T.C.
İSTANBUL ÜNİVERSİTESİ-CERRAHPAŞA
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

DOKTORA TEZİ

Kullanıcı Davranışları ve Risk Tahminlerine Dayalı Gelişmiş Kalıcı Tehdit Tespiti

Mustafa İPEKBAYRAK

DANIŞMAN
Doç. Dr. Gülsüm Zeynep GÜRKAŞ AYDIN

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği, Doktora Programı

Şubat, 2026

TEZ KABUL VE ONAYI

Mustafa İPEKBAYRAK tarafından, **Doç. Dr. Gülsüm Zeynep GÜRKAŞ AYDIN** danışmanlığında hazırlanan **"Kullanıcı Davranışları ve Risk Tahminlerine Dayalı Gelişmiş Kalıcı Tehdit Tespiti"** başlıklı bu çalışma, jürimiz tarafından **18/02/2026** tarihinde yapılan sınav sonucunda **oy birliği** ile başarılı bulunarak **Doktora Tezi** olarak kabul edilmiştir.

Tez Jürisi

		İmza	Sonuç
DANIŞMAN	Doç. Dr. Gülsüm Zeynep GÜRKAŞ AYDIN İstanbul Üniversitesi-Cerrahpaşa Bilgisayar Mühendisliği Anabilim Dalı		☒ Kabul ☐ Ret
ÜYE	Prof. Dr. Şerif BAHTİYAR İstanbul Teknik Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı		☒ Kabul ☐ Ret
ÜYE	Dr. Öğr. Üyesi Özgür Can TURNA İstanbul Üniversitesi-Cerrahpaşa Bilgisayar Mühendisliği Anabilim Dalı		☒ Kabul ☐ Ret
ÜYE	Doç. Dr. Mehmet Ali ERTÜRK İstanbul Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı		☒ Kabul ☐ Ret
ÜYE	Dr. Öğr. Üyesi Ebu Yusuf GÜVEN İstanbul Üniversitesi-Cerrahpaşa Bilgisayar Mühendisliği Anabilim Dalı		☒ Kabul ☐ Ret

Aileme ithaf ediyorum…

BÜTÇE DESTEKLERİ

Kullanıcı Davranışları ve Risk Tahminlerine Dayalı Gelişmiş Kalıcı Tehdit Tespiti

Bu tez çalışması, 100/2000 YÖK DOKTORA BURSU projesi ile desteklenmiştir.

TEŞEKKÜR

Rehberliğini her zaman hissettiğim, beni daima savunan değerli danışman hocam Doç. Dr. Gülsüm Zeynep GÜRKAŞ AYDIN hocama, beni her aşamada motive eden, desteğini esirgemeyen Prof. Dr. Muhammed Ali AYDIN hocama, samimiyetlerini ve iyi niyetlerini her zaman hissettiren Prof. Dr. Şerif BAHTİYAR ve Dr. Öğr. Üyesi Özgür Can TURNA hocalarıma içten teşekkürlerimi sunarım.

Şubat 2026 Mustafa İPEKBAYRAK

İÇİNDEKİLER

Sayfa No

1. GİRİŞ 1
2. KAVRAMSAL ÇERÇEVE 5
2.1. Nedensel güvenlik telemetrisi olarak olay köken izleri 6
2.2. Olay köken izi tabanlı APT tespitinde yöntem aileleri ve SOC işlevleri 7
2.3. Karşılaştırılabilirlik sorunu ve karar birimi 8
2.4. Kodlama şeması ve kullanılan metrikler 9
2.5. Literatür taramasından elde edilen kapsam bulguları 10
2.6. Literatür bulgularının tez tasarımına etkisi 12
2.7. Tek LP ile gözlemden bağımsız performans alt sınırı 13
2.8. Risk kademeli bütçeli operasyonel karar katmanı 15
2.9. Literatür boşluklarının tez hedefiyle ilişkisi 17
3. YÖNTEM 18
3.1. Araştırma yaklaşımı ve genel akış 19
3.2. Teorik çalışma: Problem tanımı, notasyon ve belirsizlik modeli 20
3.2.1. Aksiyonlar, gizli durumlar ve fayda (action–latent state–reward) 20
3.2.2. Gözlem modeli ve geçmişe dayalı karar kuralı varsayımı 21
3.3. Belirsizlik modeli 21
3.3.1. Önsel belirsizliği: Huber ε-kontaminasyon 21
3.3.2. Nominal getiri tablosu etrafında belirsizlik aralığı ve güvenli taban 22
3.3.3. Garanti fonksiyonunun nominal bileşeni 22
3.4. Belirsizliğe dayanıklı (robust) LP ile performans alt sınırı (sertifika) üretimi 23
3.4.1. Karar değişkenleri ve anlamları 23
3.4.2. Performans alt sınır LP'si (Primal) 23
3.4.3. Sonlu ufukta yüksek olasılıklı performans bandı 26
3.5. eriyle fayda belirsizliği aralığı ve performans alt sınır bandını daraltma: Empirik-Bernstein 26
3.6. Ampirik çalışma: Olay köken izi tabanlı UEBA karar katmanı 27

3.6.1. Deney tasarımı ve yapılandırma 27
3.6.2. Skor standardizasyonu: konformal p-değeri ve gözlemden bağımsızlık 31
3.6.3. Risk-nötr referans yöntemi (RN) 32
3.6.4. Belirsizlik altında performans alt sınırı garantileyen belirsizliğe dayanıklı (robust) yaklaşım 33
3.6.5. Nadir ama yıkıcı risk (tail risk) karşılaştırması için CVaR senaryo LP 34
3.6.6. Değerlendirme tasarımı: bütçe eşleştirme, sertifikaya uyum ve referans etikete sabitlenmiş metrikler 36
3.6.7. Duyarlılık ve stres: kademe sayısı, aralık genişliği ve gözlem kaymaları 37
3.6.8. Uygulama ve tekrar üretilebilirlik 37
4. BULGULAR 39
4.1. Tek LP ile gözlemden bağımsız ardışık savunma sertifikası 39
4.1.1. Teorik çalışma: Performans alt sınırının belirsizlik parametrelerine duyarlılığı 40
4.2. Gözlem bozulmaları altında değişmezlik (observation invariance) 41
4.3. Sonlu tur sayısında Freedman ve Azuma–Hoeffding güven bantlarının karşılaştırması 41
4.4. Sıkılık, tanılama ve verimlilik için mikro örnekler, örnekleme ve çözüm süresi 42
4.5. UEBA karar katmanı bulguları 43
4.5.1. Konformal kanıt ve risk kademeleri 43
4.5.2. Bütçe eşleştirmesi altında yaklaşım karşılaştırması 45
4.5.3. Performans alt sınırı düzeyi ve belirsizliğe dayanıklı yaklaşımın üstünlüğü 47
4.5.4. Stres testleri: dağılım kayması altında davranış 47
4.5.5. Hesaplama maliyeti: milisaniye ölçeğinde çözüm 48
4.6. Destekleyici çalışma: Sabit alarm eşiğinin alan kayması altında taşınabilir olmaması 49
5. TARTIŞMA 51
5.1. Literatürle ilişki ve değerlendirme tasarımı 51
5.2. Operasyonel çıkarımlar: ortam değişiminde kalibrasyon ve parametre yönetimi 52
5.3. Sınırlılıklar ve geçerlilik sınırları 53
5.4. Gelecek çalışmalara yön veren başlıklar 53
6. SONUÇ VE ÖNERİLER 55
6.1. Sonuçlar 55
6.2. Geleceğe yönelik öneriler 56

ŞEKİL LİSTESİ

Sayfa No

Şekil 1.1: Tezin genel araştırma akışı ve önerilen karar çerçevesi.....4

Şekil 2.1: Nedensel köken izi grafı.....7

Şekil 2.2: Sertifikanın kapsamı ve saha raporlaması.....15

Şekil 3.1: Tek LP sertifikası ve bütçeli aksiyon.....18

Şekil 3.2: Tek LP ile sertifika üretim hattı, ikili çözüm ve Freedman güven bandı.....26

Şekil 4.1: ε ve ω arttıkça performans alt sınırı $\boldsymbol{v}*$ değerinin azalması.....40

Şekil 4.2: Konformal p-değeri ve risk kademeleri üzerinden tek LP ile sertifikalı karar.....44

Şekil 4.3: UEBA skorlarından risk kademeleri ve bütçe altında LP tabanlı karar.....45

Şekil 4.4: Risk yaklaşımlarının performans karşılaştırması.....46

Şekil 4.5: IoT-23-A doğrulamada seçilen sabit alarm eşiklerinin IoT-23-A test, IoT-23-B ve ToN_IoT üzerindeki TPR değerleri (XGBoost).....50

TABLO LİSTESİ

Sayfa No

Tablo 2.1: Metrik kümeleri ve amaçları 10

Tablo 2.2: Tezde yürütülen iki çalışmanın ilgili mevcut çalışmalar ile eşleşmesi 13

Tablo 3.1: Teorik hat ile pratik hat karşılaştırması ve deneysel yapılandırma 29

Tablo 3.2: UEBA karar katmanında kullanılan nominal net fayda puanları ve operasyonel maliyet puanları 30

Tablo 4.1: ε ve ω değerlerinde performans alt sınırı $\boldsymbol{v}*$ (ortalama ± %95 güven aralığı). 40

Tablo 4.2: Gözlem bozulmaları altında ortalama getiri $\bar{r}$ ve alt sınır ihlal oranı (K=10.000) .. 41

Tablo 4.3: Tur sayısı K K için bant yarı genişlikleri (H) 42

Tablo 4.4: Performans tabanı sıkılığı, örnekleme kazanımı ve seçili çözüm süreleri (özet). ... 43

Tablo 4.5: B=0.90 ± 0.01 bütçe bandında yaklaşım karşılaştırması (havuzlanmış). 46

Tablo 4.6: Karışım-ağırlıklı performans alt sınırı seviyesi $\bar{\gamma}$ (B=0.90 ± 0.01). 47

Tablo 4.7: Dağılım kaymasında Stres testinde γ-uyum (parantez içi: nominale göre fark) 48

Tablo 4.8: Medyan çözüm süreleri (ms) – temsilî ayarlar. 49

Tablo 6.1: Seçili göstergelerle bulgu özeti 56

SİMGE VE KISALTMA LİSTESİ

Simgeler	Açıklama
V(D)	: Veri kümesi/senaryo D'de gözlenebilir saldırı aşamaları kümesi
N	: Gözlenebilir aşama sayısı (N = |V(D)|)
k	: Kapsanan/yakalanan aşama sayısı; kampanya kapsamı payı (k/N)
$\boldsymbol{A_d}$	: Savunucunun aksiyon kümesi
$\| \boldsymbol{A_d} \| \ (= \boldsymbol{m})$	: Savunucu aksiyon sayısı
$\boldsymbol{A_a}$	: Saldırganın aksiyon kümesi
$\| \boldsymbol{A_a} \| \ (= \boldsymbol{N_a})$	: Saldırgan aksiyon sayısı
$\boldsymbol{\pi_0(\theta)}$	: Nominal önsel dağılım
$\boldsymbol{\pi_b(\theta)}$	: Risk kademesi *b* için nominal durum önseli
Θ	: Gizli durum sayısı
$\mathcal{S}$	: Durum kümesi {BENIGN, ATTACK}
ω	: Fayda aralığı yarı-genişliği; hücre bazlı interval belirsizlik parametresi
$\boldsymbol{\pi_b(s)}$	: Risk kademesi b için durum önseli
ε	: Huber ε-kirlenme (kontaminasyon) parametresi
Δ	: Olasılık simpleksi (q(s) ≥ 0, ∑_s q(s) = 1)
$\mathbf{P_b}$	: Risk kademesi b için ε-kontaminasyon önsel belirsizlik kümesi
x	: Karma strateji / aksiyon dağılımı
$\mathbf{x_{b,a}}$	: Kademe b'de aksiyon a'nın seçilme olasılığı
$\mathbf{x^*}$	: LP çözümünden elde edilen optimal karma strateji
v	: Performans alt sınırı seviyesi / tur başına garanti edilen taban değer
$\mathbf{v^*}$	: Optimal performans alt sınırı değeri
$\boldsymbol{\gamma_b}$	: Kademe bazlı performans alt sınırı
$\bar{\gamma}$	: Karışım-ağırlıklı (tüm kademeler) performans alt sınırı seviyesi
$\mathbf{w_b}$	: Risk kademesi b'nin ağırlığı

$\mathbf{u_{a,s}}$	: Aksiyon a'nın durum s altındaki faydası
$\mathbf{L_{a,s}}$	: u_{a,s} için alt sınır (lower bound)
$\mathbf{U_{a,s}}$	: u_{a,s} için üst sınır (upper bound)
B	: Küresel maliyet bütçesi (beklenen maliyet üst sınırı)
Q	: Konformal risk kademeleri için p-değeri eşikleri
p(e)	: Konformal p-değeri (zararsız referansa göre standartlaştırılmış skor)
s(e)	: Dedektör/UEBA skor fonksiyonu
$\mathbf{N_{scen}}$	: CVaR için senaryo sayısı
j	: CVaR senaryo indeksi (j=1,…,N_scen)
n	: Konformal kalibrasyon (zararsız örnek) sayısı
K	: Sonlu ufuk/tur sayısı
δ	: Güven bantlarında hata olasılığı parametresi (1−δ güven)
τ	: CVaR kuyruk parametresi
$\mathbf{t_b}$	: VaR eşiği (CVaR epigraf değişkeni)
$\mathbf{m_b}$	: En kötü durum katkısını temsil eden yardımcı değişken
α	: Kademe önseli kurulurken kullanılan konformal referans eşiği
η	: Küçük örneklem kaynaklı oynaklığı sınırlayan düzleştirme sabiti

Kısaltmalar	**Açıklama**
APT	: Advanced Persistent Threat (Gelişmiş Kalıcı Tehdit)
ATT&CK	: Adversarial Tactics, Techniques, and Common Knowledge
AUC	: Area Under the Curve (Eğri Altı Alan)
CVaR	: Conditional Value at Risk (Koşullu Değer-Risk)
DARPA	: Defense Advanced Research Projects Agency
DP	: Dynamic Programming (Dinamik Programlama)
DRO	: Distributionally Robust Optimization (Dağılımsal Belirsizliğe Dayanıklı Optimizasyon)
EB	: Empirical Bernstein (Empirik Bernstein sınırı)
eCAR	: Extended Cyber Analytics Record (OpTC veri formatı)

ECE	: Expected Calibration Error (Beklenen Kalibrasyon Hatası)
FPR	: False Positive Rate (Yanlış Pozitif Oranı)
KKT	: Karush–Kuhn–Tucker (KKT koşulları)
LP	: Linear Programming (Doğrusal Programlama)
MDP	: Markov Decision Process (Markov Karar Süreci)
NCR	: OpTC "NCR Release" veri seti sürümü
OpTC	: Operationally Transparent Cyber (DARPA veri seti/programı)
POMDP	: Partially Observable MDP (Kısmi Gözlenebilir Markov Karar Süreci)
PR-AUC	: Precision–Recall AUC (Kesinlik–Duyarlılık Eğrisi Altı Alan)
RN	: Risk-Neutral (Risk-Nötr)
ROC	: Receiver Operating Characteristic (ROC eğrisi)
ROC-AUC	: ROC eğrisi altı alan
SOC	: Security Operations Center (Güvenlik Operasyon Merkezi)
TPR	: True Positive Rate (Doğru Pozitif Oranı)
UEBA	: User and Entity Behavior Analytics (Kullanıcı ve Varlık Davranış Analitiği)
VaR	: Value at Risk (Değer-Risk)
DL	Deep Learning (Derin Öğrenme)
DRL	Deep Reinforcement Learning (Derin Pekiştirmeli Öğrenme)

ÖZET

[DOKTORA TEZİ]

Kullanıcı Davranışları ve Risk Tahminlerine Dayalı Gelişmiş Kalıcı Tehdit Tespiti

[Mustafa İPEKBAYRAK]

İstanbul Üniversitesi-Cerrahpaşa

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği, Doktora Programı

[Danışman: Doç. Dr. Gülsüm Zeynep GÜRKAŞ AYDIN]

[Gelişmiş Kalıcı Tehdit (APT) saldırılarına karşı Güvenlik Operasyon Merkezlerinde (SOC) savunucu, sistem, ağ ve cihaz loglarından üretilen kanıtlar üzerinden tekrar eden kararlar alır. Ancak log kaybı, logların toplanma ve işlenme biçimindeki değişimler ve saldırganın iz silme girişimleri aynı olayın farklı görünmesine neden olabilir. Ayrıca kullanıcı ve varlık davranış analizi ile olay köken izi tabanlı sistemlerin ürettiği şüphelilik skorlarının hangi durumda hangi müdahaleye dönüşeceği çoğu zaman açık değildir. Bu tez, bu iki sorunu birlikte ele almaktadır. Amaç, gözlem hattı değişse bile geçerliliğini koruyan ve belirsizlik altında savunucunun beklenen faydası için doğrulanabilir bir taban değer, yani performans alt sınırı, üreten bir karar çerçevesi kurmaktır. Bu amaçla, önsel belirsizlik ile fayda ve maliyet belirsizliği aynı model içinde birleştirilmiş ve bunları birlikte ele alan belirsizliğe dayanıklı bir doğrusal program geliştirilmiştir. Bu program, hangi müdahalenin hangi olasılıkla seçileceğini hesaplamaktadır. Aynı zamanda bu seçimin sağlayacağı asgari performans düzeyini de sayısal olarak vermektedir. Modelin ikili çözümü, bu alt sınırı en çok zayıflatan belirsizlik kaynaklarını da göstermektedir. Önce kuramsal model kontrollü deneylerde sınanmıştır, ardından aynı yaklaşımın gerçek log verisi üzerinde çalışıp çalışmadığı DARPA OpTC NCR verisi kullanılarak test edilmiştir. Bu aşamada kullanıcı ve varlık davranış analizi ile olay köken izi

tabanlı sistemlerden elde edilen şüphelilik skorları konformal kalibrasyonla p-değerine dönüştürülmüş ve risk kademelerine ayrılmıştır. Müdahaleler, toplam bütçe kısıtı altında aynı doğrusal programla belirlenmiştir. Kuramsal deneyler, belirsizlik arttıkça garanti edilen alt sınırın öngörülebilir biçimde azaldığını, buna karşılık log kaybı ve bozulmuş log içeriği altında geçerliliğini koruduğunu göstermiştir. Değerlendirilen OpTC veri dilimlerinde ve seçilen bütçe bandında, önerilen yaklaşım risk-nötr kıyasa göre γ-uyumda yaklaşık yedi puanlık bir artış göstermiştir. Ayrıca zararsız olaylara gereksiz ağır müdahaleyi azaltmış ve ortalama faydayı büyük ölçüde korumuştur. Sonuçlar, seçilen veri ve belirsizlik modeli altında, yaklaşımın hem kuramsal olarak gerekçelendirilebildiğini hem de gerçek log verisi üzerinde ampirik olarak desteklendiğini göstermektedir.

Şubat 2026 , 90 sayfa.

Anahtar kelimeler: Gelişmiş Kalıcı Tehdit (APT), gözlemden bağımsız performans alt sınırı, belirsizliğe dayanıklı doğrusal programlama, olay köken izi, konformal kalibrasyon, UEBA karar katmanı

ABSTRACT

Ph.D. THESIS

Advanced Persistent Threat Detection Based on User Behaviors and Risk Predictions

Mustafa İPEKBAYRAK

İstanbul University-Cerrahpaşa

Institute of Graduate Studies

Department of Computer Engineering

Computer Engineering, PHD Programme

Supervisor: Assoc. Prof. Dr. Gülsüm Zeynep GÜRKAŞ AYDIN

In Security Operations Centers (SOCs), the defender makes repeated decisions against Advanced Persistent Threat (APT) attacks based on evidence derived from system, network, and device logs. However, log loss, changes in how logs are collected and processed, and attackers' attempts to erase traces may cause the same event to appear differently. In addition, it is often unclear under what conditions the suspicion scores produced by User and Entity Behavior Analytics and provenance-based systems should be translated into which response actions. This thesis addresses these two problems together. Its aim is to establish a decision framework that remains valid even when the observation pipeline changes and that produces a verifiable baseline value for the defender's expected utility under uncertainty, namely a performance floor. To this end, prior uncertainty and reward-cost uncertainty are combined within a single model, and a robust linear program is developed to handle them jointly. This program computes the probability with which each response action should be selected. At the same time, it provides a numerical value for the minimum performance level guaranteed by that choice. The dual solution of the model also identifies the sources of uncertainty that weaken this lower bound the most. The theoretical model is first examined through controlled experiments, then the same approach is then tested on real log data using the DARPA OpTC

NCR dataset to determine whether it works in practice. At this stage, the suspicion scores obtained from User and Entity Behavior Analytics and provenance-based systems are converted into p-values through conformal calibration and grouped into risk tiers. Response actions are then determined by the same linear program under a global budget constraint. The theoretical experiments show that, as uncertainty increases, the guaranteed lower bound decreases in a predictable manner, while remaining valid under log loss and degraded log content. Across the evaluated OpTC data slices and within the selected budget range, the proposed approach improves γ-compliance by approximately seven points relative to a risk-neutral benchmark. It also reduces unnecessarily severe responses to benign events while largely preserving average utility. These results show that, under the chosen data and uncertainty model, the approach is both theoretically justified and empirically supported on real log data. ⌋

February 2026, ⌈90⌋ pages.

Keywords: ⌈Advanced Persistent Threat (APT), observation-invariant performance floor, robust linear programming, provenance, conformal calibration, UEBA decision layer⌋

1. GİRİŞ

Ardışık savunma problemleri, siber izleme ve müdahale ortamlarında savunucunun birbirine benzer kararları tekrar almak zorunda kaldığı ve savunucunun planlı hamleler yapan bir saldırgana karşı kararlar aldığı durumları ifade eder. Güvenlik Operasyon Merkezlerinde (SOC) bu karar döngüsü çoğu zaman sistem günlük kayıtları (log), ağ ve cihaz logları ile farklı kaynaklardan gelen kanıtlar üzerinden yürütülür. Gelişmiş Kalıcı Tehdit (APT) saldırıları zaman içinde birbirine bağlı birçok adımdan oluştuğu ve düşük görünürlükte ilerlediği için savunucunun yanlış alarmları azaltıp gerçek saldırıyı engelleyebilmesi tekrarlı ve tutarlı bir karar sürecini gerektirir [1]–[3]. APT saldırıları veriyi kurum dışına çoğu zaman kurumun zaten izin verdiği çıkış kanalları üzerinden çıkarır. Komuta-kontrol iletişimi ve veri sızdırma, HTTPS trafiği veya izinli bulut servisleri içinde normal ağ etkinliği gibi gösterilebilir. Saldırgan görünürlüğünü azaltmak için PowerShell, WMI ve benzeri meşru yönetim araçlarını kullanır, faaliyetini normal süreçlerin arasına dağıtır ve fırsat bulduğunda iz silme girişimlerinde bulunur [1], [22], [23]. Buna ek olarak kayıt kaybı, ayrıştırıcı kayması ve zaman damgası uyuşmazlığı gibi gözlem hattı sorunları aynı olayın farklı zamanlarda farklı görünmesine neden olabilir [1], [22], [23]. Bu nedenle belirli bir log üretim ve toplama düzenine bağlı çalışan sistem güvenliği karar mekanizması, log hattı değiştiğinde veya bozulduğunda tutarsızlaşır.

Bu kırılganlığın arka planında üç alanın kesişimi gözlenmiştir. Birincisi, olay köken izi tabanlı APT tespitidir. İşletim sistemi düzeyinde süreç, dosya ve ağ varlıkları arasındaki etkileşimler yönlü ve zaman damgalı bir grafa dönüştürülerek saldırının neden-sonuç zinciri izlenebilir hâle getirilir [1], [2]. İkincisi, SOC'un operasyonel karar döngüsüdür. Tespit sistemleri bir skor üretir ancak bu skorun hangi koşulda hangi aksiyona dönüşeceği, analist kapasitesi ve müdahale maliyetiyle birlikte değerlendirilmelidir [17]–[20], [23]. Üçüncüsü, belirsizlik altında karar verme problemidir. Başlangıç varsayımlarının yanlış çıkması ve aksiyonların fayda/zararının tam bilinememesi, ortalama koşullara göre tasarlanmış olan karar kurallarının kötü durum senaryolarında düşük performanslı çalışmasına neden olabilir [82]–[84]. Bu üç alan birlikte ele alınmadığında, tespit çıktısı ile SOC kararı arasındaki bağ ya gözlem hattına sıkı biçimde bağımlı çalışır ya da belirsizlik karşısında güvenilmez olur.

Bu sorunların sahada ne ölçüde karşılık bulduğu, olay köken izi tabanlı APT tespiti literatüründe görünür biçimde belgelenmiştir. İncelenen çalışmaların önemli bir bölümü log kayıt kayıplarını ve log toplama/ayrıştırma bozulmalarını bir sınırlılık olarak raporlamakla birlikte, bu bozulmalar altında sistematik stres testleri ve doğrulanabilir dayanıklılık raporlaması çoğu zaman sınırlı kalmaktadır [1]–[3], [23]. Buna ek olarak, tespit sistemlerinin ürettiği skorların SOC iş akışında hangi aksiyona nasıl dönüşeceği çoğu çalışmada tutarlı biçimde tanımlanmamaktadır. Değerlendirme birimi ile operasyonel uyarı birimi arasındaki bu yapısal uyumsuzluk, başarım raporu ile sahadaki müdahale kalitesinin doğrudan ilişkilendirilmesini zorlaştırmaktadır [1], [3], [23]. Bu iki darboğaz, gözlem hattına bağımlı olmadan oluşturulacak bir performans güvenilirliği skorunu ve bu güvenilirlikle tutarlı, beklenen ortalama harcamayı kısıtlayacak bütçe kısıtlı bir karar mekanizması ihtiyacını birlikte doğurur.

