Online Supplement to:

Optimal Policies for Security Patch Management

Debabrata Dey
Foster School of Business, University of Washington, Seattle, WA 98195, USA

Atanu Lahiri
Jindal School of Management, University of Texas in Dallas, Richardson, TX 75080, USA

Guoying Zhang
Dillard College of Business, Midwestern State University, Wichita Falls, TX 76308, USA

Appendix A: Notational Summary

Table A1: A Summary of Important Notations

Symbol	Meaning	Remark
C_d, c_d	Disruption cost per patching cycle for unplanned patching	$c_d = \frac{\lambda C_d}{\bar{C}_e}$
C_e	Severity level of an unpatched vulnerability; also the exploitation cost per unit time for the vulnerability	$C_e \sim$ Exponential Mean $= \bar{C}_e$
C_s, c_s	Setup cost per patching cycle	$c_s = \frac{\lambda C_s}{\bar{C}_e}$
m	Threshold for the patch-based policy	
S_k	Cumulative severity of k unpatched vulnerabilities	
TC_p	Normalized total expected cost per unit time for policy p, $p \in \{$Patch, Time, Total, Emgcy, Hybrid$\}$	
X, x	Threshold for the time-based policy	$x = \lambda X$
Y, y	Threshold for the total-control policy	$y = \frac{Y}{\bar{C}_e}$
Z, z	Threshold for the emergency-control policy	$z = \frac{Z}{\bar{C}_e}$
λ	Arrival rate of patches	
μ	Arrival rate of patches with $C_e > Z$	$\mu = \lambda e^{-z}$
τ	Length of a patching cycle	

Appendix B: Proofs

Proof of Lemma 1

We first derive Equation (1).

$$\begin{aligned}
&\int_0^\infty dt_m \int_0^{t_m} dt_{m-1} \cdots \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \lambda^m e^{-\lambda\tau} \left(\sum_{i=1}^{m-1} \mathop{\mathrm{E}}_{C_e(i)} [(t_m - t_i) C_e(i)] \right) \\
&= \bar{C}_e \int_0^\infty \lambda^m e^{-\lambda t_m} dt_m \left((m-1) t_m \left(\int_0^{t_m} dt_{n-1} \cdots \int_0^{t_2} dt_1 \right) - \left(\int_0^{t_m} dt_{n-1} \cdots \int_0^{t_2} dt_1 \sum_{i=1}^{m-1} t_i \right) \right) \\
&= \bar{C}_e \int_0^\infty \lambda^m e^{-\lambda t_m} dt_m \left((m-1) t_m \frac{t_m^{m-1}}{(m-1)!} - \frac{t_m^m}{2(m-2)!} \right) \\
&= \frac{\bar{C}_e m(m-1)}{2\lambda} \int_0^\infty \frac{\lambda^{m+1} t_m^m}{m!} e^{-\lambda t_m} dt_m w = \frac{m(m-1)\bar{C}_e}{2\lambda}.
\end{aligned}$$

The last step follows from the observation that $\int_0^\infty \frac{\lambda^{m+1} t_m^m}{m!} e^{-\lambda t_m} dt_m = 1$, since the integrand is an Erlang probability density function.

Therefore, the total expected cost per cycle is $\left(C_s + C_d + \frac{m(m-1)\bar{C}_e}{2\lambda} \right)$. Dividing this by the length of the cycle, $\frac{m}{\lambda}$, and normalizing with respect to $\bar{C}_e$, we get the result. ■

Proof of Proposition 1

We first relax the constraint that m is an integer. Because $\frac{d^2}{dm^2}(TC_{\text{Patch}}) = \frac{2(c_s+c_d)}{m^3} > 0$, the cost function is convex in m, and the minimum can be found by simply solving the first order condition: $\frac{d}{dm}(TC_{\text{Patch}}) = -\frac{c_s+c_d}{m^2} + \frac{1}{2} = 0$. This results in: $\tilde{m} = \sqrt{2(c_s + c_d)}$. If $\tilde{m}$ is not an integer, the convexity of the cost function implies that the actual integer solution must be in the vicinity of $\tilde{m}$ and can be found by either rounding it up or down, i.e., $m^* = \lfloor \tilde{m} \rfloor$ or $m^* = \lfloor \tilde{m} \rfloor + 1$. Of course, $m^* = \lfloor \tilde{m} \rfloor$ if $TC_{\text{Patch}}(\lfloor \tilde{m} \rfloor) < TC_{\text{Patch}}(\lfloor \tilde{m} \rfloor + 1)$. We can simplify this condition to obtain $\tilde{m}(1 - 2\alpha) \geq \alpha(1 - \alpha)$, where $\alpha = \tilde{m} - \lfloor \tilde{m} \rfloor$. The result follows by putting all of these together. Finally, $TC_{\text{Patch}}(m^*) \geq TC_{\text{Patch}}(\tilde{m}) = \sqrt{2(c_s + c_d)} - \frac{1}{2}$. ■

