ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

AI-enhanced cyber threat simulation in a real-time economic forecasting environment

Mingming Cao[1,2,*]

[1] School of Finance and Trade, Harbin Finance University, Harbin, Heilongjiang, 150030, China

[2] School of Economics and Business Administration, Heilongjiang University Harbin, Heilongjiang, 150006, China

SUMMARY: *Economic forecasting provides support for governments to formulate rules in advance and ensure the smooth operation of the social economy. At the same time, it is also very important to identify possible network threats in the process of economic forecasting in a timely manner to ensure the security of economic data. In this paper, research is carried out from two aspects of economic forecasting and cyber threat detection. The prediction model MIDAS is constructed to predict the quarterly GDP growth rate with monthly economic data, and the prediction error is reduced by determining the weight function and parameter constraints. Design a behavioral pattern graph that contains normal and abnormal user behaviors. Combine the behavior pattern graph embedding algorithm (GraphSAGE) to achieve graph embedding and graph dimensionality reduction, and introduce Gaussian noise to enhance the effect of cyber threat retrieval.The success rate of GraphSAGE's behavior pattern graph dimensionality reduction reaches up to 97.12%, and 12 cyber threat paths hidden in user behaviors are successfully identified in an average time of 1.82ms-9.18ms.*

KEYWORDS: *MIDAS; weight function; behavioral pattern graph; GraphSAGE; economic prediction; cyber threat*

1 Introduction

Economic forecasting is based on the theoretical basis of economics, through the calculation and analysis of economic data and data for a certain period of time, through the relevant forecasting methods and techniques, so that it is possible to study and analyze the changes in economic development [1]. Economic forecasting comes from the process of production and trade of that commodity, but also from the changes in the economic decisions of companies, local and national economic development needs, and after the economic data are analyzed, the rules of economic development are demonstrated through the use of exercises to guide the economic activity and master these rules [2-5]. In today's changing economic environment with high uncertainty of economic development, traditional economic forecasting relies on the knowledge and experience of forecasters, and traditional economic forecasting does not well reflect the large number of nonlinear relationships in the economic system, which will directly affect the accuracy of economic forecasting, and it is no longer able to adapt to the progress of the current society [6-10]. And artificial intelligence (AI) technology of economic forecasting methods, for the inherent relationship between the input and output of the economic forecasting

*18545192704@163.com
https://doi.org/10.65102/is2026009
Published: 30 April, 2026

system, and to confirm the optimization and selection of the system, which will comprehensively improve the accuracy of the model prediction results [11-13].

In order to effectively combine economic forecasting with AI technology, it is necessary to overcome the shortcomings and boundaries of the previous economic forecasting, data collection and updating, and give full play to the advantages of artificial intelligence, organizing the data through sounds, images and other information as an important source of information for economic analysis [14-17]. Secondly, the use of AI technology improves the fault tolerance of the prediction system, and predictive models can be used for incomplete information or incorrect data [18, 19]. In addition, AI technology improves the data processing capacity and the speed of the system to build dynamic information models that absorb new information [20, 21]. As a result, the accuracy of economic forecasting can be reached.

However, AI technology-driven economic forecasting presents both opportunities and challenges. On the one hand, there are the ethical issues inherent in AI algorithms. Kokogho et al [22] (2024) identified multiple challenges to AI-based economic forecasting, data quality and availability, algorithmic black boxes leading to a lack of transparency in the decision-making process, algorithmic bias, and the risk of automated decision-making. Fu et al [23] (2022) reported that the digital economy transformation is characterized by capital and technological constraints on Internet digital economy development forecasting, and digital security, high-risk vulnerabilities, cyberattacks, data imbalance, and lagging laws and regulations exacerbate economic forecasting challenges. On the other hand, with the deep development of AI-enabled economic forecasting, cyber threats are gradually complicated, mainly because attackers can perform persistent lurking, use automated tools or scripts, gain unauthorized access by trying all possible combinations of passwords, keys, or credentials, and may directly bypass the target system's protection against brute-force cracking to carry out destructive attacks, such as data tampering and private data access [24-28]. Therefore, cyber threats in economic forecasts can be modeled to provide decision-making recommendations for defense strategies.