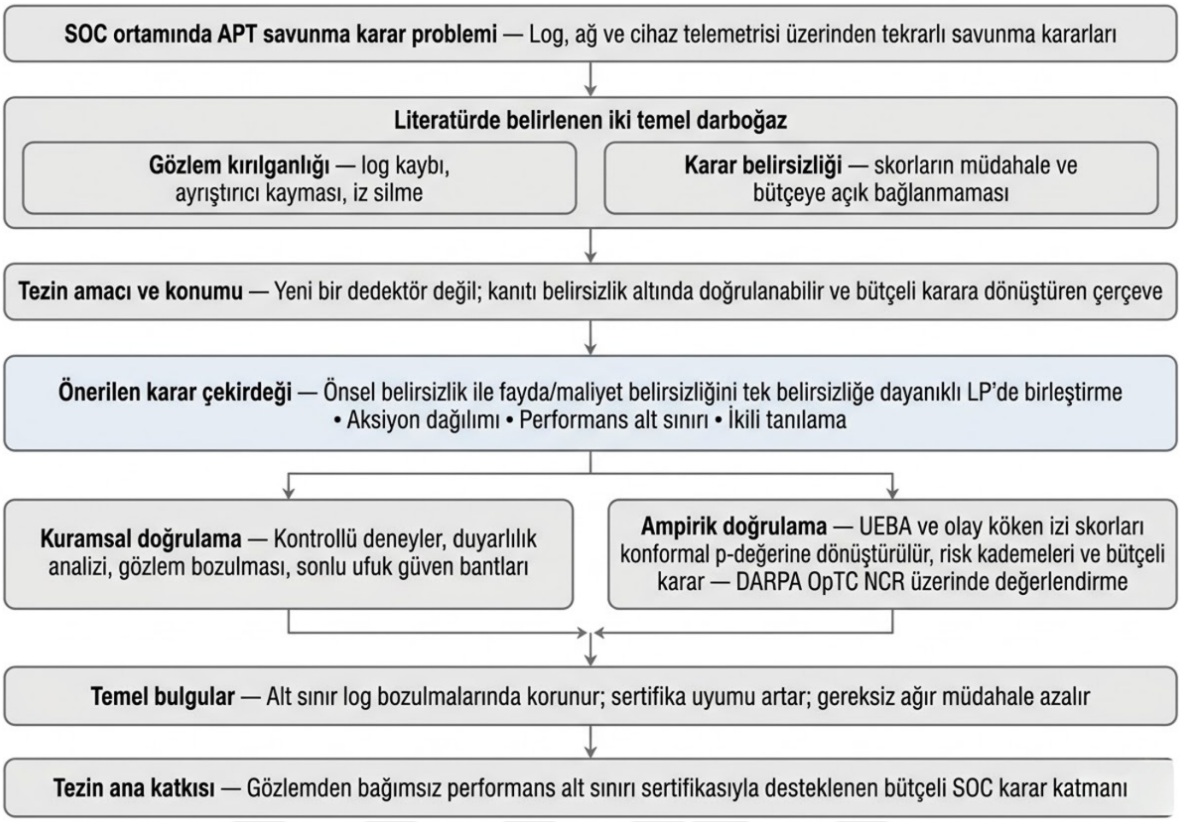
Bu tez, gözlem hattı değişse veya bozulsa bile savunucunun önceden ilan edebileceği ve sonradan doğrulanabilir bir performans alt sınırı üretmenin mümkün olduğunu göstermeyi ve bu alt sınırla tutarlı, bütçe kısıtlı bir SOC karar mekanizması kurmayı amaçlar. Bu bağlamda tez, yeni bir APT dedektörü önermekten çok, UEBA ve olay köken izi tabanlı sistemlerin ürettiği kanıtları belirsizlik altında aksiyona dönüştüren karar katmanı önerir. Çalışma, bu amacı üç araştırma sorusu etrafında ele alır. Birincisi, gözlem modeli kurmadan güvenilir bir performans alt sınırı üretmek mümkün müdür? İkincisi, böyle bir alt sınır üretilebiliyorsa, hangi koşullarda zayıfladığını sayısal olarak teşhis edip iyileştirme yönünü gösterebilir miyiz? Üçüncüsü, bu garantiyi ve teşhis mekanizmasını bütçe kısıtlı SOC kararlarına, yani "hangi olaya ne yapılacak" sorusuna, doğrulanabilir biçimde bağlayabilir miyiz?

Bu sorulara yanıt üretmek için tez, belirsizliği açık biçimde modelleyen ve tek bir belirsizliğe dayanıklı doğrusal program (LP) ile aksiyon dağılımını ve bu dağılımın performans alt sınırını birlikte üreten bir çerçeve önerir. Burada amaç, yeni bir APT dedektörü geliştirmekten çok, dedektör çıktısını belirsizlik ve bütçe kısıtları altında doğrulanabilir bir karara dönüştürmektir. Bu nedenle yöntem çekirdeğinde LP kullanılmıştır. Bunun nedeni, aksiyonların, maliyetlerin, belirsizlik kümelerinin ve garanti edilen performans alt sınırının aynı model içinde açık ve denetlenebilir biçimde yazılabilmesidir. Çerçevenin ayırt edici özelliği, üretilen garantinin gözlemi üreten mekanizmaya bağımlı olmamasıdır. Log üretim düzeni değişse, log akışı eksilse veya sinyaller karışsa bile performans alt sınırı, tanımlanan belirsizlik varsayımları altında geçerliliğini korur. Bu özellik, SOC için somut bir operasyonel kazanım

sağlar. Süreç, dosya ve ağ telemetrisinin formatı değiştiğinde karar mekanizmasını baştan kalibre etmek yerine aynı güvence taşınabilir; bütçe kararları denetlenebilir bir tabana oturur ve analist iş yükü daha öngörülebilir hâle gelir.

Bu çerçeve altında tezin temel katkıları iki tanedir. Birincisi, önsel belirsizliği ve fayda/maliyet belirsizliğini birlikte kapsayan tek bir belirsizliğe dayanıklı doğrusal program formülasyonu önerilmiş; bu formülasyon gözlem modeli kurmadan doğrulanabilir bir performans alt sınırı, savunucunun karma stratejisini ve ikili formdan türetilen tanılama mekanizmasını birlikte üretmiştir. Sonlu karar turu sayısı için varyansa duyarlı güven bantları ile alt sınır sahada raporlanabilir hâle getirilmiştir. İkincisi, bu sertifika ilkesi SOC karar katmanına taşınmış; UEBA ve olay köken izi tabanlı dedektör skorları konformal kalibrasyon ile risk kademelerine dönüştürülmüş ve küresel bütçe kısıtı altında aksiyon dağılımı belirsizliğe dayanıklı tek bir LP ile belirlenmiştir. Bu iki temel katkıyı destekleyen iki ek çalışma da tezde yer almaktadır. Olay köken izi tabanlı APT tespiti literatüründen 76 çalışma özgün bir kodlama şemasıyla incelenmiş ve gözlem kırılganlığı ile karar birimi belirsizliği iki temel darboğaz olarak belirlenmiştir. Ayrıca sabit alarm eşiğinin alan kayması altında taşınabilir olmadığı ayrı bir destekleyici çalışmayla gösterilerek konformal kanıta dayalı karar arayüzünün gerekliliği doğrulanmıştır.

Genel araştırma akışı Şekil 1.1'de özetlenmiştir. Şekil, log tabanlı APT savunmasında belirlenen iki temel darboğazdan hareketle tezin amaç ve konumunu, önerilen karar çekirdeğini, kuramsal ve ampirik doğrulama hatlarını, temel bulguları ve nihai katkıyı göstermektedir.



Şekil 1.1: Tezin genel araştırma akışı ve önerilen karar çerçevesi

Tezin geri kalanı şu şekilde düzenlenmiştir. Bölüm 2, olay köken izi tabanlı APT tespiti literatürünü değerlendirme-öncelikli bir bakışla inceler; kodlama şeması ve kapsam bulguları üzerinden iki temel darboğazı tanımlar ve tezin dayandığı kavramsal temelleri ortaya koyar. Bölüm 3, belirsizlik modelini, tek LP formülasyonunu, sonlu ufuk güven bantlarını ve UEBA karar katmanının yöntemsel adımlarını tanımlar. Bölüm 4, ardışık savunma deneyleri ile SOC karar katmanı deneylerinin bulgularını raporlar. Bölüm 5, bulguları tartışır; yaklaşımın kapsadığı alanı, sınırlılıkları ve operasyonel çıkarımları ele alır. Bölüm 6, net sonuçları özetler ve geleceğe yönelik önerileri sunar.

2. KAVRAMSAL ÇERÇEVE

Bu bölüm, log kayıtlarından türetilen olay köken izine dayalı Gelişmiş Kalıcı Tehdit (APT) tespiti literatürünü değerlendirme-öncelikli bir bakış açısıyla ele alır. Literatür taraması; literatürdeki boşlukların belirlenmesi, farklı çalışmaların kapsamlarının birbiriyle kıyaslanarak benzerlik ve farklılıkların ortaya konulması ve bu tezde geliştirilen kuramsal ve ampirik çalışmaların doğru biçimde konumlandırılabilmesi amacıyla yürütülmüştür [1]–[3]. Bu çerçevede amaç, önerilen yöntemleri yalnızca sınıflandırma doğruluğu veya ayırt edicilik gibi dar performans ölçütleri üzerinden değil; üretilen çıktının hangi kanıt biriminde temsil edildiği (düğüm, altgraf ya da tüm graf), bu çıktının Güvenlik Operasyon Merkezi (SOC) iş akışında hangi karar süreçlerine hizmet ettiği ve belirsizlik ile bütçe kısıtları altında sonuçların operasyonel olarak nasıl karşılaştırılabilir hâle getirildiği üzerinden tartışmaktır [1], [3].

Bu amaç doğrultusunda, 2018–2025 döneminde yayımlanmış ve 20 Ekim 2025 tarihine kadar derlenmiş çalışmalar toplanmış; yöntemsel adımlar, SOC bağlamında hedeflenen işlevler ve raporlama biçimleri başlıkları üzerinden kategorilendirilmişlerdir. [1], [2], [17]–[19], [23]. İncelenen çalışmaların önemli bir bölümünde, makale düzeyinde ileri sürülen kapsam iddiaları ile deneysel değerlendirmede fiilen test edilen koşulların tam olarak örtüşmediği görülmüştür; örneğin bazı çalışmalar yalnızca tek bir cihazdan toplanan iz kayıtlarıyla, saldırının sınırlı bir bölümünü kapsayan veya yalnızca çevrimdışı veri üzerinde yürütülen bir kurulumla sonuç raporlamaktadır [1], [3]. Bu nedenle makaleler arası karşılaştırılabilirliği artırmak için, tek bir performans ölçütüne (örneğin ROC-AUC/PR-AUC) odaklanmak yerine uçtan uca operasyonel zinciri görünür kılan bir metrik seti benimsenmiştir. Makalelerdeki analiz tipi, SOC fonksiyonu, raporlama birimi ve değerlendirmenin yapıldığı birim çekirdek alanlar olarak ele alınmıştır. Ayrıca deney modu (çevrimdışı/akış), kapsam (tek cihaz/çoklu cihaz), veri kaynağı, çıktı türü ve sınırlılık–dayanıklılık raporlaması sistematik biçimde şemaya dahil edilmiştir [1], [2], [23].

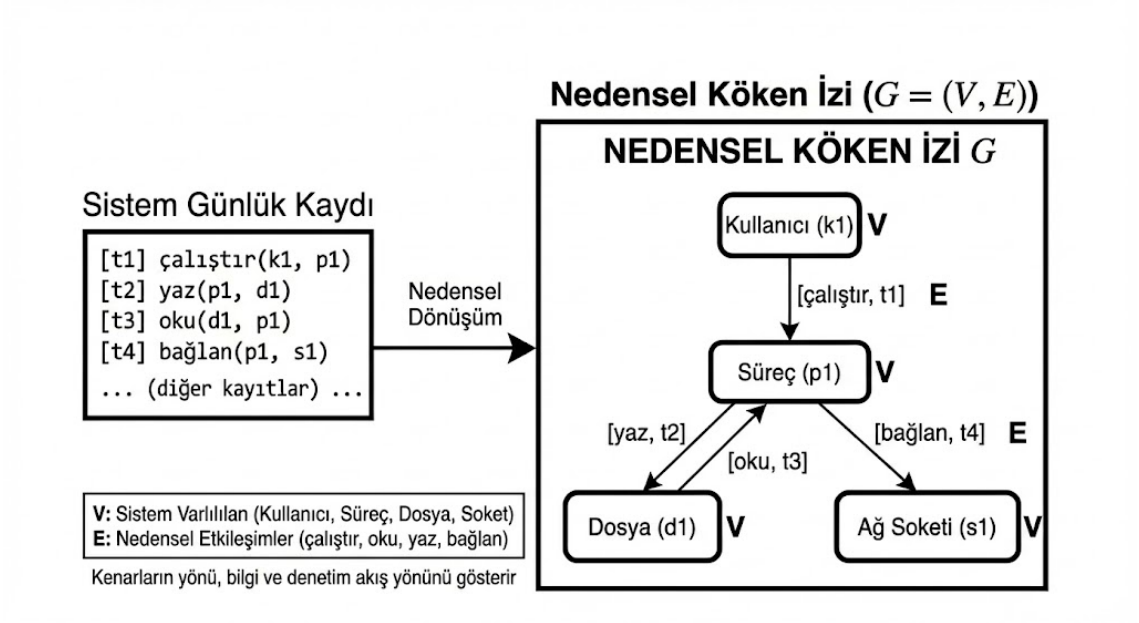
Bölüm boyunca, saldırı aşamalarının tutarlı biçimde ifade edilebilmesi için MITRE ATT&CK ile uyumlu bir terminoloji benimsenmiştir. Böylece farklı çalışmaların "hangi evreleri kapsadığı" iddiaları ortak bir payda üzerinden karşılaştırılabilir hâle getirilmiştir [1], [3], [4]. Kullanılan veri kümeleri ve yöntemler incelendiğinde, literatür çalışması tezin odaklandığı problem alanıyla örtüşmektedir. Olay köken izi (provenance) düzeyindeki tespit ve

puanlama çıktılarının SOC ortamında aksiyona dönüştürülmesinde gözlem hattındaki kayıt eksikliği, güvenlik telemetrisi heterojenliği ve log üretim/toplama hattı bozulmalarının oluşturduğu kırılganlıklar ele alınmaktadır. Bu çerçevede, gözlem koşullarındaki bu tür sapmalara rağmen geçerliliğini koruyabilen doğrulanabilir güvenli performans aralıklarının hangi varsayımlar altında kurulabildiği tartışılmaktadır. Bölümün devamında, belirsizlik altında üretilen şüphe olasılık skoru ve uyarıların operasyonel kararlara bağlanmasını mümkün kılan belirsizliğe dayanıklı karar verme yaklaşımlarının kavramsal temeli ortaya konulmaktadır [1], [3].

2.1. Nedensel güvenlik telemetrisi olarak olay köken izleri

APT saldırıları zaman içinde gerçekleşen ve birbirine bağlı birçok adımdan oluşur [1]–[3]. Bu adımlar icra edilirken saldırganlar, tespit mekanizmalarından kaçınmak (evasion) amacıyla sisteme dışarıdan özel zararlı yazılımlar indirmek yerine, çoğu zaman işletim sisteminin yasal yönetim araçlarını (PowerShell, WMI, vb.) istismar ederek kendi aktivitelerini meşru süreçlerin arasına gizlerler (Living off the Land). Bu yüzden savunma tarafının yanlış alarmları azaltıp gerçek saldırıyı engelleyebilmesi için tekil alarmlarla yetinmeyip olaylar arasındaki neden sonuç ilişkisini ortaya koyan kanıta ihtiyaç duyduğu belirtilmiştir [1]–[2].

Uygulamada olay köken izi, ham log kayıtlarının basit birleştirilmesiyle değil; bu kayıtların yönlü, tiplenmiş ve zaman damgalı bir grafa dönüştürülmesiyle temsil edilmektedir. Tipik gösterimde G=(V,E) grafında düğümler (V) süreç, dosya ve ağ varlıklarını; kenarlar (E) ise bu varlıklar arasındaki bilgi/etkileşim akışını temsil etmektedir [2], [24]. Bir varlığa ya da sisteme ilişkin olay kaydı üretilmediğinde veya toplanamadığında, o varlık gözlenen köken grafında yer almaz. Bu tezde bu durum eksik gözlem olarak ele alınmıştır. Şekil 2.1, bahsi geçen nedensel köken izi yapısını göstermektedir. Kenar yönünün bilgi akışını kodlaması sayesinde, nedensel köken verisi grafları üzerinden ileriye dönük izdüşümleme (etkinin yayılımı) ve geriye dönük iz sürme (köken analizi) işlemlerinin gerçekleştirilebildiği gösterilmiştir [24], [57].



Şekil 2.1: Nedensel köken izi grafı

Güvenlik telemetrisi hattının görünürlüğünün yakalama altyapısına ve kayıtlandırma ilkesine bağlı olduğu; kayıt kaybı, ayrıştırıcı (parser) kayması, zaman damgası uyuşmazlığı ve hedefli kurcalama/manipülasyon (tampering; ör. iz silme) gibi saha etkilerinin aynı olayın farklı zamanlarda farklı görünmesine yol açabildiği raporlanmıştır [1], [22], [23]. Bu nedenle köken analizi tabanlı tespit sistemlerinin yalnızca doğruluk açısından değil; gözlem hattındaki değişkenlik altında taşınabilirlik ve doğrulanabilirlik açısından da değerlendirilmesi gerektiği belirtilmiştir [1], [3].

2.2. Olay köken izi tabanlı APT tespitinde yöntem aileleri ve SOC işlevleri

Olay köken izi tabanlı APT tespit literatüründe yöntemlerin, aykırılık/öğrenme tabanlı tespit; akış/çevrimiçi tespit; kural/örüntü eşleme tabanlı avcılık; ve açıklanabilirlik/atıf odaklı inceleme hatları altında kümelendiği görülmüştür [1], [2]. Aykırılık ve temsil öğrenme çizgisinde, köken graflarından düğüm/kenar öznitelikleri türetilmekte; normal davranışın sayısal temsili kurulmakta ve sapmalar risk skoru veya alarm olarak raporlanmaktadır [5], [6], [41], [62].

Akış/çevrimiçi tespit hattında, gecikme kısıtları altında gerçek zamanlı skor üretimi ve alarm oluşturma hedeflenmiştir. Bu doğrultuda çalışma modunun akış kipinde (streaming) olarak ele alındığı ve sistemin sahaya yakın koşullarda işletilmeye çalışıldığı örnekler sunulmuştur [17]–[19], [65]. Öte yandan, kural ve örüntü eşleme temelli yaklaşımlarda bilinen

saldırı davranışlarının nedensel köken grafı üzerinde altgraf eşleme veya davranış hizalama yöntemleriyle aranması hedeflenmiş; çıktının çoğu zaman analiste sunulan yerel bir kanıt parçası (yol/motif) olarak raporlandığı belirtilmiştir [29], [30], [40], [52]. Açıklanabilirlik ve atıf odaklı çalışmalarda, bir alarmın gerekçesi olarak kanıt yolunun/altgrafın üretilmesi; yani "neden şüpheli" sorusuna kanıta dayalı yanıt verilmesi amaçlanmıştır [20], [23]. Olay incelemesi ve yeniden kurulum (reconstruction/storyline) hatlarında ise, köken analizi üzerinden saldırı senaryosunun yeniden kurulumu veya analistin etkileşimli soruşturmasına uygun bir hikâye çizgisinin oluşturulması hedeflenmiştir [23], [24], [66].

Son yıllarda GNN tabanlı temsil öğrenme, maskeleme/kontrastif öğrenme ve uzun dönem davranış öznitelikleri gibi tekniklerin köken analizi tabanlı APT tespitinde yaygınlaştığı raporlanmıştır [6], [11], [62], [71], [74]. Buna karşın, yöntemsel çeşitliliğe rağmen, bu çıktıların SOC iş akışında hangi karar nesnesine bağlandığının ve hangi birimde değerlendirildiğinin her çalışmada aynı açıklıkla raporlanmadığı gözlenmiştir [1], [3], [23].

2.3. Karşılaştırılabilirlik sorunu ve karar birimi

Olay köken izi tabanlı APT tespiti literatüründe karşılaştırılabilirlik sorununun, tek bir performans ölçütünün yokluğundan ziyade; gözlem hattı varsayımlarının, raporlama birimlerinin ve SOC çıktısının tutarsız biçimde tanımlanmasından kaynaklandığı değerlendirilmiştir [1]–[3]. Benzer sayısal başarı değerlerinin farklı uyarı hacmine, farklı analist iş yüküne ve farklı müdahale maliyetlerine karşılık gelebileceği; dolayısıyla "ne ölçüldüğü" kadar "hangi birimde ölçüldüğü"nün de açık biçimde yazılması gerektiği vurgulanmıştır [1], [23].

Bu kapsamda iki kavram ayrıştırılmıştır. Birinci kavram olan değerlendirme birimi, başarımın hesaplandığı değerlendirme birimini (Node/Subgraph/Graph) ifade etmektedir. İkinci kavram olan uyarı birimi ise analiste sunulan veya otomasyona bağlanan operasyonel çıktı birimini ifade etmektedir. Literatürde bu iki birimin çoğu zaman hizalı olmadığı; örneğin düğüm düzeyinde ölçülen bir ayırt edicilik değerinin, graf düzeyinde alarm üreten bir sistemin operasyonel kalitesine doğrudan karşılık gelmediği görülmüştür [1], [3], [23].

Bu tez kapsamında ayrıca karar birimi kavramı vurgulanmıştır. Karar birimi, SOC içinde aksiyonun uygulandığı nesnenin hangi düzeyde tanımlandığını göstermektedir (ör. tekil olay, risk kademesi, makine (host), altgraf veya kampanya görünümü). Karar biriminin açık

biçimde tanımlanmadığı durumda, tespit skorlarının hangi eşiğe göre nasıl bir aksiyona dönüştüğü belirsiz kalmakta; dolayısıyla ölçütler ile operasyonel çıktı arasındaki ilişki zayıflamaktadır [1], [22], [23].

Karşılaştırılabilirliği zayıflatan bir diğer eksen, deney protokollerinin ve veri bölme kurallarının açıklığıdır. Köken graflarının zaman ve sunucu (host) bağımlılığı taşıması nedeniyle zaman/sunucu/kampanya sızıntısı riskinin bulunduğu; bu riskin kontrol edilmediği durumlarda performansın olduğundan yüksek raporlanabildiği belirtilmiştir [1], [23]. Benzer protokol bağımlılığı ve sızıntı risklerinin, grafik tabanlı güvenlik analitiğinin yakın alanlarında da raporlandığı görülmüştür [80], [81].

2.4. Kodlama şeması ve kullanılan metrikler

Bu tez kapsamında, literatürdeki iddia ve kapsam uyuşmazlıklarını somutlaştırmak ve çalışmaları ortak bir düzlemde karşılaştırabilmek amacıyla, operasyonel zinciri bir metrik seti üzerinden kodlayarak değerlendirme birimi ile uyarı birimi arasındaki yapısal uyumsuzluğu ölçülebilir hâle getiren özgün bir kodlama şeması geliştirilmiştir; bu şema heterojen raporlamaları “standart birim” altında toplarken eksik verileri “uygulanamaz” (NA) olarak işaretleyerek analiz örnekleminin şeffaflığını ve sonuçların doğrulanabilirliğini güvence altına almaktadır [1], [3].

Bu tez kapsamında kullanılan metrik kümeleri ve her bir kümenin karşılaştırma içindeki işlevi Tablo 2.1’de özetlenmiştir. Özellikle Tablo 2.1’de yer alan APT evre kapsamı alanı, kampanya kapsamının yalnızca doğruluk skorlarıyla değil, saldırının gözlenebilir evrelerinin ne ölçüde kapsandığıyla birlikte değerlendirilmesini amaçlamaktadır. Bu tezde kampanya kapsamı, Campaign Recall (CR) mantığıyla ele alınmıştır. Buna göre CR = k/N olup, burada N ilgili senaryoda gözlenebilir evre sayısını, k ise en az bir doğru alarm/kanıt ile yakalanan evre sayısını ifade eder. Bu hesaplamada N değeri, MITRE ATT&CK ile uyumlu sabit 8-evrelik standart ontolojiye (Initial Access, Execution, Privilege Escalation, Command & Control, Discovery and Collection, Lateral Movement, Persistence, Exfiltration/Impact) dayanmaktadır. Aynı evre içinde üretilen birden fazla alarm k değerini artırmaz. Örneğin gözlenebilir 4 evreden 2’si doğru biçimde yakalanmışsa kampanya kapsamı CR = 2/4 = 0,50 olarak raporlanır.

Tablo 2.1: Metrik kümeleri ve amaçları

Küme	Metrikler	Amaç / kullanım
A) Korpus ve deney bağlamı	Yayın türü (konferans/dergi/çalıştay); deney modu (çevrimdışı/akış: offline/streaming); kapsam (tek makine/çoklu makine: single-host/cross-host)	Korpus profilinin ve deney bağlamının standardize edilmesi
B) Yöntem ve gözlem hattı	Analiz tipi (anomali-ML, graf örüntü, kural/etiket, hibrit: anomaly-ML, graph-pattern, rule/tag, hybrid); SOC fonksiyonu (ön değerlendirme, tehdit avcılığı, olay örgüsü, yeniden kurma, tespit, görselleştirme: triage, hunting, storyline, reconstruction, detection, visualization); çok-fonksiyonlu çalışmalarda birincil fonksiyon ataması	Yöntemin hangi SOC ihtiyacına hizmet ettiğinin ve gözlem hattı varsayımlarının kodlanması
C) Raporlama birimi ve çıktı	Değerlendirme birimi ve uyarı birimi (evaluation unit / alert unit) (düğüm/altgraf/graf: node/subgraph/graph); çıktı türü (uyarılar, düğüm puanları/sıralamaları, olay örgüleri, atfedilmiş kenarlar/kanıt altgrafları: alerts, node scores/rankings, storylines, attributed edges/evidence subgraphs); eşleme düzeyi (katı vs birleştirilmiş: strict vs consolidated)	Ölçütlerin operasyonel çıktıyla hizalanmasının doğrulanması
D) APT evre kapsamı	Metin/şekil/tablolardaki evre iddialarının ATT&CK-hizalı bir ontolojiye normalize edilmesi	Evre iddialarının ortak bir dilde raporlanması
E) Limitasyon ve dayanıklılık	Limitasyon kümeleri (veri kümesi yanlılığı, loglama boşlukları, yanlış pozitifler, ölçeklenebilirlik/bellek, kurcalama riski, çoklu makine, yorumlanabilirlik, etiket kıtlığı: dataset bias, logging gaps, false positives, scalability/memory, tamper risk, cross-host, interpretability, label scarcity); raporlanan dayanıklılık tartışmaları (kaçınma/taklit/zehirleme: evasion/mimicry/poisoning)	Saha etkilerinin ve saldırgan uyarlamasının ne ölçüde tartışıldığının görünür kılınması

Tablo 2.1'deki çekirdek alanlara ek olarak platform ve telemetri kaynağı (örn. auditd/ETW), veri kümesi tipi (örn. DARPA TC, OpTC NCR), etiket ve referans etiket (ground truth) durumu, giriş temsili (graph/node/subgraph), model ailesi (GNN, otoenkoder, sıralı modeller, kural tabanlı yaklaşımlar) ile ölçek ve kalite hedefleri (bellek, gecikme, yanlış pozitif, bütünlük/tamper riski) gibi alanlar yardımcı üstveri olarak kaydedilmiştir [1]–[3], [17]–[19], [23], [57], [112].

CR hesabında referans yapı, MITRE ATT&CK taktiklerinden türetilmiş sekiz aşamalı saldırı yaşam döngüsü ontolojisidir. Ancak incelenen makaledeki veri kümesi bu aşamaların tümünü kapsamıyorsa payda sabit 8 alınmaz. Payda, veri kümesinde gözlenebilir olan aşama sayısıdır. Bu nedenle gözlenebilir aşama sayısı 5 ise CR, k/5 olarak hesaplanır.

2.5. Literatür taramasından elde edilen kapsam bulguları

Kodlanan korpus üzerinde yapılan kapsam analizi, literatürün ağırlık merkezinin anomali temelli makine öğrenmesi (anomaly-ML), çevrimdışı deney kurgusu (offline) ve tek cihaz kapsamı (single-host) ekseninde toplandığını göstermiştir. Bu eğilimde, köken verisi grafı

(provenance graph) üzerinde temsil öğrenme, yeniden kurma ve sıralı temsil yaklaşımları kullanılarak çoğunlukla tespit (detection) çıktısı üretildiği raporlanmaktadır [1], [2], [6], [41], [62]. Yayın türü dağılımı konferans 42/76 (%55,3), dergi 33/76 (%43,4) ve çalıştay 1/76 (%1,3) olarak gözlenmiş; isimli bir çerçeve (framework) sunan ve dedektör-odaklı (detector-focused) katkı üreten çalışmaların oranı 55/76 (%72,4) düzeyindedir. Benzer biçimde anomali temelli yaklaşım içeren çalışmalar 55/76 (%72,4) iken, deney modunda çevrimdışı yürütülen çalışmalar 48/76 (%63,2), akış (streaming) içeren çalışmalar 27/76 (%35,5) olarak bulunmuştur. Kapsam açısından tek cihaz (single-host) 66/76 (%86,8) ile baskınken, çoklu cihaz (cross-host) kapsamı 10/76 (%13,2) ile sınırlı kalmıştır.

