ARIES

eightfold.ai

INNOV8 2.0

SOLUTION PROPOSAL

Team ModuleNotFound

IIT KHARAGPUR

1 Introduction

In the escalating conflict between the Phrygians and the Xernians, we face not only a powerful enemy but also a dangerous threat from within. The Phrygians, with their superior wealth and manpower, have historically exploited these advantages by luring Xernian soldiers into defection. This internal betrayal has weakened our forces time and again, leaving us vulnerable on the battlefield.

To counter this, our objective is clear: we must develop a robust system capable of **predicting potential betrayal within our ranks**. Our aim is to design a system that can anticipate possible defections, thereby **strengthening our strategic position**.

A simple model based solely on quantifiable features is insufficient, as it fails to capture the complex interplay of **psychological and social dynamics** that significantly influence a soldier's decision to defect. Therefore, we have developed a comprehensive **three-stage pipeline** that encompasses all the relevant factors, incorporating 00fquantifiable features while also considering the **influence of social circles** and **soldiers' opinions**. This approach ensures a more nuanced understanding of the dynamics at play in potential defections.

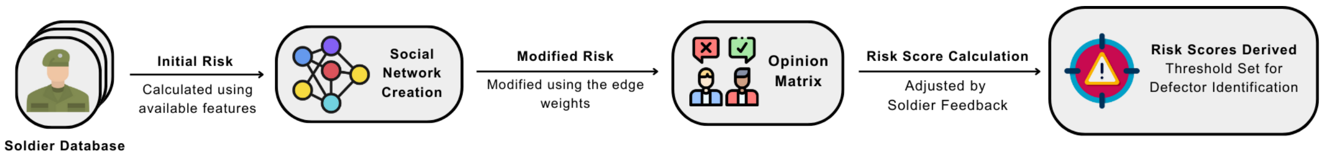

Figure 1: Multi-stage defection risk prediction pipeline

In the following sections, we will:

1. Demonstrate various techniques to quantify factors such as **greed and allure of power**, and how they can influence a soldier's decision to defect.
2. Emphasize the strong influence of **group dynamics** on the likelihood of defection.
3. Present a **synthetic dataset** developed based on the above mentioned factors.
4. Present our hypotheses regarding the relationships between various factors and defection probability.
5. Work towards building and evaluating a system capable of predicting betrayal within the army.

Through this approach, we aim to address the critical issue of soldier defection and its impact on our military strength. By developing a predictive model, we hope to mitigate the risk of future defections and maintain the integrity of our forces in the face of the Phrygian threat.

2 Challenges

In developing our predictive system for soldier defection, we face several significant challenges:

1. **Absence of Dataset:** The lack of a pre-existing dataset presents a fundamental challenge. We must:
   - Identify and **select relevant attributes** $A = \{a_1, a_2, ..., a_n\}$ that are predictive of betrayal.

- Develop a methodology to quantify abstract features, transforming **qualitative traits into numerical representations** $f : A \to \mathbb{R}^n$.

2. **Sophisticated Risk Level Predictor:** Creating an effective risk assessment model requires:
   - Developing a multi-step workflow that integrates individual traits, social dynamics, and psychological patterns.
   - Formulating a **risk function** $R(x)$ that accurately maps a soldier's attributes to a numerical risk factor.
3. **System Robustness:** Our system must be capable of:
   - Encompassing a wide variety of defection motives within a unified framework.
   - Addressing the **complexity of human behavior** by incorporating both quantifiable and abstract attributes.
   - Implementing effective anomaly detection in the context of social beings, where the definition of an "anomaly" is inherently complex.
4. **Evolving and Scalable System:** The predictive pipeline must:
   - Adapt to new sets of traits and **changing patterns of behavior** over time.
   - Maintain accuracy across varying army sizes, from small units to large forces, ensuring that $Accuracy(S) \approx constant$ for all $S \in \{S_1, S_2, ..., S_k\}$, where S_i represents armies of different sizes.
5. **Ethical Considerations:** We must address the ethical implications of:
   - Implementing a system that could potentially impact soldiers' careers and lives.
   - Ensuring fairness and **avoiding bias** in our predictive model.

Addressing these pain points is crucial for developing a reliable and effective system for predicting potential betrayals within the army. The solutions to these challenges will form the core of our innovative approach in the subsequent sections.

3 Approach

3.1 Pipeline

Our framework utilizes a multi-stage risk prediction system that incorporates both individual attributes and social dynamics to predict the likelihood of soldier defection. This approach is designed to capture the complex interplay of personal characteristics, social influences, and psychological factors that contribute to a soldier's decision to defect. Figure 2 illustrates the overall pipeline of our approach.

The pipeline consists of three main stages, each building upon the previous to refine and improve our risk assessment:

1. **Initial Risk Assessment**: Evaluates quantifiable attributes to produce a initial defection risk estimate.
2. **Social Network Influence**: Utilizes graph-based proximity modeling to analyze relationships.

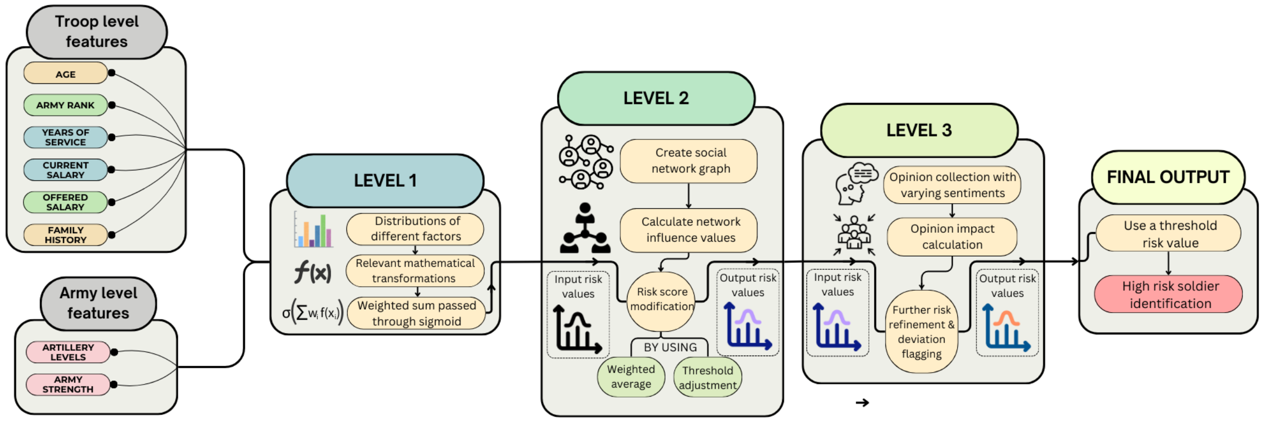

Figure 2: Flowchart of the multi-stage risk prediction pipeline

3. **Peer Sentiment Analysis**: Incorporates opinions and sentiments to provide a comprehensive view.