Scholars have improved the performance of AI economic forecasting models by continuously optimizing them against cyber threats. Hopp et al [29] (2022) performed instantaneous economic forecasting through the ability of Long Short-Term Memory Networks (LSTMs) to process a large number of input time-series features at various temporal frequencies, which performs better than the popular Dynamic Factor Forecasting models. Amini and Kalantari [30] (2024) constructed a hybrid model based on bi-directional LSTM and convolutional neural network with automatic parameter tuning to form a forecasting system for predicting daily gold price movements. Atif [31] (2025) used linear autoregressive integration of moving average and LSTM models to construct an integrated model for GDP forecasting in a fluctuating economic environment, significantly reducing the forecasting errors. Al-Karkhi and Rządkowski [32] (2025) argued that integrating cross-regional data fusion and interpretable AI can lead to more accurate, reliable, and adaptable economic forecasts by solving the problems of data imbalance, real-time data fusion, and feature selection. Ivashchenko and Ivashchenko [33] (2025) combined generative AI with long and short-term memory networks for optimizing the accuracy and computational resource consumption of financial and economic indicator forecasts and adapting them to various devices. Ravi et al [34] (2025) created an economic forecasting model based on real-time pushing of economic data, combining machine learning methods and deep learning networks, and introducing high-frequency economic indicators and sentiment analysis to reduce model error. And AI technology itself has the ability to deal with cybersecurity threats in the financial system. Soundenkar et al [35] (2024) explored in an all-round way that AI technology enhances cyber risk management in the field of financial

system, and detects, predicts, evaluates, and defends against cyberthreats through a variety of AI algorithms and tools, which shortens the response time of the system, reduces the human error, and provides cyber risk management for financial institutions to Protecting and escorting.

In recent years, both cyber threat simulation techniques and techniques to cope with threats have been deeply developed. Niu et al [36] (2017), facing the problem of simulating advanced sustainability threat processes, proposed a target complex attack network model combining dynamic attack graphs and network evolution, as well as dynamic evolution rules. Lerums et al [37] (2018) proposed a statechart-using simulation model that can effectively simulate cyber threat scenarios faced by enterprises, while obtaining cyber risks and prevention costs based on the simulation results. Moustafa et al [38] (2018) established a framework for designing cyber-attack threat intelligence techniques using generalized outlier Gaussian mixing technique based on auto-associative features for simulating and detecting web application attacks with good performance in detection rate and false alarm rate. Moskal et al [39] (2018) developed a model of cyber-attacker behavior by introducing cyber-attack scenarios and network defense simulator, which simulates the interactions between networks and different types of attackers from the perspective of attacker's capabilities, opportunities, intentions and preferences, as well as from the perspective of cyber-attack kill chain. Yeboah-Ofori and Islam [40] (2019) created a model for the supply chain network created a cybersecurity threat model, introduced discrete probability calculations to explore the propagation and chain effects of cyberattacks, and proposed computer network control measures to deal with cyberthreats. Ajmal et al [41] (2023) developed a model for the network security threats through MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) attacker simulation scheme, which develops adversary simulation strategies based on hidden attack vectors and paths, providing a commutative approach for cyber defenders. Rauf et al [42] (2025) generate cyber attack scenarios of practical significance with the help of generative AI and simulate yet-to-be-emerged cyber attack scenarios for evaluating the defense mechanisms and thereby formulate defense strategies. Choudhary et al [43] (2025) integrated Random Forest and XGBoost algorithms to design a wireless network threat simulation framework and added hyper-parameter optimization to improve the detection capability of the framework, which is conducive to threat alerts in network traffic patterns.

