Technical Document Template

Version 2.3.1

Overleaf

May 11, 2025

Abstract

Small and medium-sized enterprises (SMEs) face significant challenges in monitoring workforce performance, optimizing team collaboration, and mitigating insider cybersecurity threats due to limited resources and the lack of integrated tools. This document introduces SynapseAI, a novel AI-powered analytics platform designed to address these challenges. SynapseAI integrates with common SME collaboration and development tools to extract and analyze activity metadata. It employs Markov chain modeling for performance and task flow intelligence, graph-based analysis for collaboration mapping, and anomaly detection techniques along with NLP on metadata for cybersecurity behavior monitoring. The platform provides actionable insights through intuitive dashboards and alerts, aiming to enhance productivity, improve team dynamics, and bolster security posture while respecting employee privacy and ensuring compliance. We detail the architecture of SynapseAI, its core analytical modules, and discuss its potential impact on SME operational efficiency and security.

1 Introduction

Small and medium-sized enterprises (SMEs) constitute the backbone of many economies, yet they often struggle with operational intelligence due to resource constraints. According to Gartner, SMEs face particular challenges in three key areas: performance blind spots from outdated feedback mechanisms, inefficiencies in team collaboration (especially in remote/hybrid settings), and increasing vulnerability to insider cybersecurity threats as noted by the Ponemon Institute.

Existing solutions often address these challenges in isolation, creating silos between performance monitoring and security analytics. Moreover, many tools are designed for large enterprises with dedicated IT teams, making them too complex for SME adoption. This paper presents SynapseAI, an integrated AI-driven platform specifically tailored for SMEs that unifies workforce performance analytics with cybersecurity monitoring.

Our primary contributions include:

- The design of SynapseAI's architecture that integrates multiple data sources while preserving privacy
- Novel application of Markov chains for task flow analysis in SME contexts

- A graph-based approach to collaboration mapping using metadata from common SME tools
- An integrated anomaly detection system for cybersecurity monitoring

2 Related Work

Existing workforce analytics tools like ActivTrak and Microsoft Viva provide productivity insights but lack integration with security monitoring. Conversely, cybersecurity solutions for SMEs typically focus on signature-based threat detection rather than behavior analysis. The gap in unified platforms that combine performance, collaboration, and security analytics in an SME-friendly manner remains largely unaddressed.

3 Architecture and Methodology

3.1 Overall System Architecture

SynapseAI's architecture consists of four layers (Fig. 1):

1. Data Collection Layer with API connectors
2. Data Normalization Layer for unified processing
3. AI Analytics Core with performance, collaboration, and security modules
4. Insight Generation Layer with dashboards and alerts

3.2 Data Collection and Integration

SynapseAI integrates with popular SME tools including Slack, Jira, GitHub, and Microsoft Teams. Table 1 shows the metadata extracted from each source.

Table 1: Data Sources and Key Metadata Types

Tool	**Metadata Collected**
Slack	Timestamps, channel activity, reaction counts
Jira	Ticket transitions, assignee changes, resolution times
GitHub	Timestamps, pull requests, code review, comments
Microsoft Teams	Time stamp, meeting durations, participant lists