3.1.1 Stage 1: Initial Risk Assessment

In this foundational stage, we quantify various personal and global attributes to calculate an initial risk score for each soldier. This stage is crucial as it forms the basis for all subsequent refinements.

Feature Selection: We carefully selected a set of features that are likely to influence a soldier's loyalty and propensity for defection. These include:

- Age (a): Modeled as a normal distribution $\mathcal{N}(\mu = 35, \sigma = 7)$
- Years of service (s): Following a gamma distribution $\Gamma(k = 2, \theta = 6)$
- Rank (r): Categorical variable with probabilities [0.55, 0.25, 0.12, 0.06, 0.02] for ranks from Private to Captain
- Current salary (c): Calculated based on rank and years of service
- Enemy salary offer (e): Modeled as a multiple of current salary, $e = c \cdot U(1.1, 1.5)$
- Family members in army (f): Poisson distribution $Pois(\lambda = 1.5)$
- Artillery level of army (l): Uniform distribution $U(1, 10)$
- Army strength (s_a): Fixed value for the entire army
- Enemy army strength (s_e): Fixed value for the enemy army

The dataset statistics are provided in Figure 10 in the Appendix.

Risk Function: The **initial risk function** $f(x_i)$ for soldier i with feature vector x_i is defined as:

$$f(x_i) = \sigma\left(w_0 + \sum_{j=1}^{n} w_j \phi_j(x_i)\right) \tag{1}$$

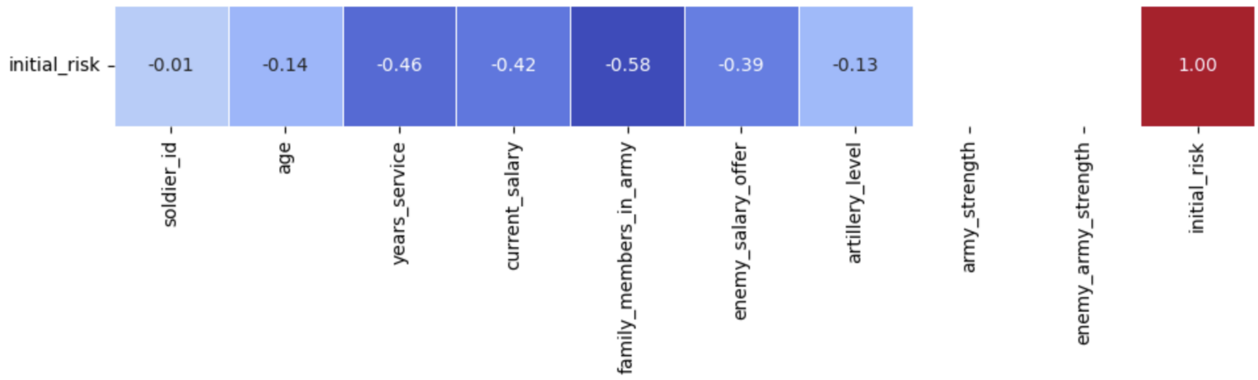

Figure 3: Correlation of attributes with initial risk

where $\sigma(z) = \frac{1}{1+e^{-z}}$ is the sigmoid function, w_j are learnable weights, and $\phi_j(x_i)$ are feature transformations (refer 7.2.1 in the Appendix for details):

This formulation allows for **non-linear relationships** between features and risk, capturing complex interactions.

Noise Injection: To account for unobserved factors and introduce a level of uncertainty, we add a small amount of **Gaussian noise** to the risk scores:

$$f'(x_i) = f(x_i) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, 0.01) \tag{2}$$

3.1.2 Stage 2: Social Network Influence

In this stage, we model the **army's social structure as a graph** and use it to refine our risk estimations based on social influences.

Consider a scenario involving three soldiers: A, B, and C. Soldier B contemplates an action that contradicts the core values and loyalty upheld by both Soldier A and Soldier C. However, driven by strong external pressures, Soldier C eventually decides to align with B's decision. Left isolated, Soldier A contemplates whether to maintain allegiance to the military or follow the path of his comrades, ultimately succumbing to the influence of his social network. Conversely, in a similar case, if Soldiers A and C remain steadfast in their loyalty, Soldier B is likely to reconsider defection, instead reinforcing his allegiance due to the positive influence of his peers. This scenario illustrates how **peer dynamics** within military units can significantly **shape individual decisions**, particularly in the context of defection.

Soldiers are organized into distinct ranks, ranging from Privates to Captains. Each rank has its own unique peer groups, which may sometimes overlap across ranks. This social structure significantly impacts decision-making, reinforcing the idea that individuals are influenced by their connections.

We apply concepts from graph theory, particularly focusing on **Small World Networks**, to construct our model. Our graph $G(V, E)$ is defined as follows:

- **Vertices** V represent soldiers, while **edges** E represent relationships between them.
- The graph exhibits a **hierarchical structure** based on five ranks: Private, Corporal, Sergeant, Lieutenant, and Captain.
- Within each rank, soldiers form **subgroups**, typically consisting of about 100 soldiers each.

- There is a high probability of connections within subgroups, with $p_{in} = 0.8$.
- Connections between subgroups of the same rank have a moderate probability of $p_{out\_within} = 0.2$.
- Connections across different ranks are less likely, with a probability of $p_{out\_across} = 0.05$.

Weights w_{ij} are assigned based on the nature of the connection:

- For connections within a subgroup: $w_{ij} \sim U(7, 10)$
- For connections between subgroups of the same rank: $w_{ij} \sim U(3, 6)$
- For connections across ranks: $w_{ij} \sim U(1, 3)$

This structure closely resembles the social networks observed in real-world organizations, where closely-knit peer groups significantly influence individual behavior. The impact of a defector within a person's close group can lead to a chain reaction, potentially influencing others to consider defection as well.