In this paper, we construct MIDAS, a real-time GDP growth forecasting model with quarterly steps. The data of monthly macroeconomic indicators are utilized to predict the GDP growth rate of the current quarter in advance, and the Almon polynomial function is selected as the weight function of the MIDAS model to constrain the parameter weights and improve the prediction accuracy. For the network threat caused by abnormal user behavior in the process of economic prediction, graph neural network is introduced to retrieve it. Under the premise of comprehensively considering the weights of the four rules, the behavior pattern graph that hides the potential relationship of user behavior is constructed. The nodes and edges in the behavior pattern graph are extracted using the behavior pattern graph embedding algorithm GraphSAGE, and the low-dimensional conversion of the graph structure is accomplished through information aggregation and other steps to reduce the difficulty of abnormal behavior identification. For the active network threat recognition of GraphSAGE, Gaussian noise is introduced in the fully connected layer to optimize the efficiency of network threat retrieval.

2 Economic predictive cyber threat detection supported by behavioral pattern graphs

2.1 GDP growth rate real-time forecasting model MIDAS construction

In this paper our forecasting of the economy focuses on the quarterly GDP growth rate, set to y_{t_q}, where $t_q\left(t_q=1,2,3,\ldots,T_q^y\right)$ represents the quarterly period, where T_q^y stands for the last quarter in which the quarterly GDP was available.GDP growth rate can also be expressed as a monthly indicator of the form $y_{t_m}=y_{t_q}\forall t_m=3t_q$, with t_m representing the month. Because GDP growth rates are only released quarterly, only the $t_m=3,6,9,\ldots,T_m^y$ era represents observable quarterly GDP growth rates, where $T_m^y=3T_q^y$. Where the monthly indicator used for analysis is denoted as x_{t_m}, $t_m=1,2,3,\ldots,T_m^x$, where T_m^x represents the last month available for the monthly indicator. Typically T_m^x is approximately T_m^y, meaning that many monthly macroeconomic indicators are published earlier than quarterly GDP. The purpose of this paper is to forecast and predict GDP in real time, because the frequency of GDP observation is once a quarter, so when the quarterly GDP data is not available, the government needs to get an estimate of the quarterly GDP growth rate to formulate the policy, which needs to get the prediction value of the GDP growth rate $y_{T_m^y+h_m|T_m^x}$, where $h_m=3h_q$, representing a forecast step of h_m months or h_q quarters, GDP is observed at T_m^y, the monthly indicator is observed at T_m^x, and $T_m^x>T_m^y=3T_q^y$.

A MIDAS model with a forecasting step of h_q quarters is:

$$y_{t_q+h_q}=y_{t_m+h_m}=\beta_0+\beta_1 b\left(L_m;\theta\right)x_{t_m+w}^{(3)}+\varepsilon_{t_m+h_m} \tag{1}$$

where $w=T_m^x-T_m^y$ represents the time lag in which the monthly indicator is more available than the quarterly GDP, $b\left(L_m;\theta\right)$ is the lag polynomial $b\left(L_m;\theta\right)=\sum_{k=0}^{K}c\left(k,\theta\right)L_m^k$, L_m is the monthly lag operator, $L_m x_{t_m}=x_{t_m-1}$.

In the MIDAS model, the selection of the weight function $c\left(k;\theta\right)$ is the key to estimating the MIDAS model because the regressor $x_{t_m}^{(3)}$ has a higher sampling frequency than y_{t_q}, and the lag terms added in the modeling tend to lead to over-parametrization, so a parameter-refined weight function is needed. There are generally four types of weight polynomial functions, Almon, Exponential Almon, Beta, and Step function, each of which can yield various types of weight functions depending on the parameters and implicitly assuming that the sum of the weights is 1.0.

