# 140509\_42.md — AI-Powered Threat Detection and Response System

**Theme:** AI for CyberSecurity & CyberSecurity for AI  
**Mission:** Detect and respond to advanced threats (including zero-day) across network, endpoint, and cloud using AI analytics and automated playbooks, while minimizing false positives.

## README (Problem Statement)

**Summary:** Develop an intelligent cybersecurity platform that uses ML to detect advanced threats, analyze attack patterns, and automate incident response.  
**Problem Statement:** Modern threats are sophisticated and evolve rapidly. Build a platform that analyzes network traffic, logs, and user behavior to detect known and zero-day attacks, provide intelligence, and automate responses with minimal false positives.

**Steps:**  
- Multi-source data ingestion (network, endpoint, cloud)  
- Anomaly detection for unusual behaviors  
- Threat classification & severity scoring  
- Automated incident response workflows  
- Threat intel integration & pattern analysis  
- Forensic analysis & automated documentation

**Suggested Data:** NetFlow, PCAP, endpoint logs, cloud audit trails, attack signatures, threat intel feeds, baselines of user activity, response playbooks.

## 1) Vision, Scope, KPIs

**Vision:** An AI-SOC assistant that reduces Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR) by >60%.

**Scope:**  
- v1: ingest logs, detect anomalies + IOCs, threat scoring, manual SOAR playbooks.  
- v2: add UEBA, graph-based lateral movement detection, predictive analytics.  
- v3: cross-tenant intelligence sharing, federated detection.

**KPIs:**  
- **MTTD median < 15 min**  
- **MTTR reduced by 60%**  
- **False Positive Rate <3% @ Recall ≥95%** for critical threats  
- **≥70% of commodity threats auto-contained**

## 2) Personas & User Stories

* **SOC Analyst L1:** Wants prioritized, contextual alerts.
* **SOC Analyst L2:** Needs deep forensic drill-downs.
* **Incident Responder:** Wants 1-click containment (quarantine hosts, disable accounts).
* **SecOps Engineer:** Needs integrations with EDR, SIEM, and ticketing tools.
* **CISO:** Needs executive dashboards on trends, risk posture, SLA compliance.

**User Stories:**  
- US-01: “As an L1, I want anomalies scored & ranked with explanations.”  
- US-07: “As a responder, I want auto-playbooks triggered with approvals.”  
- US-12: “As a CISO, I want weekly summaries of top attack tactics.”

## 3) PRD

**Capabilities:**  
1. **Ingestion:** logs & traffic (firewalls, NetFlow/PCAP, EDR/AV, IAM, cloud).  
2. **Detection:** anomaly + UEBA + sequence modeling.  
3. **Classification/Scoring:** ensemble (anomaly + signature + threat intel).  
4. **Response:** automated playbooks with human oversight.  
5. **Threat Intel:** integrate STIX/TAXII feeds; map to MITRE ATT&CK.  
6. **Forensics:** package evidence, timeline generation, auto-docs.

## 4) FRD

* **ETL:** Normalize to common schema {timestamp, src, dst, user, action, confidence}
* **UEBA:** z-score deviation; peer group analysis; impossible travel.
* **Anomaly Models:** autoencoders, isolation forests.
* **Sequence/Graph:** transformers for event sequences; Neo4j for lateral movement.
* **Signature Match:** Suricata/YARA rules.
* **Severity Scoring:** ensemble calibration → High/Med/Low.
* **Playbooks:** YAML-defined (quarantine, disable, notify).
* **Intel:** dedup, enrich with TTPs, IOC to ATT&CK mapping.
* **Forensic Store:** immutable, WORM compliant.

## 5) NFRD

* **Scale:** 100k events per second.
* **Latency:** End-to-end < 2 s P95.
* **Reliability:** 99.95% uptime.
* **Security:** FIPS-compliant crypto, RBAC, PII masking.
* **Auditability:** Full chain-of-custody for evidence.

## 6) Architecture (Logical)

[Sensors: NetFlow, EDR, Cloud]  
 -> [Collectors] -> [Kafka Bus]  
 -> [Stream ETL] -> [Feature Store]  
 |  
 ------------------------------------------------  
 | | |  
 [Anomaly Detectors] [UEBA Models] [Signature Engines]  
 | | |  
 ------------------------------------------------  
 -> [Threat Scoring Engine]  
 |  
 [SOAR / Playbooks]  
 |  
 [Case Mgmt + Forensics DB]  
 |  
 [Dashboards + Threat Intel]

## 7) HLD

* **Ingestion:** Kafka + Flink ETL.
* **Detection Engines:** Microservices (PyTorch models, Suricata, rule engines).
* **UEBA:** features on top of feature store (Redis/Feast).
* **SOAR:** automation engine (playbooks as YAML/JSON).
* **Case Mgmt:** Elastic + Kibana dashboards.

## 8) LLD Examples

**UEBA:**  
- Profile mean/variance per user (logins/hour).  
- Alert if z-score > 3 or “impossible travel” (geo/time delta).

**Sequence Model:**  
- Input: event sequences per host/user.  
- Model: transformer w/ masked prediction.  
- Output: probability of malicious tactic.

**Playbook (YAML):**

playbook: quarantine\_host  
trigger: {severity: High, entity: host}  
steps:  
 - isolate\_network: {agent: edr, host\_id: $host}  
 - disable\_account: {idp: okta, user: $user}  
 - notify: {channel: soc-alerts, msg: "Host $host quarantined"}  
 - log\_case: {case\_id: $id}

## 9) Pseudocode

for event in ingest\_stream:  
 features = featurize(event)  
 scores = [detector.predict(features) for detector in detectors]  
 severity = calibrate(scores, check\_intel(event))  
 if severity >= threshold:  
 case = create\_case(event, severity)  
 if auto\_allowed(case): execute\_playbook(case)  
 else: alert(case)

## 10) Data & Evaluation

* **Data:** CIC-IDS, UNSW-NB15, CTU-13, KDD-Cup, red-team simulations.
* **Metrics:** ROC/PR AUC, precision/recall, MTTD, MTTR, alert fatigue reduction.
* **Eval Strategy:** offline training on labeled attacks + online shadow deployment.

## 11) Security & Governance

* Logs hashed & signed; WORM forensic store.
* RBAC with least privilege.
* PII anonymization before model ingestion.
* Compliance with ISO 27001, NIST 800-53, GDPR.

## 12) Observability & Cost

* Metrics: EPS, precision/recall, playbook SLA, containment %
* Tracing: OpenTelemetry from collector → case mgmt.
* Cost: GPU reserved only for heavy models, autoscale during surges.

## 13) Roadmap

* **M1 (4w):** Ingest + baseline detectors + manual SOAR.
* **M2 (8w):** Add UEBA + auto-playbooks + threat intel integration.
* **M3 (12w):** Sequence/graph models + full forensic suite.
* **M4 (16w):** Predictive analytics + cross-tenant correlation.

## 14) Risks & Mitigations

* **Model drift:** retrain with online feedback.
* **Alert overload:** calibrate + adaptive thresholds.
* **Automation loops:** require approvals for High-severity.
* **Adversarial evasion:** ensemble detectors + honeypot traps.