To account for these social influences, we update the risk scores through an iterative process:

$$f^{(t+1)}(x_i) = (1 - \alpha) f^{(t)}(x_i) + \alpha \sum_{j \in N(i)} \frac{w_{ij} f^{(t)}(x_j)}{\sum_{k \in N(i)} w_{ik}}$$

In this formula, $f^{(t)}(x_i)$ represents the risk score of soldier i at time t, α is a parameter controlling the influence of network connections, and $N(i)$ denotes the neighbors of soldier i in the graph.

By integrating social structures into our risk assessments, we recognize that decisions are often shaped by the actions and influences of close peers, reinforcing the need to consider social networks in military contexts.

3.1.3 Stage 3: Peer Sentiment Analysis

This stage **incorporates peer opinions** to further refine our risk assessments. Soldiers' opinions provide valuable **insight into group sentiment**, revealing underlying tensions, dissatisfaction, or loyalty shifts that directly influence defection risk. By weighing these **subjective viewpoints**, we enhance our model's ability to detect emerging threats that purely structural or quantifiable data might overlook, making our predictions more precise and contextually informed.

Opinion Generation: For each soldier, we generate textual opinions (which map to a decimal sentiment value between -1.0 to 1.0) about a subset of other soldiers (e.g., 30 opinions per soldier). Some sample text opinions have been given in Section 7.3 of the Appendix. The sentiment of these opinions is influenced by the difference in risk scores between the opinion giver and receiver:

$$P(\text{positive opinion}) = \sigma(b(f(x_j) - f(x_i))) \tag{3}$$

where b is a scaling factor controlling the strength of the relationship between risk difference and opinion sentiment.

Sentiment Scoring: We assign sentiment scores s_{ji} to each opinion, ranging from -1 (highly negative) to 1 (highly positive).

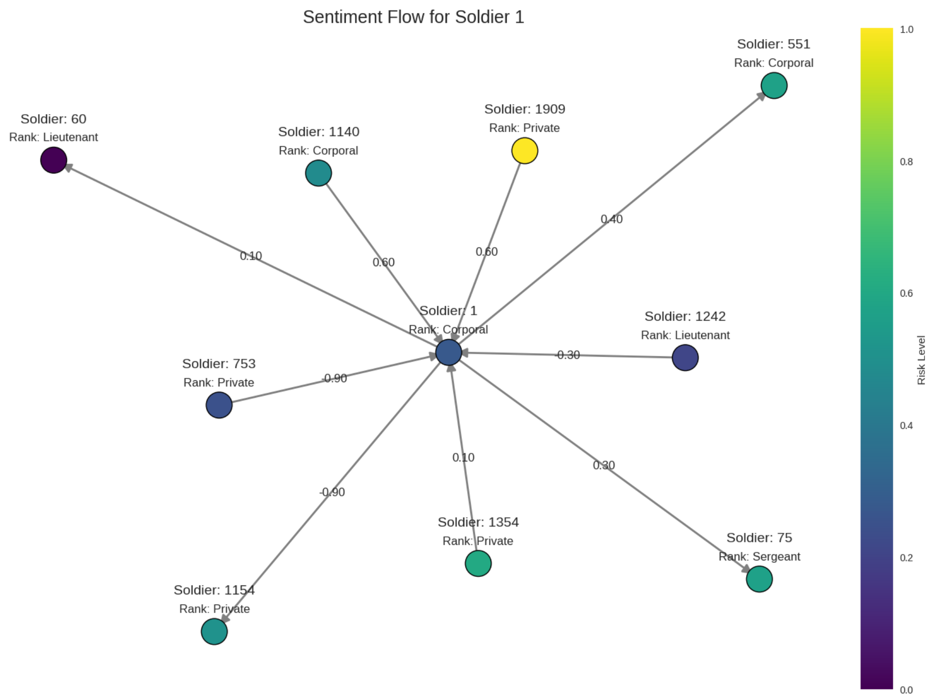

Figure 4: Sample of sentiment flow for a single soldier

Risk Update: We update the risk scores based on the aggregated sentiments:

$$f'''(x_i) = f''(x_i) + g\frac{\sum_{j \neq i} w_j s_{ji}}{\sum_{j \neq i} w_j} \tag{4}$$

where $w_j = 1 - f''(x_j)$ is a weight that gives more importance to opinions from low-risk soldiers, and g is a scaling factor controlling the overall impact of sentiments.

By iteratively refining our risk assessments, we aim to achieve a high degree of accuracy in predicting potential defections.

3.2 Approach Analysis

Our multi-stage pipeline for predicting soldier defection risk represents a comprehensive approach that combines individual attributes, social dynamics, and peer influences. This section analyzes the design choices, justifications, benefits, and potential limitations of our solution.

3.2.1 Design Choices and Justifications

Feature Selection and Distribution: The selection of features and their corresponding distributions was based on a combination of real-life modeling of analogous workplace/army scenarios and reasonable assumptions where empirical data was not available. For instance:

- Age is modeled as a normal distribution ($\mathcal{N}(\mu = 35, \sigma = 7)$), reflecting the typical age range in military service.
- Years of service follows a gamma distribution ($\Gamma(k = 2, \theta = 6)$), capturing the right-skewed nature of service length.
- Rank probabilities ([0.55, 0.25, 0.12, 0.06, 0.02] for Private to Captain) reflect the pyramid structure of military hierarchies.

These choices aim to create a plausible representation of a military force, balancing realism with computational tractability.

Risk Function Formulation: The risk function incorporates both linear and non-linear components to capture complex relationships between features. The use of a sigmoid function to bound the risk score between 0 and 1 ensures interpretability and comparability across soldiers.

Social Network Model: The hierarchical structure of the social network, with higher connection probabilities within ranks and subgroups, reflects the typical organizational structure of military units. This design choice allows us to model the stronger influences within peer groups while still accounting for cross-rank interactions.

Opinion Generation and Sentiment Analysis: The incorporation of peer opinions adds a layer of psychological realism to our model. By allowing opinions to be influenced by the risk scores of both the giver and receiver, we capture the tendency for individuals to form judgments based on perceived similarities or differences.

