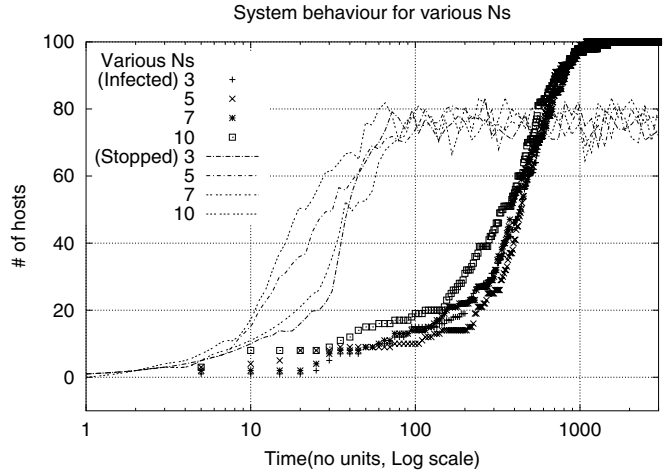

Fig. 7. Larger Ns do not contribute much to improve performance due to the small number of dimensions to the state, $x_k \in \{P, \overline{P}\}$. fp = fn = 0.1, $DD = 0.9$, $DF = 0.1$

N?". Fig. 7 shows the performance of the system for various values of N while holding the ratio of τ_2/τ_1 constant at 30. The set of sigmoidal curves that increase monotonically trace the growth of the worm, while the other set of curves trace the number of nodes that are shut-down at any given instant. We notice that there is no appreciable slowing of the worm with increased values of N – all the worm growth curves are bunched up together. This is due to the small number of dimensions to the state, $x_k \in \{P, \overline{P}\}$. A larger observation space does not contribute much to improve the performance of the system.

6 Conclusion

This section concludes this paper by reflecting on the strengths and weaknesses of the approach discussed so far. Assumptions are identified. Arguments are made for the choice of certain design and evaluation decisions. Where appropriate, future directions are provided to address the limitations identified.

6.1 Limitations and Redress

There are several topics in this paper yet to be addressed. There are issues to be addressed from three different perspectives – one, problems that would arise during the practical adoption of this model; two, in the evaluation; and three, in the model itself.

Adoption Impediments: This is a collaborative response model. As with any collaborative effort, there is a host of issues such as privacy, confidentiality, non-repudiation, etc, that will need to be addressed during practical adoption. Thankfully, these are issues for which there are solutions available already

through cryptography and IPSEC. In a co-operative or collaborative environment, we expect these issues to be either easily resolved or already addressed. Regardless, co-operation amongst various entities on the Internet such as amongst different corporate networks pose more legal, political, and economic problems than technical. In such cases where sharing anomaly information with networks outside of the corporation is not feasible, applying this response model within the corporate network itself can provide valuable protection.

Assigning realistic values to τ_1 and τ_2 is another major impediment to adoption. However, that is a separate problem requiring independent study. There are indeed prior work that attempt to assign costs to various responses that can be used [14,4]. Whereas, this paper focusses on optimizing those costs for overall operation of a system.

Evaluation Issues: Integral and faithful scaling down of the Internet is a difficult problem [22], which makes evaluating worm defenses more so [8]. At one extreme we have realistic evaluation possible only on the Internet, which is infeasible. At the other extreme, we have pure mathematical models. In between these two extremes, we have simulations such as the one used in this paper and emulation as used in one of our previous studies for worm detection [9].

With the availability of data about Internet traffic during worm outbreaks, it may be possible to evaluate the defense model on a network testbed such as Emulab [23] or DETER [3] by replaying the traffic for a scaled down version of the Internet. Such an experiment would need the available data to be carefully replayed with tools such as TCP Replay,TCP Opera [12], etc. This is a future task. Nevertheless, such emulation experiments can only scale up to a certain level and after that we would have to resort to mathematics or simulations to extrapolate the results to Internet scales.

We avoid emulation experiments by choice. Emulations can provide details about exploit behavior, traffic patterns, etc. As important as those issues are, they lie outside the scope of our present interest and are considered for later study. Focus for this paper is primarily on the cost optimization models. As mentioned in the previous paragraph, experiment population sizes are limited in emulations while simulations can support larger number of nodes. Given that stochastic processes are involved in our model, we need a large population to achieve fidelity in results. Consequently, simulations form a natural choice for our experiments.