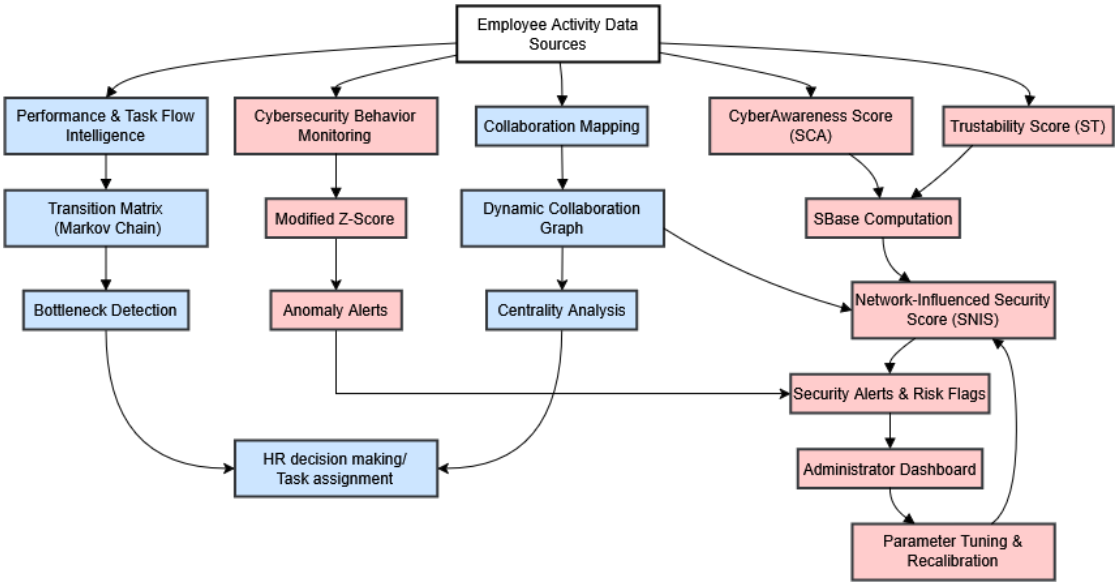

Figure 1: SynapseAI System Architecture

3.3 AI-Powered Analytical Modules

3.3.1 Performance & Task Flow Intelligence

We model task workflows as Markov chains where states represent task stages and transitions capture handoff probabilities. The transition matrix P is defined as:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{1}$$

where p_{ij} represents the probability of transitioning from state i to j. Bottlenecks are identified as states with high incoming but low outgoing transition probabilities.

The transition matrix P reveals hidden bottlenecks through:

$$\text{BottleneckScore}_i = 1 - \sum_{j=1}^{n} P_{ij} \cdot \mathbb{I}_{\{P_{ij} > \theta\}} \tag{2}$$

where $\theta = 0.1$ is the minimum significant transition probability. This score highlights nodes with weak or few strong outgoing transitions, flagging them as possible bottlenecks.

3.3.2 Collaboration Mapping

Using graph theory, we construct dynamic collaboration networks where nodes represent team members and edges represent interaction frequencies. Key metrics include:

$$\text{Centrality}(v) = \frac{1}{\sum_{u \neq v} d(u,v)} \tag{3}$$

where $d(u,v)$ is the shortest path between nodes u and v, it indicates the cohesiveness of the team and how strongly connected they are.

$$\textbf{Community}(Q) = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \tag{4}$$

where

A_{ij} : Adjacency matrix entry; 1 if edge between i and j, else 0

k_i, k_j : Degrees of nodes i and j, respectively

m : Total number of edges in the graph

$\delta(c_i, c_j)$: 1 if nodes i and j are in the same community, else 0

$\frac{k_i k_j}{2m}$: Expected number of edges between i and j in a random graph

Community measures the ratio between the actual graph connections and expected connections of a random graph, if community is high, then it reflects strong bonds. This can aid in finding subcommunities with excellent team dynamics that form organically

$$\textbf{Bridge}(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{5}$$

v : The node (person)

σ_{st} : Total number of shortest paths from node s to node t

$\sigma_{st}(v)$: Number of those paths that pass through node v

$\sum_{s \neq v \neq t}$: Sum over all distinct pairs (s,t) excluding v

This is essentially betweenness centrality, which helps identify bridge players that hold together the team structure and are key for successful collaboration.