3.2.2 Strengths of the Approach

- **Comprehensive Risk Assessment:** By incorporating multiple stages, our approach captures various factors influencing defection risk, from personal attributes to social dynamics and peer perceptions.
- **Adaptability:** The pipeline's modular nature allows for easy adjustment of parameters or even the addition of new stages to accommodate different army sizes, structures, or specific operational contexts.
- **Interpretability:** Each stage of the pipeline provides insights into different aspects of risk, aiding in targeted interventions. For instance, the social network stage can highlight influential nodes, while the sentiment analysis can reveal hidden tensions.
- **Robustness:** The multi-stage approach reduces the impact of errors or biases in any single stage. By refining risk assessments through multiple lenses, we aim to achieve a more reliable final risk score.
- **Realistic Modeling:** Our approach attempts to replicate real-life scenarios by incorporating complex social interactions and psychological factors, going beyond simple attribute-based predictions.

3.2.3 Theoretical Underpinnings

The pipeline's design is grounded in several theoretical frameworks:

- **Social Influence Theory:** The social network stage draws on concepts from social influence theory, recognizing that individuals' attitudes and behaviors are shaped by their social connections.
- **Cognitive Dissonance Theory:** The opinion generation mechanism in the sentiment analysis stage aligns with cognitive dissonance theory, as individuals are more likely to form positive opinions about those similar to themselves.
- **Risk Homophily:** The tendency for risk scores to converge within social groups reflects the principle of homophily, where similar individuals are more likely to associate and influence each other.

3.2.4 Future Directions

To address these limitations and further enhance the model, several avenues for future work are proposed:

- **Incorporating Temporal Dynamics:** Extending the model to capture changes in risk over time, possibly through the use of time series analysis or dynamic network models.
- **Ethical Framework:** Developing a comprehensive ethical framework for the deployment and use of such a system, including guidelines for data handling, interpretation of results, and intervention strategies.

In conclusion, our multi-stage pipeline offers a nuanced and comprehensive approach to assessing defection risk in military contexts. By combining individual attributes, social network analysis, and peer perceptions, we aim to provide a more accurate and actionable risk assessment tool. However, the ethical implications and potential limitations of such a system must be carefully considered and addressed in its development and deployment.

4 Key Findings and Experiments

4.1 Initial Risk Variation with Quantifiable Features

Our analysis of initial risk scores reveals significant **correlations with various quantifiable attributes of soldiers**. Figure 5 illustrates these relationships.

Key observations from these plots include:

1. **Rank and Initial Risk:** There is a clear inverse relationship between rank and initial risk. Privates exhibit the highest median risk and the widest distribution, while Captains show the lowest median risk with the narrowest distribution. This suggests that higher ranks are associated with lower defection risk, possibly due to increased loyalty, better compensation, or more invested career progression.
2. **Enemy Army Strength:** As the enemy army strength increases relative to our fixed strength of 2000, the average initial risk of our soldiers also increases. This indicates that perceived threat levels significantly influence defection risk, with soldiers more likely to consider defection when facing a stronger enemy.
3. **Family Members in Army:** There is a notable decrease in initial risk as the number of family members serving in the army increases. Soldiers with no family members show the highest median risk, while those with three or more family members exhibit the lowest risk. This suggests that family ties within the army serve as a strong loyalty enhancing factor.
4. **Years of Service:** The relationship between years of service and initial risk is complex. There's a general trend of decreasing risk with increased service years, particularly evident in the first 10-15 years. However, the relationship becomes less clear for very long service periods, possibly due to factors like burnout or increased exposure to enemy incentives over time.

These findings underscore the multifaceted nature of defection risk and highlight the importance of considering various personal and environmental factors in our risk assessment model. The strong influence of rank and family ties suggests that targeted loyalty programs and family-oriented policies could be effective in reducing defection risk. Additionally, the impact of relative army strength emphasizes the need for morale-boosting strategies when facing numerically superior enemies.