The first, the Almon polynomial, has the basic form:

$$c\left(k;\theta\right)=\frac{\theta_0+\theta_1 k+\theta_2 k^2+\cdots+\theta_Q k^Q}{\sum_{k=1}^{K}\left(\theta_0+\theta_1 k+\theta_2 k^2+\cdots+\theta_Q k^Q\right)} \tag{2}$$

The second, the exponential Almon polynomial, has the specific form:

$$c(k;\theta)=\frac{\exp\left(\theta_0+\theta_1 k+\theta_2 k^2+\cdots+\theta_Q k^Q\right)}{\sum_{k=1}^{K}\exp\left(\theta_0+\theta_1 k+\theta_2 k^2+\cdots+\theta_Q k^Q\right)} \tag{3}$$

Exponential Almon lag polynomials are a very flexible form of weights and different shapes of weight functions can be obtained with only few parameters. This type of polynomial is often used in real-world research because it can construct a variety of different weighting functions, and at the same time ensure that the number of weights is positive while giving the equation the excellent property of zero approximation error.

The third kind, Beta polynomial, is a β polynomial function with only two parameters, similar to the exponential Almon polynomial, and it is also capable of constructing a variety of forms of weight functions. Its specific form can be expressed as follows:

$$c(k,\theta_1,\theta_2)=\frac{f\left(k/K,\theta_1;\theta_2\right)}{\sum_{k=1}^{K} f\left(k/K,\theta_1;\theta_2\right)} \tag{4}$$

Among them,

$$f\left(x,a;b\right)=\frac{x^{a-1}\left(1-x\right)^{b-1}\Gamma\left(a+b\right)}{\Gamma\left(a\right)\Gamma\left(b\right)} \tag{5}$$

$$\Gamma\left(a\right)=\int_0^{\infty} e^{-x}x^{a-1}dx \tag{6}$$

The fourth, the step function polynomial, is expressed in the concrete form:

$$c\left(\theta_1,\ldots,\theta_p\right)=\theta_1 I_{i\in[a_0,a_1]}+\sum_{p=2}^{p}\theta_p I_{i\in\left[a_{p-1},a_p\right]} \tag{7}$$

where $a_0=1<a_1<\cdots<a_p=N$, $I_{i\in\left[a_{p-1},a_p\right]}=\begin{cases}1, & a_{p-1}\le i\le a_p\\ 0, & a_{p-1}\ge i,\ i\ge a_p\end{cases}$. The polynomials thus constructed are a series of off-walks, and the number of steps p can be defined as needed for the model.

In macroeconomic forecasting using MIDAS, for the weight function most of the second exponential Almon polynomial function is chosen and the two parameters $\theta_1\le 300.0$ and $\theta_2\le 0.0$ are constrained to satisfy the form of the weights needed for the macroeconomic analysis and forecasting; and for the number of lagging steps, which relates to the how many periods of high-frequency data are needed to predict low-frequency data in model estimation and forecasting. Generally a large lag order represents that the impulse response function of GDP can be approximated to the monthly indicator, the more complete the available information, but then the model will be plagued by uncertainty due to unlimited lags, and the parameters to be estimated are subsequently increased, which in turn affects the estimation and prediction of the model. By optimizing the parameter vector θ in the weight function in the MIDAS regression equation through a nonlinear estimation method, the optimal length of the lag order is obtained from the graph of the weight function plotted by the parameter estimation $\hat{\theta}$, and it is argued that the lag order so determined is entirely data-driven and therefore optimal. In practical macroeconomic applications, the AIC and BIC criteria are often used as selection

criteria for the lag order.

2.2 Insider threat detection and research based on graph neural networks

2.2.1 Behavioral pattern diagram design

The government's economic forecasting work is often handed over to the enterprise responsible for the successful bidding, and due to the influence of factors such as the large scale of the enterprise, unclear rights and responsibilities of the employees, and more and more diversified means of cyberattacks, the link of real-time economic forecasting work faces high cyberthreat risks. In order to reduce the risk of economic data leakage and improve the security of real-time economic forecasting, the corresponding behavioral data of users accessing economic data in the enterprise are collected and a behavioral pattern diagram is constructed. At the same time, because the structure of the behavior pattern graph is large and the computational cost is high, the detection efficiency of cyber threats is improved by means of dimensionality reduction embedding of the behavior pattern graph.