Proof of Lemma 2

We start with the derivation of Equation (2).

$$\begin{aligned}
&\sum_{n=1}^\infty \int_0^X dt_n \int_0^{t_n} dt_{n-1} \cdots \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \lambda^n e^{-\lambda X} \left(\sum_{i=1}^{n} \mathop{\mathrm{E}}_{C_e(i)} [(X - t_i) C_e(i)] \right) \\
&= \bar{C}_e \sum_{n=1}^\infty \lambda^n e^{-\lambda X} \left(nX \left(\int_0^X dt_n \int_0^{t_n} dt_{n-1} \cdots \int_0^{t_2} dt_1 \right) - \left(\int_0^X dt_n \int_0^{t_n} dt_{n-1} \cdots \int_0^{t_2} dt_1 \sum_{i=1}^{n} t_i \right) \right) \\
&= \bar{C}_e \sum_{n=1}^\infty \lambda^n e^{-x} \left(nX \frac{X^n}{n!} - \frac{X^{n+1}}{2(n-1)!} \right) = \frac{\lambda \bar{C}_e X^2}{2}.
\end{aligned}$$

The last step follows from the fact that $\sum_{n=1}^\infty \frac{(\lambda X)^{n-1} e^{-\lambda X}}{(n-1)!}$ is identically equal to one, as it is the sum of all the probability masses from a Poisson distribution.

Now, since there is no disruption cost under this policy, the total expected cost per cycle is simply $\left(C_s + \frac{\lambda \bar{C}_e X^2}{2} \right)$. Dividing this by the length of the cycle, X, and normalizing with respect to $\bar{C}_e$, we get the desired result. ■

Proof of Proposition 2

The first order condition in this case is $\frac{d}{dx}(TC_{\text{Time}}) = -\frac{c_s}{x^2} + \frac{1}{2} = 0$, which can be solved to obtain $x^* = \sqrt{2c_s}$. Substituting for x^*, we get $TC_{\text{Time}}(x^*) = \sqrt{2c_s}$. Since $\frac{d^2}{dx^2}(TC_{\text{Time}}) = \frac{2c_s}{x^3} > 0$, the second order condition is also satisfied. ■

Proof of Lemma 3

First, note that:

$$\sum_{i=1}^{n-1}(S_i - S_{i-1})(\sigma - t_i) = \sum_{i=1}^{n-2} S_i(t_{i+1} - t_i) + S_{n-1}(\sigma - t_{n-1}).$$

We will now prove the result by induction. It is, of course, easy to see that the claim is true for $n = 2$, since:

$$\Phi_2(\sigma, R) = \int_0^\sigma dt_1 \int_0^R dS_1(S_1(\sigma - t_1)) = \frac{(\sigma R)^2}{4}.$$

Assuming that the above claim is true for $n = k - 1$, we get:

$$\Phi_{k-1}(\sigma, R) = \frac{(\sigma R)^{k-1}}{2(k-1)!(k-3)!}.$$

Also, by definition:

$$\begin{aligned}
\Phi_{k-1}(\sigma, R) &= \int_0^\sigma dt_{k-2} \cdots \int_0^{t_2} dt_1 \int_0^R dS_{k-2} \cdots \int_0^{S_2} dS_1 \left(\sum_{i=1}^{k-3} S_i(t_{i+1} - t_i) + S_{k-2}(\sigma - t_{k-2}) \right) \\
&= \int_0^\sigma dt_{k-2} \cdots \int_0^{t_2} dt_1 \int_0^R dS_{k-2} \cdots \int_0^{S_2} dS_1 \left(\sum_{i=1}^{k-3} S_i(t_{i+1} - t_i) \right) \\
&\quad + \int_0^\sigma dt_{k-2} \cdots \int_0^{t_2} dt_1 \int_0^R dS_{k-2} \cdots \int_0^{S_2} dS_1 S_{k-2}(\sigma - t_{k-2}) \\
&= \int_0^\sigma dt_{k-2} \cdots \int_0^{t_2} dt_1 \int_0^R dS_{k-2} \cdots \int_0^{S_2} dS_1 \left(\sum_{i=1}^{k-3} S_i(t_{i+1} - t_i) \right) + \frac{(\sigma R)^{k-1}(k-2)}{(k-1)!(k-1)!}.
\end{aligned}$$