Raporlama birimi ve birim uyumuna ilişkin analiz, değerlendirme tanecik düzeyini (evaluation granularity) raporlayan çalışmaların 74/76 olduğunu; bunların 33/76'sında düğüm (node), 26/76'sında tüm graf (graph) ve 15/76'sında altgraf (subgraph) düzeyinin kullanıldığını göstermiştir. Buna karşılık, uyarı birimi (alert unit) ile değerlendirme birimini (evaluation unit) eşleyebilen veya uyarı birimini açıkça raporlayan çalışmaların oranı düşüktür; 40/76 (%52,6) çalışmada uyarı birimi eşlenememiş ya da raporlanmamıştır. Uyarı birimi eşlenebilen 36 çalışmada uyarı birimi dağılımı 14'ünde grafik, 11'inde altgraf ve 11'inde düğüm düzeyindedir. Uçtan uca akışın, yani analiz türü → SOC işlevi → uyarı çıktısı zincirinin çıkarılabildiği alt-küme bu 36 çalışmayla sınırlı kalmış; bu alt-kümede en sık gözlenen akış anomali temelli makine öğrenmesi → tespit → grafik düzeyinde uyarı biçiminde olup 9/36 (%25,0) oranında raporlanmıştır. Değerlendirme birimi ile uyarı birimi arasındaki köşegen uyum (Graph–Graph, Node–Node, Subgraph–Subgraph) 29/70 (%41,4) olarak hesaplanmış; en yaygın uyumsuzlukların düğüm düzeyi değerlendirme ile grafik düzeyi uyarı (22/70) ve grafik düzeyi değerlendirme ile düğüm düzeyi uyarı (13/70) biçiminde ortaya çıktığı görülmüştür.

Çıktı türleri ve dayanıklılık/sınırlılık raporlaması incelendiğinde, çoklu etiketli olarak sayılan çıktı türlerinin uyarılar (alerts) 43, düğüm puanları (node scores) 27, olay örgüleri (storylines) 25 ve atfedilmiş kenarlar (attributed edges) 8 biçiminde dağıldığı belirlenmiştir. Sınırlılık başlıkları içinde en sık raporlanan temalar veri kümesi yanlılığı (dataset bias) 20, yanlış pozitifler (false positives) 17 ve kayıt boşlukları (logging gaps) 17 olarak öne çıkmaktadır. Açık dayanıklılık testlerinin sınırlı kaldığı; kaçınma (evasion) 9, taklit (mimicry) 6 ve zehirleme (poisoning) 4 başlıklarının düşük sıklıkla raporlandığı görülmüştür [103], [104]. Bu bulgular, literatürde operasyonel zincirin çoğu çalışmada baştan sona ve doğrulanabilir biçimde raporlanmadığını göstermektedir. Özellikle değerlendirme birimi ile uyarı birimi

arasındaki hizasızlığın, ROC/PR gibi ölçütlerin doğrudan SOC çıktısına taşınmasını zayıflattığı; ayrıca kayıt bütünlüğü ve kayıt boşlukları gibi saha koşulları sıkça sınırlılık olarak belirtilmesine karşın, sistematik dayanıklılık testlerinin görece az raporlandığı değerlendirilmektedir [1], [23], [46].

2.6. Literatür bulgularının tez tasarımına etkisi

Yürütülen kapsamlı literatür analizi, olay köken izi tabanlı APT tespiti alanında çözülmesi gereken iki kritik operasyonel darboğazı görünür kılmıştır.

Birincisi, gözlem kırılganlığıdır: İncelenen çalışmaların önemli bir bölümü, saha koşullarındaki kayıt kayıplarını (logging gaps), log toplama/ayrıştırma bozulmalarını ve bütünlük risklerini bir sınırlılık olarak raporlamakta; ancak bu bozulmalar altında sistematik stres testleri ve doğrulanabilir belirsizliğe dayanıklılık raporlaması çoğu zaman sınırlı kalmaktadır. Bu bulgu, tezin ilk tasarım hedefini belirlemiştir: Telemetri kalitesi değişse dahi geçerliliğini koruyan ve doğrulanabilir biçimde raporlanabilen bir performans alt sınırı garantisi (certificate) üretmek. Bu ihtiyaç, tezin teorik katkısı olan Tek LP ile Sertifikalı Gözlemden Bağımsız Performans Alt Sınırı modelini motive etmiştir.

İkincisi, uyarı/karar birimi belirsizliğidir: Literatürde çalışmaların kayda değer bir kısmında üretilen uyarının (alert unit) hangi operasyonel nesneye karşılık geldiği açık değildir; ayrıca değerlendirme birimi ile uyarı birimi arasındaki geçişler çoğu zaman doğrulanabilir biçimde yazılmadığından, skorların analist iş yükü ve müdahale kapasitesiyle nasıl hizalanacağı belirsiz kalmaktadır. Bu boşluk, tezin ikinci tasarım hedefini şekillendirmiştir: Skorları soyut birer çıktı olmaktan çıkarıp, SOC'un kapasite/bütçe kısıtları altında uygulanabilir aksiyonlara ve açık bir karar birimine bağlayan bir karar katmanı kurmak. Bu ihtiyaç, tezin ampirik katkısı olan Risk Kademeli Bütçeli Karar Katmanı (UEBA uyumlu operasyonel karar kuralı) ile karşılanmıştır.

Tez çalışmalarının ilgili mevcut çalışmaların çekirdek alanlarıyla eşleşmesi Tablo 2.2'de özetlenmiştir.

Tablo 2.2: Tezde yürütülen iki çalışmanın ilgili mevcut çalışmalar ile eşleşmesi

Çalışma	Analiz tipi	SOC fonksiyonu	Karar/alert birimi	Değerlendirme birimi	Çıktı türü
Teorik çalışma	Belirsizliğe dayanıklı optimizasyon tabanlı; ardışık turlar	Garanti üretimi / doğrulanabilirlik	Tur bazında karar kuralı/aksiyon dağılımı	Tur başına fayda (per-round)	Performans alt sınırı: ($x^\star$, $v^\star$)
Pratik (ampirik) çalışma	Köken verisi +UEBA kanıtı; konformal kalibrasyon + belirsizliğe dayanıklı LP	Triage/decision layer (bütçeli aksiyon seçimi)	Risk kademesi (tier)	Kademe bazında fayda ve uyum	Policy + certificate (γ); adherence ve bütçe denetimi

2.7. Tek LP ile gözlemden bağımsız performans alt sınırı

Bu bölümde, belirsizlik altında karar vermeye ilişkin yaklaşımlar tek bir çerçevede ele alınmaktadır. Önerilen çerçeve, tek bir doğrusal programla hem savunucunun aksiyon dağılımını hem de bu dağılımın garanti ettiği performans alt sınırını üretmektedir. Bu tezin kuramsal katkısı, önsel belirsizlik ile getiri belirsizliğini aynı doğrusal program içinde birleştirerek karar dağılımı ile performans alt sınırını birlikte üreten bu formülasyondur. Bu katkının biçimsel tanımı Bölüm 3.4.2'de Denklem (3.5), bu formülasyonda kullanılan garanti fonksiyonunun kapalı biçimi ise Denklem (3.6) ile verilmektedir. Yaklaşımda önsel belirsizlik Huber tipi ε-kontaminasyon ile, getiri belirsizliği ise aralık ya da çokyüzlü belirsizlik kümeleriyle modellenmektedir. Amaç, yalnızca ortalama koşullarda iyi sonuç vermek değil, belirsizlik altında da doğrulanabilir bir performans alt sınırı sağlamaktır [82]–[84], [86]–[92], [97]–[100].

Önerilen yaklaşım, klasik belirsizliğe dayanıklı dinamik programlama ve Markov karar süreci modellerinden, garantiyi belirli bir gözlem modeline bağlamaması bakımından ayrılmaktadır. Kayıt kaybı, ayrıştırıcı kayması ve benzeri bozulmaların sık görüldüğü saha koşullarında gözleme bağımlı güvenceler kırılganlaşabilmektedir. Bu nedenle tezde garanti üretimi, tek bir durağan doğrusal program üzerinden kurulmuştur. Böylece tanımlanan belirsizlik sınıfı altında, gözlem hattı değişse bile geçerliliğini koruyan bir performans alt sınırı elde edilmektedir [89]–[92], [114].

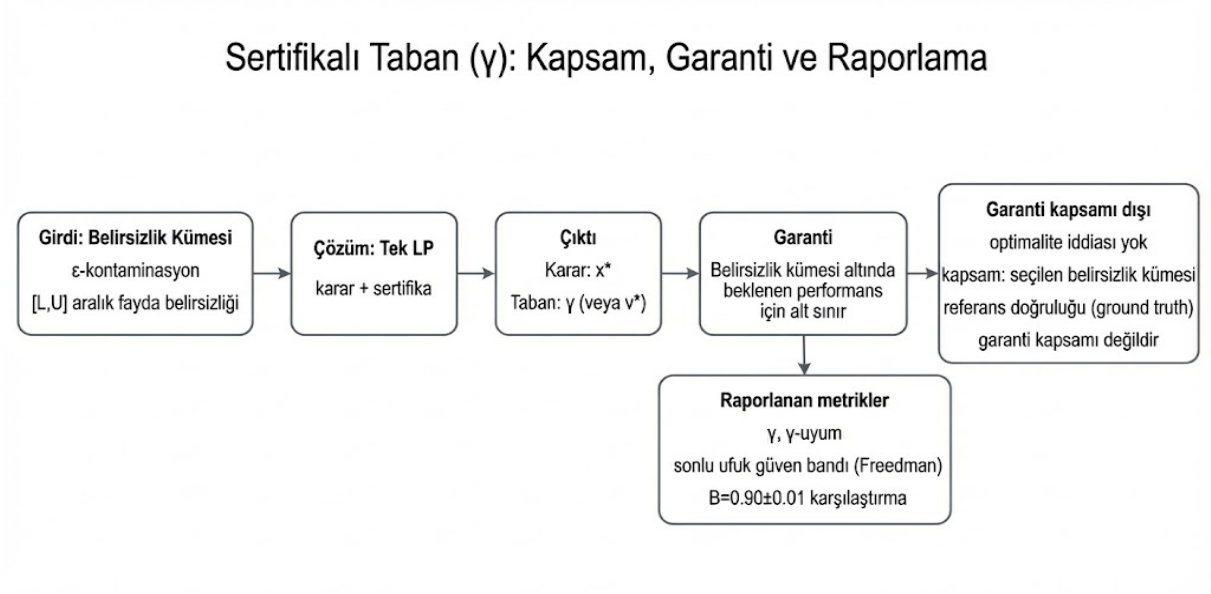
Bu nedenle elde edilen sonuç, tur bazında gözlemden bağımsız bir referans performans tabanı olarak yorumlanabilir. Belirsizlik modelinin ayrıntıları Bölüm 3.3'te, LP formülasyonu ise Bölüm 3.4.2'de verilmektedir.

Bu formülasyonun bir başka avantajı, LP'nin dual biçiminin performans alt sınırını zayıflatan belirsizlik bileşenlerini görünür kılmasıdır. Böylece ölçüm ve etiketleme çabasının hangi alanlara öncelikle yöneltilmesi gerektiği sayısal olarak değerlendirilebilmektedir [83], [108] (bkz. Bölüm 3.4).

Markov Karar Süreci (MDP) yaklaşımı, kararları zaman içinde durum ve eylem geçişleri üzerinden ele alır. Amacı, uzun vadeli performansı eniyileyen karar kuralları üretmektir [113], [115]. Belirsizliğe dayanıklı MDP yaklaşımında ise geçiş olasılıkları ve faydalar tam olarak bilinmez. Bu durumda olası modellerden oluşan bir küme tanımlanır ve en kötü duruma karşı güvenli kalan karar kuralı aranır [90]–[93]. Kısmi gözlemlenebilir ortamlarda (POMDP) durum doğrudan izlenemez. Bu nedenle karar, gözlemlerden türetilen inanç dağılımına ve gözlem modeline bağlıdır. Ancak log üretim veya toplama düzeni değiştiğinde ya da kayıt kaybı yaşandığında bu bağımlılık ciddi bir kırılganlık yaratır [114]. Tek LP ile performans alt sınırı yaklaşımı bu yüzden en iyi gözlem bağımlı karar kuralını aramaz. Bunun yerine, gözlem bozulsa bile çökmeyen ve tanımlanan belirsizlik sınıfı altında altına düşülmeyecek bir referans taban üretmeyi amaçlar. Gözlem kalitesi yüksek olduğunda daha iyi kararlar alınabileceği kabul edilir. Buna rağmen asıl hedef, gözlem değişse bile karar katmanının asgari güvenceyi korumasıdır.

Stratejik savunmada saldırgan uyarlanabilir bir rakiptir. Bu nedenle savunucunun kararı, tek taraflı bir eniyileme problemi olarak ele alınamaz. Karar, saldırganın tepkisini de içeren oyun teorik bir etkileşim içinde değerlendirilmelidir. Güvenlik oyunları literatürü, savunucunun sınırlı kaynaklarını hedeflere olasılıksal olarak dağıtmasını karma stratejilerle modeller ve saldırganın en iyi tepkisini de hesaba katar. Bu yaklaşım sahada ölçeklenebilir yöntemler sunar [95]–[97]. Getiriler en iyi ve en kötü durum aralıklarıyla verildiğinde ya da saldırgan davranışı belirsiz olduğunda, belirsizlik altında Stackelberg çözümleri geliştirilmiştir. Tek bir doğrusal programdan elde edilen performans alt sınırı değerinin ikili yapısı da bu çözümlerle doğal biçimde ilişkilendirilebilir [98]–[100]. Veri geldikçe modeli güncelleyen çevrimiçi öğrenme yaklaşımları da uyarlanabilir rakiplere karşı en iyi sabit eyleme göre fark garantisi sağlar. Ancak bu garanti, tezde hedeflenen her koşulda korunacak performans alt sınırı fikriyle aynı şey değildir [94]. Bu nedenle performans alt sınırı yalnızca asimptotik bir sonuç olarak bırakılmamıştır. Belirli bir karar turu sayısı için yüksek olasılıklı güven bantlarıyla birlikte raporlanmıştır. Bu bantların biçimsel tanımı Bölüm 3.5'te verilmiştir [100]–[102]. Güvenlik ortamında saldırganın kaçınma davranışı göstermesi ve örnekleme düzenini etkilemeye

çalışması hem tespit hem karar katmanında belirsizliğe dayanıklı değerlendirme ihtiyacını daha da artırır [103], [104].



Şekil 2.2: Sertifikanın kapsamı ve saha raporlaması.

Şekil 2.2, üretilen sertifikanın hangi koşullar altında geçerli olduğunu göstermektedir. LP, karar kuralı x* ile taban değer γ/v*'yi birlikte üretmektedir. Ancak bu taban yalnızca tanımlanan belirsizlik kümesi içinde geçerlidir. Saha sapmaları bu kümenin dışına çıktığında garanti kapsamı da aşılmış olur. Benzer şekilde referans etiketlerdeki hata, sertifikanın değil ölçümün sınırlılığıdır. Bu nedenle sahada karar kuralı, taban değer, sertifika uyumu ve sonlu tur güven bantları birlikte raporlanmalıdır.

2.8. Risk kademeli bütçeli operasyonel karar katmanı

UEBA (Kullanıcı ve Varlık Davranış Analizi) ve olay köken izi temelli tespit hatları sahada genellikle bir tespit skoru üretir. Operasyonel sorun bu skorun hangi koşulda hangi aksiyona dönüşeceğini net ve işlenebilir biçimde tanımlamaktır. SOC ortamında karar yalnızca "skor yükseldi" gerekçesine bırakılamaz. Analiz zamanı, sistem yükü ve hizmet kesintisi gibi maliyetler müdahale kapasitesini sınırlar. Bu yüzden karar bütçe altında verilmelidir [17]-[20], [23], [57], [65]. Bu bağlamda sık kullanılan sabit eşikle karar verme yaklaşımı zamanla sorun çıkarır. Gözlem ve görünürlük değiştikçe aynı skor farklı anlamlar taşıyabilir. Ortama özgü fayda ve zarar yapısı da aynı müdahalenin farklı ortamlarda farklı maliyetler doğurmasına yol

açar. Sonuç çoğu zaman aşırı uyarı yükü veya atlanan olaylar (yanlış negatifler) olur [1], [22], [23], [46].

Bu katman, tespit skorunu aksiyona iki aşamada dönüştürür. İlk aşamada, loglardan veya güvenlik telemetrisinden üretilen tespit skorları zararsız referans verisi üzerinde kalibre edilir ve konformal p-değerine çevrilir. İkinci aşamada bu p-değerleri risk kademelerine ayrılır. Böylece farklı tespit sistemlerinden gelen skorlar ortak bir risk dilinde ifade edilir. Sonuç olarak sabit eşiklerin yarattığı ölçek bağımlılığı azalır; tespit teknolojisi değişse bile karar katmanının doğrudan girdisi olan konformal p-değeri ve risk kademeleri tutarlılığını korur [8], [9].

İkinci aşamada aksiyon seçimi, bütçe kısıtı ile belirsizliği birlikte ele alır. Her risk kademesi için eylemsizlikten karantinaya uzanan müdahalelerin maliyet ve fayda yapıları aralık değerleriyle modellenir. Bu aralıklar, ortama ve zamana bağlı değişkenliği doğrudan temsil eder. Ardından toplam harcamayı sınırlayan bütçe altında, kademeler arasında uygulanacak aksiyon karışımı bir optimizasyon problemi olarak belirlenir [17]-[20], [23], [57], [65]. Böylece SOC'un kapasite sınırı kararın dışındaki bir pratik kısıt olmaktan çıkar ve karar mekanizmasının parçası hâline gelir.

Bu tasarım 2.7'deki sertifika yaklaşımıyla uyumludur. Amaç yalnızca ortalama koşullarda beklenen faydayı artırmak değil, tanımlanan belirsizlik sınıfı içinde korunacak asgari performans alt sınırını sayısal ve doğrulanabilir biçimde raporlamaktır [82]–[84], [86]–[92]. Bu nedenle karar kuralı, sadece "hangi aksiyon seçildi" sorusuna yanıt vermez; aynı zamanda belirsizlik altında altına düşülmeyeceği güvence altına alınan bir performans alt sınırı ifadesiyle birlikte sunulur [82]–[89].

Sahaya uyarlanan kısım ise kararın girişinin yapılandırılmasıdır. Tek bir genel karar yerine risk kademelerine koşullu karar üretilir ve karar bütçe kısıtıyla bağlanır. Dedektör skorları p-değerine dönüştürülerek kararın dayandığı kanıt daha taşınabilir bir forma getirilir [8], [9], [17]-[20], [23], [57], [65]. Böylece 2.7'deki gözlemden bağımsız performans alt sınırı fikri, skor akışı üzerinde uygulanabilir bir karar kuralı diline aktarılır. Bu dönüşüm sabit eşik kullanımının kırılganlığını azaltır ve kararları iş yükü ile maliyet disiplini içinde doğrulanabilir kılar [1], [22], [23], [46].

Değerlendirme iki ekseni birlikte taşır. Referans etiketlere (ground truth) dayalı ölçütler, zararsız etkinliğe (benign) ağır müdahale oranı gibi doğrudan operasyonel zarar riskini görünür

kılar. Sertifika temelli ölçütler ise belirsizlik altında kararın asgari performansını alt sınırlandıran sertifikayı öne çıkarır. Sertifikanın sahada ne ölçüde sağlandığını izlemek için sertifika uyum (adherence) ölçütü de bu grupta yer alır. Bu iki ölçüt ailesi, 2.7'de tanımlanan gözlemden bağımsız performans alt sınırı fikrinin saha koşullarında izlenebilir ve raporlanabilir olmasını destekler [82]–[84], [86]–[92].

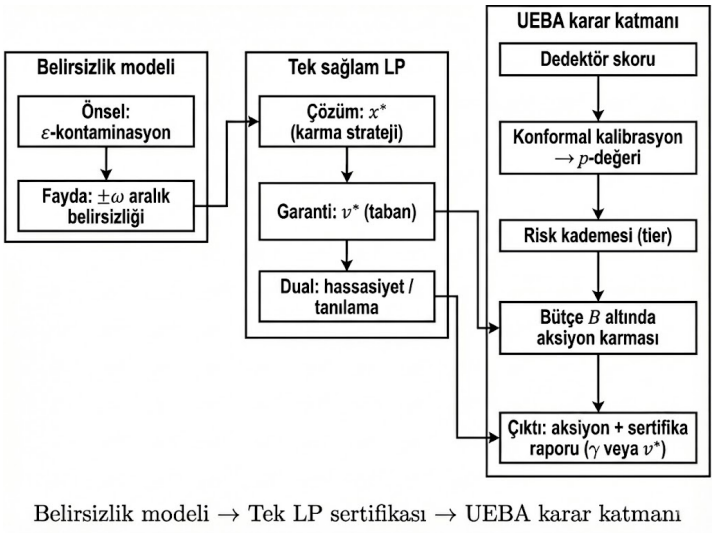
2.9. Literatür boşluklarının tez hedefiyle ilişkisi

Literatür taraması, sistem olay köken izi temelli APT tespiti çalışmalarında iki tekrarlayan ekseni birlikte görünür kılmıştır. Bir tarafta karşılaştırılabilirlik sorunu öne çıkmaktadır: çalışmalar düğüm, altgraf veya tüm graf düzeyinde farklı kanıt birimleri üretmesine rağmen, bu çıktıların SOC iş akışında hangi operasyonel karara karşılık geldiği çoğu zaman tutarlı biçimde açıklanmamakta; aşama dili, ölçüm paydası ve iş yükü/bütçe varsayımları açık yazılmadığında kapsam iddialarının pratik yorumu zayıflamaktadır [1]–[3], [23]. Diğer tarafta saha koşullarında gözlem hattı değişkendir; kayıt kaybı, log üretim farklılığı ve dağılım kayması gibi koşullar altında sonuçların taşınabilirliği çoğu çalışmada sistematik biçimde gösterilmemekte ve sabit eşiklere dayalı uygulamalar uzun soluklu savunmada kırılganlaşabilmektedir [1], [22], [23], [46]. Bu tablo, önce gözlem hattına bağlanmadan geçerli kalan bir güvence ihtiyacını doğurarak Tek LP ile sertifikalı gözlemden bağımsız performans alt sınırı yaklaşımını gerekçeleştirmiştir. Bununla birlikte, bu çekirdeğin sahada anlamlı olabilmesi için tespit çıktılarının SOC'un müdahale repertuvarına bütçe ve kapasite kısıtları altında bağlanması gerekir. Bu gereklilik tespit skoru akışını bütçe altında risk kademelerine göre aksiyona çeviren operasyonel karar katmanının temelini oluşturur. Tez literatürdeki telemetri değişkenliği ve karşılaştırılabilirlik sorununa iki aşamalı bir çözüm sunar. Çekirdek katman gözlemden bağımsız ve doğrulanabilir bir performans alt sınırı güvencesi sağlarken, saha katmanı skor akışını taşınabilir bir kanıt formuna dönüştürür. Bütçe disiplinini karar mekanizmasıyla bütünleştirerek aynı belirsizlik dili üzerinden raporlanabilir güvenceyle uyumlu operasyonel sonuçlar üretir [1]–[3], [23], [82]–[84], [86]–[92].

3. YÖNTEM

Bu bölüm, tezin ana iddiasını sahada uygulanabilir bir karar mekanizmasına dönüştüren yöntemi baştan sona tanımlamaktadır. Yöntem iki ana bölümden oluşur. İlk bölümde, gözlem modeli kurulmadan önsel belirsizlik ε-kontaminasyon ile, aksiyonların fayda ve zararlarına ilişkin belirsizlik ise aralık belirsizliği ve alt zarf yaklaşımı ile ifade edilir. Bu tanımlar altında tek bir belirsizliğe dayanıklı doğrusal program çözülür ve savunucunun her tur uygulayacağı durağan karma strateji x* ile bu stratejinin tanımlı belirsizlik kümesi içinde garanti ettiği taban değer v* birlikte elde edilir. İkinci bölümde, UEBA ve olay köken izi tabanlı sistemlerin ürettiği farklı ölçeklerdeki skorlar zararsız referans verisi üzerinde konformal kalibrasyonla p-değerine dönüştürülür. Ardından olaylar az sayıda risk kademesinde toplanır. Karar katmanı, küresel bütçe kısıtı altında her risk kademesi için uygulanacak aksiyon dağılımını doğrusal programlama ile seçer ve kararı performans alt sınırıyla birlikte raporlar.

Şekil 3.1, yöntemin uçtan uca akışını özetler. Çekirdek hatta, ε-kontaminasyon ve getirilerin ($\bar{R}$) etrafında $\pm\omega$ sapma bandı içinde değişebildiği varsayımı altında tek bir belirsizliğe dayanıklı LP çözülür. Bu çözüm (x^*,v^*) ile ikili (dual) tanılama sinyallerini verir. Ampirik hatta ise dedektör skorları konformal p-değerlerine dönüştürülür, risk kademelerine ayrılır ve küresel bütçe B altında her kademe için aksiyon dağılımı seçilir.



Şekil 3.1: Tek LP sertifikası ve bütçeli aksiyon.

3.1. Araştırma yaklaşımı ve genel akış

Yöntemin çekirdeği, ardışık savunma kararlarında belirsizlik varken savunucunun her tur için doğrulanabilir bir alt sınır performans garantisini (sertifika) nasıl üreteceği problemidir. Tezin temel iddiası, gözlem hattı bozulsa veya değişse bile gözlem mekanizmasını modellemek zorunda kalmadan böyle bir garanti üretilebileceğidir. Bu iddia, Tek LP ile Gözlemden Bağımsız Belirsizliğe Dayanıklı Ardışık Savunma (Observation-Invariant Robust Sequential Defense via a Single Linear Program) hattı altında gerçekleştirilir. Gizli-durum önselindeki sapmalar ε-kontaminasyon (ε-contamination) ile, aksiyon–saldırgan–durum hücrelerindeki fayda belirsizliği ise aralık (interval) biçiminde ifade edilir. Ardından savunucunun durağan karma stratejisi (stationary mixed strategy) ve bu stratejinin belirsizlik sınıfı altında garanti ettiği en büyük taban değer tek bir belirsizliğe dayanıklı doğrusal program çözümüyle hesaplanır [82]–[84]. Burada "durağan karma strateji", savunucunun aksiyonları her tur aynı olasılıklarla seçmesi, "sertifika", tanımlı belirsizlik kümesi içinde kalındığı sürece beklenen performansın altına düşmeyeceği tabanın sayısal ve doğrulanabilir ifadesidir.

Bu çekirdeğin sahada doğrudan kullanılabilmesi için, ek bir çıktı olarak olay köken izi tabanlı UEBA için gözlemden bağımsız karar katmanı geliştirilmiştir. Bu katman, konformal kanıta dayalı tek bir belirsizliğe dayanıklı doğrusal program çerçevesinde tanımlanmıştır. Bu hat, UEBA ve benzeri sistemlerin ürettiği farklı ölçeklerdeki skorları önce konformal kalibrasyon (conformal calibration) ile p-değerine (p-value) dönüştürür. p-değeri, bir olayın "normal referansa göre ne kadar sıra dışı" olduğunu skorun biriminden ve dağılımından bağımsız biçimde temsil eden taşınabilir bir kanıttır. Ardından p-değerleri az sayıda risk kademesinde (risk tiers) toplanır ve her seviye için uygulanacak aksiyon dağılımı, küresel bir maliyet bütçesi (global cost budget) altında doğrusal programlama ile belirlenir [109], [110]. Böylece "hangi aksiyonun ne sıklıkla uygulanacağı" sorusu, bütçe kısıtı korunurken belirsizlik altında raporlanabilir bir performans alt sınırı hedefiyle birlikte çözülür.

