Allocating Scarce Security Attention: A Decision-Analytic Framework for Alert Policies

John Pierman

Abstract. Problem statement. Security operations centers receive more alerts than analysts can investigate, and several behavioral detectors compete for the same attention. Evaluating detectors one at a time cannot settle the allocation, because detectors may reach the same harmful instances and interact through the queue rules. **Methodology.** We fix the queue mechanics that do not concern allocation and represent each alternative by a budget-feasible detector-allocation vector, which the fixed mechanics turn into an executable queue policy. **Results.** The feasible allocations are the whole-number divisions of the review budget among the detectors. Detector-level coverage curves do not determine the coverage of a divided budget, but the joint distribution of the ranks at which each detector first reaches the same harmful instance does. Coverage converts to residual loss through type-specific harm, and expected utility or expected loss selects the preferred allocation. **Practical implications.** Organizations should specify the queue mechanics, record each harmful instance's rank under every detector, and compare feasible integer allocations. Sensitivity and value-of-information analysis then direct further evidence collection to the uncertainties that could change the preferred allocation.

Key words: decision analysis; cybersecurity; alert policies; attention allocation; value of information

1 Introduction

Security operations centers (SOCs) receive more alerts than analysts can investigate, so analyst attention is a binding resource. A *detector* is an automated rule or model that scores events, and its scores rank the events for review. It may be a statistical rule, such as a

threshold on failed logins per hour, or a machine learning model. Such a model may be population-based, scoring each event against the behavior of all users or hosts, or entity-based, scoring each event against the past behavior of the same user or host. At low prevalence, false alarms consume much of the review capacity (Axelsson 2000, Ndichu et al. 2025). An event matters only if the combined queue rules place it within the review limit, and a detector is valuable when it reaches consequential behavior that the rest of the portfolio misses, so detector performance in isolation cannot settle how attention should be divided.

This paper treats attention allocation as a prescriptive decision-analysis problem. The decision maker is the security leader or organizational body responsible for the review queue. The *alternative* is a feasible allocation vector specifying the share of a common review budget assigned to each detector. The remaining queue mechanics are fixed before allocations are compared, and together with an allocation vector they induce a complete, executable queue policy. The mix of harmful instances and their consequences are uncertain, and the baseline decision selects the allocation with greatest expected utility.

The main contribution is this allocation-level representation and a characterization of the evidence it requires. Detector-specific coverage curves omit whether detectors reach the same harmful instances. For fixed nonrefilling queue mechanics, the coverage of every feasible allocation is determined by where each harmful instance first appears in each detector's ranked list, recorded jointly across detectors. Evaluation must therefore keep these positions together for the same instance, because separately they lose the overlap that decides whether dividing the budget adds coverage or duplicates it.

2 Structuring the alert-allocation decision

2.1 Decision maker, alternatives, and timing

Fix one operating period, such as an analyst-day. Let S denote a monitored source, such as a network sensor or a host, identity, or application log observed by the SOC. An event

from S is a vector $\boldsymbol{e}_i^S = (\tau_i, \boldsymbol{z}_i^S)$, where τ_i is its timestamp and $\boldsymbol{z}_i^S$ is its observed-attribute vector. The realized event history is the timestamp-ordered sequence $E_S = (\boldsymbol{e}_1^S, \ldots, \boldsymbol{e}_N^S)$, and $E_{S,i} = (\boldsymbol{e}_1^S, \ldots, \boldsymbol{e}_i^S)$ is the history available when event i is scored.

Let $\mathcal{K} = \{1, \ldots, K\}$ index the detectors. Detector k returns a real-valued score for each event or abstains:

$$y_{ik} = d_k(E_{S,i}) \in \mathbb{R} \cup \{\mathrm{NA}\}, \tag{1}$$

where NA denotes abstention. The queue mechanics other than the allocation, namely event eligibility, abstention, ties, grouping, deduplication, refill, and timing, are fixed before allocations are compared. This paper calls them the *fixed mechanics*. Review capacity is assigned in whole queue positions, so allocations act on ranks rather than on scores. The eligibility, abstention, and tie rules order the events scored by detector k into one list by decreasing score. Let $\rho_k(i) \in \mathbb{N} \cup \{\infty\}$ be the position of event i in that list, with $\rho_k(i) = \infty$ when event i is ineligible or detector k abstains on it. Let $\widetilde{Y}(E_S)$ collect the scores, the ranks, and the event metadata that the fixed mechanics use.