Therefore:

$$\begin{aligned}
&\int_0^\sigma dt_{k-2} \cdots \int_0^{t_2} dt_1 \int_0^R dS_{k-2} \cdots \int_0^{S_2} dS_1 \left(\sum_{i=1}^{k-3} S_i(t_{i+1} - t_i) \right) \\
&\qquad = \frac{(\sigma R)^{k-1}}{2(k-3)!(k-1)!} - \frac{(\sigma R)^{k-1}(k-2)}{(k-1)!(k-1)!} = \frac{(\sigma R)^{k-1}(k-2)(k-3)}{2(k-1)!(k-1)!}.
\end{aligned}$$

We can now rewrite $\Phi_k(\sigma, R)$ as:

$$\Phi_k(\sigma, R) = \int_0^\sigma dt_{k-1} \cdots \int_0^{t_2} dt_1 \int_0^R dS_{k-1} \cdots \int_0^{S_2} dS_1 \left(\sum_{i=1}^{k-3} S_i(t_{i+1} - t_i) + S_{k-2}(t_{k-1} - t_{k-2}) + S_{k-1}(\sigma - t_{k-1}) \right).$$

We now simplify each of the three terms from above:

$$\int_0^\sigma dt_{k-1}\cdots\int_0^{t_2} dt_1\int_0^R dS_{k-1}\cdots\int_0^{S_2} dS_1\left(\sum_{i=1}^{k-3} S_i(t_{i+1}-t_i)\right)$$
$$= \int_0^\sigma dt_{k-1}\int_0^R dS_{k-1}\frac{(\sigma R)^{k-1}(k-2)(k-3)}{2(k-1)!(k-1)!} = \frac{(\sigma R)^k(k-2)(k-3)}{2k!k!},$$

$$\int_0^\sigma dt_{k-1}\cdots\int_0^{t_2} dt_1\int_0^R dS_{k-1}\cdots\int_0^{S_2} dS_1 S_{k-2}(t_{k-1}-t_{k-2}) = \frac{(\sigma R)^k(k-2)}{k!k!}, \text{ and}$$

$$\int_0^\sigma dt_{k-1}\cdots\int_0^{t_2} dt_1\int_0^R dS_{k-1}\cdots\int_0^{S_2} dS_1 S_{k-1}(\sigma-t_{k-1}) = \frac{(\sigma R)^k(k-1)}{k!k!}.$$

Combining them, we obtain:

$$\Phi_k(\sigma,R) = \frac{(\sigma R)^k(k-2)(k-3)}{2k!k!} + \frac{(\sigma R)^k(k-2)}{k!k!} + \frac{(\sigma R)^k(k-1)}{k!k!} = \frac{(\sigma R)^k}{2k!(k-2)!},$$

which proves the induction hypothesis. ■

Proof of Lemma 4

We first find the expected cycle time $\mathrm{E}[\tau]$. We first note that:

$$\int_0^{t_n} dt_{n-1}\cdots\int_0^{t_2} dt_1 = \frac{t_n^{n-1}}{(n-1)!} \quad\text{and}\quad \int_0^Y dS_{n-1}\cdots\int_0^{S_2} dS_1 = \frac{Y^{n-1}}{(n-1)!},$$

which can be substituted to obtain:

$$\mathrm{E}[\tau] = \sum_{n=1}^{\infty}\int_0^\infty dt_n\int_0^{t_n} dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_Y^\infty dS_n\int_0^Y dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} t_n f\, dS_1$$
$$= \sum_{n=1}^{\infty}\int_0^\infty \frac{(\lambda t_n)^n e^{-\lambda t_n}}{(n-1)!}dt_n\int_Y^\infty \frac{1}{\bar{C}_e}e^{-\frac{S_n}{\bar{C}_e}}dS_n \times \frac{\left(\frac{Y}{\bar{C}_e}\right)^{n-1}}{(n-1)!} = \sum_{n=1}^{\infty}\frac{\Gamma(n+1)}{\lambda(n-1)!}e^{-y}\frac{y^{n-1}}{(n-1)!}.$$

Since $\Gamma(n+1)=n!$, we can reduce the above to get $\mathrm{E}[\tau] = \frac{y+1}{\lambda}$.