An issue to be studied is the behavior of this model in the face of false alarms and isolated intrusions. For example, consider one and only participant raising an alarm for an isolated event and several other participants choosing the S control. We would like to know when these participants would apply the C control. Trivially, we can set a time-out for the defense to be turned-off. However, the time-out should be chosen carefully and probably be dynamic to guard against exposing oneself to slow-worm attacks.

In our experiments we have showed only one worm operating at a time. While this might seem like a limitation of the model, it is not. As mentioned in Sect. 3.1, there is an *anomaly vector* associated with each suspected worm incident. When

multiple worms operate simultaneously, each will be associated with a different *anomaly vector*. In operational practice, we expect a different process to be associated with each *anomaly vector* so multiple worms can be handled independently and concurrently.

Limitations and Extensions to the Model: When there is a cost to sampling packets, this model can be extended to optimally stop the sampling process and declare either that there is a worm or that there is no worm – essentially a distributed worm detector. Interestingly, this extension would lead us to the distributed Sequential Hypothesis Testing that we discussed in our previous paper [9].

One of the assumptions in our model is that the worm is a random scanning worm. This model will not work against more intelligent worms such as hit-list or flash worms but will likely be moderately successful against sub-net scanning worms [19]. Evaluating and extending the model against such other kinds of worms is a future work.

Integrity of the sensors, and absence of wilful malfeasance are assumed in our model. After all, in the real world we do assume the safety and security of the firewalls and IDSes we use. Nevertheless, if a few of the sensors are compromised by the attackers, we expect the stochastic nature of our model to act as a cushion absorbing some of the ill-effects. This needs to be evaluated. If numerous sensors are affected, our assumption about collaboration is not valid any more and the results are undefined.

Actions such as C and S if applied frequently could lead to a very unstable system. We need to evaluate this factor in light of ambient anomaly levels in different environments. This is a problem with the model itself. However, this can be alleviated to some extent during adoption in various ways. For example, the set of response options, $\{C, S\}$, can be made larger by introducing several levels of reduced functionality. This will however increase the complexity of the DP formulation but can be tolerated as we solve the formulation offline.

When all participants behave identically each participant knows exactly how the others will behave. In such a scenario, each one can make a better decision about the optimal control to be applied taking into account the others' behavior. For example, if participant A determines that the optimal policy to be applied is S, it now knows that all other participants will also apply the same control. Then, there is no need for A to apply S. Instead A could apply C as there is no opportunity for a worm to spread when all others participants are stopped. The problem now enters the realm of *game theory*.

6.2 Strengths

One question that needs to be answered for any defensive technique is this: "If the attacker knows about the approach being used for defense, will s/he be able to write a new generation of worms that can overcome the defense?"

There are two different goals that an attacker with knowledge about our system can try to achieve. One, try to circumvent the defense and spread the worm. Two, trick the defense into over-reacting.

The second goal cannot be achieved because of the dynamic and self-regulating nature of our approach, which is based on the current environmental conditions as depicted in Fig. 3. The attacker may force our system to react to an initial stimulus that is not a true worm, but once the stimulus has reduced, the defence pulls back too. If the sensors are compromised, however, the results are undefined as mentioned in the previous section. However, compromising sensors are extraneous to the model and is not a tenable argument against the model.

To achieve the first goal, the worm needs to either spread very slowly such that information about anomalous incidents are forgotten by the participants, or attack pre-selected victims that may not be alerted by its peers. However, since the alerts are shared with randomly chosen peers while the worm is spreading, there can be no effective pre-selection that can overcome the defense. Whereas a slow spreading worm might be successful to a certain extent.

Nevertheless, we believe that a slow spreading worm can be identified by other means such as manual trouble-shooting prompted by the ill-effects of the worm; unless the worm installs a time-bomb that is set to trigger after the worm has spread to most vulnerable nodes. We also believe that such slow worms will be circumvented by routine maintenance patches – most worms we know so far have exploited only known, but unpatched, vunlerabilities.

Moreover, there is a heightened awarness about security issues amongst the information technology community than ever before. Laws related to data security are being tightened and enforced more vigorosly than in the past. Patch generation and deployment techniques have advanced tremendously recently. In such an environment, we expect that steps to patch or workaround known vulnerabilities will be taken with more urgency than ever before effectively thwarting extremely slow worms discussed in the preceeding paragraphs.

Thus, the worm has a very narrow window between spreading too slow and spreading too fast – the window where our response mechanism works to thwart the worm. In conclusion, to answer the question above, knowledge of our approach does not provide much value to the attacker or new generation of worms.

6.3 Summary

This paper presents a novel control-theoretic approach toward worm response. We showed how to formalize a response strategy as a Dynamic Programming problem and solve this formulation to yield a practically applicable response solution. This formalization has been one of the key contributions of this paper. We show how this model severely curtails the worm options available to attackers. Several interesting directions in which this work could be extended are identified.