Let $B \in \mathbb{N}$ be the review budget, and let $\mathcal{Q}_B(E_S)$ contain the ordered lists of at most B distinct events from E_S, including the empty list. Let $\mathcal{D}_B$ be the finite, nonempty set of detector-allocation vectors satisfying the budget and operational constraints. Each $\alpha \in \mathcal{D}_B$, together with the fixed mechanics, induces a complete policy

$$p_\alpha : \ \widetilde{Y}(E_S) \longmapsto Q_\alpha(E_S) \in \mathcal{Q}_B(E_S). \tag{2}$$

Thus α identifies the alternative and p_α is its operational implementation. Changing B or any fixed mechanic defines a different policy family rather than another value of α, and a detector is an input to the allocation, not an alternative by itself.

2.2 Threat types, consequences, beliefs, and preferences

Let $\mathcal{T} = \{1, \ldots, m\}$ be a finite set of threat types, such as credential spraying, account takeover, or lateral movement, described by observable behavior and scope rather than by attacker intent. A harmful instance is an event, or an episode of related events, that exhibits one type and carries that type as its label. Several types can occur in the same period. For every allocation $\alpha \in \mathcal{D}_B$ and type $t \in \mathcal{T}$, let $Z_{\alpha t}$ denote the uncertain consequences of implementing p_α for a harmful instance of type t, such as timely detection, delay, residual harm, workload, privacy burden, and response disruption. The allocation-by-type expected utility is

$$U_{\alpha t} := \mathbb{E}[u(Z_{\alpha t})], \tag{3}$$

where the expectation averages the remaining variation among instances of type t. These utilities accommodate risk attitude and multiple consequence attributes when they are decision-relevant (Keeney and Raiffa 1976).

Let $\Delta_m = \{q \in [0,1]^m : \sum_{t=1}^m q_t = 1\}$ be the set of probability vectors over the m types, and let $q \in \Delta_m$ be the decision maker's assessed type mix, so q_t is the expected share of type-t instances among the harmful instances of the period. Two probability levels appear throughout: q_t concerns the type of an instance, whereas $U_{\alpha t}$ and every coverage probability below average variation among the instances of one type. The expected utility of an allocation per harmful instance, and the optimal allocation, are

$$EU(\alpha) = \sum_{t=1}^{m} q_t U_{\alpha t}, \qquad \alpha^* \in \arg\max_{\alpha \in \mathcal{D}_B} EU(\alpha), \tag{4}$$

and p_{α^*} is the induced optimal policy. Since the number of harmful instances in the period does not depend on the allocation, the per-instance expectation ranks allocations as the period total would.

3 Constructing the feasible allocation set

3.1 Allocation vectors and the quota simplex

Attention can be allocated through detector quotas, score or rank fusion, thresholds, or learned routing rules. This paper studies detector quotas under the fixed mechanics, a transparent and operationally common family.

Suppose all B queue positions are assigned among the K detectors. Let

$$\Delta_K = \left\{ \alpha = (\alpha_1, \ldots, \alpha_K) \in [0,1]^K : \sum_{k=1}^{K} \alpha_k = 1 \right\} \tag{5}$$

be the allocation simplex. Its kth vertex assigns the entire budget to detector k, and its points are shares of the budget, not probability distributions over detectors. Because queue positions are indivisible, the implementable shares at budget B are the lattice points

$$\Delta_K^{(B)} = \left\{ \alpha \in \Delta_K : B\alpha \in \mathbb{N}_0^K \right\}, \qquad b(\alpha) = B\alpha, \tag{6}$$

so $b_k(\alpha)$ is the integer quota of detector k. The lattice contains

$$\left| \Delta_K^{(B)} \right| = \binom{B+K-1}{K-1} \tag{7}$$

candidate quota vectors, so the optimization is finite but grows combinatorially in B and K. Operational restrictions may exclude some lattice points, and the feasible set is $\mathcal{D}_B \subseteq \Delta_K^{(B)}$. Assigning all B positions does not force the realized queue to contain B alerts, because abstention and nonrefilling deduplication can leave capacity unused. Deliberate reserve capacity can be represented by an additional reserve coordinate or by the truncated set $\{b \in \mathbb{N}_0^K : \sum_k b_k \leq B\}$.