We now consider the per-cycle expected exploitation cost. The cost arising from the exploitation of patch i arriving at t_i is simply $(t_n-t_i)C_e(i) = (S_i-S_{i-1})(t_n-t_i)$. Summing this over all possible values of i and taking expectation over all possible patch arrival patterns, the per-cycle expected exploitation cost can be written as:

$$\mathrm{E}\left[\sum_{i=1}^{n-1}(S_i-S_{i-1})(t_n-t_i)\right] = \sum_{n=2}^{\infty}\int_0^\infty dt_n\int_0^{t_n} dt_{n-1}\cdots\int_0^{t_2} dt_1\int_Y^\infty dS_n\int_0^Y dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} f\, dS_1\left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(t_n-t_i)\right)$$
$$= \sum_{n=2}^{\infty}\int_0^\infty dt_n\int_Y^\infty f\Phi_n(t_n,Y)dS_n = \sum_{n=2}^{\infty}\int_0^\infty dt_n\int_Y^\infty f\frac{(t_nY)^n}{2n!(n-2)!}dS_n \quad \text{[from Lemma 3]}$$
$$= \frac{Y^2}{2\lambda\bar{C}_e}\sum_{n=2}^{\infty}e^{-y}\frac{y^{n-2}}{(n-2)!} = \frac{\bar{C}_e y^2}{2\lambda}.$$

The last step follows from the observations that $y = \frac{Y}{\bar{C}_e}$ and the sum is identically equal to 1 as it represents the sum of all the probability masses from a Poisson distribution. Therefore, the total expected cost per cycle is $\left(C_s + C_d + \frac{\bar{C}_e y^2}{2\lambda}\right)$. Dividing this by the length of the cycle, $\frac{y+1}{\lambda}$, and normalizing with respect to $\bar{C}_e$, we get the desired result. ■

Proof of Proposition 3

The first order condition is: $\frac{d}{dy}(TC_{\text{Total}}) = \frac{y}{y+1} - \frac{2(c_s+c_d)+y^2}{2(y+1)^2} = 0$. Since $y+1 \neq 0$, we can multiply both sides by $2(y+1)^2$ and simplify to obtain $y^* = \sqrt{1+2(c_s+c_d)} - 1$. Since $\frac{d^2}{dy^2}(TC_{\text{Total}}) = \frac{1+2(c_s+c_d)}{(y+1)^3} > 0$, the second order condition holds as well. In order to obtain the optimal cost, we simply substitute for y^* to obtain the desired result. ■

Proof of Lemma 5

To obtain the normalized total cost per unit time, we divide the per-cycle total expected cost with the expected cycle length and normalize it with respect to $\bar{C}_e$:

$$TC_{\text{Emgcy}}(Z) = \frac{C_s + C_d + \frac{\lambda \bar{C}_e}{\mu^2}(1 - e^{-z} - ze^{-z})}{\bar{C}_e \frac{1}{\mu}} = \frac{\mu}{\lambda}(c_s + c_d) + \frac{\lambda}{\mu}\left(1 - e^{-z} - ze^{-z}\right).$$

The result follows from the fact that $\mu = \lambda e^{-z}$. ■

Proof of Proposition 4

The first order condition is: $\frac{d}{dz}(TC_{\text{Emgcy}}) = e^z - 1 - e^{-z}(c_s + c_d) = 0$, which can be solved to obtain $z^* = \ln(\frac{1+\sqrt{1+4(c_s+c_d)}}{2})$. Since the second order condition—$\frac{d^2}{dz^2}(TC_{\text{Emgcy}}) = e^z + e^{-z}(c_s + c_d) > 0$—is also satisfied, we simply substitute for z^* to obtain the cost at optimality. ■

Proof of Theorem 1

We know that $TC_{\text{Total}}(y^*) = \sqrt{1+2(c_s+c_d)} - 1$ and $TC_{\text{Patch}}(m^*) \geq \sqrt{2(c_s+c_d)} - \frac{1}{2}$. Therefore, the difference between the two can be written as:

$$\Lambda = TC_{\text{Patch}}(m^*) - TC_{\text{Total}}(y^*) \geq \frac{1}{2} + \sqrt{2(c_s+c_d)} - \sqrt{1+2(c_s+c_d)}.$$

If $c_s + c_d > \frac{9}{32}$, it can be easily shown that this $\Lambda > 0$, implying that the total-control policy performs better. Therefore, to complete the proof, we only need to consider the case where $c_s + c_d \leq \frac{9}{32}$. In that case, however, $m^* = 1$ and $TC_{\text{Patch}}(m^*) = c_s + c_d$. As a result, $\Lambda = 1 + (c_s + c_d) - \sqrt{1+2(c_s+c_d)} = \frac{(c_s+c_d)^2}{1+(c_s+c_d)+\sqrt{1+2(c_s+c_d)}} \geq 0$.