Bu bütünleşik yaklaşımın çıktıları sahada izlenebilir ve doğrulanabilir üç tür bilgi üretir. Savunucu için, her karar anında veya her risk kademesinde uygulanacak aksiyonların olasılıklarını veren bir aksiyon dağılımı (policy) elde edilir. Bu aksiyon dağılımının tanımlanan belirsizlik kümesi altında garanti ettiği alt sınır performans, sertifika (certificate / floor) olarak raporlanır. Buna ek olarak, LP'nin ikili (dual) değişkenlerinden türeyen hassasiyet (sensitivity) sinyalleri, garantiyi en fazla düşüren belirsizlik bileşenlerini işaretler. Bu sinyaller, veri toplama

ve modelleme çabasının getiri tablosundaki (aksiyon, durum) girişlerinden hangilerine yoğunlaştırılması gerektiğine karar vermeyi destekler.

3.2. Teorik çalışma: Problem tanımı, notasyon ve belirsizlik modeli

İlk çalışmada sistem, belirli bir log hattı veya ayrıntılı bir gözlem modeli üzerinden değil, karar üretimi için gerekli olan soyut bileşenler üzerinden temsil edilmektedir. Bu soyut model; savunucunun aksiyon kümesi A_d, saldırganın aksiyon kümesi A_a, doğrudan gözlenemeyen gizli durumlar kümesi Θ ve bu durumlar altında tanımlı fayda yapısından oluşur. Belirsizlik, gizli durum önselindeki ε-kontaminasyon ve fayda tablosundaki aralık belirsizliği ile modellenir. Gözlem mekanizması bu sertifikanın parçası değildir; bu nedenle performans alt sınırı belirli bir observation kernel'e bağlanmaz. Tek bir doğrusal program çözümü, bu soyut model üzerinde savunucunun durağan karma stratejisini x^*ve tur başına garanti edilen performans alt sınırını $v^{\backslash *}$birlikte üretir.

3.2.1. Aksiyonlar, gizli durumlar ve fayda (action–latent state–reward)

Savunucunun seçebileceği aksiyonlar kümesi A_d ile gösterilir. Bu küme pratikte "hiçbir şey yapma", "inceleme", "kısıtlama" ve "karantinaya alma" gibi seçenekleri içerebilir. Saldırganın seçebileceği hamleler ise A_a aksiyon kümesiyle ifade edilir. Ortamın gerçek koşulu, savunucunun doğrudan gözleyemediği ayrık bir gizli durum uzayı Θ (latent state space) ile temsil edilir; örneğin Θ "zararsız", "saldırı", "yüksek riskli iç tehdit" gibi durumları kapsayabilir.

Bu tezde iki ayrı ölçek kullanılmaktadır. Teorik Tek-LP bölümünde ödül $[0, 1]$ aralığında normalize edilmiştir. Bu bölümde LP'nin ürettiği v^*, seçilen stratejinin tanımlı belirsizlik kümesi altında garanti ettiği beklenen fayda alt sınırını gösterir. UEBA karar katmanında ise Tablo 3.2'de verilen değerler net fayda puanı ve maliyet puanı olarak tanımlanmıştır; bu değerler olasılık değildir ve negatif olabilir. Bu nedenle bu hatta faydanın alabileceği alt ve üst sınırlar sırasıyla Lve Uile gösterilir. Karar katmanında risk kademesi $b$için performans alt sınırı γ_b, tüm kademeler için karışım-ağırlıklı performans alt sınırı ise $\bar{\gamma}$ile gösterilir. Ortalama fayda gözlenen gerçekleşmeyi, performans alt sınırı ise garanti edilen taban değeri ifade eder. Her iki hatta belirsizlik, ε-kontaminasyon varsayımıyla modellenir. Bu varsayım, nominal önselin büyük ölçüde geçerli kaldığını, küçük bir εpayının ise sahada farklı bir dağılıma kayabileceğini kabul eder.

Notasyonun netliği için boyutlar da sabitlenir: $m := |A_d|$, $n := |A_a|$, $S := |\Theta|$. Θ üzerindeki olasılık simpleksi (probability simplex) $\Delta(\Theta) := \{p \in \mathbb{R}_+^S : \mathbf{1}^\top p = 1\}$ şeklindedir. Vektörler, aksi belirtilmedikçe savunucu aksiyonları i üzerinden indekslenir; yani bir "fayda satırı" m-boyutlu bir vektör olarak düşünülebilir.

3.2.2. Gözlem modeli ve geçmişe dayalı karar kuralı varsayımı

Sahada gözlemler çoğu zaman loglardan ve loglardan türetilen kanıtlardan oluşur; dropped logs, parser drift ve label swaps gibi bozulmalar bu hattı etkileyebilir. Bu nedenle bu tez, gözlemi üreten mekanizmayı (observation kernel) açıkça P ile modellemek zorunda değildir. Teknik olarak P, (a_t, θ_t) ve rastgele getiri gerçekleşmelerinden gözlemlere giden bir üretim süreci olarak tanımlanabilir. Ancak sertifikasyon probleminde P hiçbir adımda kullanılmaz. Bu noktada "gözlemden bağımsızlık (observation invariance)", kararların gözlem kullanmaması değil, garantinin log gözlem mekanizmasına bağlanmaması anlamına gelir.

Savunucu karar kuralının geleceği bilmeyen / yalnızca geçmiş veriye dayalı (non-anticipating) olması, kararların yalnızca geçmiş ve mevcut gözlemlerden türetilen bilgiye dayanması demektir. Savunucu gelecekte gerçekleşecek olayları bilerek karar vermez. Bu varsayım SOC gerçekliğiyle uyumludur. Sertifikasyonun kritik noktası şudur: sertifika, gözlem mekanizmasının ayrıntılı modelini (observation kernel) gerektirmeyecek şekilde üretilir; ancak geçerliliği, sahadaki sapmaların tanımlanan belirsizlik kümesi içinde kalması koşuluyla taşınabilir kalır [82]–[84].

3.3. Belirsizlik modeli

3.3.1. Önsel belirsizliği: Huber ε-kontaminasyon

Gözlem yapılmadan önce, sistemin hangi gizli durumda olabileceğine ilişkin nominal bir önsel dağılım $\pi_0 \in \Delta(\Theta)$ varsayılır. Ancak bu önselin hatalı olabileceği kabul edilir ve belirsizlik, Huber'ın ε-kontaminasyon (ε-contamination) modeliyle ifade edilir [82]:

$$\Pi(\varepsilon) = \{(1-\varepsilon)\pi_0 + \varepsilon q : q \in \Delta(\Theta)\}, \qquad \varepsilon \in [0,1] \tag{3.1}$$

Sezgisel anlamı: ε, nominal önsel π_0'ın sahada ne ölçüde "yanlış çıkabileceğine" izin verdiğimizi belirleyen bir tolerans payıdır. ε = 0.05 demek, önsel kütlenin 1−ε = 0.95'lik kısmının π_0'a "sabitlenmiş" kaldığını; kalan en fazla %5'lik kısmın ise, sahada beklenmeyen sınıf oranları, yanlış kalibrasyon veya ortam kayması gibi etkileri temsil edecek biçimde Θ

üzerinde serbestçe yeniden dağıtılabileceğini kabul etmek demektir. Bu yeniden dağılım "en kötü durum" bakışıyla ele alındığından, serbest bırakılan ε payı garantiyi aşağı çekecek şekilde konumlanabilir. ε küçük seçildiğinde sertifika daha iyimser görünür (taban daha yüksek çıkar). Fakat gerçek sistemdeki önsel sapma bu toleransı aşıyorsa, garanti tanımlanan belirsizlik sınıfının dışında kalabilir. ε büyüdükçe sertifika daha temkinli hâle gelir (taban düşer) ancak daha geniş önsel kaymalarını kapsar. Bu nedenle ε, serbest bir hiperparametre değil; önsel kaymasına karşı kabul edilen sapma toleransını kodlayan bir risk ayarı olarak okunmalıdır.

3.3.2. Nominal getiri tablosu etrafında belirsizlik aralığı ve güvenli taban

Getiri (ödül/fayda) her aksiyon–saldırgan hamlesi–gizli durum hücresinde tam olarak bilinemeyebilir. Bu belirsizlik, nominal getiri tablosu $\bar{R}(i,a,\theta)$ etrafında, hücre-bazlı yarı genişlik $\omega(i,a,\theta) \geq 0$ ile tanımlanan aralık (kutu/box) belirsizliği olarak modellenir:

$$R(i,a,\theta) \in [\, \bar{R}(i,a,\theta) - \omega(i,a,\theta), \bar{R}(i,a,\theta) + \omega(i,a,\theta) \,] \tag{3.2}$$

Belirsizliğe dayanıklı yaklaşımda gereksiz iyimserlikten kaçınmak için alt zarf (lower envelope) kullanılır. Alt zarf, (3.2) aralığındaki olası değerler içinden en düşük olanı esas alır; yani getiri tam bilinmiyorken "olabilecek en iyi" değeri değil, garantiyi koruyan en kötü değeri temel alır.

$$c_{a,\theta}(i) := \bar{R}(i,a,\theta) - \omega(i,a,\theta) \tag{3.3}$$

Çekirdek (Single-LP) hatta ödüller [0,1] ölçeğinde normalize edildiği için, aralık tanımı bu ölçekle tutarlı seçilir. Pratikte bu, ω değerlerinin $\bar{R} - \omega \geq 0$ koşulunu sağlayacak şekilde belirlenmesi veya eşdeğer biçimde alt zarfın $c := \max\{0, \bar{R} - \omega\}$ olarak alınmasıyla sağlanır. Böylece hem LP formülasyonu hem de sonlu-ufuk güven bantları için kullanılan sınırlılık varsayımı aynı ölçek üzerinde korunur.

Burada $c_{a,\theta}$, saldırgan aksiyonu a ve gizli durum θ altında, savunucunun karma stratejisiyle garanti edilebilecek en düşük hücre-bazlı katkıyı temsil eden alt sınır satırıdır. Aralıklar büyüdükçe (yani belirsizlik arttıkça) bu alt zarf küçülür; sertifikanın da doğal olarak düşmesi beklenir [83], [108].

3.3.3. Garanti fonksiyonunun nominal bileşeni

Huber ε-kontaminasyon modelinde toplam olasılık kütlesinin $1-\varepsilon$kadarlık bölümü nominal önsel $\pi_0$üzerinde kalır. Garanti fonksiyonunun ilk bileşeni bu nominal bölümden gelir.

Her saldırgan aksiyonu $a$için, alt-zarf getirilerinin $\pi_0$altındaki beklenen değeri alınarak nominal katkı tanımlanır. Bu katkı, belirsizlik yokken değil, getiri belirsizliği alt-zarf ile güvenli hâle getirildikten sonra hesaplanan güvenli beklenen değerdir. Dolayısıyla burada tanımlanan terim, garanti fonksiyonunun nominal parçasını temsil eder; kontaminasyonla serbest bırakılan εpayının en kötü durum etkisi ise bir sonraki adımda ayrıca eklenir.

$$E_a := (1-\varepsilon)\sum_{\theta\in\Theta}\pi_0(\theta)\,c_{a,\theta} \in \mathbb{R}^m \tag{3.4}$$

Burada $c_{a,\theta}$, saldırgan aksiyonu a ve gizli durum θ altında savunucu aksiyonlarına ait alt-zarf getiri vektörüdür. E_a ise bu alt-zarfın nominal önsel altındaki beklenen katkısını gösterir. Başka bir ifadeyle E_a, garanti fonksiyonunun yalnızca nominal kısmıdır. Nihai garanti fonksiyonu, bu nominal katkıya ek olarak kontaminasyon payının yöneldiği en kötü gizli durum katkısını da içerir.

3.4. Belirsizliğe dayanıklı (robust) LP ile performans alt sınırı (sertifika) üretimi

Bu tezin temel amacı, tek bir belirsizliğe dayanıklı doğrusal program (robust linear program) ile hem bir aksiyon dağılımı hem de bu aksiyon dağılımının performans alt sınırını (sertifikasını) üretmektir.

3.4.1. Karar değişkenleri ve anlamları

Savunucunun durağan karma stratejisi $x \in \Delta(A_d)$ olsun. x_i, savunucunun aksiyon i'yi uzun vadede hangi olasılıkla uyguladığını gösterir. Performans alt sınırı (sertifika) $v \in \mathbb{R}$, tur başına garanti edilmek istenen performans alt sınırıdır (floor / certificate). Her saldırgan aksiyonu için m_a yardımcı değişkeni tanımlanır; m_a, kontaminasyon kütlesinin yığılabileceği "en kötü gizli durum" etkisini doğrusal tutmak için kullanılır.

3.4.2. Performans alt sınır LP'si (Primal)

Aşağıdaki LP, v'yi mümkün olduğunca büyütürken, bu değerin tüm saldırgan aksiyonları ve tüm gizli durumlar altında geçerli bir alt sınır olmasını zorunlu kılar.

$$\begin{aligned} &\max_{x,v,\{m_a\}} v \\ &\text{koşuluyla} \quad v - \langle E_a, x\rangle - \varepsilon m_a \le 0, && \forall a \in A_a, \\ &\qquad\qquad m_a - \langle c_{a,\theta}, x\rangle \le 0, && \forall a \in A_a,\ \forall \theta \in \Theta, \\ &\qquad\qquad 1^\top x = 1, && x \ge 0, \end{aligned} \tag{3.5}$$

Bu LP'nin amacı, belirsizlik altında savunucunun hangi aksiyonları hangi olasılıklarla seçeceğini ve bu kararın garanti ettiği en düşük performans düzeyini birlikte belirlemektir. Bu formülasyonda garanti fonksiyonu iki bileşenden oluşur. İlki, Bölüm 3.3.3'te tanımlanan nominal katkıdır. İkincisi ise Huber ε-kontaminasyon altında dikkate alınan en kötü gizli durum katkısıdır. Bu nedenle sabit karma strateji $x$ve saldırgan aksiyonu $a$için garanti değeri, bu iki katkının toplamı olarak yazılır. LP'de $m_a$yardımcı değişkeni, en kötü gizli durum terimini doğrusal biçimde ifade etmek için kullanılır. İlk kısıt satırı da $v$değişkenini bütün saldırgan aksiyonları için geçerli ortak alt sınır olarak tanımlar. Bu yüzden v, bir sınıflandırma doğruluğu ya da ortalama başarı ölçütü değildir; seçilen aksiyon dağılımı için belirsizlik altında garanti edilen en düşük performans düzeyidir. Sonuç olarak LP, savunucunun aksiyonları hangi olasılıklarla seçeceğini belirlerken, bu kararın garanti ettiği alt sınırı da birlikte hesaplar. Çözüm sonunda x^*optimal aksiyon dağılımını, v^*ise bu dağılım için elde edilen nihai performans alt sınırını verir.

LP içinde çalışan garanti fonksiyonu φ_a(x;ε) dir. (3.6)

$$\varphi_a(x;\varepsilon) = (1-\varepsilon)\sum_{\theta\in\Theta}\pi_0(\theta)\langle c_{a,\theta}, x\rangle + \varepsilon \min_{\theta\in\Theta}\langle c_{a,\theta}, x\rangle$$

Bu ifade, Huber ε-kontaminasyon belirsizlik kümesinden doğrudan elde edilir. Çünkü Π(ε) = {(1−ε)π_0 + εq} yapısında toplam olasılık kütlesinin $1-\varepsilon$ bölümü nominal önsel π_0'da kalırken, ε bölümü en kötü gizli duruma yönelebilir. Bu nedenle garanti fonksiyonu, nominal önsel katkısı ile en kötü durum katkısının toplamı biçimine iner. Getiriler [R−ω, R+ω] aralığında modellendiği için güvenli alt zarf olarak $c = R - \omega$ alınır. LP'nin yaptığı iş, bu garanti fonksiyonunun en küçük değerini maksimize etmektir. Burada gözlem çekirdeği Pkullanılmadığından, elde edilen v^* değeri, aynı belirsizlik modeli geçerli kaldığı sürece gözlem hattındaki değişimlerden bağımsız taşınabilir bir performans alt sınırı sağlar [82]–[84]. Bu noktada formülasyonun nominal modelin genellemesi olduğu açıkça görülür. ε = 0 olduğunda önsel belirsizliği ortadan kalkar. ω = 0 olduğunda getiri aralıkları kapanır ve alt zarf c = R olur. Bu durumda (3.5) numaralı robust LP, nominal LP'nin özel durumuna indirgenir. Dolayısıyla burada sadeleştirilen kısım belirsizliğin kendisi değil, gözlem çekirdeği P'nin açıkça modellenmemesidir.

3.4.2.1. İkili çözümün (Dual) yorumu ve duyarlılık analizi

LP'nin ikili tarafı, belirsizliğin "nasıl zarar verdiğini" açıklayan bir teşhis (diagnostic) aracı sağlar. İkili değişkenler, saldırganın aksiyonlar üzerinde bir karışım (mix) seçtiği ve her aksiyon için kontaminasyon gizli durumlar arasında paylaştırdığı bir yapı gibi okunur. Bu okuma, dağılımsal belirsizliğe dayanıklı Stackelberg (distributionally robust Stackelberg) yorumunu verir [94], [97], [98]. Pratik açıdan önemli sonuç şudur: ikili çözümden elde edilen sinyaller, "hangi saldırgan aksiyonları ve hangi durum/hücre belirsizlikleri tabanı en çok düşürüyor?" sorusuna sayısal yanıt üretir.

Doğrusal program SciPy üzerinden çözülmüş ve çözümleyici olarak HiGHS çözücüsü (HiGHS solver) kullanılmıştır. HiGHS, büyük ölçekli doğrusal programları hızlı ve sayısal olarak kararlı biçimde çözmeye odaklanan bir optimizasyon paketidir. Çözücü, asal çözümün yanında kısıtların ikili çarpanlarını (dual multipliers) da üretmektedir. SciPy, bu bilgiyi çıktı nesnesinde eşitsizlik kısıtlarının marjinal değerleri (marginals) olarak raporlamaktadır.

Bu çalışmada ikili tanılama sinyalleri, modelde yer alan iki kısıt ailesi üzerinden elde edilmektedir. Birincisi, her saldırgan hamlesi a için yazılan sertifika kısıtlarıdır. Bu kısıtlara karşılık gelen ikili çarpanlar, taban değeri en çok hangi a hamlesinin baskıladığını göstermektedir. Bu çarpanlar toplamı 1 olacak biçimde normalize edildiğinde λ elde edilmektedir. λ, saldırgan hamleleri üzerinde oluşan bir ağırlık dağılımı olarak yorumlanmaktadır.

İkincisi, her a altında en olumsuz gizli durumu θ seçtiren kısıtlardır. Bu etkiyi doğrusal biçimde temsil etmek amacıyla her a için yardımcı değişken m_a tanımlanmış ve m_a'nın tüm θ seçenekleri için ilgili değerin altında kalması zorunlu kılınmıştır. Bu nedenle bu kısıtlar min-θ kısıtları olarak adlandırılmıştır. Bu kısıtlara karşılık gelen ikili çarpanlar, ε ile serbest bırakılan kontaminasyon payının hangi θ hücrelerine yığıldığında tabanın daha fazla düştüğünü işaret etmektedir. Bu sinyal $q(\theta \mid a)$ ile gösterilmektedir. Sonuç olarak ikili çözüm, tabanı en çok düşüren a hamlelerini ve bu hamleler altında kritikleşen (a, θ) durum çiftlerini sayısal olarak belirginleştirmektedir.

Bu tezde ikili çözüm temelli duyarlılık bilgisi iki yerde kullanılır. Birincisi, kontaminasyon seviyesi ε arttığında sertifikanın nasıl düşeceğini yerel olarak öngörmek için kullanılır. Sertifika $v^*(\varepsilon)$ beklenen şekilde azalan bir fonksiyondur (kontaminasyon arttıkça garanti zorlaşır). İkincisi, aralık yarı genişlikleri ω(i, a, θ) büyüdüğünde sertifikanın hangi

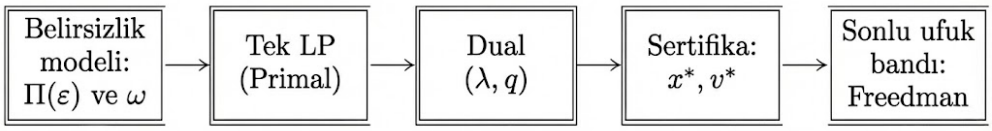
hücreler üzerinden daha hızlı düştüğünü görmek için kullanılır. Böylece daha fazla veri toplayarak aralıkları daraltma imkânı varsa, ölçüm çabasının "garantiye en çok etki eden hücrelere" yöneltilmesi mümkün olur [83], [108]. Bu, taban değer performansı için harcanan eforun rastgele değil, hedefe dönük olmasını sağlar.

3.4.3. Sonlu ufukta yüksek olasılıklı performans bandı

LP'nin ürettiği x^*aksiyon dağılımı olarak uygulandığında, her tur elde edilen normalize edilmiş tur faydası $r_t \in [0,1]$ ve K tur sonundaki ortalama

$$\bar{r}_K = \frac{1}{K}\sum_{t=1}^{K} r_t \tag{3.7}$$

şeklindedir. Şekil 3.2'de aşamaları gösterilen bu süreçte, sertifika v^* esasen "beklenen" (expectation) bir taban değerdir. Bu nedenle tez, K sonlu iken $\bar{r}_K$'nin v^*etrafında nasıl yoğunlaştığını yüksek olasılıklı bir bantla (high-probability band) raporlanabilir hâle getirir. Freedman eşitsizliği (Freedman inequality), varyansa duyarlı (variance-adaptive) bir alt bant sağlar ve düşük varyanslı rejimlerde gereksiz temkinliliği azaltır [100]. Pratikte bu bant, δ(bandın ihlal edilmesine izin verilen hata olasılığı) seçildiğinde, 'K turda en az $1-\delta$olasılıkla şu performansın altına düşmemeliyim' türü operasyonel olarak anlaşılır bir güvenceye çevrilebilir [100].



Şekil 3.2: Tek LP ile sertifika üretim hattı, ikili çözüm ve Freedman güven bandı.

3.5. eriyle fayda belirsizliği aralığı ve performans alt sınır bandını daraltma: Empirik-Bernstein

Aralık belirsizliği $\omega(i,a,\theta)$ her zaman el ile seçilmek zorunda değildir. Fayda katsayıları veriden tahmin edilebiliyorsa, fayda tablosundaki her (i, a, θ) hücresi için tek taraflı ampirik Bernstein güven bantları kullanılarak ωveriyle kurulabilir. Temel fikir, aynı hücreye ilişkin daha fazla örnek toplandıkça belirsizliğin azalmasıdır. Buna bağlı olarak ωküçülür ve LP'nin ürettiği performans alt sınırı sıkılaşır. Hücre sayısı çok olduğunda, tüm hücreler için aynı anda geçerli bir güvence vermek gerekir. Bu amaçla birlik bağı (union bound) ile hata

olasılığı hücreler arasında paylaştırılır. Böylece belirsizlik kümesi veriyle daraltılabilir ve sertifika adım adım iyileştirilebilir [101], [102]. Sonuç olarak performans alt sınırı, sabit bir tablo olmaktan çıkar ve veri biriktikçe sıkılaşan bir güvenceye dönüşür. Tezde ayrıca ikili duyarlılık sinyalleri kullanılarak ölçüm çabasının en kritik hücrelere yönlendirilebileceği gösterilmiştir. Bu yaklaşım, sınırlı ölçüm bütçesi altında sertifikayı en çok iyileştirmesi beklenen (i, a, θ) hücrelerini seçip aralıkları hedefe dönük olarak daraltmaya dayanır.

3.6. Ampirik çalışma: Olay köken izi tabanlı UEBA karar katmanı

Ampirik saha çalışmasındaki UEBA karar katmanı, Bölüm 3.4'te verilen genel tek-LP formülasyonunun iki durumlu ve saldırgan aksiyonu boyutunun fayda aralıkları içine eklendiği özel bir halidir. Teorik modelde ayrı bir boyut olarak yazılan saldırgan aksiyonu kümesi A_a, pratik hatta açıkça modellenmemiş; saldırgan hamlelerinin fayda üzerindeki etkisi $[L_{a,\theta}, U_{a,\theta}]$ aralıklarına gömülmüştür. Bu nedenle UEBA karar katmanındaki LP, genel formülasyonla tutarlı fakat daha düşük boyutlu bir karar katmanı olarak okunmalıdır.

UEBA ve benzeri dedektörler sayısal skorlar üretir; ancak SOC ortamında asıl ihtiyaç, bu skorları bütçe kısıtı altında uygulanabilir aksiyonlara dönüştüren açık bir karar kuralıdır. Bu bölümde, dedektör skorlarının konformal p-değerlerine dönüştürülmesi, risk kademelerine ayrılması ve bu kademeler için aksiyon dağılımlarının tek bir doğrusal programla belirlenmesi açıklanmaktadır. İş akışı Tablo 3.1'de özetlenmiştir.

Bu çalışmada olay köken izi verisi, dedektör skorlarının üretildiği temel veri kaynağıdır. Karar katmanı doğrudan ham graf topolojisini değil, bu veriden türetilen p-değerlerini ve risk kademelerini kullanmaktadır. Böylece dedektörün iç yapısından bağımsız bir karar katmanı kurulmaktadır. Köken izi graflarının topolojik ve nedensel özelliklerinin belirsizlik kümelerine doğrudan yansıtılması ise gelecek çalışmanın konusu olarak bırakılmıştır.

3.6.1. Deney tasarımı ve yapılandırma

Karar katmanının ana deneysel değerlendirmesi DARPA OpTC NCR eCAR veri seti üzerinde gerçekleştirilmiştir [112]. OpTC, bir SOC ağındaki süreç, dosya ve ağ nesneleri arasındaki etkileşimleri eCAR formatında zaman damgalı olay zincirleri olarak kaydeder ve gerçekçi bir APT değerlendirme ortamı sunar. Tezdeki ana bulgular bu veri seti üzerinden raporlanmıştır. Sabit eşik taşınabilirliğini inceleyen ek deney ise IoT-23 ve ToN_IoT veri kümeleri üzerinde ayrıca yürütülmüştür. Veri kümesi, dosya kimliği temelinde birbirini

dışlayan iki değerlendirme dilimine ayrılmıştır: OpTC-A ve OpTC-B. Bunlar ayrı veri setleri değil, aynı veri setinin iki ayrık dilimidir. Hiçbir dosya iki dilimde birden yer almamaktadır.

Bu ayrım bir eğitim/test bölünmesi değildir. Amaç, önerilen yöntemin aynı yakalama altyapısından elde edilen iki bağımsız dilimde tutarlı sonuç verip vermediğini sınamaktır. Her dilimde ve her seed için en çok 80.000 satır taranmıştır. Bu satırlardan 10.000 satırlık bir alt küme seçilmiştir. Aynı seed kullanıldığında aynı satır kümesi yeniden elde edilmiştir. Bu sayede deneyler aynı örneklem üzerinde tekrarlanabilmiştir. Seçilen satırlar nesne–eylem çiftlerine ayrıştırılmıştır. Biçimsel olarak bozuk satırlar elenmiştir. Bir sisteme ya da varlığa ilişkin olay kaydı üretilmediğinde ya da toplanamadığında, ilgili varlık gözlenen köken grafında yer almaz. Bu durum eksik gözlem olarak değerlendirilmiştir. Konformal kalibrasyonda yalnızca zararsız satırlar kullanılmıştır. Değerlendirme aşamasında ise zararsız ve kötü niyetli olayların birlikte bulunduğu satırlar kullanılmıştır. Kalibrasyon ve değerlendirme için kullanılan satırlar bilerek ayrık tutulmuştur. Böylece kalibrasyon verisinin değerlendirme sonuçlarına sızması önlenmiştir.

Bu ön işleme sürecinin amacı, platforma özgü ayrıntıları öğrenmek değil, her olayı ortak bir kanıt temsiline dönüştürmektir. Böylece karar katmanı, olayları ham skor ölçeğine göre değil, normal davranışa göre ne kadar sıra dışı olduklarına göre değerlendirmektedir.