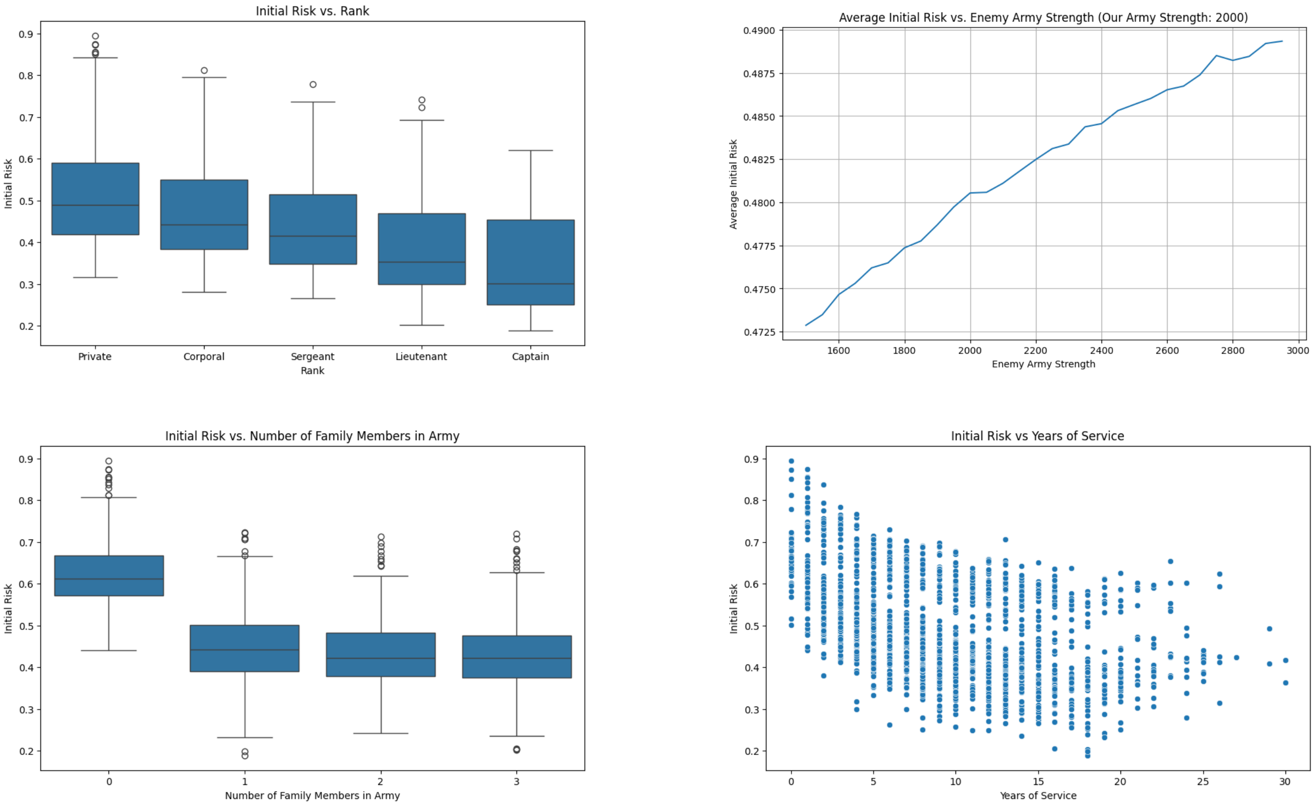

Figure 5: Variation of initial risk with quantifiable attributes

4.2 Influence of Social Factors and Group Dynamics

The social network structure within the army plays a crucial role in the propagation of defection risk. To visualize these dynamics, we constructed a sample network graph of 100 soldiers.

Figure 6 illustrates the complex social dynamics within our simulated army:

- **Node Colors:** Represent the modified risk scores after network influence, ranging from purple (low risk) to yellow (high risk).
- **Edge Thickness:** Indicates the strength of connections between soldiers, with thicker lines representing stronger relationships.
- **Rank Clusters:** Clear groupings are visible for different ranks, with Privates forming the largest and most interconnected cluster.
- **Inter-rank Connections:** Thinner edges between rank clusters show the limited but existing connections across different ranks.
- **Risk Distribution:** Higher-ranking officers (Captains, Lieutenants) generally show lower risk scores (darker colors), while lower ranks exhibit more varied risk levels.

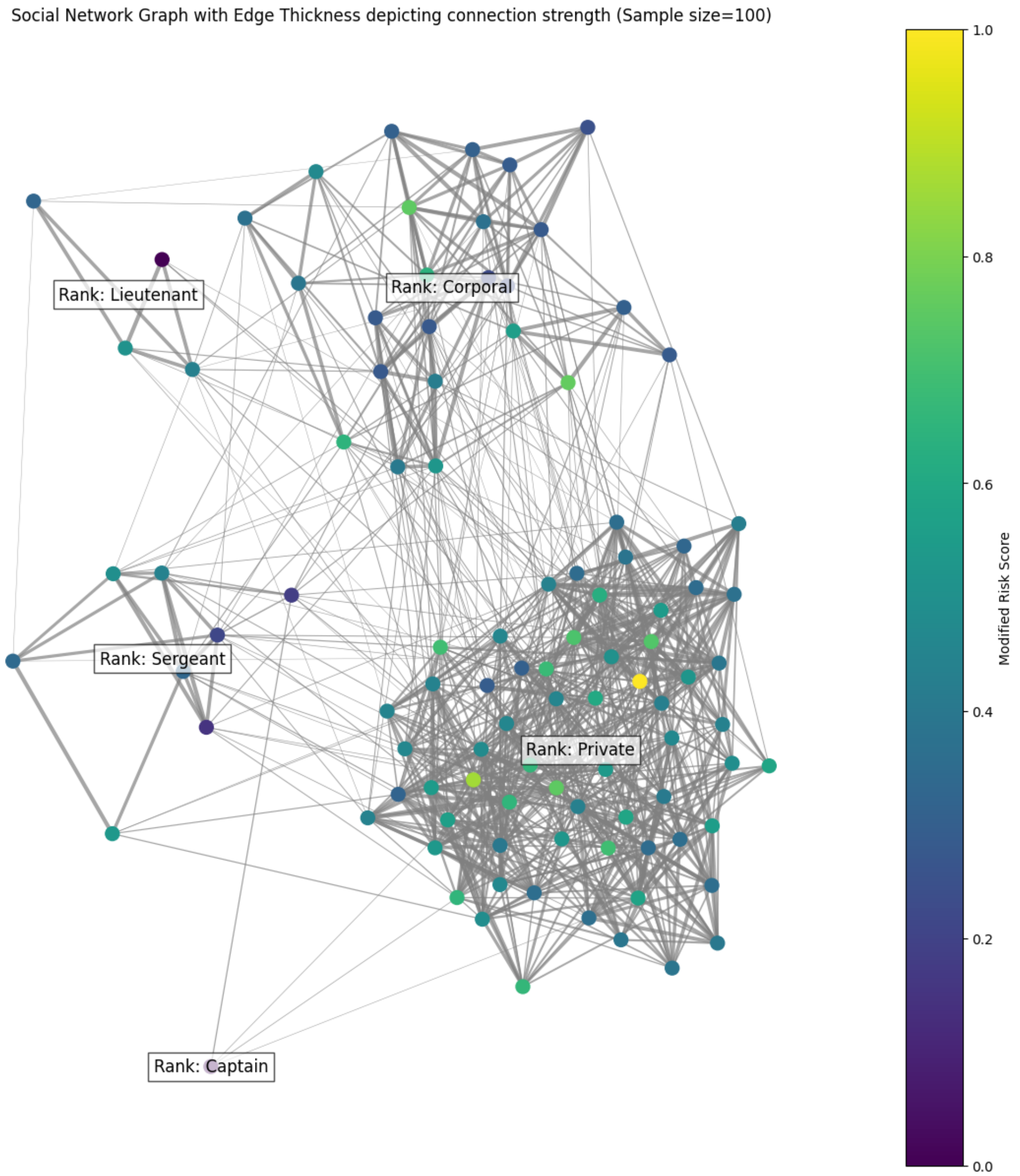

Figure 6: Social Network Graph depicting connection strengths

This visualization underscores the importance of rank and social connections in influencing defection risk. It demonstrates how our model captures the hierarchical nature of military structures and the potential for risk propagation through social ties. The dense connections among Privates suggest that peer influence may be particularly strong at lower ranks, while the sparse connections of higher-ranking officers indicate their relative isolation and potentially lower susceptibility to network effects.

4.3 Risk Evolvement through the Levels

Our analysis of risk evolution through the different stages of our model revealed significant insights into the dynamics of defection risk.