Similarly when the emergency-control policy is compared to the total-control policy, the cost difference can be written as:

$$\Lambda = TC_{\text{Emgcy}}(z^*) - TC_{\text{Total}}(y^*) = \sqrt{1+4(c_s+c_d)} - \sqrt{1+2(c_s+c_d)} - \ln\left(\frac{1+\sqrt{1+4(c_s+c_d)}}{2}\right).$$

It is easy to see that $\Lambda = 0$ at $Q = c_s + c_d = 0$. Furthermore, for $Q > 0$,

$$\frac{\partial \Lambda}{\partial Q} = \frac{2Q}{\sqrt{1+2Q}(1+\sqrt{1+4Q})}\left(\frac{1}{1+\sqrt{1+2Q}} - \frac{1}{\sqrt{1+2Q}+\sqrt{1+4Q}}\right) > 0.$$

Therefore, $\Lambda \geq 0$. ■

Proof of Theorem 2

We know that $TC_{\text{Time}}(x^*) = \sqrt{2c_s}$ and $TC_{\text{Total}}(y^*) = \sqrt{1+2(c_s+c_d)} - 1$. Clearly, time-based policy would perform better if $TC_{\text{Time}}(x^*) < TC_{\text{Total}}(y^*)$. However, since the costs are non-negative, this is equivalent to $(TC_{\text{Time}}(x^*)+1)^2 < (TC_{\text{Total}}(y^*)+1)^2$. Simplifying, we obtain $c_d > \sqrt{2c_s}$. ■

Proof of Lemma 6

Since $H_1(a,x) = \sum_{n=0}^{\infty} \frac{x^n}{n!(n+a-1)!}$, we get:

$$H_{1x}(a,x) = \frac{\partial H_1(a,x)}{\partial x} = \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!(n+a-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!(k+(a+1)-1)!} = H_1(a+1,x).$$

Also, $H_2(a,b,x,y) = \sum_{n=b+1}^{\infty} \frac{e^{-y}y^{n-b-1}\Gamma(n+a,x)}{(n-1)!(n-b-1)!}$. First we note that $\frac{\partial\Gamma(n+a,x)}{\partial x} = -e^{-x}x^{n+a-1}$. Therefore:

$$\begin{aligned}
H_{2x}(a,b,x,y) &= \frac{\partial H_2(a,b,x,y)}{\partial x} = -\sum_{n=b+1}^{\infty} \frac{e^{-y}y^{n-b-1}e^{-x}x^{n+a-1}}{(n-1)!(n-b-1)!} = -e^{-(x+y)}x^{a+b}\sum_{n=b+1}^{\infty}\frac{(xy)^{n-b-1}}{(n-1)!(n-b-1)!} \\
&= -e^{-(x+y)}x^{a+b}\sum_{k=0}^{\infty}\frac{(xy)^k}{k!(k+(b+1)-1)!} = -e^{-(x+y)}x^{a+b}H_1(b+1,xy).
\end{aligned}$$

Finally:

$$\begin{aligned}
H_{2y}(a,b,x,y) &= \frac{\partial H_2(a,b,x,y)}{\partial y} = \sum_{n=b+1}^{\infty} \frac{e^{-y}y^{n-b-1}\Gamma(n+a,x)}{(n-1)!(n-b-1)!} \\
&= \sum_{n=b+2}^{\infty} \frac{e^{-y}y^{n-b-2}\Gamma(n+a,x)}{(n-1)!(n-b-2)!} - \sum_{n=b+1}^{\infty} \frac{e^{-y}y^{n-b-1}\Gamma(n+a,x)}{(n-1)!(n-b-1)!} \\
&= H_2(a,b+1,x,y) - H_2(a,b,x,y),
\end{aligned}$$

which completes the proof. ■

Proof of Lemma 7

We first consider the expected cycle time:

$$\begin{aligned}
\mathrm{E}[\tau] &= \sum_{n=1}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1 \int_0^{\infty} dS_n \int_0^{\min\{Y,S_n\}} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} XfdS_1 \\
&\quad + \sum_{n=1}^{\infty}\int_0^X dt_n \int_0^{t_n} dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_Y^{\infty} dS_n\int_0^Y dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} t_n f dS_1.
\end{aligned}$$