Acknowledgments. We would like to thank Branislav Kveton of Intel Research and Jed Crandall of University of New Mexico for providing early critique on the work.

References

1. Anagnostakis, K.G., et al.: A cooperative immunization system for an untrusting internet. In: Proc. of IEEE ICON, October 2003, pp. 403–408 (2003)
2. Anagnostakis, K.G., Greenwald, M.B., Ioannidis, S., Keromytis, A.D.: Robust reactions to potential day-zero worms through cooperation and validation. In: Katsikas, S.K., López, J., Backes, M., Gritzalis, S., Preneel, B. (eds.) ISC 2006. LNCS, vol. 4176, pp. 427–442. Springer, Heidelberg (2006)
3. Bajcsy, R., et al.: Cyber defense technology networking and evaluation. Commun. of the ACM 47(3), 58–61 (2004)
4. Balepin, I., Maltsev, S., Rowe, J., Levitt, K.: Using specification-based intrusion detection for automated response. In: Vigna, G., Krügel, C., Jonsson, E. (eds.) RAID 2003. LNCS, vol. 2820, pp. 136–154. Springer, Heidelberg (2003)
5. Bertsekas, D.P., Shreve, S.E.: Stochastic Optimal Control: The Discrete Time Case. Academic Press, N.Y (1978)
6. Bertsekas, D.P.: Dynamic Programming and Optimal Control, 3rd edn., vol. 1. Athena Scientific (2005)
7. Cai, M., Hwang, K., Kwok, Y.-K., Song, S., Chen, Y.: Collaborative internet worm containment. IEEE Security and Privacy 4(3), 34–43 (2005)
8. Cheetancheri, S.G., et al.: Towards a framework for worm defense evaluation. In: Proc. of the IPCCC Malware Workshop on Swarm Intelligence, Phoenix (April 2006)
9. Cheetancheri, S.G., Agosta, J.M., Dash, D.H., Levitt, K.N., Rowe, J., Schooler, E.M.: A distributed host-based worm detection system. In: Proc. of SIGCOMM LSAD, pp. 107–113. ACM Press, New York (2006)
10. Costa, M., et al.: Vigilante: end-to-end containment of internet worms. In: Proc. of the SOSP, pp. 133–147. ACM Press, New York (2005)
11. Dash, D., Kveton, B., Agosta, J.M., Schooler, E., Chandrashekar, J., Bachrach, A., Newman, A.: When gossip is good: Distributed probabilistic inference for detection of slow network intrusions. In: Proc. of AAAI, AAAI Press, Menlo Park (2006)
12. Hong, S.-S., Felix Wu, S.: On Interactive Internet Traffic Replay. In: Valdes, A., Zamboni, D. (eds.) RAID 2005. LNCS, vol. 3858, pp. 247–264. Springer, Heidelberg (2006)
13. Kim, H.-A., Karp, B.: Autograph: Toward automated, distributed worm signature detection. In: Proc. of the USENIX Security Symposium (2004)
14. Lee, W., Fan, W., Miller, M., Stolfo, S.J., Zadok, E.: Towards cost-sensitive modeling for intrusion detection and response. J. of Computer Security 10(1,2) (2002)
15. Malan, D.J., Smith, M.D.: Host-based detection of worms through peer-to-peer cooperation. In: Proc. of the WORM, pp. 72–80. ACM Press, New York (2005)
16. Newsome, J., Karp, B., Song, D.: Polygraph: Automatically generating signatures for polymorphic worms. In: Proc. of the IEEE Symposium on Security and Privacy, pp. 226–241. IEEE, Los Alamitos (2005)
17. Singh, S., Estan, C., Varghese, G., Savage, S.: Automated worm fingerprinting. In: Proc. of OSDI, San Francisco, CA (December 2004)
18. Sidiroglou, S., Keromytis, A.D.: Countering network worms through automatic patch generation. IEEE Security and Privacy 3(6), 41–49 (2005)
19. Staniford, S., Paxson, V., Weaver, N.: How to 0wn the Internet in Your Spare Time. In: Proc. of the Summer USENIX Conf., Berkeley, August 2002. USENIX (2002)