Tablo 3.1'deki "Aksiyonlar ve maliyetler" satırı maliyet vektörünü ve aralık genişliğini vermektedir; ancak belirsizliğe dayanıklı LP'nin ve risk-nötr referansın girdi olarak kullandığı nominal fayda matrisi ayrıca gösterilmelidir. Tablo 3.2, deneylerde kullanılan iki durumlu (zararsız/saldırı) nominal fayda ve maliyet değerlerini sunmaktadır.

Tablo 3.1: Teorik hat ile pratik hat karşılaştırması ve deneysel yapılandırma

Parametre	Teorik (Single LP Sertifika)	Pratik (UEBA / DARPA OpTC)
Veri kaynağı	D1–D4 sentetik / yarı-sentetik kurulumlar	DARPA OpTC NCR, sistem eCAR logları
Önsel belirsizlik (ε)	0,00–0,50 aralığı tarama	0,02 (küçük-kontaminasyon çalışma noktası)
Getiri belirsizliği yarı-genişliği (ω)	ω ∈ {0,00; 0,05; 0,10}	ω ∈ {0.05, 0.10, 0.15, 0.20}
Bütçe kısıtı (B)	Kısıtsız ya da normalize; odak v* tabanı	0,90 (±0,01), beklenen maliyet üst sınırıyla eşleme
Girdi temsili	Doğrudan olasılık + getiri matrisleri	Konformal p-değeri + risk kademeleri
Temel çıktı	v* (tur başına garanti performans alt sınırı)	γ (sertifika düzeyi) + x* (aksiyon dağılımı)
Değerlendirme odağı	Parametre duyarlılığı; gözlem bozulmasına dayanıklılık; Freedman/Azuma bantları	γ-uyum; ortalama fayda; zararsız olayda ağır müdahale oranı
Çözücü	scipy.optimize.linprog (HiGHS)	scipy.optimize.linprog (HiGHS)
Risk eşikleri (Q)	—	0,01 / 0,05 / 0,20 / 0,50 (5 risk kademesi)
Konformal anlamlılık (α)	—	0,05
Düzleştirme sabiti (η)	—	0,20
Dosya örnekleme	—	Her OpTC dilimi ve her seed için 3 zararsız + 3 değerlendirme dosyası
Tarama / seçim limiti	—	80 000 / 10 000 satır
Seed değerleri	—	42, 17, 101 vb.

Bu tezde belirsizlik iki ayrı parametre ile gösterilmiştir. ε, nominal durum önselinin Huber ε-kontaminasyon altında ne kadar sapabileceğini gösterir. ω ise nominal fayda/getiri değerinin etrafındaki aralık belirsizliğinin yarı-genişliğini gösterir. Teorik ve pratik hat aynı iki parametreyi kullanır. Teorik bölümde bu parametreler duyarlılık analizi için taranır; pratik bölümde ise karar katmanının çalışma noktası olarak seçilir. Bu nedenle sayısal değerler farklı olabilir; ancak anlamları aynıdır. Bu tezde fayda/getiri belirsizliği her yerde ω ile, güven bantlarındaki hata olasılığı ise yalnızca δ ile gösterilmiştir.

Pratik hatta kullanılan çalışma parametreleri rastgele seçilmemiştir. Bu parametrelerin tümü belirsizlik parametresi değildir. Q, α ve s, konformal kalibrasyon ve risk kademeleme adımında kullanılan parametrelerdir. ε, önsel belirsizliğini; ω ise nominal fayda tablosu etrafındaki aralık belirsizliğinin yarı-genişliğini gösterir. B ise beklenen maliyet bütçesidir.

Seçim dört temele dayanmaktadır: ortak ve dedektörden bağımsız bir risk arayüzü kurmak, kademe sayısını duyarlılık analiziyle belirlemek, veri işleme yükünü sınırlayarak deneyleri tekrar üretilebilir tutmak ve tüm yöntemleri aynı maliyet düzeyinde karşılaştırmak. Bu nedenle ham dedektör skoru yerine yalnızca zararsız kalibrasyon verisinden elde edilen konformal p-değeri kullanılmıştır. P-değerleri, ortama özgü elle ayarlı alarm eşiklerine bağlı kalmamak ve karar optimizasyonunu kararlı ve yorumlanabilir tutmak için sabit Q={0,01; 0,05; 0,20; 0,50} eşikleriyle risk kademelerine ayrılmıştır. Varsayılan kademe sayısı beştir; çünkü 3, 5 ve 8 kademe üzerinde yapılan duyarlılık incelemesinde kazanımın büyük bölümü beş kademede elde edilmiştir. α=0,05, kademe önsellerinin kurulmasında kullanılan referans eşiğidir. η=0,20, kademe içi oranı genel şüphelilik oranına doğru küçülterek küçük örneklem kaynaklı oynaklığı sınırlamak için kullanılmıştır. ε=0,02, küçük kontaminasyon düzeylerinde dayanıklılık kazancının büyük bölümünün sınırlı fayda kaybıyla elde edildiğini gösteren duyarlılık bulgusuna dayanarak seçilmiştir. ω=0,15 ise $\omega \in$ {0,05; 0,10; 0,15; 0,20} aralığındaki taramada yöntemin davranışını temsil eden varsayılan çalışma noktası olarak alınmıştır. B=0,90 ise en iyi bütçe olduğu için değil, bütün yöntemleri aynı beklenen maliyet düzeyinde karşılaştırmak için eşlenmiş harcama hedefi olarak sabitlenmiştir. Veri tarafında her dilim ve her seed için 3 zararsız ve 3 değerlendirme dosyası seçilmiş, dosya başına en çok 80.000 satır taranmış ve 10.000 satırlık alt örnek kullanılmıştır. Bu sınırlar, G/Ç yükünü kontrol etmek, deneyleri aynı örnekleme kuralıyla tekrar üretilebilir tutmak ve kalibrasyon ile değerlendirme akışlarını ayrık korumak için belirlenmiştir.

Tablo 3.2: UEBA karar katmanında kullanılan nominal net fayda puanları ve operasyonel maliyet puanları

Aksiyon (a)	Zararsız (s = benign)	Saldırı (s = attack)	Maliyet puanı k_a
Hiçbir şey yapma (NoOp)	0,00	−10,00	0,0
İnceleme (Investigate)	−0,05	0,80	0,8
Kısıtlama (Throttle)	−2,00	5,78	3,0
Karantina (Quarantine)	−7,00	11,55	5,0

Tablo 3.2'deki nominal değerler saha ölçümlerinden türetilmemiş, göreli operasyonel tercihleri temsil eden nominal net fayda puanlarıdır. Bu nedenle bu tablo üzerinden raporlanan sonuçlar mutlak operasyonel kazanç tahmini olarak değil, sabit bir tercih yapısı altında yöntemler arası göreli karşılaştırma olarak yorumlanmalıdır. Bu nedenle maliyet sütunu, aksiyonların göreli operasyonel yükünü gösteren operasyonel maliyet puanları; fayda sütunları

ise net operasyonel kazanç veya kaybı gösteren net operasyonel fayda puanları olarak yorumlanmalıdır. Bu değerler gerçek para, süre veya olasılık değildir. Tabloda bazı hücrelerin negatif olması beklenen bir durumdur; bu değerler zararsız bir olaya gereksiz sert müdahale uygulanması ya da saldırıya karşı yetersiz aksiyon alınması gibi durumlarda oluşan operasyonel kaybı göstermektedir. Farklı durum uzayları ve ölçeklerde benzer yapısal sonuçların korunması, yöntemin seçilen katsayılara aşırı duyarlı olmadığını göstermektedir. Bununla birlikte, ölçüme dayalı fayda tabloları gelecekteki çalışmalar için önemini korumaktadır.

3.6.2. Skor standardizasyonu: konformal p-değeri ve gözlemden bağımsızlık

Farklı tespit sistemleri farklı ölçeklerde skor üretebildiği için tek bir sabit eşik kullanımı kırılgan olabilir. Bu nedenle bu tezde yalnızca zararsız örneklere dayalı indüktif konformal tahmin yaklaşımı benimsenmiştir. Bu çalışmada veri kaynağı eCAR loglarıdır. Bu loglar önce nesne–eylem çiftlerine ayrıştırılmış, ardından bu çiftlerden türetilen uyumsuzluk skorları konformal p-değerine dönüştürülmüştür. Karar katmanının doğrudan girdisi ham log değil, bu loglardan türetilen konformal p-değeri ve risk kademesidir.

Her olay, nesne–eylem çifti olarak $e = (o, op)$ biçiminde temsil edilir. Zararsız kalibrasyon kümesinde bu çiftin kaç kez gözlendiği $freq(o, op)$ ile gösterilir. $freq(o, op)$ büyüdükçe bu çiftin normal akışta sık görüldüğü, küçüldükçe ise nadir görüldüğü anlaşılır. Karar katmanını izole değerlendirmek amacıyla minimal bir uyumsuzluk skoru kullanılmıştır.

$$s(e) = -\log\left(freq(o, op) + 1\right) \tag{3.8}$$

Bu skor, yeni bir UEBA veya provenance tespiti önerisi olarak değil; karar katmanının farklı ölçekli skorları konformal p-değeri arayüzüne nasıl bağladığını göstermek için kullanılan en yalın örnekleme olarak değerlendirilmelidir. İlk olarak, "+1" terimi hiç görülmeyen çiftler için $freq(o, op) = 0$olsa bile $\log(0)$oluşmasını engeller. İkinci olarak, log dönüşümü sayım değerlerini sıkıştırır; frekanslar geniş bir aralıkta değişse bile skor daha yönetilebilir bir ölçeğe taşınır. Üçüncü olarak, baştaki eksi işareti sıralamayı "nadir olan daha büyük skor alsın" yönünde çevirir: $freq(o, op)$küçükse $s(e)$göreli olarak büyür, $freq(o, op)$büyükse küçülür. Burada kritik nokta şudur: konformal geçerlilik yalnızca bu skorların sıralamasına bağlıdır. Bu nedenle belirleyici olan skorun mutlak ölçeği değil, zararsız referans içindeki konumudur.

Zararsız örnek kalibrasyon kümesindeki skorlar $\{s(e_\ell)\}_{\ell=1}^{n}$ile gösterilirse, olay $e$için konformal p-değeri aşağıdaki gibi tanımlanır:

$$p(e) = \frac{1+|\{\ell \in \{1,\dots,n\}: s(e_\ell) \geq s(e)\}|}{n+1} \tag{3.9}$$

Böylece $p(e) \in (0,1]$ elde edilir. Konformal p-değeri, bir olayın zararsız referansa göre ne kadar sıra dışı olduğunu gösteren standartlaştırılmış bir ölçüdür. p-değeri küçüldükçe olay daha şüpheli kabul edilir. Bu dönüşüm, farklı tespit sistemlerinden gelen skorların ortak bir ölçekte ifade edilmesini sağlar.

Bu bölümde kullanılan gözlemden bağımsızlık (observation invariance) kavramı, gözlemin yok sayılması anlamına gelmez. Buradaki anlam, sertifika üretiminin belirli bir gözlem modeli ya da gözlem çekirdeği (observation kernel) üzerine kurulmamasıdır. Başka bir deyişle karar katmanı loglardan türetilen kanıtı kullanır; ancak üretilen sertifika belirli bir gözlem üretim mekanizmasına veya tek bir skor ölçeğine bağlanmaz.

Bu çalışmada her olay için elde edilen konformal p-değeri, eşit yüzdelik dilimler (kantil) temelli eşikler kullanılarak risk kademelerine ayrılmıştır. $Q = \{0{,}01,\ 0{,}05,\ 0{,}20,\ 0{,}50\}$ eşikleri kullanılarak her olay uygun kademeye atanmıştır. Kademe ağırlığı w_b, ilgili kademeye düşen olayların toplam olaylara oranıdır. Bir kademeye hiç olay düşmezse $w_b = 0$ alınır. Kademe içi nominal durum önseli π_b, iki durumlu $\{BENIGN, ATTACK\}$ uzayında p-değerlerinden türetilmiştir. $ATTACK$ olasılığı, ilgili kademede αeşiğinin altında kalan p-değerlerinin oranı esas alınarak ve genel saldırı oranı ile düzleştirilerek hesaplanmıştır. $BENIGN$ olasılığı ise kalan olasılık kütlesi olarak tamamlanmıştır.

3.6.3. Risk-nötr referans yöntemi (RN)

Karar katmanı, farklı hedef fonksiyonlarının etkisini karşılaştırmak için önce risk-nötr bir kıyas programı kurmaktadır. Bu yaklaşım, belirsiz fayda aralıklarını orta nokta değerleriyle temsil etmekte ve karar kuralını ortalama davranış altında belirlemektedir. Böylece belirsizliğe dayanıklı yaklaşım için açık bir karşılaştırma zemini elde edilmektedir.

Bu çalışmada durum kümesi $\mathcal{S}$={BENIGN, ATTACK} olarak alınır. Risk kademesi b ve aksiyon a için belirsiz fayda aralığı

$$u_{a,\theta} \in [L_{a,\theta}, U_{a,\theta}] \tag{3.10}$$

ile verilir. Uygulamada bu aralık nominal fayda ū etrafında simetrik seçilir: L=ū−ω, U=ū+ω. RN kıyasında orta nokta fayda olup, kademe-bazlı nominal önsel altında beklenen orta nokta fayda

$$E_{b,a}^{mid} := \sum_{\theta \in S} \pi_b(\theta)\, u_{a,\theta}^{mid} \tag{3.11}$$

şeklinde tanımlanır. Karar değişkeni $x_{b,a}$, kademe b için durağan (stationary) aksiyon dağılımını gösterir. RN programı, kademe ağırlıkları w_b ile ağırlıklandırılmış toplam beklenen orta nokta faydayı, her kademede bir olasılık dağılımı olacak şekilde ve küresel maliyet bütçesini aşmadan maksimize eder:

$$\max_{x} \sum_{b=1}^{B_{\text{tiers}}} w_b \sum_{a \in A} E_{b,a}^{mid}\; x_{b,a} \tag{3.12}$$

koşuluyla

$$\sum_{a \in A} x_{b,a} = 1, x_{b,a} \geq 0, \sum_{b} w_b \sum_{a} k_a\, x_{b,a} \leq B. \tag{3.13}$$

RN çözümü, belirsizliği ortalama değerlerle temsil ettiği için yüksek ortalama fayda verebilir. Ancak $\pi_b$tahmin hatası ve fayda aralıklarındaki yanlış kalibrasyon karşısında kırılganlaşabilir. Bu nedenle RN, ana yöntem için kıyas ölçütü olarak kullanılmıştır.

3.6.4. Belirsizlik altında performans alt sınırı garantileyen belirsizliğe dayanıklı (robust) yaklaşım

Ana hedef, her risk kademesi için doğrulanabilir bir performans alt sınırı garantisi (certificate) üretmektir. Önsel (prior) hatası Huber tipi ε-kontaminasyon (ε-contamination) ile modellenir. Buna göre her kademe için gerçek durum dağılımı, nominal π_b etrafındaki belirsizlik kümesinden seçilir:

$$P_b = \{\, q := (1-\varepsilon)\pi_b + \varepsilon r, \in \Delta(S)\,\}, \varepsilon = 0.02. \tag{3.14}$$

Bu bölümde amaç, her risk kademesi için doğrulanabilir bir performans alt sınırı üretmektir. Önsel belirsizlik Huber tipi ε-kontaminasyon ile modellenmiş, pratik OpTC kurulumunda ε=0,02 çalışma noktası olarak seçilmiştir. Bu değer ölçülmüş bir fiziksel sabit değil, sertifika uyumu ile ortalama fayda arasındaki dengeyi gözeten operasyonel bir ayardır.

Fayda belirsizliği $[L_{a,\theta}, U_{a,\theta}]$ aralıklarıyla verildiğinde, dayanıklı sertifika, "nominal önsel altında alt-uç beklenti" ile "en kötü duruma karşı koruma"yı tek ifadede birleştirir. Önce nominal önsel altında alt-uç beklenti

$$E_{b,a}^{lb} := \sum_{\theta \in S} \pi_b(\theta)\, L_{a,\theta} \tag{3.15}$$

olarak tanımlanır. Ardından, kademe $b$için aksiyon dağılımının performans alt sınırı garantisi

$$\gamma_b(x_b) = (1-\varepsilon) \sum_{a \in A} E_{b,a}^{lb}\ x_{b,a} + \varepsilon \min_{\theta \in S} \sum_{a \in A} L_{a,\theta}\ x_{b,a} \tag{3.16}$$

şeklinde alınır. İlk terim, nominal π_b (b risk kademesi için nominal durum dağılımı) altında "alt-uç" beklentiyi temsil eder; ikinci terim ise ε oranındaki kayma payı için en kötü duruma karşı koruma sağlar.

Belirsizliğe dayanıklı program, aynı simplex (geçerli olaslılık dağılımı) ve bütçe kısıtları altında kademeler arası ağırlıklı tabanı büyütür:

$$\max_{x} \sum_{b=1}^{B_{\text{tiers}}} w_b\ \gamma_b(x_b)\ \ ise: \sum_{a} x_{b,a} = 1, x_{b,a} \geq 0, \sum_{b} w_b \sum_{a} k_a\, x_{b,a} \leq B. \tag{3.17}$$

Bu ifade içindeki min operatörü nedeniyle parça-parça doğrusal bir terim içerir. Uygulamada bu terim m_b epigraf değişkeni ile doğrusal biçime çevrilir ve tek bir LP olarak çözülür. Çözümün çıktısı, her kademe için aksiyon dağılımı x_b^*ve buna karşılık gelen sertifika düzeyi γ_b^*'dir. Bu değerler elle verilmemekte; bütçe ve belirsizlik modeli altında LP tarafından üretilmektedir.

Bu sertifikanın yorumu şudur: Kademe b için önsel $q \in P_b$olacak biçimde kaymış olsa ve faydalar her hücrede [L, U] aralığında kalsa bile, seçilen x_b karmasının en kötü durumdaki beklenen faydası $\gamma_b(x_b)$ tabanının altına düşmez; kademeler arası birleşik taban değer, $\sum_b w_b\, \gamma_b(x_b)$ ile raporlanır.

3.6.5. Nadir ama yıkıcı risk (tail risk) karşılaştırması için CVaR senaryo LP

Ortalama faydayı azami kılan RN yaklaşımı, nadir fakat çok düşük fayda üreten durumlara karşı yeterince temkinli olmayabilir. Bu nedenle nadir yıkıcı riski (tail risk) doğrudan hedefleyen bir kıyas olarak CVaR (Conditional Value at Risk) tabanlı bir program kurulur [111]. Amaç, her risk kademesinde "en kötü senaryolar" altında da kabul edilebilir performans sağlayan daha korumacı bir aksiyon dağılımı elde etmektir.

CVaR kıyası, her kademe b için Monte Carlo senaryoları üzerinden tanımlanır. Her senaryo j'da önce durum $\theta_{b,j} \sim \pi_b$ örneklenir. Durum uzayı $\mathcal{S} = \{\text{BENIGN,ATTACK}\}$ olarak alınır. Ardından her aksiyon a için o senaryoda gerçekleşebilecek fayda, belirsizlik aralığından örneklenir:

$$u_{b,j,a} \sim \text{Uniform}(L_{a,\theta_{b,j}}, U_{a,\theta_{b,j}}). \tag{3.18}$$

Kademe b'de seçilen aksiyon dağılımı x_b altında senaryo-toplam fayda

$$U_{b,j}(x_b) := \sum_{a \in A} u_{b,j,a}\ x_{b,a} \tag{3.19}$$

olarak tanımlanır. Kuyruğu büyütmek için "alt kuyruk CVaR", epigraf değişkenleri ile doğrusal bir biçime indirgenir. Her kademe için bir eşik t_b(VaR seviyesi) ve her senaryo için $\zeta_{b,j} \geq 0$ artıkları tanımlanarak

$$\zeta_{b,j} \geq t_b - U_{b,j}(x_b), \zeta_{b,j} \geq 0 \tag{3.20}$$

kısıtları altında aşağıdaki hedef maksimize edilir:

$$\max_{x,t,\zeta} \sum_{b=1}^{B_{\text{tiers}}} w_b \left[t_b - \frac{1}{(1-\tau)N_scen} \sum_{j=1}^{N_{scen}} \zeta_{b,j} \right]. \tag{3.21}$$

Bu amaç fonksiyonu, her kademede en kötü $(1-\tau)$ dilimdeki ortalama faydayı yükseltir. Dolayısıyla düşük fayda üreten nadir senaryolara karşı daha temkinli karar kurallarını öne çıkarır. Çözüm, RN ve robust programlarda olduğu gibi her kademe için simpleks koşulları $\sum_a x_{b,a} = 1, x_{b,a} \geq 0$ ve tüm kademeleri kapsayan küresel bütçe kısıtı $\sum_b w_b \sum_a k_a\, x_{b,a} \leq B$ altında elde edilir.

CVaR'de kuyruk parametresi τ, ayrı bir doğrulama diliminde $\tau \in \{0{,}70;\ 0{,}80;\ 0{,}90;\ 0{,}95\}$ ızgarası üzerinde seçilmiştir. Her τiçin LP çözülmüş, politika sertifika uyumu ile değerlendirilmiş ve eşitlik durumunda aşırı tutuculuğu önlemek için daha yüksek τtercih edilmiştir. Bu süreç sonunda $\tau^* = 0{,}90$ elde edilmiştir.

3.6.6. Değerlendirme tasarımı: bütçe eşleştirme, sertifikaya uyum ve referans etikete sabitlenmiş metrikler

Değerlendirme, farklı yöntemlerin ürettiği aksiyon dağılımlarını aynı bütçe düzeyi altında karşılaştıracak şekilde tasarlanmıştır. Böylece yöntemler arasındaki farkın bütçeden değil, hedef fonksiyonundan kaynaklanması sağlanmıştır. Bu amaçla tüm raporlama aynı maliyet vektörü altında ve $B = 0{,}90 \pm 0{,}01$ bütçe bandında yürütülmüştür.

Bu çerçevede iki ana çıktı raporlanmaktadır: performans alt sınırı seviyesi γve sertifikaya uyum. İlk ölçüt, karar kuralının belirsizlik altında garanti ettiği tabanı gösterir. Bu tezde γ-uyum, teoremdeki en kötü durum garantisinin doğrudan olasılık yorumu değil; seçilen politikanın Monte Carlo ile örneklenen senaryolarda kendi γ tabanını ne sıklıkla aştığını gösteren ampirik bir tanı ölçütüdür. Böylece değerlendirme, yalnızca ortalama başarıya değil, karar kuralının ilan ettiği tabanı ne ölçüde koruduğuna da odaklanır.

Sertifikaya uyum Monte Carlo örnekleme ile ölçülmüştür. Burada "ihlal", güvenlik ihlali değil, performans alt sınırı ihlalidir. Her risk kademesi için $S = 1000$tekrar üretilmiştir. Her tekrarda durum π_b'den örneklenmiş, aksiyon faydaları ilgili $[L, U]$aralıklarından çekilmiş ve seçilen aksiyon dağılımı altında gerçekleşen fayda hesaplanmıştır. Gerçekleşen fayda $U_{b,i}$ ilgili kademe için hesaplanan $\gamma_b(x_b)$ değerinin altında kaldığında senaryo-bazlı ihlal, üstünde veya eşit kaldığında ise tabanın ampirik olarak sağlandığı bir tekrar olarak işaretlenmiştir. Sertifikaya uyum ise ihlal oluşmayan tekrarların oranı olarak raporlanmıştır.

$$\text{Uyum} \coloneqq \sum_{b=1}^{B_{\text{tiers}}} w_b \cdot \frac{1}{S_{MC}} \sum_{i=1}^{S_{MC}} 1\left[U_b^{(i)} \geq \gamma_b(x_b)\right]. \tag{3.22}$$

Bu ölçüm, belirsizlik modeli altında karar kuralının ilan ettiği tabanı ne ölçüde koruduğunu doğrudan gösterir.

Değerlendirmenin yalnızca model içi ölçütlerle sınırlı kalmaması için etiket temelli ek ölçütler de raporlanmıştır. Bunların başında zararsız akışta ağır müdahale oranı gelmektedir. Bu ölçüt, sertifika düzeyi yüksek görünen bir karar kuralının pratikte zararsız olaylara gereksiz müdahalede bulunup bulunmadığını gösterir.

Bu doğrultuda zararsız akışta ağır müdahale oranı (benign strong action rate) tanımlanır ve zararsız örneklerde sert aksiyonlara (THROTTLE, QUARANTINE) giden olasılık kütlesi ölçülür:

$$FI_{\text{benign}} \coloneqq \sum_{b=1}^{B_{\text{tiers}}} w_b^{\text{benign}} \cdot \left(x_{b,\text{THROTTLE}} + x_{b,\text{QUARANTINE}}\right), \tag{3.23}$$
$$w_b^{\text{benign}} = \Pr\left[b(x) = b \mid \text{BENIGN}\right], \sum_b w_b^{\text{benign}} = 1.$$

Buna ek olarak, aynı bütçe eşleştirme koşulları altında ortalama fayda da raporlanmıştır. Bu değer uygulamada orta nokta faydalar (midpoint utilities) üzerinden hesaplanmıştır.

Tüm istatistikler farklı rastgelelik tohumları (seed) üzerinden hesaplanmış ve güven aralıklarıyla birlikte verilmiştir.

3.6.7. Duyarlılık ve stres: kademe sayısı, aralık genişliği ve gözlem kaymaları

Yöntemin dayanıklılığı, tek bir ayarda yüksek performans göstermesinden çok, koşullar değiştiğinde tutarlı davranmasına bağlıdır. Bu nedenle iki eksende duyarlılık analizi yürütülmüştür: risk kademe sayısı $B_{\text{tiers}} \in \{3,5,8\}$ ve fayda aralığı genişliği $\omega \in \{0.05, 0.10, 0.15, 0.20\}$. Ayrıca önsel kayması, kademe kütlesi kayması ve eksik gözlem gibi stres senaryoları altında sertifika uyumunun nasıl değiştiği incelenmiştir.

Kademe sayısı (number of tiers): $B_{\text{tiers}} \in \{3, 5, 8\}$ olacak şekilde değiştirilir. Böylece daha kaba veya daha ince risk kademesinin karar dağılımına etkisi izlenir.

Fayda aralığı genişliği (interval width): belirsizlik düzeyi $\omega \in \{0{,}05; 0{,}10; 0{,}15; 0{,}20\}$ tablosuna taranır. Böylece [L, U] aralıklarının genişlemesine karşı karar yapısının performans alt sınırı garantisi ve uyum davranışı değerlendirilir.

Bunlara ek olarak, sahadaki bozulmaları taklit eden stres senaryoları uygulanır: önselin ve kademe kütlelerinin kayması, olayların bir kısmının maskelenmesi (missingness) ve kademelere düşen oranların değişmesi gibi durumlarda, sertifika tanımı sabit tutularak karar kuralının performansının ne ölçüde kontrollü biçimde düştüğü raporlanır. Bu bağlamda log kaydı bulunmayan varlıklar karar katmanına doğrudan kanıt sağlamaz. Bu durum eksik gözlem olarak ele alınır.

3.6.8. Uygulama ve tekrar üretilebilirlik

Tüm deneyler Python ortamında yürütülmüş, doğrusal programlar scipy.optimize.linprog içindeki HiGHS çözücüsü ile çözülmüştür. Tekrar üretilebilirlik için örneklenen dosyalar, kademe eşikleri, kademe ağırlıkları w_b, kademe önsel olasılıkları π_b, çözücü çıktıları, aksiyon dağılımları $x_{b,a}$ve tüm değerlendirme metrikleri kaydedilmiştir.

Operasyonel akış şöyledir: eCAR logları okunur ve nesne–eylem çiftlerine ayrıştırılır. Zararsız referans üzerinden konformal p-değerleri hesaplanır, olaylar risk kademelerine ayrılır ve bütçe bandı altında LP çözülerek x_b^*ile γ_b^*elde edilir. Yeni bir olay geldiğinde önce olayın risk kademesi belirlenir, ardından karar ilgili kademe için önceden hesaplanmış aksiyon dağılımından verilir. Sistem tepkiyi değil, olayın risk düzeyini belirler; karar da bu düzeye karşılık gelen dağılımdan üretilir. LP yalnızca bütçe, εveya fayda tablosu güncellendiğinde yeniden çözülür.