3.2 Joint ranks characterize allocation coverage

Fix a type t and represent a harmful instance of that type by the nonempty set J of indices of its events in the history, a single event or an episode spanning several. Detector k first reaches the instance at the best position that any of its events holds in detector k's list,

$$R_{kt} = \min_{i \in J} \rho_k(i) \in \mathbb{N} \cup \{\infty\}, \tag{8}$$

with $R_{kt} = \infty$ when no event of the instance appears in that list. This *first-harmful rank* is the number of queue positions detector k needs on its own to reach the instance, so a quota of b_k positions for detector k reaches the instance exactly when $R_{kt} \leq b_k$. Across harmful instances of type t these ranks vary, and all probabilities in this subsection are conditional on type t.

A quota vector b selects from each detector k the events with $\rho_k(i) \leq b_k$, takes their union, deduplicates without refill, and reviews every retained event within the declared horizon. The instance is covered exactly when some detector reaches it within quota. Writing $R_t = (R_{1t}, \ldots, R_{Kt})$ for the positions at which each detector first reaches the same instance, the coverage probability is

$$G_t(b) = \mathbb{P}\left(\bigcup_{k=1}^{K} \{R_{kt} \leq b_k\} \,\middle|\, t \right). \tag{9}$$

For $K = 2$, let $F_{kt}(r) = \mathbb{P}(R_{kt} \leq r \mid t)$ be the probability that detector k alone reaches an instance within r positions, and let $H_t(r, s) = \mathbb{P}(R_{1t} \leq r, R_{2t} \leq s \mid t)$ be the probability that both detectors reach it within their quotas. Inclusion-exclusion gives coverage under quotas (r, s) as

$$G_t(r, s) = F_{1t}(r) + F_{2t}(s) - H_t(r, s). \tag{10}$$

The detector-level curves F_{1t} and F_{2t} do not determine the overlap H_t, so they do not determine coverage under any divided budget. Two detectors with identical curves may reach

the same instances or different ones, and only the overlap tells which.

For any K, the complement of the coverage event in Equation (9) is the event that every detector places the instance beyond its quota, so

$$G_t(b) = 1 - \mathbb{P}(R_{1t} > b_1, \ldots, R_{Kt} > b_K \mid t). \tag{11}$$

The joint distribution of R_t therefore determines coverage for every nonrefilling quota vector, and the scores of Equation (1) enter only through the ranks they induce. Operational evaluation should record, for each harmful instance, its first-harmful rank under every detector together with its label and timing. Separate precision, recall, ROC, AUC, or coverage summaries discard the dependence across detectors that divided budgets require. Figure 1 places this evidence requirement in the decision path.