20. Wang, K., Cretu, G., Stolfo, S.J.: Anomalous payload-based worm detection and signature generation. In: Proc. of RAID. ACM Press, New York (2005)
21. Wang, K., Stolfo, S.J.: Anomalous payload-based network intrusion detection. In: Proc. of RAID, September 2004. ACM Press, New York (2004)
22. Weaver, N., Hamadeh, I., Kesidis, G., Paxson, V.: Preliminary results using scale-down to explore worm dynamics. In: Proc. of WORM, pp. 65–72. ACM Press, New York (2004)
23. White, B., et al.: An integrated experimental environment for distributed systems and networks. In: OSDI, Boston, December 2002, pp. 255–270. USENIX (2002)
24. Zou, C.C., Gao, L., Gong, W., Towsley, D.: Monitoring and early warning for internet worms. In: Proc. of the CCS, pp. 190–199. ACM Press, New York (2003)

A DP Example

We provide a short, classical inventory control example to help readers unfamiliar with DP to formulate a DP problem. This is an example from Bertsekas [6].

Consider the problem of stocking store shelves for N days. The state of the system is denoted by the quantity (x_k) of a certain item available on the store shelves at the beginning of a day. Shelves are stocked(with u_k units) at day break while demand(w_k) for the item is stochastic during the day. Both w_k and u_k are non-negative. There is no change overnight. It is clear that this system evolves according to:

$$x_{k+1} = \max(0, x_k + u_k - w_k)\,.$$

While there is an upper bound of, say, 2 units on the stock that can be on the shelves, demand in excess of stocks is lost business. Say, the storage costs for a day is $(x_k + u_k - w_k)^2$ implying penalty for both lost business and for excess inventory at the end of the day. Assuming the purchase cost incurred by the store is 1 per unit stock, the operating cost per day is

$$g_k(x_k, u_k, w_k) = u_k + (x_k + u_k - w_k)^2\,.$$

The terminal cost at the end of N days is assumed to be 0. Say the planning horizon N is 3 days and the initial stock $x_0 = 0$. Say, the demand w_k has the same probability distribution for all three days and is given by

$$p(w_k = 0) = 0.1 \qquad p(w_k = 1) = 0.7 \qquad p(w_k = 2) = 0.2\,.$$

The problem now is to determine the *optimal policy* for reordering of stocks so as to minimize the total operational cost. Applying (3), the DP algorithm for this problem is

$$J_k(x_k) = \min_{\substack{0 \le u_k \le 2 - x_k \\ u_k = 0,1,2}} \underset{w_k}{E} \left\{ u_k + (x_k + u_k - w_k)^2 + J_{k+1}(x_{k+1}) \right\}\,, \qquad (17)$$

where $k = 0, 1, 2$, and x_k, u_k, w_k can take the values of $0, 1, 2$ while the terminal condition $J_3(x_3) = 0\,.$

Now starting with $J_3(x_3) = 0$ and solving (17) backwards for $J_2(x_k)$, $J_1(x_k)$ and $J_0(x_k)$ for $k = 0, 1, 2$, we find that the *optimal policy* is to reorder one unit if the shelves are empty and nothing otherwise.

B Applying Bayes' Rule

The probabilities, $P(x_1 = \overline{P} \,|\, I_1, C)$ for (12) can be calculated using Bayes' rule and (6–8). We show the calculations for one of them here for exposition.

$$
\begin{aligned}
&P(x_1 = \overline{P} \,|\, G,\, G,\, S) \\
&\quad = \frac{P(x_1 = \overline{P},\, G,\, G,\, |\, S)}{P(G,\, G,\, |\, S)} \\
&\quad = \frac{\sum\limits_{i=\{P,\overline{P}\}} P(G|x_0 = i) \cdot P(x_0 = i) \cdot P(G|x_1 = \overline{P}) \cdot P(x_1 = \overline{P}|x_0 = i, u_0 = S)}{\sum\limits_{i=\{P,\overline{P}\}} \sum\limits_{j=\{P,\overline{P}\}} P(G|x_0 = i) \cdot P(x_0 = i) \cdot P(G|x_1 = j) \cdot P(x_1 = j|x_0 = i, u_0 = S)} \\
&\quad = \frac{(\overline{\text{fp}} \cdot \overline{\lambda} \cdot \text{fn} \cdot 0) + (\text{fn} \cdot \lambda \cdot \text{fn} \cdot 1)}{(\overline{\text{fp}} \cdot \overline{\lambda} \cdot \overline{\text{fp}} \cdot 1) + (\overline{\text{fp}} \cdot \overline{\lambda} \cdot \text{fn} \cdot 0) + (\text{fn} \cdot \lambda \cdot \overline{\text{fp}} \cdot 0) + (\text{fn} \cdot \lambda \cdot \text{fn} \cdot 1)}
\end{aligned}
$$