The construction of behavior pattern graph needs to dig out the potential connection between user behaviors, this paper formulates four rules from the temporal relationship, behavioral logic and organizational relationship, and defines four types of edges, while different types of edges are given different weights. The following elaborates the specific formulation method of the four rules:

1) Rule 1: The log data of the day is connected in chronological order with weight set to $w1=1.0$, the purpose of this rule is to build the daily behavior of the user on the host device.

2) Rule 2: The log data of the day is connected by the same type, with the weight set to $w1=1.0$, the purpose of this rule is to strengthen the connection between the same type of operation on the device.

3) Rule 3: log data of each day is connected by log sessions with weight set to $w1=1.0$, by defining the log data connected by rule 1 as user sessions, and then connecting the logs at the beginning and the end of each user session, the purpose of this rule is to explore the potential connection of log data of multiple days.

Figure 1 shows the user's personal behavior pattern. After multidimensional analysis to obtain the above three rules, the behavioral pattern graph captures the user's daily behavior in time series, strengthens the association of the behavioral patterns through operation types, and then connects the log sessions across days to mine potential long-term abnormal behaviors.

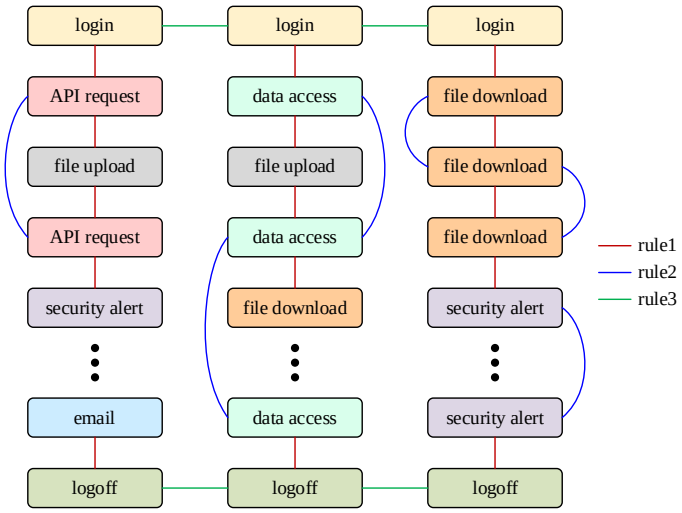

Figure 1: User's Personal Behavior Pattern

4) Rule 4: Connect the log data according to the enterprise organization relationship, the weight is set to $w2 = 1/2$, and extract the similarity of abnormal behavioral patterns in the same organization by understanding the user's position and responsibility in the organization, the purpose of this rule is to represent the relationship between user organizations by setting the edges. Figure 2 shows the structure of user organization relationship.

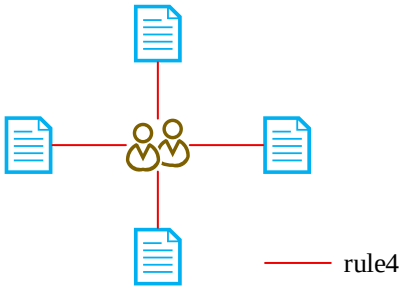

Figure 2: User Organization Relationship

In constructing the user behavior pattern map, the relatively small weight set for this side of the corporate organizational relationship is mainly based on the consideration of its low influence on the user behavior pattern. The enterprise organizational relationship reflects the user's position and responsibilities in the organizational structure, which helps to determine the scope of the user's authority and the resources he or she may be involved in. However, the organizational relationship itself does not directly reflect the specific behavioral patterns of users, especially when facing complex insider threat detection scenarios, too much emphasis on the weight of the organizational relationship may lead to biased judgments on the actual behavior of users. Therefore, assigning a smaller weight to organizational relationships can avoid over-reliance on the assumption of organizational structure and reduce misjudgments that may be caused by organizational changes or overlapping roles.