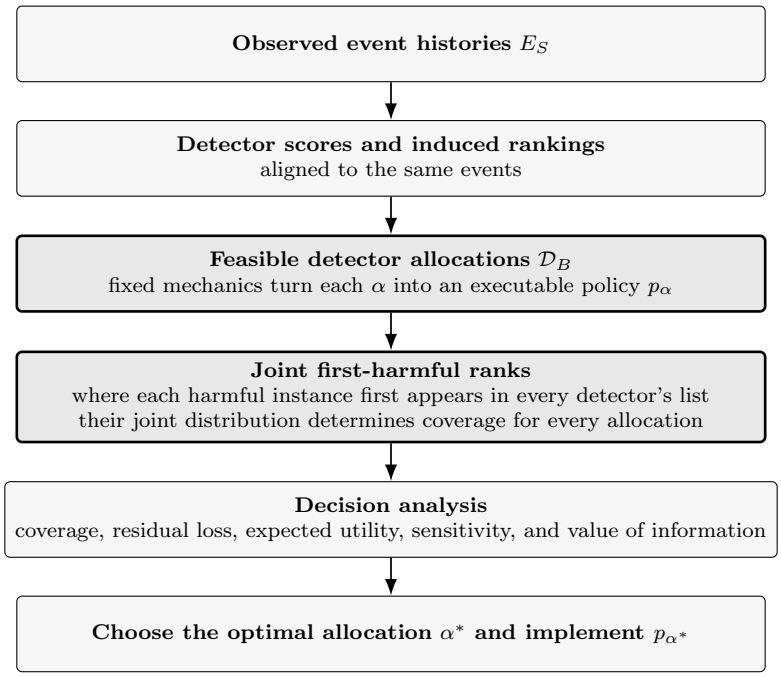


Figure 1: The decision path from observed event histories to the optimal allocation.

3.3 Candidate allocations and evaluation

When direct search over integer quotas is burdensome, a continuous relaxation over Δ_K can locate a promising region, but the relaxed optimum is not itself implementable. Nearby lattice points must be evaluated under the fixed mechanics, because rounding can change overlap, coverage, and loss. Candidate construction and final evaluation should use separate evidence, or the analysis should account for data reuse. An allocation that another feasible allocation matches or beats for every type at no greater operational cost can be dropped, although coverage dominance may disappear once harms, delays, or privacy costs enter the consequence model.

4 Core framework: coverage, loss, and decision analysis

4.1 Coverage and residual loss

Coverage is the principal consequence in the quota analysis. Under the mechanics of Section 3.2, the coverage probability of allocation α for type t is

$$C_{\alpha t} := G_t(b(\alpha)), \tag{12}$$

the probability that p_α places a harmful instance of type t in the review queue within the declared horizon. If covered and missed outcomes have utilities one and zero, with other consequences fixed, then $U_{\alpha t} = C_{\alpha t}$.

Coverage can also be expressed as loss. Let $h_t \geq 0$ be the harm of a missed instance of type t. If timely coverage averts all modeled harm, the residual loss for a type-t instance is

$$L_{\alpha t} = h_t(1 - C_{\alpha t}), \tag{13}$$

and the expected residual loss per harmful instance and its minimizer are

$$EL(\alpha) = \sum_{t=1}^{m} q_t L_{\alpha t}, \qquad \alpha_L^* \in \arg\min_{\alpha \in \mathcal{D}_B} EL(\alpha). \tag{14}$$

The loss model assumes linear harm, full harm avoidance after timely coverage, and no separate false-alert or delay cost. When those assumptions fail, the utility model in Equation (3) should include the additional consequences directly.

4.2 Adversarial threat selection

The type mix q may be estimated from historical frequencies, threat intelligence, or expert judgment, and it then describes the expected mix of harmful instances before the allocation is chosen. This fixed-belief model is appropriate when the mix is uncertain but not selected in response to the defender's allocation. Under it, randomizing over allocations is unnecessary, since a mixture cannot exceed the expected utility of its best component.

The problem becomes adversarial when an attacker can respond to the defender's allocation by shifting toward types that the allocation covers poorly. Let $a \in \Delta_m$ be the type mix chosen by the attacker, let the residual losses $L_{\alpha t}$ of Equation (13) be the zero-sum payoff, and let $\Delta(\mathcal{D}_B)$ be the probability simplex over feasible allocations. If the defender draws an allocation from $x \in \Delta(\mathcal{D}_B)$, and the attacker knows x but not the realized draw before acting, the robust allocation problem is

$$v^* = \min_{x \in \Delta(\mathcal{D}_B)} \max_{a \in \Delta_m} \sum_{\alpha \in \mathcal{D}_B} \sum_{t=1}^{m} x_\alpha a_t L_{\alpha t}. \tag{15}$$

Here the equilibrium type mix is endogenous rather than forecast, and randomization can improve the defender's guarantee by preventing the attacker from exploiting one known allocation. If the realized allocation is observed before the attacker acts, or either side commits first, the corresponding sequential game applies instead.