The first term in the expression above can be written as:

$$\begin{aligned}
&\sum_{n=1}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_0^{\infty} dS_n \int_0^{\min\{Y,S_n\}} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} XfdS_1\\
=\;& \sum_{n=1}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_0^{Y} dS_n \int_0^{S_n} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} XfdS_1\\
&+\sum_{n=1}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_Y^{\infty} dS_n \int_0^{Y} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} XfdS_1\\
=\;& \sum_{n=1}^{\infty}\int_X^{\infty} dt_n \frac{X^n}{(n-1)!}\int_0^{Y} fdS_n\frac{S_n^{n-1}}{(n-1)!}+\sum_{n=1}^{\infty}\int_X^{\infty} dt_n \frac{X^n}{(n-1)!}\int_Y^{\infty} fdS_n\frac{Y^{n-1}}{(n-1)!}\\
=\;& \sum_{n=1}^{\infty}\frac{e^{-x}x^n\left((n-1)!+e^{-y}y^{n-1}-\Gamma(n,y)\right)}{\lambda(n-1)!(n-1)!},
\end{aligned}$$

where, as before, $y=\frac{Y}{C_e}$ and $x=\lambda X$. Moving on to the second term in the expression for $\mathrm{E}[\tau]$, we get:

$$\begin{aligned}
&\sum_{n=1}^{\infty}\int_0^{X} dt_n \int_0^{t_n} dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_Y^{\infty} dS_n \int_0^{Y} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} t_nfdS_1\\
=\;& \sum_{n=1}^{\infty}\int_0^{X} dt_n \frac{{t_n}^n}{(n-1)!}\int_Y^{\infty} fdS_n\frac{Y^{n-1}}{(n-1)!}=\sum_{n=1}^{\infty}\frac{e^{-y}y^{n-1}(n!-\Gamma(n+1,x))}{\lambda(n-1)!(n-1)!}.
\end{aligned}$$

Combining, we get:

$$\begin{aligned}
\mathrm{E}[\tau] &= \sum_{n=1}^{\infty}\left(\frac{e^{-x}x^n\left((n-1)!+e^{-y}y^{n-1}-\Gamma(n,y)\right)}{\lambda(n-1)!(n-1)!}+\frac{e^{-y}y^{n-1}(n!-\Gamma(n+1,x))}{\lambda(n-1)!(n-1)!}\right)\\
&= \frac{1}{\lambda}\left(x+y+1+e^{-(x+y)}xH_1(1,xy)-xH_2(0,0,y,x)-H_2(1,0,x,y)\right),
\end{aligned}$$

where $H_1(a,x)$ and $H_2(a,b,x,y)$ are as defined in (4).

We now consider the per-cycle expected exploitation cost. The cost arising from the exploitation of patch i arriving at $t_i<\tau$ is simply $(\tau-t_i)C_e(i)=(S_i-S_{i-1})(\tau-t_i)$. Summing this over i and taking expectation over all possible patch arrival patterns, the per-cycle expected exploitation cost can be written as:

$$\begin{aligned}
&\mathrm{E}\left[\sum_{i=1}^{n-1}(S_i-S_{i-1})(\tau-t_i)\right]\\
=\;& \sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_0^{\infty} dS_n \int_0^{\min\{Y,S_n\}} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} fdS_1\left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(X-t_i)\right)\\
&+\sum_{n=2}^{\infty}\int_0^{X} dt_n \int_0^{t_n} dt_{n-1}\cdots\int_0^{t_3} dt_2\int_0^{t_2} dt_1\int_Y^{\infty} dS_n \int_0^{Y} dS_{n-1}\cdots\int_0^{S_3} dS_2\int_0^{S_2} fdS_1\left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(t_n-t_i)\right).
\end{aligned}$$