This setting enables the behavioral pattern graph to focus more on the user's actual operational behaviors and behavioral sequences, rather than making inferences based solely on individual behaviors or organizational hierarchies. By balancing the weights of different edges in the above method, it can reveal abnormal patterns in user behaviors, optimize the information of organizational structure, and improve the overall detection accuracy and response effect, especially for the detection of potential insider threat scenarios.

2.2.2 Behavioral Pattern Map Embedding

1) Method Design

The structure of the behavior pattern graph constructed by the above method contains a large number of nodes and edges, and the analysis using these behavior pattern graphs directly faces the problems of high computational complexity and excessive dimensionality. Therefore, in order to detect anomalies in user behavior, this study performs graph embedding on the constructed behavioral pattern graphs to transform the complex graph structure into low-dimensional vectors, which is used to extract information from enterprise log data. By analyzing these graph embedding vectors, it provides effective data for the subsequent graph neural network detection model, and then identifies abnormal patterns to support the enterprise's security monitoring and log analysis needs.

2) Algorithm Flow

Based on the user behavior pattern graph defined above, this paper adopts GraphSAGE as a feature extractor to learn the embedding of each node, and the behavior pattern graph embedding algorithm mainly includes the following steps:

Step 1: Neighbor Sampling: In user behavior pattern graph data, each node has a large number of neighbors, in order to reduce the computational complexity, GraphSAGE adopts a neighbor sampling strategy to sample only some of the neighbors of each node. Specifically, for node V, K neighbor nodes are randomly selected from its neighbor set $N(v)$, which can effectively reduce the computation. The sampled neighbors can be a fixed number or selected based on a certain probability distribution.

Step 2: Information Aggregation: The core step of GraphSAGE is the aggregation of neighbor node information. By aggregating the neighbor features of each node, the node can effectively fuse the structural information from its neighbors. Among them, Mean Aggregation (MA) is to average the features of neighboring nodes to generate a comprehensive feature vector, the formula is shown in (8).

$$h_v^{(k)} = \text{mean}\left(h_u^{(k-1)} \middle| u \in \mathcal{N}(v)\right) \tag{8}$$

Step 3: Node embedding update: After completing the neighbor information aggregation, GraphSAGE first aggregates the neighbor information of the node, and then combines the aggregation results with the node's own features, and then obtains the new embedding representation of the node through the multilayer perceptron by performing a nonlinear transformation, as shown in the formula in (9).

$$h_v^{(k)} = \sigma\left(W^{(k)} \cdot \text{concat}\left(h_v^{(k-1)}, \text{agg}\left(h_u^{(k-1)}, u \in \mathcal{N}(v)\right)\right) + b^{(k)}\right) \tag{9}$$

where $h_v^{(k)}$ denotes the embedding representation of node v in the k th layer, agg denotes the aggregation operation of the neighboring node's features, σ is the activation function (the ReLU function is used in this chapter), and $W^{(k)}$ and $b^{(k)}$ are the weight matrices and the bias terms learned during the training process.

Step 4: Loss Function and Training: The model optimization mechanism of GraphSAGE updates the network weights by minimizing the loss function, and the cross-entropy loss function is often used for model training in the node classification task, as shown in the formula (10).

$$\mathcal{L} = -\sum_{v \in V} y_v \log\left(\hat{y}_v\right) \tag{10}$$

where y_v denotes the true label of node v and $\hat{y}_v$ is the model output prediction result. Through optimization methods such as gradient descent, GraphSAGE updates the weight matrices and bias terms of all layers to finally obtain a low-dimensional vector representation of each node.