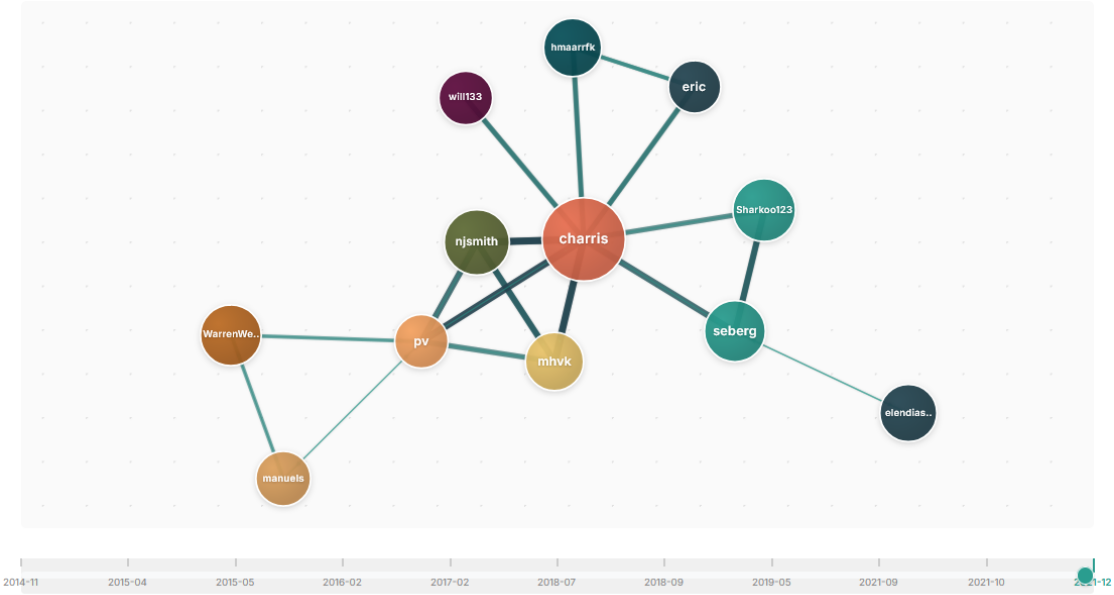

Figure 2: Sample Collaboration Graph

3.3.3 Cybersecurity Behavior Monitoring

Anomaly detection uses a modified z-score approach:

$$z = \frac{x - \mu}{\sigma} \tag{6}$$

where x is the observed value, μ is the 30-day rolling mean, and σ is the standard deviation. This approach helps identify statistically unusual behavior patterns indicative of potential insider threats, activities with $|z| > 3$ will trigger alerts.

4 Integration of Cybersecurity KPIs

To effectively integrate cybersecurity into the employee network analysis, we propose a set of Key Performance Indicators (KPIs). These KPIs are designed to be quantifiable and provide actionable insights for administrators regarding the

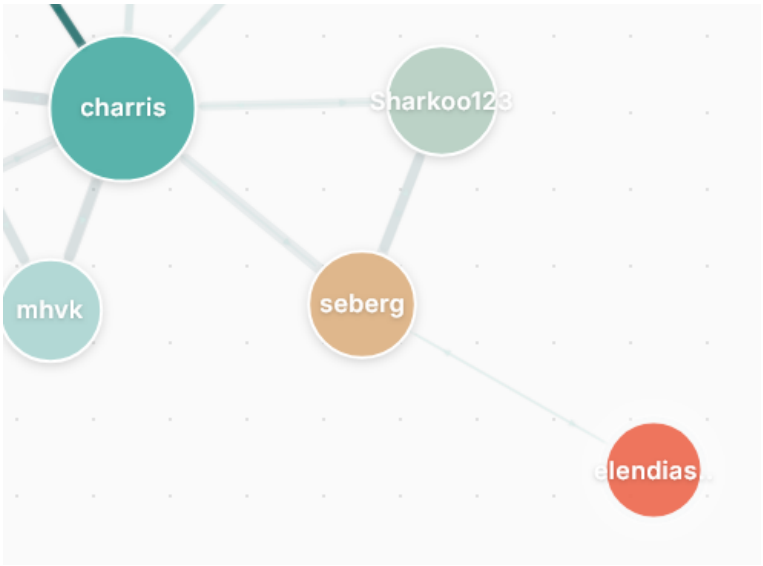

Figure 3: Anomaly Detection Example (A node exibiting unusual isolation from the community)

organization's security posture at an individual and network level. An overview of these core metrics is presented in Table 2.

Table 2: Summary of Core Cybersecurity KPIs

Metric (Symbol)	**Name**	**Significance/Purpose**
S_{CA}	CyberAwareness Score	Measures an employee's understanding of and adherence to cybersecurity best practices.
S_T	Trustability Score	Reflects the potential risk associated with an employee's access and behavior, considering factors like access privileges and policy adherence.
S_{NIS}	Network-Influenced Security Score	Leverages the graph structure to show how an employee's security is affected by their connections, highlighting systemic vulnerabilities.

4.1 CyberAwareness Score (S_{CA})

The CyberAwareness Score quantifies an employee's understanding and adherence to cybersecurity best practices.

4.1.1 Data Sources for S_{CA}

Table 3: Data Sources for CyberAwareness Score

Data Source	**Description**
Phishing Simulation	Tracks click-through rates, reporting rates, and credential entry attempts on simulated phishing emails.
Cybersecurity Training & Scores	Records completion of mandatory training modules and scores achieved on associated quizzes.
Incident Reporting (Positive)	Measures frequency of *valid* reports of suspicious emails/activities, indicating proactive awareness.
Knowledge-Based Quizzes/Surveys	Periodically assesses understanding of company security policies and best practices.