4.3 Sensitivity analysis

A recommendation should come with the assessments that sustain it. For allocations $\alpha, \beta \in \mathcal{D}_B$, α is preferred under type mix q when

$$\sum_{t=1}^{m} q_t(U_{\alpha t} - U_{\beta t}) \geq 0. \tag{16}$$

For two types, write $q = (\theta, 1-\theta)$, where $\theta \in [0,1]$ is the share of type one, and equality in Equation (16) gives a switching threshold in θ. With more types, fixed utilities partition Δ_m into regions on which the same allocation is preferred. Varying the type mix, harms, delay penalties, coverage values, or budget locates the boundaries that can change the recommendation and the assessments that are immaterial within a credible range (French 2003).

4.4 Value of information

Sensitivity analysis says whether an assessment could change the preferred allocation. Value of information (VOI) says how much a specified information source is expected to improve the choice. Here useful information may concern the joint rank distribution that determines coverage, the type mix q, or the utilities $U_{\alpha t}$. Let the random variable I take values in the set $\mathcal{I}$ of possible results from a proposed replay study, exercise, telemetry assessment, or expert elicitation, under a stated predictive distribution. If observing I updates the coverage, type-mix, or utility model while $\mathcal{D}_B$ and the fixed mechanics are unchanged, the expected value of sample information is

$$EVSI(I) = \mathbb{E}_I \left[\max_{\alpha \in \mathcal{D}_B} EU(\alpha \mid I) \right] - \max_{\alpha \in \mathcal{D}_B} EU(\alpha). \tag{17}$$

Holding $\mathcal{D}_B$ and the mechanics fixed matters because an intervention that changes either one changes the decision problem instead of informing it. EVSI is zero when no possible

result would change the allocation, however wide the uncertainty, and information is worth acquiring only when its value exceeds acquisition, delay, privacy, and operational costs. EVSI is measured in utility units, so a monetary buying price or certainty equivalent requires additional preference assumptions (Howard 1966, Abbas and Hazen 2024).

5 Conclusion

Scarce analyst attention should be allocated through a decision model, not inferred from detector scores alone. The detector-allocation vector is the alternative, and fixed queue mechanics turn each feasible vector into an executable policy. Detector scores enter only through the ranks they induce, and identical detector-level coverage can support different optimal allocations, because coverage under a divided budget depends on where the same harmful instance first appears in every detector's list. Recording these joint ranks is the evidence requirement of the framework. Coverage is then evaluated directly or translated into residual loss when types carry different harms.

SOC teams should fix the mechanics, define the feasible allocation set, characterize each allocation's consequences from joint-rank evidence, and assess the type mix and preferences that expected utility requires. Sensitivity analysis identifies the assessments that can change the optimal allocation, and value of information determines whether further evidence about joint ranks, the type mix, or consequences is worth acquiring. The model abstracts workload, delay, and analyst accuracy into the allocation-by-type utilities, and richer consequence models and empirical validation are its natural extensions.

References

Abbas AE, Hazen GB (2024) On the value of information across decision problems. *Decision Anal.* 22(1):1–13. https://doi.org/10.1287/deca.2024.0187.

Axelsson S (2000) The base-rate fallacy and the difficulty of intrusion detection. *ACM Trans. Inform. System Security* 3(3):186–205. https://doi.org/10.1145/357830.357849.

French S (2003) Modelling, making inferences and making decisions: The roles of sensitivity analysis. *TOP* 11:229–251. https://doi.org/10.1007/BF02579043.

Howard RA (1966) Information value theory. *IEEE Trans. Systems Sci. Cybernetics* 2(1):22–26. https://doi.org/10.1109/TSSC.1966.300074.

Keeney RL, Raiffa H (1976) *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* (Wiley, New York).

Ndichu S, Ban T, Ozawa S, Takahashi T, Inoue D (2025) Alert fatigue in security operations centres: Research challenges and opportunities. *ACM Comput. Surveys* 57(11):1–38. https://doi.org/10.1145/3723158.