Sonuç olarak bu karar katmanı, dedektör skorlarını ortak bir risk diline çevirir ve bunları bütçe altında raporlanabilir, denetlenebilir aksiyonlara dönüştürür.

Varsayılan çalışma noktası olarak $Q = \{0{,}01; 0{,}05; 0{,}20; 0{,}50\}$ ile beş risk kademesi, $\alpha = 0{,}05$, $\eta = 0{,}20$, $\varepsilon = 0{,}02$, $\omega = 0{,}15$ ve $B = 0{,}90$ kullanılmıştır. Bu değerler serbest bir performans aramasıyla değil, yöntem tasarımı, duyarlılık incelemeleri, tekrar üretilebilirlik ve bütçe eşleştirmesi gözetilerek belirlenmiştir. Kademe sayısı ve belirsizlik genişliği için yapılan taramalar ayrıca raporlanmıştır. Yalnızca CVaR'nin τparametresi ayrı bir doğrulama diliminde ayarlanmıştır.

Beklenen aksiyon maliyeti bütçesi $B = 0{,}90$ olarak alınmış, değerlendirme $| E[cost] - 0{,}90 | \leq 0{,}01$ bandında yürütülmüştür. Örnekleme için 42, 17 ve 101 gibi sabit rastgelelik tohumu (seed) değerleri kullanılmış, böylece tekrar üretilebilirlik sağlanmıştır. CVaR için $\tau^* = 0{,}90$ayrı bir doğrulama diliminde seçilmiştir.

4. BULGULAR

Bu bölümde tez kapsamında yürütülen iki çalışmadan elde edilen bulgular, her bir çalışmanın kendi deneysel düzeni korunacak şekilde ve açık bir ayrım yapılarak sunulmuştur. Teorik hat D1 scaled robust game, D2 observation stress, D3 3×3×3 micro tightness ve D4 NSL-KDD-derived semi-empirical oyunlarını kapsar; saha karar katmanı ise DARPA OpTC NCR eCAR logları üzerinde raporlanır. Ek olarak, önerilen karar katmanının neden sabit alarm eşiğine dayanmadığını göstermek için, alan kayması altında eşik taşınabilirliğini inceleyen destekleyici ek çalışma sunulmuştur (Bölüm 4.6). İlk olarak ardışık savunma probleminde, tek bir doğrusal program (LP) çözümü ile üretilen gözlemden bağımsız sertifikanın sayısal davranışı ve doğrulama testleri raporlanmıştır. İkinci olarak aynı sertifika fikrinin, UEBA (User and Entity Behavior Analytics) çıktılarının konformal kanıt (conformal evidence) ile risk kademelerine dönüştürülmesi ve küresel bütçe kısıtı altında tek bir belirsizliğe dayanıklı (robust) LP ile eyleme çevrilmesi üzerinden SOC karar katmanında nasıl çalıştığı, DARPA OpTC NCR köken izi verisi üzerinde ampirik olarak gösterilmiştir.

4.1. Tek LP ile gözlemden bağımsız ardışık savunma sertifikası

Birinci çalışmada, savunucu ile saldırganın tekrarlı olarak etkileştiği ve çevresel durumun (latent durum) tam gözlenemediği ardışık savunma probleminde, savunucunun garanti edebileceği en yüksek sabit performans alt sınırı tek bir LP üzerinden oluşturulmuştur. LP'nin temel çıktıları, savunucunun sabit karma stratejisi (savunma eylemlerini hangi olasılıklarla seçeceğini belirleyen dağılım) ve tur başına beklenen getiri tabanı v*'dir. Sertifika gözlemden bağımsızdır. Başka bir ifadeyle, verilen performans alt sınırı belirli bir log üretim ve toplama düzenine bağlı değildir ve POMDP (kısmi gözlemlenebilir Markov karar süreci; gözlemlerle gizli duruma ilişkin olasılık dağılımını güncelleyerek karar veren model) türü güncellemelere ihtiyaç duymadan da geçerlidir. Bu bağımsızlık, karar kuralının gözlem kullanmaması değil, garantinin gözlem mekanizmasını açıkça modellemeye bağlanmaması anlamına gelir. Belirsizlik modeli iki kaynağı birlikte ele alır: önsel dağılımda Huber ε-kirlenmesi ve getirilerin hücre bazında aralık belirsizliği (nominal getirinin etrafında ±ω yarı-genişlikli aralıklar).

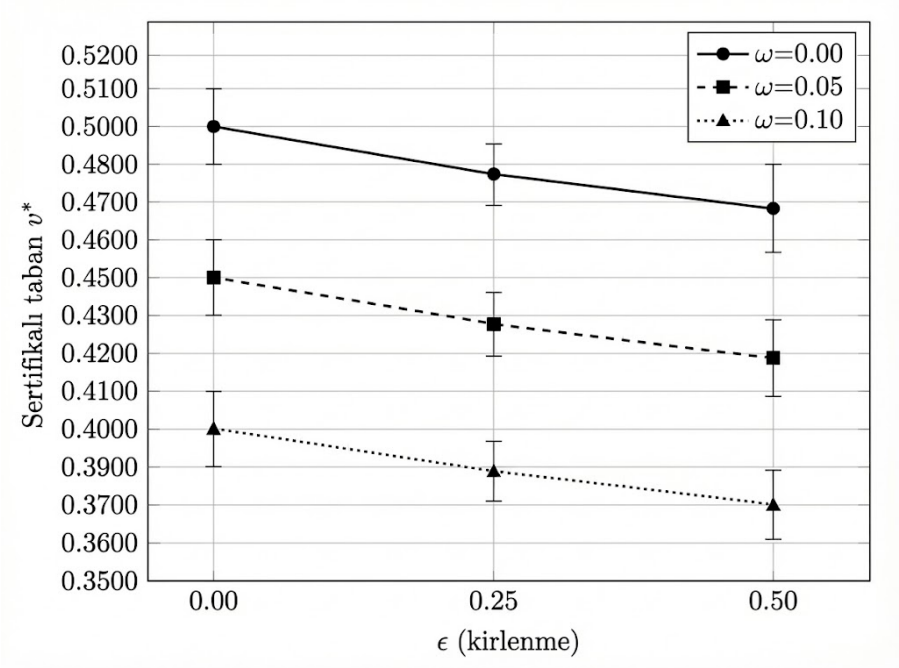
4.1.1. Teorik çalışma: Performans alt sınırının belirsizlik parametrelerine duyarlılığı

Sentetik olarak üretilen problem örneklerinde performans alt sınırı v^* değerinin, hem ε hem de ω arttıkça düzenli biçimde azaldığı gözlenmiştir. Bu azalış, sertifikanın pratik ayar parametreleri (ε ve ω) seçilirken öngörülebilir bir davranış sergilediğine işaret etmektedir. Ölçülen eğilim, ε'ye göre neredeyse doğrusal olup Tablo 4.1'deki örnek deney setinde örnek bir deney tablosunda v^*'nin ε-eğimi yaklaşık −0.054 olarak raporlanmıştır ($\partial v^*/\partial \varepsilon \approx -0.054$). Bu değer, sertifikanın ikili (dual) çözümünden türetilen duyarlılık (subgradient) öngörüleri ile tutarlıdır. ε toleransı büyüdükçe garanti edilen tabanın sistematik olarak düştüğünü gösterir.

Tablo 4.1: ε ve ω değerlerinde performans alt sınırı $\boldsymbol{v}^*$ (ortalama ± %95 güven aralığı).

ε	ω=0.00	ω=0.05	ω=0.10
0.00	0.5000 ± 0.0007	0.4500 ± 0.0007	0.4000 ± 0.0007
0.25	0.4850 ± 0.0006	0.4350 ± 0.0006	0.3850 ± 0.0006
0.50	0.4731 ± 0.0009	0.4231 ± 0.0009	0.3731 ± 0.0009

Yukarıdaki tabloda değerler, belirsizlik düzeyi artırıldığında tabanın sistematik olarak aşağı kaydığını göstermektedir. Örneğin ω sabit tutulduğunda ε=0.00'dan ε=0.50'ye geçişte v^* yaklaşık 0.027 puan düşmektedir; ω artırıldığında ise aynı ε seviyesinde taban daha düşük bir bantta konumlanmaktadır. Bu parametrik değişim ve azalma eğilimi Şekil 4.1 üzerinde de görselleştirilmiştir.



Şekil 4.1: ε ve ω arttıkça performans alt sınırı $\boldsymbol{v}^*$ değerinin azalması

v^*'nin ε arttıkça düzenli biçimde azalması, ε'nın nominal önsel $\pi_0$etrafında izin verilen sapma payını belirleyen bir ayar parametresi olduğunu gösterir. ε Büyüdükçe, model nominal önsel π_0'dan daha büyük sapmaları olası kabul eder ve bu genişletilmiş belirsizlik altında çözülen LP daha düşük bir v^*tabanı üretir. Bununla birlikte gerçek önsel sapma, kullanılan ε değerinin kapsadığı sapma sınırını aşarsa, v^*değişmeden kalsa bile bu sertifika artık o koşullar için geçerli bir garanti olarak yorumlanmamalıdır.

4.2. Gözlem bozulmaları altında değişmezlik (observation invariance)

Çalışmanın ayırt edici bulgusu, gözlem hattında bozulma olsa bile tek LP ile hesaplanan performans alt sınırı v*'nin yeniden hesaplanmadan geçerliliğini korumasıdır. Bu amaçla gözlem akışına üç tür bozulma uygulanmıştır. Kayıt kaybı (miss) bazı logların düşmesini, sinyal yer değiştirmesi (swap) ise log semantiğinin karışmasını temsil eder. Kayıp ve yer değiştirmenin beraber geçerli olduğu durum (both) iki bozulmanın birlikte uygulanmasıdır. Her üç durumda da gözlenen ortalama getiri v* alt sınırının üzerinde kalmış ve alt sınır ihlal oranı sıfır olarak raporlanmıştır. Bu sonuç alt sınırın belirli bir log toplama düzenine bağlı olmadığını pratikte doğrular; log hattı değişse bile aynı sertifika taşınabilir.

Tablo 4.2: Gözlem bozulmaları altında ortalama getiri r̄ ve alt sınır ihlal oranı (K=10.000)

Bozulma türü	Ortalama getiri r̄ (± %95 GA)	Alt sınır ihlal oranı
Beraber	0.4979 ± 0.0015	0.000
Kayıp	0.4966 ± 0.0020	0.000
Yer değiştirmiş	0.4971 ± 0.0023	0.000

Tablo 4.2'deki sonuçlar gösteriyor ki gözlem bozulmaları altında ortalama getiri r̄ ve yüzde 95 güven aralıklarında alt sınır ihlal oranı tüm bozulma türlerinde sıfırdır.

4.3. Sonlu tur sayısında Freedman ve Azuma–Hoeffding güven bantlarının karşılaştırması

Sertifika yalnızca asimptotik (tur/örnek sayısı çok büyüdüğünde geçerli) bir taban olarak bırakılmamıştır. Sonlu tur sayısı (K) için, gözlenen ortalama getiri $\bar{r}$'nin performans alt sınırı değer v^*'ın altına düşmesine ilişkin belirsizliği nicelleştiren yüksek olasılıklı güven bantları türetilmiştir. Buna göre, seçilen güven düzeyi $1-\delta$ altında $\bar{r}$ için bant yarı-genişliği $H(K,\delta)$ hesaplanır ve $\bar{r}$'nin v^* çevresindeki sapması bu yarı-genişlik ile sınırlandırılır.

Bu amaçla iki farklı sapma sınırı birlikte raporlanmıştır. İlkinde, martingal toplamları için koşullu varyans bilgisini de kullanan, varyansa duyarlı bir sınır veren Freedman eşitsizliği temel alınarak H_F elde edilmiştir. İkincisinde, artımların yalnızca aralıkla sınırlı olmasına dayanıp en kötü durum sapma garantisi veren Azuma–Hoeffding eşitsizliği kullanılarak H_A hesaplanmıştır. Freedman bandı, getirilerin oynaklığının düşük olduğu rejimlerde daha dar bantlar üretebilir; oynaklığın yüksek olduğu rejimlerde ise Azuma–Hoeffding bandına yakın ya da ondan daha geniş olabilir. İncelenen temsilî ayarlarda her iki bant da Karttıkça yaklaşık $O(1/\sqrt{K})$ ölçeğinde daralmaktadır.

Tablo 4.3: Tur sayısı K K için bant yarı genişlikleri (H)

Tur Sayısı (K)	**Freedman H_F**	**Azuma H_A**
1000	0.0567	0.0387
10000	0.0175	0.0122

Tablo 4.3, tur sayısı arttıkça güven bandının hızla daraldığını gösterir. Yani daha fazla tur gözlendiğinde ortalama getiri daha az oynar. Bu nedenle performans taban değer yalnızca teorik bir iddia değildir. Belirli bir tur sayısında ortalamanın bu tabandan en fazla ne kadar sapabileceği sayısal olarak verilebilir.

4.4. Sıkılık, tanılama ve verimlilik için mikro örnekler, örnekleme ve çözüm süresi

Küçük boyutlu 3×3×3 mikro örneklerde, tek LP sertifikasının değeri ile belirsizliğe dayanıklı dinamik programlama (DP) temelli referans değerin tam olarak çakışması (sıfır fark) sertifikanın bu ölçeklerde sıkı olduğunu göstermektedir; yani performans taban sınırı, gereksiz bir gevşeklikten kaynaklanan performans kaybı oluşturmadan, ulaşılabilir en iyi tabanı vermektedir. Buna ek olarak, LP'nin ikili (dual) çözümünden elde edilen duyarlılık sinyalleri, hangi durum-eylem hücrelerinde belirsizliğin performans alt sınırını daha fazla sınırladığını görünür kılmıştır. Bu bilgi, ölçüm veya örnekleme çabasının hangi hücrelere önceliklendirilmesi gerektiğini açıklayan pratik bir tanılama mekanizması olarak kullanılabilmektedir. Tablo 4.4, elde edilen bu sonuçları özetlemektedir.

Tablo 4.4: Performans tabanı sıkılığı, örnekleme kazanımı ve seçili çözüm süreleri (özet).

Test	Ölçüt	Sonuç
Mikro sıkılık (3×3×3)	DP referansı − LP sertifikası	0.0 (ör. v^*: 0.3, 0.325, 0.3375)
İkili (dual) çözüme dayalı örnekleme	v^* (100.000 örnek)	rehberli: 0.385291; eşit dağılımlı: 0.384741
Çözüm süresi (sentetik)	tek LP (örnek ayarlar)	450–520 ms

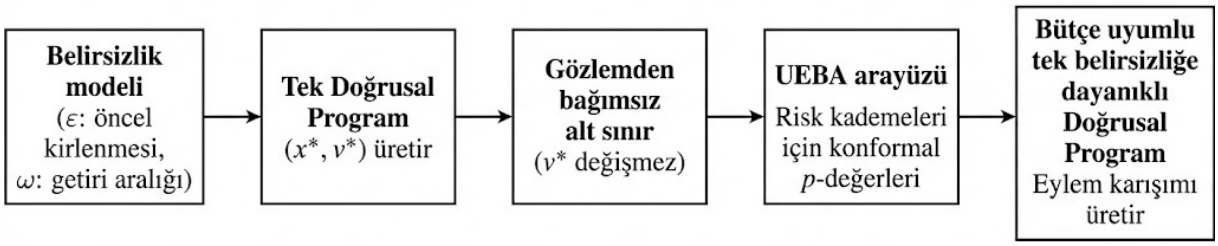
İkili (dual) çözüme dayalı örnekleme sonuçları, iyileşmenin sayısal olarak küçük fakat tutarlı olduğunu göstermektedir. Bu tür bir kazanım, özellikle yüksek boyutlu tablolarda ölçüm maliyeti arttığında anlamlı hâle gelir; çünkü aynı örnekleme bütçesiyle daha sıkı aralıklar elde edilerek v^* tabanı yükseltilebilir.

4.5. UEBA karar katmanı bulguları

4.5.1. Konformal kanıt ve risk kademeleri

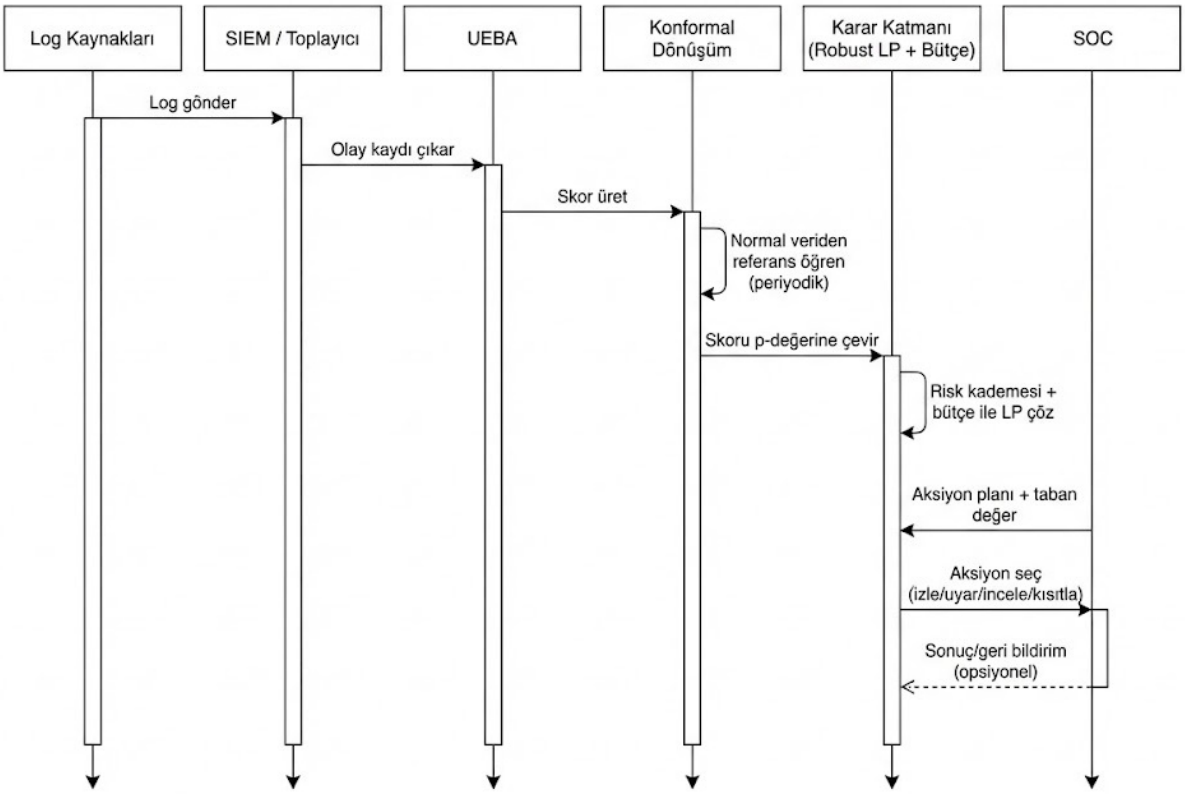
Bu çalışmada sertifika fikri, operasyonel bir SOC karar problemine taşınmıştır. Amaç, farklı UEBA bileşenlerinin üretebileceği heterojen skorları (ölçekleri ve dağılımları farklı olabilen puanlar) ortak ve taşınabilir bir arayüze dönüştürmek ve bu arayüz üzerinden bütçeli eylem seçimini tek bir LP ile çözmektir. Bunun için zararsız (normal) veri üzerinde konformal kalibrasyon yapılarak her olay için p-değeri elde edilmiştir (Sistem önce saldırı olmayan kayıtlara bakarak "bu skor ne kadar sıra dışı sayılır?" ölçüsünü öğrenir; sonra her yeni olay için, skorun normale göre ne kadar aykırı olduğunu 0 ile 1 arasında bir değerle ifade eder. Bu değer küçüldükçe olay daha şüpheli kabul edilir. p-değerleri birkaç risk kademesine (tier) ayrılmıştır. Ardından her risk kademesi için sabit bir karma eylem dağılımı hesaplanmış; küresel maliyet bütçesi altında, aralık (interval) fayda belirsizliği ve ε-kirlenmeli önsel belirsizliğini hesaba katan belirsizliğe dayanıklı LP, kademeler arası karışım ağırlıklarıyla tanımlanan performans alt sınırını (γ) maksimize edecek şekilde çözülmüştür.

UEBA skorlarının konformal p-değerine dönüştürülmesinden başlayıp risk kademeleri ve bütçe (B) altında aksiyon dağılımı seçimine uzanan operasyon akışı aşağıda Şekil 4.2 üzerinde özetlenmiştir.



Şekil 4.2: Konformal p-değeri ve risk kademeleri üzerinden tek LP ile sertifikalı karar.

SOC ortamında her gün çok sayıda olay kaydı gelir ve UEBA sistemi her olay için bir skor üretir. Önce saldırı olmadığı bilinen normal kayıtlar kullanılarak skorların normalde nasıl davrandığı öğrenilir (zararsız kalibrasyon) ve her olayın skoru normale göre 0 ile 1 arasında bir değere dönüştürülür (konformal p-değeri), bu değer küçüldükçe olay daha şüpheli sayılır. Sonra bu değerler sıralanır ve olaylar birkaç risk kademesine ayrılır (eşit dilimlere ayırma, B_{tiers}), her kademenin akış içindeki payı da bir ağırlık olarak belirlenir (w_b). Ardından tek bir doğrusal program çözülür (belirsizliğe dayanıklı LP). Bu çözüm, toplam müdahale maliyetini bütçe içinde tutar (küresel bütçe B). Her risk kademesi için hangi müdahalenin ne sıklıkla uygulanacağını verir (aksiyon dağılımı $x_{b,a}$) ve aynı anda her kademe için bir performans alt sınırı değeri üretir (kademe sertifikası γ_b). Bu değerlerin ağırlıklı toplamı olan genel taban değer de hesaplanır (genel sertifika $\gamma = \sum_b w_b \gamma_b$). Sistem sahada çalışırken hem ilan edilen tabanın ne kadar sık sağlandığı izlenir (γ-uyum) hem de yeterli veri biriktiğinde tabanın etrafındaki dalgalanmanın sınırını veren bir bant raporlanır (Freedman bandı). Şekil 4.3, bu süreci log skorların konformal p-değerlerine çevrilmesi, risk kademelerine ayrılması ve bütçe altında tek bir belirsizliğe dayanıklı LP ile kademe bazında aksiyon dağılımı ile sertifika çıktılarının üretilmesi olarak özetler.



Şekil 4.3: UEBA skorlarından risk kademeleri ve bütçe altında LP tabanlı karar

4.5.2. Bütçe eşleştirmesi altında yaklaşım karşılaştırması

Değerlendirmeler, Bölüm 3.6.1'de tanımlanan birbirini dışlayan iki OpTC değerlendirme dilimi üzerinde yürütülmüştür: OpTC-A ve OpTC-B. Her dilim için 50 bağımsız veri beslemesi (seed) kullanılmıştır.

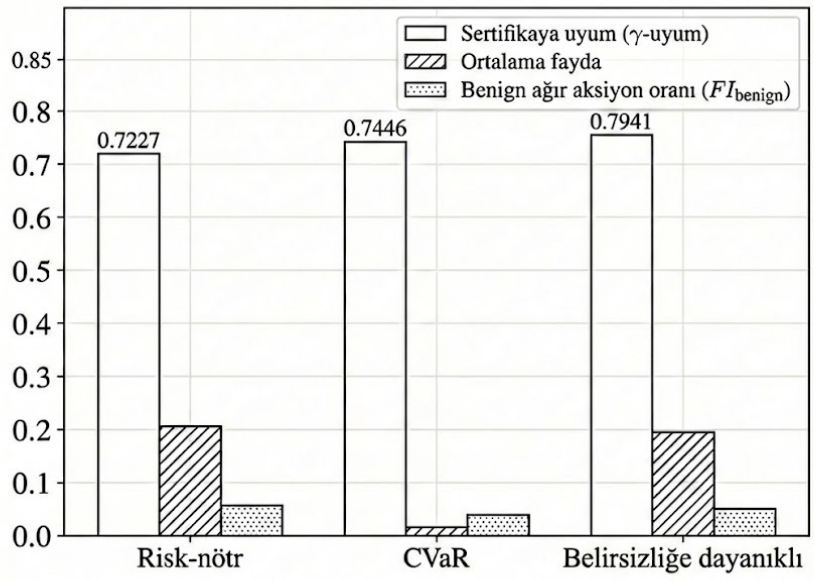
Tüm aksiyon dağılımları için beklenen maliyet, B=0.90 hedefi etrafında iki taraflı ±0.01 bant ile eşleştirilmiştir; böylece yaklaşım karşılaştırmaları aynı harcama düzeyinde gerçekleştirilmiştir. Karşılaştırılan yaklaşımlar üç gruptur. Risk nötr LP orta nokta faydayı en yüksek yapacak şekilde seçer (orta nokta fayda her hücredeki alt ve üst fayda sınırlarının ortalamasıdır ve tek bir tahmini fayda tablosu gibi kullanılır). CVaR yaklaşımı nadir ama yıkıcı kayıpları sınırlamayı hedefler (CVaR en kötü sonuçların belli bir yüzdesindeki ortalama kaybı sınırlayan kuyruk riski ölçüsüdür). Belirsizliğe dayanıklı LP belirsizlik altında performans alt sınırını en yüksek yapacak şekilde seçer (fayda aralıkları ve önsel sapması içinde en kötü durumda bile sağlanacağı garanti edilen tabanı büyütür). Bu tasarımda ampirik taban sağlanma oranı (γ-uyum), her aksiyon dağılımının kendi ürettiği performans alt sınırının (γ) belirsizlik modeli altında ne kadar sıklıkla sağlandığını ifade eder. Kriter, aksiyon dağılımının kendi ilan ettiği tabanı ne kadar sık sağladığını ölçen ve doğrulanabilen bir güvenilirlik metriğidir. Bu

tasarımda CVaR parametresi τ, ayrık bir doğrulama diliminde belirlenmiştir. ε, ω ve risk kesimleri Q tüm değerlendirmelerde sabit tutulmuş, sonuçlar yalnızca ayrık değerlendirme dilimlerinde raporlanmıştır.

Tablo 4.5: B=0.90 ± 0.01 bütçe bandında yaklaşım karşılaştırması (havuzlanmış).

Yaklaşım	Sertifikaya uyum (γ-uyum)	Ortalama fayda	Zararsız olaylara ağır aksiyon uygulama oranı	Beklenen maliyet
Risk-nötr	0.7227 ± 0.0148	0.2060 ± 0.0246	0.0574 ± 0.0030	0.909
CVaR	0.7446 ± 0.0095	0.0162 ± 0.0118	0.0402 ± 0.0011	0.890
Belirsizliğe dayanıklı	0.7941 ± 0.0230	0.1944 ± 0.0252	0.0513 ± 0.0024	0.906

Tablo 4.5'te belirsizliğe dayanıklı yaklaşımın risk-nötr yaklaşıma göre γ-uyumu 0.7227'den 0.7941'e yükselttiği görülmektedir; bu, ihlal oranının 0.2773'ten 0.2059'a düştüğü anlamına gelir. Bu artış OpTC-A ve OpTC-B dilimlerinde ayrı ayrı yapılan çiftli karşılaştırmalarda da tutarlı çıkmış ve istatistiksel olarak anlamlı bulunmuştur. Aynı bütçe bandında ortalama fayda büyük ölçüde korunmuştur. Buna karşılık zararsız olaylarda ağır müdahaleyi temsil eden kısıtlama ve karantina türü aksiyonların oranı 0.0574'ten 0.0513'e düşmüştür. CVaR yaklaşımı γ-uyumu 0.7446'ya çıkarabilse de ortalama faydayı 0.0162'ye kadar indirmiştir. Bu sonuç, kuyruk riskini azaltmaya odaklanan yöntemlerde görülen tipik fayda kaybı ile uyumludur. Şekil 4.4 bu üç yaklaşımı sertifikaya uyum (γ-uyum), ortalama fayda ve zararsız örnek ağır aksiyon oranı bakımından görsel olarak karşılaştırmaktadır.