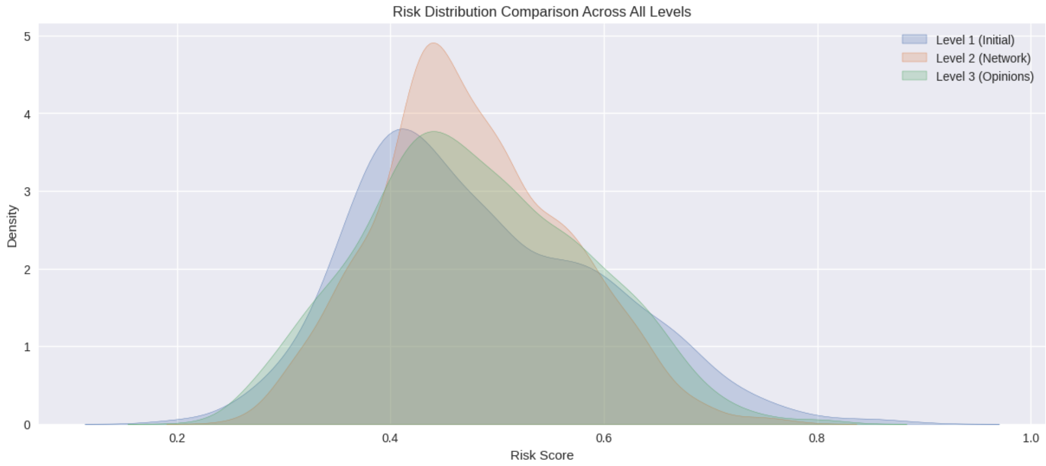

Figure 7: Risk Progression over assessment stages

Figure 7 illustrates the shift in risk score distribution from the initial assessment to the post-network influence stage. Key observations include:

- The initial risk distribution (blue) shows a bimodal pattern with peaks around 0.4 and 0.6, indicating two distinct risk groups in the initial assessment.
- The modified risk distribution (pink) demonstrates a more centralized, unimodal distribution, suggesting a convergence effect due to social network influences.
- The variance of risk scores decreases after applying social network influences, indicating a moderating effect on extreme risk values.

This convergence aligns with social influence theory, where individuals' attitudes tend to shift towards the average of their social group.

4.4 Rank-Based Analysis of Risk Changes

We examined how military rank influences the changes in risk scores from the initial assessment to the post-network influence stage.

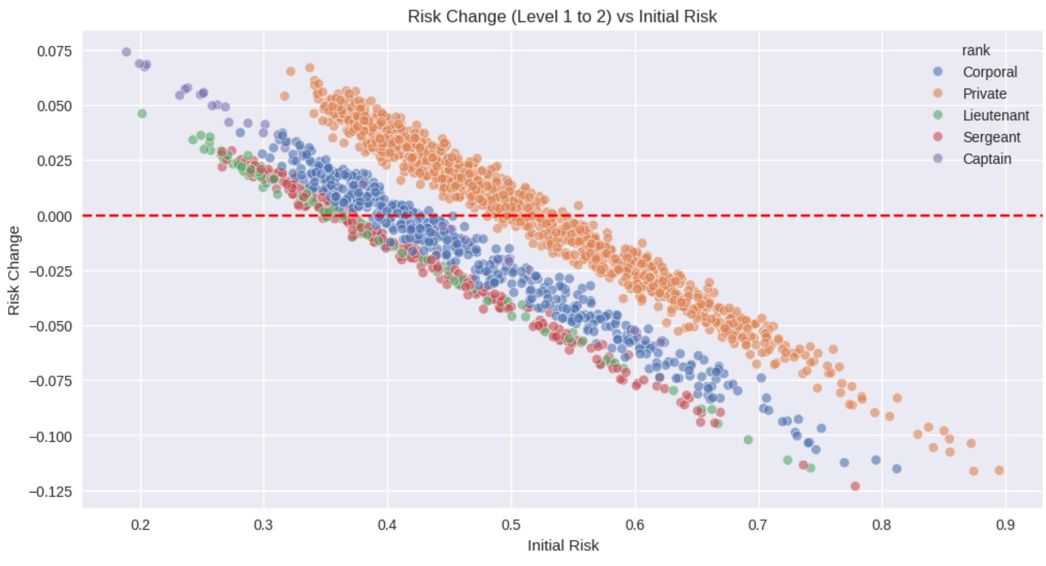

Figure 8: Risk change (Level 1 to 2) vs initial iisk by rank

Figure 8 provides valuable insights:

- Distinct clusters are visible for each rank, indicating that rank significantly influences how social factors affect risk.
- Privates generally have higher initial risk values but show a tendency for risk reduction after social modeling.
- Captains typically start with lower risk values but exhibit slight increases in risk after social influences are considered.
- Middle ranks (Corporals, Sergeants, Lieutenants) show more varied patterns, generally converging towards the center.

This analysis highlights the importance of considering rank in risk assessment and demonstrates how our model captures the nuanced effects of social influence across different levels of the military hierarchy.

4.5 Impact of Enemy Dominance on Risk Scores

In this experiment, we varied the enemy dominance parameters—specifically strength and salary offered—across three scenarios:

- **Moderately Powerful**
- **Considerably Powerful**
- **Overpowered**

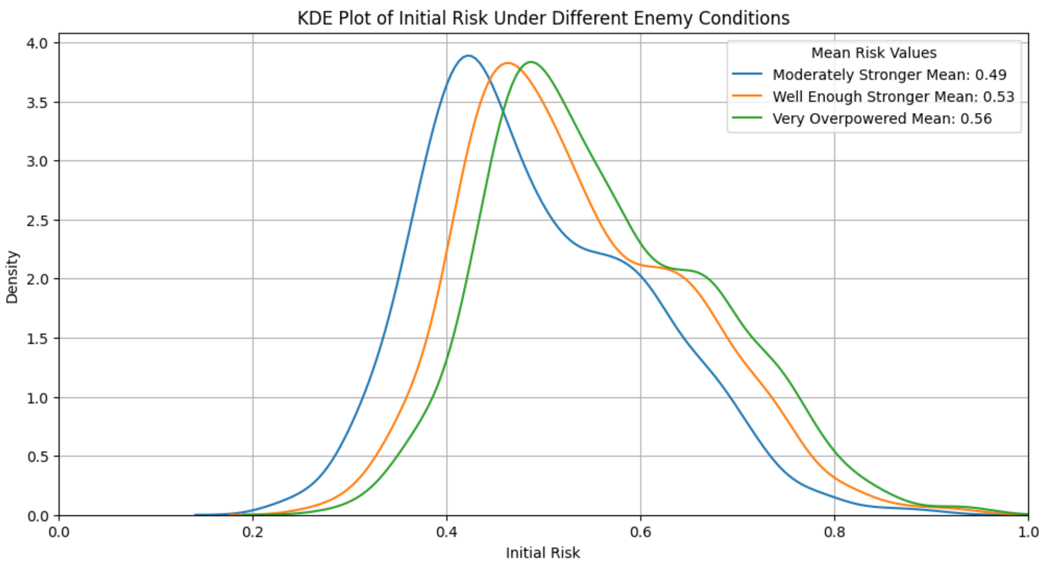

Figure 9: Impact of enemy dominance on initial risk scores

Our observations reveal several key trends:

- The mean of the initial risk curve **shifts towards the right** as enemy dominance increases, indicating an overall rise in initial risk scores.
- Notably, after a **certain maximum threshold**, there is no further increase in the mean risk scores. For instance, a 20 times larger army produces the same impact on risk as a 100 times larger army.

This finding suggests that, beyond a specific level of enemy dominance, additional increase in army size do not significantly alter perceived risk. This plateau indicates that factors other than sheer numbers may influence risk assessment.

These results can be attributed to the overwhelming presence of the enemy, reinforcing our hypothesis that greater enemy dominance correlates with higher initial risk assessments. However, the diminishing returns at high levels of dominance highlight the need for military strategies to consider both the capabilities of the enemy and the limits of risk perception. Consequently, adaptive risk management frameworks must be employed to effectively respond to evolving threats.

5 Prospective Enhancements and Scalability

5.1 The CLAP Framework

We came up with an innovative strategy, **Covert Loyalty Assessment Protocol (CLAP)** which evokes the idea of trusted soldiers working covertly within the ranks, silently gathering intelligence and testing the loyalty of their peers, all while remaining loyal to the cause themselves. It also hints at their dual role—both observing and subtly influencing the behaviour of others.

What this framework involves:

- **Identification:** Identifying troops with the **least risk factors** in their respective ranks and entrusting them with the **job of finding out potential defectors**. The main focus will be on areas where defection risk is high based on your defection prediction system.
- **Monitor:** Regularly collect **intelligence reports** from these soldiers, looking for **suspicious behaviour** within the army including **secret meetings, change in behaviour, negative talk** etc.
- **Use of Reverse Psychology:** The main highlight of CLAP is the engagement of these "trusted" soldiers in reverse psychology tactics. Acting as **undercover spies**, they subtly **influence potential defectors** by pretending to sympathise with their disloyal intentions. This technique manipulates the natural responses of those who might be inclined to betray the clan, pushing them to **reveal their true intentions** without overt coercion, while providing insight into who remains loyal under pressure.

In a way, this adds a human layer of validation to the defection prediction model, making it more dynamic and adaptive to real-time shifts in morale and loyalty.

5.2 System Scalability and Robustness

5.2.1 Addressing Computational Complexity

Our multi-stage approach provides a comprehensive framework for risk assessment. However, for large-scale military organizations, the graph-based stages can become computationally intensive. To mitigate this, we have implemented several optimizations:

- Utilization of **single look-ups** to reduce redundant calculations
- Implementation of **weighted sampling** methods to efficiently process large networks

- **Limitation of network depth** to 4 levels, balancing comprehensiveness with computational efficiency. Moreover it logically makes sense since opinions of people further distant in the social network wont have much truth behind them.

These optimizations have yielded significant performance improvements. For instance, in our initial implementation, generating and assigning opinions and computing the opinion impact for all soldiers in a 2000-person army network (considering up to 4th-level connections) required approximately **48 minutes** of computation time. This was understandable, given that a single soldier's network encompassed an average of 900 individuals up to the 4th connection level. However, after implementing our optimizations, we reduced the computation time to approximately **90 seconds**, marking a substantial efficiency gain.

5.2.2 Mitigating Error Propagation

Our multi-stage pipeline approach necessitates careful consideration of error propagation. In such systems, initial stage computation errors can potentially compound over time, especially as the scale of the dataset increases. To address this, we have implemented several design choices to ensure that the scale of the army does not lead to exponential error growth:

- **Normalization of Risk Scores**: At each stage, risk scores are normalized to maintain a consistent scale, preventing drift over multiple iterations.
- **Hierarchical Network Structure**: The social network is modeled to reflect realistic military hierarchies, with connection probabilities adjusted based on rank proximity. This structure helps contain error propagation within rank-based subgroups.
- **Damping Factor in Network Influence**: The incorporation of a damping factor (α) in the network influence calculation (Stage 2) limits the extent to which extreme risk values can propagate through the network.
- **Weighted Opinion Aggregation**: In the peer sentiment analysis stage (Stage 3), opinions are weighted based on the risk score of the opinion giver, reducing the impact of potentially erroneous extreme opinions.
- **Stochastic Elements**: The introduction of controlled randomness, such as the noise injection in the initial risk calculation and the probabilistic nature of opinion generation, helps prevent overfitting to specific patterns and maintains model generalizability across different army sizes.
- **Scalable Feature Distributions**: Key features like age, years of service, and rank probabilities are modeled using distributions that maintain their characteristics regardless of army size, ensuring consistency in the initial risk assessment stage.

These design choices collectively ensure that our risk assessment model maintains its integrity and predictive power across varying army sizes. By keeping the initial feature distributions constant and scaling army and enemy army sizes proportionally, we observe that risk value distributions remain consistent, regardless of the absolute size of the forces involved. This scalability and robustness are crucial for the practical application of our model in real-world military scenarios of varying scales.

6 Conclusion

To conclude, we synthesized a comprehensive soldier dataset and developed an advanced risk assessment model, gaining valuable insights into defection risk factors. Our **multi-stage pipeline** for predicting soldier defection risk combines **individual attribute analysis, social network modeling, and peer sentiment evaluation**.