We can simplify the first term as:

$$\begin{aligned}
&\sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \int_0^{\infty} dS_n \int_0^{\min\{Y,S_n\}} dS_{n-1}\cdots\int_0^{S_3} dS_2 \int_0^{S_2} f dS_1 \left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(X-t_i)\right)\\
&= \sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \int_0^{Y} dS_n \int_0^{S_n} dS_{n-1}\cdots\int_0^{S_3} dS_2 \int_0^{S_2} f dS_1 \left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(X-t_i)\right)\\
&\quad +\sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_0^X dt_{n-1}\cdots\int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \int_Y^{\infty} dS_n \int_0^{Y} dS_{n-1}\cdots\int_0^{S_3} dS_2 \int_0^{S_2} f dS_1 \left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(X-t_i)\right)\\
&= \sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_0^{Y} dS_n \Phi_n(X,S_n) f + \sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_Y^{\infty} dS_n \Phi_n(X,Y) f\\
&= \sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_0^{Y} dS_n \frac{(XS_n)^n}{2n!(n-2)!} f + \sum_{n=2}^{\infty}\int_X^{\infty} dt_n \int_Y^{\infty} dS_n \frac{(XY)^n}{2n!(n-2)!} f \quad \text{[from Lemma 3]}\\
&= \sum_{n=2}^{\infty}\frac{\bar{C}_e e^{-x}x^n\left(n!+e^{-y}y^n-\Gamma(n+1,y)\right)}{2\lambda(n-2)!n!}.
\end{aligned}$$

We now consider the second term:

$$\begin{aligned}
&\sum_{n=2}^{\infty}\int_0^{X} dt_n \int_0^{t_n} dt_{n-1}\cdots\int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \int_Y^{\infty} dS_n \int_0^{Y} dS_{n-1}\cdots\int_0^{S_3} dS_2 \int_0^{S_2} f dS_1 \left(\sum_{i=1}^{n-1}(S_i-S_{i-1})(t_n-t_i)\right)\\
&= \sum_{n=2}^{\infty}\int_0^{X} dt_n \int_Y^{\infty} dS_n \Phi_n(t_n,Y) f = \sum_{n=2}^{\infty}\int_0^{X} dt_n \int_Y^{\infty} dS_n \frac{(t_nY)^n}{2n!(n-2)!} f \quad \text{[from Lemma 3]}\\
&= \sum_{n=2}^{\infty}\frac{\bar{C}_e e^{-y}y^n(n!-\Gamma(n+1,x))}{2\lambda(n-2)!n!} = \frac{\bar{C}_e y^2}{2\lambda} - \sum_{n=2}^{\infty}\frac{\bar{C}_e e^{-y}y^n\Gamma(n+1,x)}{2\lambda(n-2)!n!}.
\end{aligned}$$

Combining, the expected per-cycle exploitation cost can be expressed as:

$$\begin{aligned}
\mathrm{E}\left[\sum_{i=1}^{n-1}(S_i-S_{i-1})(\tau-t_i)\right] &= \frac{\bar{C}_e y^2}{2\lambda} + \sum_{n=2}^{\infty}\frac{\bar{C}_e\left(e^{-x}x^n\left(n!+e^{-y}y^n-\Gamma(n+1,y)\right)-e^{-y}y^n\Gamma(n+1,x)\right)}{2\lambda n!(n-2)!}\\
&= \frac{\bar{C}_e}{2\lambda}\left(x^2\left(1-H_2(0,2,y,x)\right)+y^2\left(1-H_2(0,2,x,y)\right)+e^{-(x+y)}x^2y^2H_1(3,xy)\right).
\end{aligned}$$

Finally, we consider the probability that a patching cycle ends as a result of the severity trigger:

$$\begin{aligned}
\Pr[\tau<X] &= \sum_{n=1}^{\infty}\int_0^{X} dt_n \int_0^{t_n} dt_{n-1}\cdots\int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \int_Y^{\infty} dS_n \int_0^{Y} dS_{n-1}\cdots\int_0^{S_3} dS_2 \int_0^{S_2} f dS_1\\
&= \sum_{n=1}^{\infty}\int_0^{X} dt_n \frac{t_n{}^{n-1}}{(n-1)!}\int_Y^{\infty} dS_n \frac{Y^{n-1}}{(n-1)!} f = \sum_{n=1}^{\infty}\frac{e^{-y}y^{n-1}((n-1)!-\Gamma(n,x))}{(n-1)!(n-1)!}\\
&= 1-\sum_{n=1}^{\infty}\frac{e^{-y}y^{n-1}\Gamma(n,x)}{(n-1)!(n-1)!} = 1-H_2(0,0,x,y).
\end{aligned}$$