Şekil 4.4: Risk yaklaşımlarının performans karşılaştırması.

Yeni gelen APT için saldırıya özgü ayrı bir parametre vektörü öğrenilmez. Yeni olayın dedektör skoru önce zararsız kalibrasyon kümesine göre konformal p-değerine dönüştürülür. Bu p-değeri daha sonra önceden tanımlı risk kademelerinden birine atanır. Kademe ağırlığı ve kademe önseli akıştaki sayımlardan elde edilir. Q, ε, ωve B ise tüm sistem için ortak kullanılan parametrelerdir. Dolayısıyla burada yapılan işlem yeni APT'ye özel parametre öğrenmek değil, yeni kanıtı ortak karar katmanına yerleştirmektir.

4.5.3. Performans alt sınırı düzeyi ve belirsizliğe dayanıklı yaklaşımın üstünlüğü

$\bar{\gamma}$ bir olasılık değildir. $\bar{\gamma}$, risk kademelerine göre ağırlıklandırılmış garanti edilen net fayda alt sınırıdır ve Tablo 3.2'deki net fayda puanlarıyla aynı ölçektedir. Buna karşılık γ-uyum bir orandır ve [0,1] aralığındadır. Bu nedenle $\bar{\gamma}$ negatif olabilir. Negatif değer, yöntemin hata verdiğini değil, belirsizlik altındaki en kötü durumda garanti edilebilen net faydanın sıfırın altında kaldığını gösterir. Bunun nedeni, fayda tablosunda bazı hücrelerin zaten negatif olmasıdır; örneğin, zararsız olayda karantina −7.00, saldırı durumunda hiçbir şey yapmama −10.00'dur. CVaR yaklaşımı kuyruktaki kötü durumları öne çıkardığı için bu olumsuz hücrelerin etkisi büyüyebilir ve $\bar{\gamma}$ negatif çıkabilir. Tablo 4.6, belirsizliğe dayanıklı yaklaşımın $\bar{\gamma}$ düzeyini A diliminde 0.075'ten 0.102'ye, B diliminde 0.109'dan 0.130'a ve havuzda 0.092'den 0.116'ya çıkardığını göstermektedir.

Tablo 4.6: Karışım-ağırlıklı performans alt sınırı seviyesi $\bar{\gamma}$ (B=0.90 ± 0.01).

Dilim	Risk-nötr $\bar{\gamma}$	CVaR $\bar{\gamma}$	Belirsizliğe dayanıklı $\bar{\gamma}$
A	0.075 ± 0.025	-0.028 ± 0.013	0.102 ± 0.023
B	0.109 ± 0.030	-0.009 ± 0.018	0.130 ± 0.028
Havuz	0.092 ± 0.020	-0.019 ± 0.011	0.116 ± 0.018

4.5.4. Stres testleri: dağılım kayması altında davranış

Karar katmanı, operasyonel ortamda beklenen dağılım kaymalarına karşı da sınanmıştır. Önselin daha saldırgan veya daha zararsız bir çalışma koşuluna kaydırılması, risk kademelerinin karışım ağırlıklarının bozulması ve bir kısmı maskelenmiş olay akışı gibi kontrollü bozulmalar (pertürbasyonlar) altında, Tablo 4.7'de özetlendiği üzere, belirsizliğe dayanıklı yaklaşım uyumda daha dengeli bir bozulma göstermiştir.

Stres testlerinde taban değerleri sabit tutulmuştur. Bu nedenle gözlenen düşüş, aksiyon dağılımının verdiği alt sınırın küçülmesi anlamına gelmez. Düşüş, dağılım kayması (distribution shift) varken bu alt sınırın ne kadar sık sağlanabildiğinin azalmasını gösterir.

Tablo 4.7: Dağılım kaymasında Stres testinde γ-uyum (parantez içi: nominale göre fark)

Senaryo	Risk-nötr γ-uyum	Belirsizliğe dayanıklı γ-uyum
Nominal senaryo (no shift)	0.7276 (+0.0000)	0.7889 (+0.0000)
Önsel yukarı (prior up)	0.7813 (+0.0537)	0.8404 (+0.0516)
Önsel aşağı (prior down)	0.6739 (-0.0537)	0.7358 (-0.0531)
Kütle kayması (mass shift)	0.7276 (+0.0000)	0.7889 (+0.0000)
Eksik gözlem (missing)	0.6970 (-0.0306)	0.7627 (-0.0261)

4.5.5. Hesaplama maliyeti: milisaniye ölçeğinde çözüm

Üç yaklaşım da doğrusal program olarak çözüldüğünden, çözüm süreleri milisaniye düzeyinde kalmıştır. Tablo 4.8'de görüldüğü üzere, $B = 0.90$bütçe düzeyi ve beş risk kademesi için medyan çözüm süresi risk-nötr yaklaşımda 4 ms, önerilen belirsizliğe dayanıklı yaklaşımda 5 ms, CVaR yaklaşımında ise 68 ms'dir. Sekiz risk kademesinde bu süreler sırasıyla 4 ms, 3 ms ve 167 ms olarak ölçülmüştür. CVaR formülasyonu, senaryo sayısına bağlı ek değişkenler nedeniyle daha yüksek hesaplama maliyeti üretmektedir. Buna karşın, üç yaklaşım da periyodik yeniden çözüm için pratik düzeydedir. Böylece bütçe hedefi ya da fayda/maliyet tabloları değiştiğinde aksiyon dağılımı kısa sürede yeniden üretilebilmektedir.

Burada verilen süreler, uçtan uca log toplama, özellik çıkarımı ve skor üretim hattının tamamını değil, yalnızca karar katmanındaki doğrusal programın çözüm süresini göstermektedir. Nitekim yeni bir olay geldiğinde LP'nin her olay için yeniden çözülmesi gerekmemekte; olay önce ilgili risk kademesine atanmakta, ardından karar o kademe için önceden hesaplanmış aksiyon dağılımından verilmektedir. Dolayısıyla tezdeki hız iddiası, tüm tespit hattının değil, karar katmanının düşük gecikmeli ve sahada uygulanabilir olmasının ampirik olarak desteklenmesine dayanmaktadır.

Çözüm süreleri Google Colab üzerinde yürütülen deneylerden elde edilmiştir. Ölçümlerin alındığı oturumda Python 3.12.12 ve Linux 6.6.113+ (x86_64, glibc 2.35) kullanılmış; oturum tarafından 2 mantıksal CPU, 1 fiziksel çekirdek ve 12.67 GB RAM raporlanmıştır. Ampirik OpTC NCR deneylerinde endpoint-only eCAR logları kullanılmış; her seed için 3 benign ve 3 evaluation dosyası örneklenmiş, dosya başına en çok 80.000 satır taranmış ve 10.000 olay kabul edilmiştir. Buna göre tek bir dilimde işlenen veri hacmi üst sınır olarak yaklaşık 60.000 olay düzeyindedir. Bu bağlamda Tablo 4.8'de raporlanan süreler, belirli bir deney ortamında elde edilmiş ampirik çalışma süreleri olarak değerlendirilmelidir.

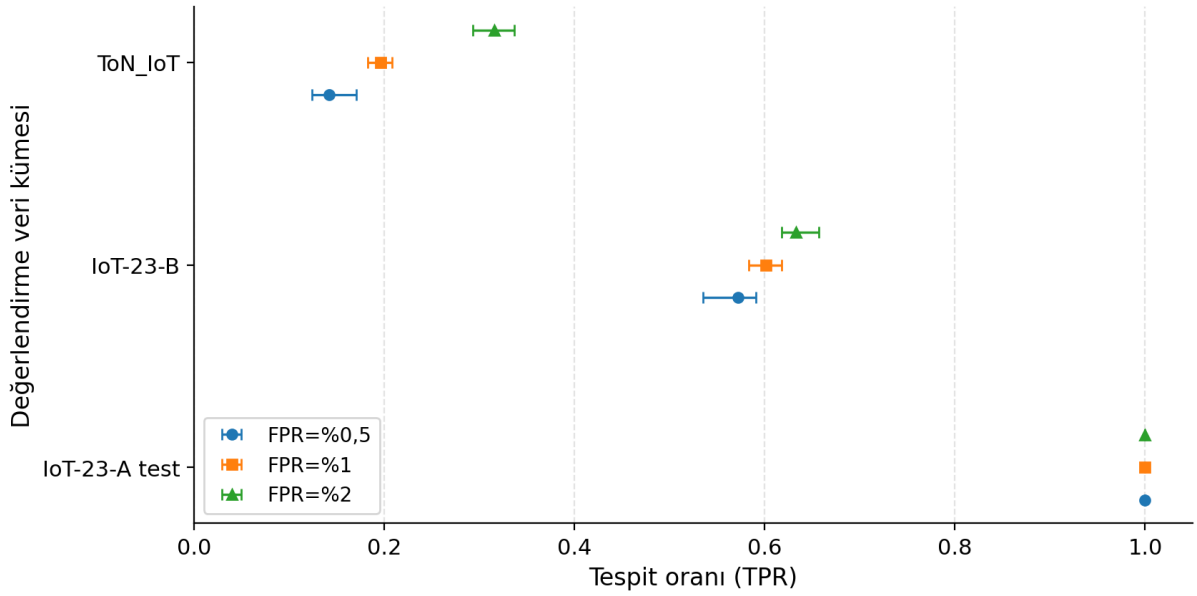
Tablo 4.8: Medyan çözüm süreleri (ms) – temsilî ayarlar.

Yaklaşım	|kademeler| = 5 (B=0.90)	|kademeler| = 8 (B=0.90)	Not
Risk-nötr	4	4	HiGHS çözücüsü
Belirsizliğe Dayanıklı	5	3	HiGHS çözücüsü
CVaR	68	167	N_scen=200 senaryo/tier

4.6. Destekleyici çalışma: Sabit alarm eşiğinin alan kayması altında taşınabilir olmaması

Bu bölüm, ana katkının doğrudan testi değil; sabit alarm eşiği yaklaşımının alan kayması altında neden yetersiz kaldığını göstermek üzere sunulan yardımcı bir doğrulama çalışmasıdır. Burada raporlanan sonuçlar, önerilen karar katmanının gerekliliğini motive eden ikincil kanıt olarak okunmalıdır. Bu destekleyici çalışma, sabit alarm eşiklerinin alan kayması altında nasıl bozulduğunu göstermek için yürütülmüştür [118]. Deneylerde IoT-23 ve ToN_IoT veri kümeleri kullanılmıştır [124], [125]. IoT-23 verisi, aynı cihazın ve aynı zararlı yazılım ailesinin iki parçaya birden düşmemesi sağlanacak biçimde IoT-23-A ve IoT-23-B olarak ayrılmıştır [124]. IoT-23-A içinde zaman sırasına göre eğitim, doğrulama ve test bölümleri oluşturulmuştur. ToN_IoT ise dış alan değerlendirmesi için kullanılmıştır [125].

Her örnek, 300 saniyelik bir ağ trafiği penceresi olarak alınmıştır. Model girdisi olarak, bu pencerelerden çıkarılan tablosal ağ özellikleri kullanılmıştır. Ölçekleme yalnızca IoT-23-A eğitim bölümünde öğrenilmiş; aynı dönüşüm, IoT-23-A doğrulama, IoT-23-A test, IoT-23-B ve ToN_IoT üzerinde değiştirilmeden uygulanmıştır. Eksik özellikler, IoT-23-A eğitim bölümündeki ortalama değerlerle doldurulmuştur. Tespit modeli olarak L2 düzenlemeli lojistik regresyon ve XGBoost kullanılmıştır. Ancak bu bölümde raporlanan ana sayısal sonuçlar XGBoost modeline aittir. Lojistik regresyon modeli de aynı yönde sonuçlar vermiştir.



Şekil 4.5: IoT-23-A doğrulamada seçilen sabit alarm eşiklerinin IoT-23-A test, IoT-23-B ve ToN_IoT üzerindeki TPR değerleri (XGBoost).

Skorlar önce Platt ölçekleme, ardından izotonik regresyon ile kalibre edilmiştir [121]–[123]. Alarm eşikleri, izotonik kalibrasyon sonrası skorlar üzerinde, yalnızca IoT-23-A doğrulama bölümündeki negatif örneklerden FPR = %0,5, %1 ve %2 hedefleri için seçilmiştir. Bu eşikler daha sonra yeniden ayarlanmadan IoT-23-A test, IoT-23-B ve ToN_IoT kümelerine aktarılmıştır. IoT-23-A test bölümünde TPR@%0,5/%1/%2 değerleri 1,000/1,000/1,000 olarak elde edilmiştir. Aynı eşikler IoT-23-B'ye aktarıldığında bu değerler 0,572/0,601/0,633'e, ToN_IoT'e aktarıldığında ise 0,142/0,196/0,316'ya düşmüştür. Buna karşılık IoT-23-B'de ROC-AUC = 0,953 ve PR-AUC = 0,965 düzeyinde yüksek kalmıştır [119], [120]. İzotonik kalibrasyon sonrasında ECE değerleri IoT-23-A, IoT-23-B ve ToN_IoT için sırasıyla 0,003, 0,014 ve 0,077 olarak raporlanmıştır [121]–[123]. Bu bulgular, genel sıralama başarısı yüksek kalsa bile sabit alarm eşiğinin alan kayması altında aynı çalışma noktasını korumadığını göstermektedir. Şekil 4.5'te XGBoost için IoT-23-A doğrulama bölümündeki negatif örneklerden FPR = %0,5, %1 ve %2 hedefleriyle seçilen sabit alarm eşiklerinin IoT-23-A test, IoT-23-B ve ToN_IoT kümelerindeki TPR karşılıkları gösterilmektedir. Yatay hata çubukları %95 bootstrap güven aralıklarını göstermektedir.

5. TARTIŞMA

Bu bölüm, bulguların tezde tanımlanan üç araştırma sorusu açısından ne anlama geldiğini tartışmaktadır. İlk soru, gözlem modeli kurulmadan doğrulanabilir bir performans alt sınırı üretmenin mümkün olup olmadığıdır. İkinci soru, bu alt sınırı zayıflatan belirsizlik bileşenlerinin sayısal olarak ayrıştırılıp ayrılamayacağıdır. Üçüncü soru ise bu yaklaşımın bütçe kısıtı altında uygulanabilir bir SOC karar mekanizmasına dönüştürülüp dönüştürülemeyeceğidir.

Ardışık savunma çalışması, gözlem hattı modellenmeden performans alt sınırı üretilebildiğini; ikili çözümün ise bu alt sınırı zayıflatan kanalları görünür kıldığını göstermiştir. UEBA karar katmanı çalışması da aynı çerçevenin, konformal kanıt ve risk kademeleri üzerinden bütçe kısıtlı operasyonel kararlara aktarılabildiğini ortaya koymuştur. Böylece tespit, değerlendirme ve karar adımları ortak bir belirsizlik dili altında birleştirilmiştir.

Alan kayması altında elde edilen sonuçlar, sabit eşiklerin farklı ortamlara doğrudan taşınmasının güvenilir olmadığını göstermektedir. Bu nedenle düzenli kalibrasyon bakımı, yöntemin saha geçerliliği açısından temel bir işletim gereksinimidir.

5.1. Literatürle ilişki ve değerlendirme tasarımı

Bu tez, olay köken izi tabanlı APT tespitinde temel sorunlardan birinin karar biriminin belirsizliği olduğunu göstermektedir. Değerlendirme birimi ile operasyonel çıktı birimi açık biçimde ayrılmadığında, aynı performans ölçütü farklı iş yükleri ve müdahale maliyetleriyle ilişkilendirilebilmektedir. Tez bu soruna, karar birimini açıkça tanımlayan ve aksiyonla ilişkilendirilebilir kanıt düzeyine dayalı bir raporlama çerçevesi önererek yanıt vermektedir.

Bu çerçevede birincil değerlendirme dili, belirsizlik altında performans alt sınırı ve sertifika uyumu gibi ölçütlere dayanmaktadır. Bu ölçütler, AUC ya da PR-AUC gibi klasik performans özetlerinin ötesine geçerek, karar kuralının kendi taahhüdünü ne ölçüde sağladığını da görünür kılmaktadır. Bu nedenle karar katmanı, tespit ile operasyonel müdahale arasındaki boşluğu dolduran gerekli bir ara katman olarak konumlanmaktadır.

Bu tezde DL ve DRL yöntemleriyle doğrudan kıyas yapılmamıştır. Bunun temel nedeni, bu tezde çözülen problemin yeni bir dedektör eğitmekten çok, üretilen skoru belirsizlik altında aksiyona dönüştürmek olmasıdır. DL ve DRL ile güvenilir bir kıyas yapabilmek için ya gerçekçi bir etkileşim simülasyonu kurmak ya da aynı olaya farklı aksiyonlar uygulanmış olsaydı ortaya çıkacak sonuçları ek varsayımlarla tahmin etmek gerekir. Bu tür bir kurgu, karar kuralını gözlemi üreten mekanizmaya ve simülasyon varsayımlarına daha sıkı bağlar. Bu ise tezin temel amacı olan gözlemden bağımsız ve doğrulanabilir bir performans alt sınırı üretme hedefiyle uyumlu değildir. Ayrıca gerçek SOC ortamında deneme-yanılma temelli öğrenme süreçleri operasyonel risk taşır. Bu nedenle DL ve DRL kıyası kapsam dışında bırakılmıştır.

5.2. Operasyonel çıkarımlar: ortam değişiminde kalibrasyon ve parametre yönetimi

Bu karar katmanında, yeni gelen APT için saldırıya özgü ayrı bir parametre vektörü öğrenilmez. Saha uygulamasında kullanılan başlıca ayarlar Q, ε, ω ve B'dir. Burada Q, p-değerinden risk kademesine geçişte kullanılan eşikler kümesini; ε, önsel belirsizliğini; ω, nominal fayda tablosu etrafındaki aralık belirsizliğinin yarı-genişliğini; B ise beklenen harcama düzeyini gösterir. ε ve ω belirsizlik parametreleridir; Q ve B ise karar katmanının işletim ayarlarıdır. Bu ayarlar fiziksel sabitler değil, duyarlılık analizleriyle desteklenen operasyonel çalışma noktalarıdır. Teorik bölümde ε ve ω taranır; pratik bölümde ise çalışma noktası olarak seçilir. Bu nedenle sayısal değerlerin farklı olması yöntemsel bir tutarsızlık değildir.

SOC ortamında uygun yaklaşım, dar bir belirsizlik kümesi ile başlayıp performans alt sınırı, sertifika uyumu, ortalama fayda ve zararsız akışta ağır müdahale oranını birlikte izlemektir. Bu göstergelerde kalıcı bozulma görüldüğünde ε ve gerekirse ω güncellenerek belirsizlik kümesi gerçek saha koşullarıyla daha uyumlu hâle getirilebilir. Bu güncelleme performans alt sınırını düşürebilir; ancak daha gerçekçi bir güvence sağlar.

Karar katmanında yeni olay önce zararsız referansa göre konformal p-değerine dönüştürülür, ardından önceden tanımlı risk kademelerinden birine atanır. Zararsız davranış dağılımı değiştikçe konformal kalibrasyonun yenilenmesi gerekir.

Her iki çalışmada da LP çözümünün milisaniye ölçeğinde elde edilmesi, bütçe veya fayda tablosu güncellendiğinde karar katmanının hızlı biçimde yeniden çözülmesini mümkün kılar. Bu pratiklik, yöntemin kademeli olarak devreye alınmasını ve saha içinde yönetilebilir olmasını destekler.

5.3. Sınırlılıklar ve geçerlilik sınırları

Bu tezde hedeflenen çıktı, en yüksek ortalama fayda değil, belirsizlik altında da korunabilen bir performans alt sınırıdır. Bu nedenle, yüksek kaliteli ve durağan ortamlarda gözleme daha sıkı bağlanan bazı yöntemler daha yüksek ortalama getiri sağlayabilir. Buna karşılık, APT gibi uzun süreli ve değişken tehditlerde asıl ihtiyaç kötü koşullarda da geçerliliğini koruyan bir taban güvencedir.

İkinci önemli sınırlılık, garantinin tanımlanan belirsizlik kümesiyle sınırlı olmasıdır. Saldırganın uyarlanabilir tepkisi, aksiyonların zamana yayılan yan etkileri ve ölçümlerin bağımlılık yapısı mevcut formülasyonda doğrudan modellenmemektedir. Bu nedenle garanti, ε ve ω ile tanımlanan belirsizlik kümesinin saha koşullarını yeterince temsil ettiği ölçüde anlamlıdır.

UEBA karar katmanının da kendine özgü sınırlılıkları vardır. Bu katman, loglardan şüphelilik skoru üreten tespit sistemini iyileştirmez; yalnızca bu skorları ortak bir risk diline çevirerek bütçe altında aksiyona bağlar. Yöntemin geçerliliği ayrıca zararsız kalibrasyon akışının bütünlüğüne bağlıdır. Bu akış bozulursa konformal p-değerleri ve risk kademeleri anlamını yitirebilir. Bu nedenle kalibrasyon verisinin korunması saha kurulumunun temel bir parçasıdır.

Son olarak, mevcut formülasyon durağandır. Zaman bağımlı müdahaleler, bütçe devri, analist geri bildirimi ve olay düzeyinde daha ayrıntılı etiketler kapsam dışındadır. Benzer biçimde, Tablo 3.2'deki fayda değerleri temsili olup saha ölçümlerine dayanmamaktadır. Ölçülmüş maliyetler ve daha temiz etiketlerle yapılacak doğrulamalar yöntemin dış geçerliliğini güçlendirecektir.

5.4. Gelecek çalışmalara yön veren başlıklar

Değerlendirme tarafında, olay ve vaka düzeyinde daha ayrıntılı referans etiketlere ihtiyaç vardır. Bu etiketlerin uzman incelemesiyle doğrulanması ve aksiyon maliyetlerinin standart biçimde kaydedilmesi, performans alt sınırı ile SOC iş yükünün daha sağlıklı ilişkilendirilmesini sağlayacaktır.

Karar katmanının dinamik bir yapıya genişletilmesi de önemlidir. Canlı sistemlerde bütçe tek bir karar anına bağlı kalmaz; vardiya, gün ve farklı kontrol noktaları arasında yönetilir.

Kullanılmayan kapasitenin devri ve analist geri bildiriminin modele katılması, yöntemi saha gerçekliğine yaklaştıracaktır.

Belirsizlik modelinin bağımsız aralık varsayımının ötesine taşınması bir diğer araştırma alanıdır. Köken izi graflarındaki nedensel ve topolojik yapı belirsizlik kümelerine dâhil edilebilirse, garanti saldırı yollarını daha doğrudan temsil edebilir ve gereksiz tutuculuk azaltılabilir.

Sahaya entegrasyon açısından ise ε, ω ve B parametrelerinin kademeli biçimde devreye alınması; fayda-zarar tablosunun saha ölçümleriyle güncellenmesi; konformal kalibrasyonun periyodik olarak yenilenmesi ve model çıktılarının olay düzeyindeki dış ölçütlerle doğrulanması önerilmektedir.

6. SONUÇ VE ÖNERİLER

6.1. Sonuçlar

Bu tez, gözlem hattı değiştiğinde dahi geçerliliğini koruyan bir performans alt sınırı üretmeyi ve bu alt sınırla uyumlu bütçe kısıtlı bir SOC karar mekanizması kurmayı amaçlamıştır. Bulgular, bu amacın hem kuramsal hem ampirik olarak karşılandığını göstermiştir.

Birinci çalışmada, önsel belirsizlik ile fayda/maliyet belirsizliğini birlikte ele alan tek bir belirsizliğe dayanıklı doğrusal program geliştirilmiştir. Bu program, savunucunun aksiyon dağılımını ve performans alt sınırını birlikte üretmektedir. Deneylerde, kayıt düşmesi, sinyal yer değiştirmesi ve birleşik bozulma senaryolarında gözlenen ortalamalar alt sınırın üzerinde kalmış; performans alt sınırının gözlem hattı değişimlerine rağmen taşınabildiği gösterilmiştir. Belirsizlik parametreleri arttıkça alt sınırın öngörülebilir biçimde azalması ve ikili çözümün kritik belirsizlik bileşenlerini göstermesi, parametrelerin nasıl ayarlanabileceğini de görünür kılmıştır. Sonlu tur sayısı için türetilen güven bantları da alt sınırın sonlu örneklem koşullarında raporlanmasını mümkün kılmıştır.

İkinci çalışmada, UEBA karar katmanı kapsamında dedektör skorları konformal kalibrasyon ile p-değerine dönüştürülmüş, risk kademelerine ayrılmış ve aksiyon dağılımları bütçe kısıtı altında tek bir doğrusal programla belirlenmiştir. $B = 0.90 \pm 0.01$ bütçe bandında belirsizliğe dayanıklı yaklaşım, risk-nötr yönteme kıyasla sertifika uyumunu artırmış, ihlal oranını azaltmış ve zararsız akışta ağır müdahale oranını düşürmüştür. Ayrıca karar katmanının milisaniye ölçeğinde çözülebilmesi, yöntemin sahada uygulanabilir olduğunu göstermiştir.

Literatür incelemesi ve alan kayması deneyleri de bu çerçeveyi desteklemiştir. İncelenen çalışmaların önemli bir bölümünde değerlendirme birimi ile uyarı biriminin hizalanmadığı görülmüş; sabit alarm eşiklerinin farklı alanlara taşındığında güvenilirliğini kaybettiği gösterilmiştir. Sonuç olarak önerilen çerçeve, gözlem hattına bağımlı olmayan doğrulanabilir bir performans tabanı sunmakta ve bu tabanı bütçe kısıtlı, denetlenebilir karar üretimiyle birleştirmektedir. Tablo 6.1 bu bulguların sayısal özetini sunmaktadır.

Tablo 6.1: Seçili göstergelerle bulgu özeti

Alan	Koşul/Parametre	Özet sonuç
Performans alt sınırı (v*)	$\omega = 0.05$, $\varepsilon = 0.00, 0.25, 0.50$	v*: 0.450 → 0.435 → 0.423 (ε arttıkça v* azalır)
Gözlem bozulması	kayıt düşmesi (miss), sinyal yer değiştirme (swap), birleşik (both)	v* değişmez, gözlenen ortalamalar alt sınırı ihlal etmez
UEBA karar katmanı	bütçe bandı $B = 0.90 \pm 0.01$	γ-uyum: 0.723 → 0.794; ihlal oranı: 0.277 → 0.206; zararsız davranışta ağır müdahale azalır; ortalama fayda büyük ölçüde korunur

6.2. Geleceğe yönelik öneriler

Bu tezden çıkan başlıca öneriler dört başlıkta toplanabilir. İlk olarak, ε, ω ve B parametreleri saha uygulamasında kademeli olarak devreye alınmalı ve performans alt sınırı, sertifika uyumu, ortalama fayda ile zararsız akışta ağır müdahale oranı birlikte izlenmelidir. İkinci olarak, fayda-zarar tablosu uzman yargısına ek olarak analist zamanı, müdahale maliyeti ve hizmet sürekliliği ölçümleriyle düzenli biçimde güncellenmelidir. Üçüncü olarak, konformal kalibrasyon kayan pencere yaklaşımıyla periyodik olarak yenilenmeli ve dağılım kayması göstergeleri ile izlenmelidir. Son olarak, karar katmanı bütçe devri, analist geri bildirimi ve çoklu kontrol noktalarını kapsayan dinamik bir yapıya genişletilmeli; belirsizlik kümelerine köken izi graflarının nedensel ve topolojik yapısı dâhil edilmelidir.