Furthermore, we proposed the **CLAP framework** to further strengthen the validation of the computed defection risk score, ensuring accuracy and adaptability. Our design also ensures scalability, making it applicable across armies of varying sizes.

While promising, it's crucial to apply this predictive tool ethically, remembering these are risk assessments, not definitive judgments. Future work should focus on real-world validation and adapting to evolving military dynamics. By combining this system with traditional loyalty-building measures, we can strengthen our forces against the threat of Phrygian-induced defections.

7 Appendix

7.1 Dataset

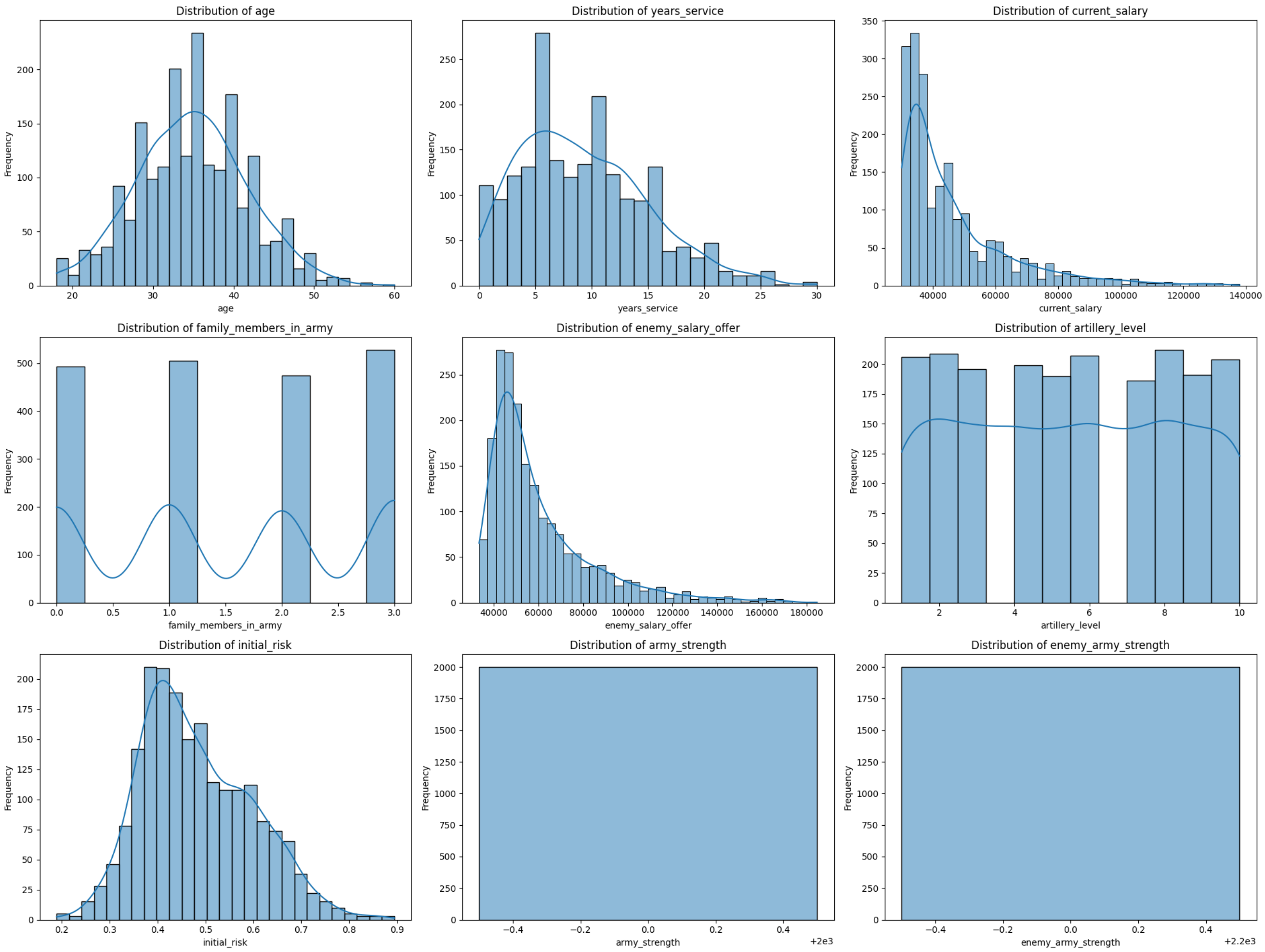

Figure 10: Distribution of various quantifiable attributes in the synthesized dataset

7.2 Implementation Details

7.2.1 Feature transformations for initial risk assessment

- $\phi_1(x_i) = 1 - \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(a_i-\mu)^2}{2\sigma^2}}$ (Age factor)
- $\phi_2(x_i) = \tanh\left(\frac{e_i - c_i}{c_i}\right)$ (Salary difference factor)
- $\phi_3(x_i) = e^{-s_i/5}$ (Service years factor)
- $\phi_4(x_i) = e^{-3f_i}$ (Family factor)
- $\phi_5(x_i) = 1 - \frac{ranks.index(r_i)}{len(ranks)-1}$ (Rank factor)
- $\phi_6(x_i) = \frac{s_e}{s_a + s_e}$ (Army strength factor)
- $\phi_7(x_i) = 1 - \frac{l_i}{10}$ (Artillery level factor)

7.3 Positive and Negative Opinion Statements with Sentiment Values

7.3.1 Positive Opinions

Opinion Statement	Sentiment Value
"I trust them completely"	1.0
"They're incredibly loyal"	0.9
"They always put the team first"	0.8
"They have a strong sense of duty"	0.7
"They're reliable in tough situations"	0.6
"They show great potential"	0.5
"They're improving every day"	0.4
"They work well with others"	0.3
"They follow orders well"	0.2
"They're punctual and disciplined"	0.1

7.3.2 Negative Opinions

Opinion Statement	Sentiment Value
"I don't trust them at all"	-1.0
"They seem disloyal"	-0.9
"They often complain about our mission"	-0.8
"They lack discipline"	-0.7
"They're unreliable under pressure"	-0.6
"They show little commitment"	-0.5
"They struggle to follow orders"	-0.4
"They don't work well in a team"	-0.3
"They're often late or unprepared"	-0.2
"They seem disinterested in their duties"	-0.1