The total expected cost per cycle is $\mathrm{E}\left[\sum_{i=1}^{n-1}(S_i-S_{i-1})(\tau-t_i)\right]+(C_s+\Pr[\tau<X]\,C_d)$, dividing which

by the length of the cycle, $\mathrm{E}[\tau]$, and normalizing with respect to $\bar{C}_e$, we get the desired result. ■

Proof of Proposition 5

We first show that (5) represents the first order conditions for this optimization problem. Since Lemma 6 ensures that $\frac{\partial N}{\partial x}$ and $\frac{\partial D}{\partial x}$ exists, the first order condition with respect to x can written as:

$$\frac{\partial TC_{\text{Hybrid}}(x,y)}{\partial x} = \frac{1}{D^2}\left(D\frac{\partial N}{\partial x} - N\frac{\partial D}{\partial x}\right) = 0.$$

Since $N, D > 0$, we can multiply both sides by $\frac{D^2}{N}$ and rewrite the first order condition as:

$$\frac{1}{N}\frac{\partial N}{\partial x} - \frac{1}{D}\frac{\partial D}{\partial x} = 0 \quad \Rightarrow \quad \frac{\partial \ln N}{\partial x} - \frac{\partial \ln D}{\partial x} = 0.$$

The first order condition with respect to y follows in an analogous manner. Since (x_H^*, y_H^*) satisfies the first order conditions in (5), the optimal solution is obtained at that point. ■

Proof of Theorem 3

Let $\psi = \Pr[\tau < X]$, $\psi' = 1 - \psi$, and $\theta = \lambda\,\mathrm{E}[\tau]$. Then, for a hybrid policy,

$$\frac{\partial^2\left(TC_{\text{Hybrid}}\right)}{\partial c_d \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial\left(TC_{\text{Hybrid}}\right)}{\partial c_d}\right) = \frac{\partial\left(\frac{\psi}{\theta}\right)}{\partial y} < 0, \tag{B1}$$

because, trivially, ψ is decreasing in y and θ is increasing. Similarly,

$$\frac{\partial^2\left(TC_{\text{Hybrid}}\right)}{\partial c_d \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial\left(TC_{\text{Hybrid}}\right)}{\partial c_d}\right) = \frac{\partial\left(\frac{\psi}{\theta}\right)}{\partial x} = \frac{\partial\left(\frac{1}{\theta}\right)}{\partial x} - \frac{\partial\left(\frac{\psi'}{\theta}\right)}{\partial x} > 0. \tag{B2}$$

This is because θ is increasing in x and ψ' is decreasing, which together mean that $\frac{1}{\theta}$ declines less rapidly than does $\frac{\psi'}{\theta}$.

Now, we apply the implicit function theorem. Since (x_H^*, y_H^*) is a minimum obtained from

$$\frac{\partial\left(TC_{\text{Hybrid}}\right)}{\partial x} = \frac{\partial\left(TC_{\text{Hybrid}}\right)}{\partial y} = 0,$$

from the implicit function theorem, it is clear that the first order reactions to changing c_d would be exactly the opposite of the cross-partials above; so, x_H^* would be reduced, and y_H^* increased. However, since x_H and y_H are strategic substitutes when it comes to trading off the exploitation cost with the setup cost, the second order effects—the effect of reducing x_H on y_H and that of increasing y_H on x_H—would further reinforce the first order effects. Hence, the signs of $\frac{dx_H^*}{dc_d}$ and $\frac{dy_H^*}{dc_d}$ should be exactly the opposites of the signs of $\frac{\partial^2\left(TC_{\text{Hybrid}}\right)}{\partial c_d \partial x}$ and $\frac{\partial^2\left(TC_{\text{Hybrid}}\right)}{\partial c_d \partial y}$, respectively. It thus follows from (B1) and (B2) that x_H^* is monotonically decreasing in c_d, while y_H^* is increasing. (Such monotonicity is also expected, since a higher c_d means disruptions are more costly, which, in turn, implies that time-based triggers are relatively more appealing while severity-based triggers are less so.)

It is easy to see that, when $c_d \to \infty$, in optimality, the hybrid policy degenerates to the time-based policy, implying $x_H^* = x^*$. At the other extreme, when $c_d = 0$, it degenerates to the total-control policy, so $x_H^* \to \infty$. It then follows from the monotonicity shown above that $x_H^* \geq x^*$. Since an analogous argument applies to y_H^*, $y_H^* \geq y^*$. ■