KAYNAKLAR

[1] Zipperle, M., Gottwalt, F., Chang, E. et al., 2022, Provenance-based intrusion detection systems: A survey, *ACM Computing Surveys*, 55(7), 1-36.

[2] Lv, Y., Qin, S., Zhu, Z. et al., 2022, A review of provenance graph based APT attack detection: Applications and developments, 7th IEEE International Conference on Data Science in Cyberspace (DSC), 498-505.

[3] Zhang, B., Gao, Y., Kuang, B. et al., 2025, A survey on advanced persistent threat detection: A unified framework, challenges, and countermeasures, *ACM Computing Surveys*, 57(3), 1-36.

[4] MITRE Corporation, 2025, Enterprise ATT&CK matrix, https://attack.mitre.org/matrices/enterprise/, [Ziyaret tarihi: 25 Eylül 2025].

[5] Wang, S., Wang, Z., Zhou, T. et al., 2022, ThreaTrace: Detecting and tracing host-based threats in node level through provenance graph learning, *IEEE Transactions on Information Forensics and Security*, Cilt 17, 3972-3987.

[6] Rehman, M. U., Ahmadi, H. and Hassan, W. U., 2024, FLASH: A comprehensive approach to intrusion detection via provenance graph representation learning, IEEE Symposium on Security and Privacy (SP), 3552-3570.

[7] Kapoor, M., Melton, J., Ridenhour, M. et al., 2021, PROV-GEM: Automated provenance analysis framework using graph embeddings, 20th IEEE International Conference on Machine Learning and Applications (ICMLA), 1720-1727.

[8] Chen, T., Dong, C., Lv, M. et al., 2022, APT-KGL: An intelligent APT detection system based on threat knowledge and heterogeneous provenance graph learning, *IEEE Transactions on Dependable and Secure Computing*, 1-15.

[9] Akbar, K. A., Wang, Y., Ayoade, G. et al., 2023, Advanced persistent threat detection using data provenance and metric learning, *IEEE Transactions on Dependable and Secure Computing*, 20(5), 3957-3969.

[10] Li, Z., Cheng, X., Sun, L. et al., 2021, A hierarchical approach for advanced persistent threat detection with attention-based graph neural networks, *Security and Communication Networks*, Cilt 2021, 9961342.

[11] Jia, Z., Wang, X., Xiong, Y. et al., 2022, AISLE: Self-supervised representation learning for the investigation of advanced persistent threat, 7th IEEE International Conference on Data Science in Cyberspace (DSC), 360-366.

[12] Yue, H., Li, T., Wu, D. et al., 2024, Detecting APT attacks using an attack intent-driven and sequence-based learning approach, *Computers & Security*, Cilt 140, 103748.

[13] Li, Y., Qiao, W., Zhu, Y. et al., 2025, DCASI: A sequence-based attack investigation method using DTW contrastive learning, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1-5.

[14] Benabderrahmane, S., Hoang, N., Valtchev, P. et al., 2024, Hack me if you can: Aggregating autoencoders for countering persistent access threats within highly imbalanced data, *Future Generation Computer Systems*, Cilt 160, 926-941.

[15] Benabderrahmane, S., Valtchev, P., Cheney, J. et al., 2025, APT-LLM: Embedding-based anomaly detection of cyber advanced persistent threats using large language models, 13th International Symposium on Digital Forensics and Security (ISDFS).

[16] Lu, S., Chi, B., Zhou, T. et al., 2025, STAE-APT: An APT detection method based on long-term behavioral features from provenance graphs, 4th International Symposium on Computer Applications and Information Technology (ISCAIT), 1834-1841.

[17] Han, X., Pasquier, T. F. J.-M., Bates, A. et al., 2020, UNICORN: Runtime provenance-based detector for advanced persistent threats, Network and Distributed System Security Symposium (NDSS).

[18] Xiong, C., Zhu, T., Dong, W. et al., 2022, CONAN: A practical real-time APT detection system with high accuracy and efficiency, *IEEE Transactions on Dependable and Secure Computing*, 19(1), 551-565.

[19] Zhu, T., Yu, J., Xiong, C. et al., 2023, APTSHIELD: A stable, efficient and real-time APT detection system for Linux hosts, *IEEE Transactions on Dependable and Secure Computing*, 20(6), 5247-5264.

[20] Welter, F., Wilkens, F. and Fischer, M., 2023, Tell me more: Black box explainability for APT detection on system provenance graphs, IEEE International Conference on Communications (ICC), 3817-3823.

[21] Liu, Q., Bao, K. and Hagenmeyer, V., 2025, Commander: A robust cross-machine multi-phase advanced persistent threat detector via provenance analytics, *Journal of Information Security and Applications*, Cilt 91, 104057.

[22] Wang, Y., Liu, H., Li, Z. et al., 2024, Combating advanced persistent threats: Challenges and solutions, *IEEE Network*, 38(6), 324-333.

[23] Cheng, Z., Lv, Q., Liang, J. et al., 2024, KAIROS: Practical intrusion detection and investigation using whole-system provenance, IEEE Symposium on Security and Privacy (SP), 3533-3551.

[24] Hossain, Md. N., Milajerdi, S. M., Wang, J. et al., 2017, SLEUTH: Real-time attack scenario reconstruction from COTS audit data, 26th USENIX Security Symposium (USENIX Security 17), 487-504.

[25] Jia, J., Yang, L., Zhou, L. et al., 2025, Embedding more knowledge: Strategic graph masking based advanced persistent threats detection, IEEE/ACM International Symposium on Quality of Service (IWQoS), 1-10.

[26] Yoon, S.-S., Shin, D.-H. and Euom, I.-C., 2025, DeepICS: Deep causal relationship modeling for multi-source log-based anomaly detection in industrial control systems, 55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume (DSN-S), 245-246.

[27] Mahmoud, M., Mannan, M. and Youssef, A., 2023, APT-Hunter: Detecting advanced persistent threats in early stages, *Digital Threats: Research and Practice*, 4(1), 11.

[28] Li, J., Li, T., Zhang, R. et al., 2023, APM: An attack path-based method for APT attack detection on few-shot learning, IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 10-19.

[29] Milajerdi, S. M., Eshete, B., Gjomemo, R. et al., 2019, POIROT: Aligning attack behavior with kernel audit records for cyber threat hunting, ACM SIGSAC Conference on Computer and Communications Security (CCS), 1813-1830.

[30] Cheng, Z., Dai, R., Wang, L. et al., 2023, GHunter: A fast subgraph matching method for threat hunting, 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 1014-1019.

[31] Aly, A., Iqbal, S., Youssef, A. et al., 2024, MEGR-APT: A memory-efficient APT hunting system based on attack representation learning, *IEEE Transactions on Information Forensics and Security*, Cilt 19, 5257-5271.

[32] Jiao, J., Min, S., Guo, D. et al., 2024, FPGAA: A multi-feature provenance graph for the accurate alert system, *IEEE Access*, Cilt 12, 149617-149632.

[33] Nazari, H., Yazdinejad, A., Dehghantanha, A. et al., 2024, P3GNN: A privacy-preserving provenance graph-based model for autonomous APT detection in software-defined networking, Workshop on Autonomous Cybersecurity, 34-44.

[34] Peng, Z., Shan, C. and Hu, C., 2024, An anomaly detection method based on meta-path and heterogeneous graph attention network, 5th International Conference on Computer Engineering and Application (ICCEA), 137-140.

[35] Peng, Z.-H., Hu, C.-Z. and Shan, C., 2025, Anomaly detection for advanced persistent threats with graph node embedding, *Journal of Information Science and Engineering*, 41(3), 713-728.

[36] Song, Q., Chen, T., Zhu, T. et al., 2025, APT detection via hypergraph attention network with community-based behavioral mining, *Applied Sciences*, 15(11), 5872.

[37] Hu, Z., Liu, L., Li, H. et al., 2025, CCLog: Actionable APT forensics via fused log semantics and provenance graph topology, *Computer Networks*, Cilt 272, 111660.

[38] Bi, P., Wang, Q., Xu, Y. et al., 2025, Pegasus: Accelerating provenance graph-based intrusion detection methods, 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2436-2441.

[39] Viana, C. M. C., Ferreira, C. H. G., Murai, F. et al., 2024, Devil in the noise: Detecting advanced persistent threats with backbone extraction, IEEE Symposium on Computers and Communications (ISCC).

[40] Li, J., Zhang, R., Liu, J. et al., 2022, LogKernel: A threat hunting approach based on behaviour provenance graph and graph kernel clustering, *Security and Communication Networks*, Cilt 2022, 4577141.

[41] Berrada, G., Cheney, J., Benabderrahmane, S. et al., 2020, A baseline for unsupervised advanced persistent threat detection in system-level provenance, *Future Generation Computer Systems*, Cilt 108, 401-413.

[42] Wang, L., Fang, L. and Hu, Y., 2025, A dynamic provenance graph-based detector for advanced persistent threats, *Expert Systems with Applications*, Cilt 265, 125877.

[43] Liu, J., Yan, J., Jiang, Z. et al., 2022, A graph learning approach with audit records for advanced attack investigation, IEEE Global Communications Conference (GLOBECOM), 897-902.

[44] Jiang, J., Chu, H. and Tian, D., 2026, A novel host-based intrusion detection approach leveraging audit logs, *Future Generation Computer Systems*, Cilt 174, 107995.

[45] Li, H., Yang, C., Zha, B. et al., 2024, A real-time APT attack detection scheme based on fusion provenance graph in private clouds, International Conference on Networking and Network Applications (NaNA), 490-495.

[46] Liu, H., An, B., Yin, Y. et al., 2024, A trust evaluation and concept drift-based approach for dynamic APT evasion detection, IEEE Cyber Science and Technology Congress (CyberSciTech), 544-547.

[47] Xu, F., Zhao, Q., Liu, X. et al., 2025, Advanced persistent threat detection via mining long-term features in provenance graphs, *Frontiers of Computer Science*, 19(10), 1910809.

[48] Cui, M., Jiang, Z., Yao, Y. et al., 2025, AGLHunter: Automated threat hunting using in-context learning-enhanced LLM, 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 1812-1819.

[49] Xu, L., Zhao, Z.-C., Zhao, D. et al., 2025, AJSAGE: A intrusion detection scheme based on jump-knowledge connection to GraphSAGE, *Computers & Security*, Cilt 150, 104263.

[50] Chen, M., Zhu, K., Liu, Z. et al., 2025, ATDetector: A thorough pipeline to detect diverse APT behaviors, 8th International Conference on Advanced Algorithms and Control Engineering (ICAACE), 2074-2087.

[51] Wang, Z., Wang, Y., Yan, W. et al., 2025, Autumn: An unsupervised APT detection via detailed process-level analysis, 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 600-605.

[52] Mei, R., Yan, H.-B., Han, Z.-H. et al., 2021, CTSCOPY: Hunting cyber threats within enterprise via provenance graph-based analysis, IEEE 21st International Conference on Software Quality, Reliability and Security (QRS), 28-39.

[53] Yan, N., Wen, Y., Chen, L. et al., 2022, DeepRo: Provenance-based APT campaigns detection via GNN, IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 747-758.

[54] Patil, R., Muneeswaran, S., Sachidananda, V. et al., 2023, E-Audit: Distinguishing and investigating suspicious events for APTs attack detection, *Journal of Systems Architecture*, Cilt 144, 102988.

[55] Shashikumar, D. R., Madhura, N., Reddy, S. L. et al., 2025, Efficient host threat detection using the ProcSAGE method, International Conference on Emerging Smart Computing and Informatics (ESCI).

[56] Wu, L., Xie, Y.-L., Zhao, S.-X. et al., 2025, Efficient intrusion detection via heterogeneous graph attention networks and parallel provenance analysis, *Computer Networks*, Cilt 270, 111552.

[57] Milajerdi, S. M., Gjomemo, R., Eshete, B. et al., 2019, HOLMES: Real-time APT detection through correlation of suspicious information flows, IEEE Symposium on Security and Privacy (SP), 1137-1152.

[58] Wang, L., Shen, X., Li, W. et al., 2025, Incorporating gradients to rules: Towards lightweight, adaptive provenance-based intrusion detection, Network and Distributed System Security Symposium (NDSS).

[59] Li, H.-W., Liu, P.-T., Lin, B.-W. et al., 2024, IPMES: A tool for incremental TTP detection over the system audit event stream, 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 265-266.

[60] Yu, Z., Zheng, B., Ma, X. et al., 2025, KanIDS: An intrusion detection system for all kernel interactions based on Kolmogorov-Arnold network, 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 1307-1312.

[61] Saha, B., Rani, N. and Shukla, S. K., 2024, MAD: A meta-learning approach to detect advanced persistent threats using provenance data in industrial IoT, Annual Computer Security Applications Conference Workshops (ACSAC Workshops), 201-207.

[62] Jia, Z., Xiong, Y., Nan, Y. et al., 2024, MAGIC: Detecting advanced persistent threats via masked graph representation learning, 33rd USENIX Security Symposium (USENIX Security 24), 5197-5214.

[63] Wang, Z., Du, D., Qi, Y. et al., 2025, MPKAN: APT attack detection on audit logs via graph semantic enhancement, 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2605-2610.

[64] Zhu, T., Ying, J., Chen, T. et al., 2025, Nip in the bud: Forecasting and interpreting post-exploitation attacks in real-time through cyber threat intelligence reports, *IEEE Transactions on Dependable and Secure Computing*, Cilt 22, 1431-1447.

[65] Li, S., Dong, F., Xiao, X. et al., 2024, NODLINK: An online system for fine-grained APT attack detection and investigation, Network and Distributed System Security Symposium (NDSS).

[66] Jiang, B., Bilot, T., El Madhoun, N. et al., 2025, ORTHRUS: Achieving high quality of attribution in provenance-based intrusion detection systems, 34th USENIX Security Symposium (USENIX Security 25), 7173-7192.

[67] Yan, W., Wu, W., Bi, B. et al., 2025, PanThreat: Global resource-based anomaly detection for APTs, 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 135-140.

[68] Wang, Z., Li, H., Qi, Y. et al., 2025, PathWatcher: A path-based behavior detection method for attack detection and investigation, *Computers & Security*, Cilt 157, 104563.

[69] Zhou, F., Chang, B., Wen, Y. et al., 2023, ProDE: Interpretable APT detection method based on encoder-decoder architecture, IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS), 340-347.

[70] Wang, T., Tang, W., Su, Y. et al., 2025, A provenance graph-based deep learning framework for advanced persistent threat detection in edge computing, *Applied Sciences*, 15(16), 8833.

[71] Ren, J. and Geng, R., 2025, Provenance-based APT campaigns detection via masked graph representation learning, *Computers & Security*, Cilt 148, 104159.

[72] Huang, Z. and Wang, P., 2025, RAS-GNN: Reconstructing APT attack scenario using graph neural network, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1-5.

[73] Zhou, F., Chang, B., Wen, Y. et al., 2023, Representation-enhanced APT detection using contrastive learning, IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 1-9.

[74] Weng, Z., Zhang, W., Zhu, T. et al., 2025, RT-APT: A real-time APT anomaly detection method for large-scale provenance graph, *Journal of Network and Computer Applications*, Cilt 233, 104036.

[75] Shen, F., Perigo, L. and Curry, J. H., 2023, SR2APT: A detection and strategic alert response model against multistage APT attacks, *Security and Communication Networks*, Cilt 2023, 6802359.

[76] Sun, J., Dai, H., Sun, B. et al., 2025, Research on APT detection model based on graph attention network and auto-encoder, *Engineering Letters*, 33(7), 2634-2642.

[77] Jiang, K., Gao, Z., Zhang, S. et al., 2025, Sylph: An unsupervised APT detection system based on the provenance graph, *Information*, 16(7), 566.

[78] Liu, H., Wang, Y., Su, Z. et al., 2024, TRACEGADGET: Detecting and tracing network level attack through federal provenance graph, IEEE International Conference on Communications (ICC), 2713-2718.

[79] Lv, M., Gao, H., Qiu, X. et al., 2024, TREC: APT tactic/technique recognition via few-shot provenance subgraph learning, ACM SIGSAC Conference on Computer and Communications Security (CCS).

[80] Yuce, M., Ertürk, M. A. and Aydın, M. A., 2024, Misbehavior detection with spatio-temporal graph neural networks, *Computers and Electrical Engineering*, Cilt 116, 109198.

[81] Yuce, M., Ertürk, M. A. and Aydın, M. A., 2025, Misbehavior detection with collective perception in V2X networks: A survey, *Transactions on Emerging Telecommunications Technologies*, 36(10), e70267.

[82] Huber, P. J., 1964, Robust estimation of a location parameter, *Annals of Mathematical Statistics*, 35(1), 73-101.

[83] Ben-Tal, A. and Nemirovski, A., 1998, Robust convex optimization, *Mathematics of Operations Research*, 23(4), 769-805.

[84] Delage, E. and Ye, Y., 2010, Distributionally robust optimization under moment uncertainty with application to data-driven problems, *Operations Research*, 58(3), 595-612.

[85] Namkoong, H. and Duchi, J. C., 2016, Learning models with distributionally robust optimization, arXiv:1610.01057.

[86] Mohajerin Esfahani, P. and Kuhn, D., 2018, Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations, *Mathematical Programming*, 171(1-2), 115-166.

[87] Bertsimas, D., Gupta, V. and Kallus, N., 2018, Data-driven robust optimization, *Mathematical Programming*, 167(2), 235-292.

[88] Rahimian, H. and Mehrotra, S., 2019, Distributionally robust optimization: A review, arXiv:1908.05659.

[89] Iyengar, G. N., 2005, Robust dynamic programming, *Mathematics of Operations Research*, 30(2), 257-280.

[90] Nilim, A. and El Ghaoui, L., 2005, Robust control of Markov decision processes with uncertain transition matrices, *Operations Research*, 53(5), 780-798.

[91] Wiesemann, W., Kuhn, D. and Rustem, B., 2013, Robust Markov decision processes, *Mathematics of Operations Research*, 38(1), 153-183.

[92] Xu, H. and Mannor, S., 2010, Distributionally robust Markov decision processes, Advances in Neural Information Processing Systems (NeurIPS).

[93] Auer, P. and Cesa-Bianchi, N., 2002, The nonstochastic multiarmed bandit problem, *SIAM Journal on Computing*, 32(1), 48-77.

[94] Tambe, M., 2011, *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*, Cambridge University Press, Cambridge.

[95] Pita, J., Jain, M., Marecki, J. et al., 2008, ARMOR: A deployed system for airport security, Twenty-Third AAAI Conference on Artificial Intelligence (IAAI).

[96] Jain, M., Brown, M., Tambe, M. et al., 2012, PROTECT: A deployed game theoretic system to protect the ports of the United States, *AI Magazine*, 33(4), 96-110.

[97] Pita, J., Jain, M., Tambe, M. et al., 2009, Robust solutions to Stackelberg games: Addressing human adversaries, 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS).

[98] Kiekintveld, C., Jain, M., Tsai, J. et al., 2009, Computing optimal randomized resource allocations for security, 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS).

[99] Fang, F., Stone, P. and Tambe, M., 2015, When security games go green: Designing conservation strategies via PAWS, International Conference on Autonomous Agents and Multiagent Systems (AAMAS).

[100] Freedman, D. A., 1975, On tail probabilities for martingales, *The Annals of Probability*, 3(1), 100-118.

[101] Audibert, J.-Y. and Munos, R., 2009, Exploration-exploitation tradeoff using variance estimates: Theoretically efficient algorithms, 22nd Conference on Learning Theory (COLT).

[102] Maurer, A. and Pontil, M., 2009, Empirical Bernstein bounds and sample variance penalization, 22nd Conference on Learning Theory (COLT).

[103] Huang, L., Joseph, A. D., Nelson, B. et al., 2011, Adversarial machine learning, 4th ACM Workshop on Security and Artificial Intelligence (AISec).

[104] Biggio, B., Fumera, G. and Roli, F., 2014, Security evaluation of pattern classifiers under attack, *IEEE Transactions on Knowledge and Data Engineering*, 26(4), 984-996.

[105] Tavallaee, M., Bagheri, E., Lu, W. et al., 2009, A detailed analysis of the KDD Cup 99, IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA).

[106] Moustafa, N. and Slay, J., 2015, UNSW-NB15: A comprehensive data set for network intrusion detection systems, Military Communications and Information Systems Conference (MilCIS).

[107] Jajodia, S., Ghosh, A. K., Swarup, V., Wang, C. and Wang, X. S. (ed.), 2011, *Moving Target Defense: Creating Asymmetric Uncertainty for Cyber Threats*, Springer.

[108] Wiesemann, W., Kuhn, D. and Sim, M., 2014, Distributionally robust convex optimization, *Operations Research*, 62(6), 1358-1376.

[109] Vovk, V., Gammerman, A. and Shafer, G., 2005, *Algorithmic Learning in a Random World*, Springer.

[110] Shafer, G. and Vovk, V., 2008, A tutorial on conformal prediction, *Journal of Machine Learning Research*, Cilt 9, 371-421.

[111] Rockafellar, R. T. and Uryasev, S., 2000, Optimization of conditional value-at-risk, *Journal of Risk*, 2(3), 21-41.

[112] Arantes, R., Weir, C., Hannon, H. and Kulseng, M., 2021, Operationally Transparent Cyber (OpTC) – NCR Release, DARPA veri seti yayını.

[113] Puterman, M. L., 1994, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons.

[114] Kaelbling, L. P., Littman, M. L. and Cassandra, A. R., 1998, Planning and acting in partially observable stochastic domains, *Artificial Intelligence*, 101(1-2), 99-134.

[115] Sutton, R. S. and Barto, A. G., 2018, *Reinforcement Learning: An Introduction*, 2. bs., MIT Press.

[116] Basel Committee on Banking Supervision, 2019, *Explanatory Note on the Minimum Capital Requirements for Market Risk*, Bank for International Settlements.

[117] Howard, S. R., Ramdas, A., McAuliffe, J. and Sekhon, J., 2021, Time-uniform, nonparametric, nonasymptotic confidence sequences, *The Annals of Statistics*, 49(2), 1055-1080.

[118] Axelsson, S., 2000, The base-rate fallacy and the difficulty of intrusion detection, *ACM Transactions on Information and System Security*, 3(3), 186-205.

[119] Fawcett, T., 2006, An introduction to ROC analysis, *Pattern Recognition Letters*, 27(8), 861-874.

[120] Davis, J. and Goadrich, M., 2006, The relationship between precision-recall and ROC curves, International Conference on Machine Learning (ICML), 233-240.

[121] Guo, C., Pleiss, G., Sun, Y. and Weinberger, K. Q., 2017, On calibration of modern neural networks, International Conference on Machine Learning (ICML), 1321-1330.

[122] Platt, J., 1999, Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, *Advances in Large Margin Classifiers*, MIT Press, 61-74.

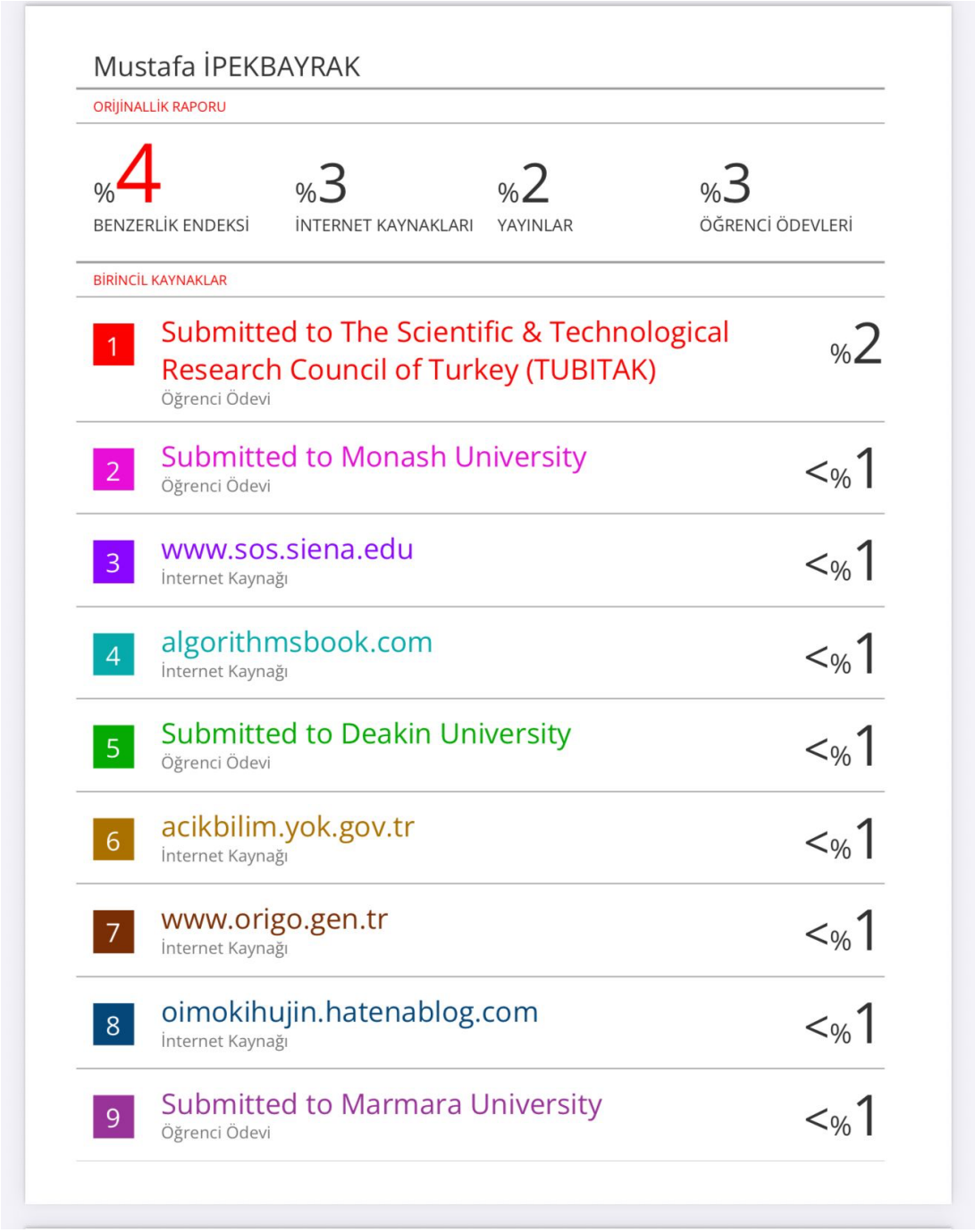
[123] Zadrozny, B. and Elkan, C., 2002, Transforming classifier scores into accurate multiclass probability estimates, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 694-699.

[124] Garcia, S., Parmisano, A. and Erquiaga, M. J., 2020, IoT-23: A labeled dataset with malicious and benign IoT network traffic (version 1.0.0), Zenodo.

[125] Booij, T. M., Chiscop, I., Meeuwissen, E., Moustafa, N. and den Hartog, F. T. H., 2022, ToN_IoT: The role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets, *IEEE Internet of Things Journal*, 9(1), 485-496.

[126] Ovadia, Y. et al., 2019, Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift, Advances in Neural Information Processing Systems (NeurIPS), 13991-14002

İNTİHAL RAPORU İLK SAYFASI



Mustafa İPEKBAYRAK

ORİJİNALLİK RAPORU

%4	%3	%2	%3
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	Submitted to The Scientific & Technological Research Council of Turkey (TUBITAK) Öğrenci Ödevi	%2
2	Submitted to Monash University Öğrenci Ödevi	<%1
3	www.sos.siena.edu İnternet Kaynağı	<%1
4	algorithmsbook.com İnternet Kaynağı	<%1
5	Submitted to Deakin University Öğrenci Ödevi	<%1
6	acikbilim.yok.gov.tr İnternet Kaynağı	<%1
7	www.origo.gen.tr İnternet Kaynağı	<%1
8	oimokihujin.hatenablog.com İnternet Kaynağı	<%1
9	Submitted to Marmara University Öğrenci Ödevi	<%1