4.1.2 Quantification of S_{CA}

Table 4: Quantification Scheme for CyberAwareness Score

Event/Metric	**Points (Example Values)**
Phishing Simulation (No Clicks/Reported Validly)	$+X_{NCR}$
Phishing Simulation (Clicked, No Data Entered)	$+X_{CND}$
Phishing Simulation (Clicked, Data Entered)	$-X_{CDE}$
Training Module Completed (per module)	$+A_{TM}$
Quiz Score $> 80\%$	$+B_{QS}$
Valid Security Incident Reported	$+C_{VIR}$

The final CyberAwareness Score for employee i is calculated as:

$$S_{CA_i} = \sum_j w_j \cdot \text{metric}_{ij}$$

where w_j is the weight for metric j, and metric_{ij} is employee i's value for metric j.

4.2 Trustability Score (S_T)

This KPI reflects potential risk associated with an employee's access and behavior.

4.2.1 Data Sources for S_T

Table 5: Data Sources for Trustability Score

Data Source	Description
Access Level & Privileges	Sensitivity of accessible data/systems; higher access implies greater potential impact if compromised.
Security Policy Adherence	Documented instances of adherence/non-adherence to critical security policies (clean desk, password management, etc.).
Incident History (Attributable)	Past security incidents directly attributable to employee actions.
MFA Adoption & Use	Tracks adoption and consistent use of multi-factor authentication.
Data Handling Practices	Observed practices regarding sharing/storage of sensitive information.

4.2.2 Quantification of S_T

- Risk tiers associated with access privileges
- Policy adherence scoring (e.g., $+X_{MFA}$ for MFA usage, $-Y_{PV}$ for violations)
- Temporal decay factor for historical incidents:

$$I_k \cdot e^{-\lambda t_k}$$

  where I_k is initial impact, t_k is time since incident k, and $\lambda > 0$ is decay rate.

4.3 Network-Influenced Security Score (S_{NIS})

This score reflects how an employee's security posture is influenced by network connections.

4.3.1 Base Score (S_{Base})

$$S_{\text{Base}_i} = f(S_{CA_i}, S_{T_i})$$

or derived from endpoint security status.

4.3.2 Influence Mechanism

- **Contagion Model**: Node's S_{NIS} negatively impacted by low scores of direct connections
- **Vulnerability Propagation**: Highly connected nodes with low scores pose disproportionate risk

4.3.3 Calculation of S_{NIS}

Let $N(i)$ be neighbors of node i, S_{Th} the security threshold:

$$S_{NIS_i} = S_{\text{Base}_i} - \left(\frac{\sum_{j \in N(i)} \max(0, S_{Th} - S_{\text{Base}_j})}{|N(i)| + \epsilon} \cdot I_F \right)$$

where:

- $\epsilon = 10^{-6}$ prevents division by zero
- $I_F \in [0, 1]$ is influence factor

4.4 Parameter Tuning and Empirical Validation

Key parameters requiring calibration:

- CyberAwareness point values (X_{NCR}, X_{CDE}, etc.)
- Trustability weights (X_{MFA}, Y_{PV}) and decay rate (λ)
- Network parameters (S_{Th}, I_F)

Values must be determined through empirical experimentation and continuous monitoring.

5 Implementation Considerations

The prototype implements the data pipeline using GitHub issues dataset as a baseline and NetworkX for graph analysis,the phishing simulation using Local LLM and local mailing server MailHog and demonstrates dynamic score propagation, The frontend uses D3.js for visualizations and highlighting problematic nodes and anomalous graph sections.

6 Future Work

Future work will focus on expanding the system to accommodate remote work dynamics and implementing additional graph neural network-based predictive capabilities. We also plan to integrate explainable layers to those models for enhanced decision making confidence.

7 Privacy and Ethical Considerations

SynapseAI employs several privacy-preserving techniques:

- Analysis limited to metadata (no content inspection)
- Granular access controls
- Compliance with GDPR and CCPA through data minimization

8 Conclusion

SynapseAI demonstrates the feasibility and value of integrating workforce performance and cybersecurity analytics for SMEs. Our approach shows particular promise in identifying workflow inefficiencies while simultaneously monitoring for security threats. Future work will expand tool integrations and develop predictive capabilities.

References

[1] Gartner, "SME Workforce Analytics Trends," 2022.

[2] Ponemon Institute, "2022 Insider Threat Report," 2022.

[3] A. Author, "Markov Models for Workflow Analysis," Journal of Process Improvement, 2020.

[4] B. Researcher, "Graph-Based Collaboration Metrics," ACM CSCW, 2021.

[5] C. Security, "Anomaly Detection in User Behavior," IEEE S&P, 2019.