2.2.3 Algorithmic improvements

The traditional GraphSAGE algorithm uses the $\epsilon - greedy$ algorithm to balance the contradiction between "exploration and utilization". The agent selects the random strategy to explore the environment based on the probability value ϵ and selects the action with the maximum value $Q(s,a;\theta)$ through the algorithm strategy based on the probability value $1-\epsilon$. This method relies on the rationality of the ϵ setting and has poor robustness in large-scale action Spaces.

NoisyNet addresses this problem by introducing Gaussian noise in the fully connected layer

of GraphSAGE and adding randomness to the output layer of the neural network to stimulate the intelligences to explore the environment. Suppose the original linear output layer is $y = b + Wx$, and the output layer after adding noise becomes:

$$y = b + b_{noisy} \odot \varepsilon^b + \left(W + W_{noisy} \odot \varepsilon^w\right) x \tag{11}$$

where $\varepsilon^b, \varepsilon^w$ represents the random noise and $\odot$ is the dot product. In this way the action selection does not require the use of the $\epsilon - greedy$ algorithm, but instead relies directly on the output $Q(s, a; \varepsilon; \psi)$ of the neural network to select the action with the largest Q value, where ε is the random noise added to the neural network and ψ is the neural network parameter.

3 Behavioral Pattern Graph-based Network Threat Simulation and Detection Practice

3.1 Behavioral Pattern Map Construction and Behavioral Data Analysis

3.1.1 Attack Behavior Dataset Construction and Identification

Collect the economic data access behaviors of enterprise users when making economic predictions, including normal and abnormal behaviors, and establish a data set. By calculating and identifying the hidden relationships of various behaviors in the behavioral dataset, a behavioral pattern diagram is constructed. Table 1 shows the recognition effects of eight types of access behaviors. The eight types of access behaviors include "write", "read", "copy", etc. Among these behaviors, there are normal and abnormal parts. The recognition accuracy of the behavioral pattern diagram for the eight types of access behaviors ranges from 83.13% to 92.96%, the recall rate is from 80.54% to 97.38%, and the F1 value is from 82.04% to 94.13%. All three identification indicators have reached over 80%, indicating a good identification effect on user data access behavior.

Table 1: Recognition Effect of 8 Types of Access Behaviors

Access Behavior	Precision (%)	Recall Rate (%)	F1 score (%)
Write	85.81	86.74	85.79
Read	83.13	94.52	89.17
Exec	91.02	92.15	92.06
Unlink	85.09	80.54	82.31
Send	89.51	97.38	92.98
Recv	92.75	94.54	92.98
Connect	92.96	96.61	94.13
Fork	85.34	84.63	82.04
Exit	88.53	85.72	85.92

3.1.2 Normal Network Behavior vs. Threat Behavior

The network normal behavior and network threat behavior are divided into two kinds of behavioral features according to the recognition results, and they are trained separately, so as to obtain two sets of network behavior patterns about different network behavior features, which

are network normal behavior pattern and network threat behavior pattern. Figure 3 shows the difference between the two network behavior patterns. The probability of network normal behavior is [0.059,0.081], and the probability of network threat behavior is [0.064,0.086], and the side-by-side comparison reveals that the probability of network threat behavior is not much different from the probability of network normal behavior. This is related to the fact that network threat behaviors can be disguised as normal behaviors, and the specific characteristics of network threat behaviors need to be carefully analyzed. For example, the probability of occurrence of network normal behavior 7 and network threat behavior 7 is 0.077 and 0.076 respectively, and with almost the same probability of occurrence, the network threat behavior will try to modify the economic data of a certain month, while the network normal behavior just reads the data.

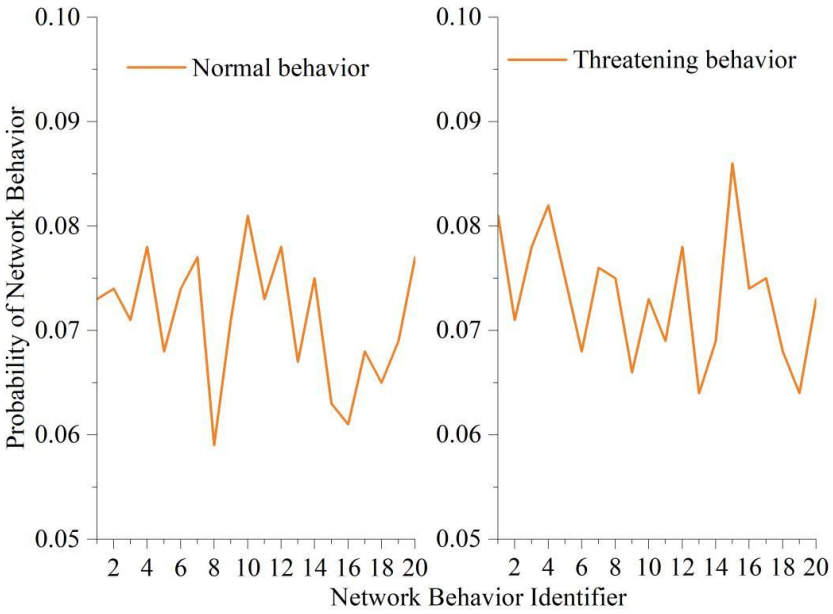

Figure 3: The differences between two types of network behavior patterns

3.2 Cyber Threat Modeling and Detection in Economic Forecasting

3.2.1 Cyber Threat Path Generation

In order to test the effectiveness of cyber threat detection in an economic forecasting environment based on behavioral pattern graphs and the behavioral pattern graph embedding algorithm GraphSAGE, the cyber threat behaviors in 3.1 are used to generate threat paths for attacking the MIDAS model, and to ensure that these threat paths have a certain attack success rate. Table 2 shows the information and probability distribution of the generated cyber threat paths. 20 cyber threat behaviors were randomly combined to generate a total of 12 cyber threat paths with lengths ranging from 3 to 5. Among them, the success rate of cyber threat path 9 reaches 0.1587, and the attack success rates of paths 2, 3, 5, 6, and 11 are also above 0.1. The higher the success rate of the cyber threat paths, the more violent the attack on the economic forecasting model MIDAS, and the higher the likelihood of economic data outsourcing.

Table 2: Network threat path information and probability distribution

Threat path	Length	Probability of success
P1(1-3-9)	3	0.0405
P2(1-20-8)	3	0.1241
P3(1-2-7)	3	0.1412
P4(2-5-3-6)	4	0.0436
P5(2-9-10-18)	4	0.1317
P6(3-10-14-16)	4	0.1495
P7(4-20-19-15)	4	0.0342
P8(3-2-6-10-20)	5	0.0548
P9(5-4-9-11-17)	5	0.1587
P10(6-19-20-16-15)	5	0.0495
P11(2-5-8-10-18)	5	0.1413
P12(4-6-7-13-16)	5	0.0605

3.2.2 Comparison of Threat Behavior Dimension Reduction Effectiveness

Behavior pattern graphs have a large structure and complex computation, in order to better detect network threat behaviors, the behavior pattern graph embedding algorithm GraphSAGE is used to downsize the threat behaviors, which facilitates the discovery of the hidden threat paths. The state transfer probability metric algorithm is chosen as a comparison algorithm to study the effect of different algorithms for dimensionality reduction of threat behaviors. Figure 4 shows the comparison of the threat behavior downgrading effect between GraphSAGE and the state transfer probability metric algorithm.The success rate of network threat behavior downgrading of GraphSAGE reaches 90.84%-97.12%, while that of the state transfer probability metric algorithm is only 76.47%-82.65%. The difference between the two types of algorithms is about 14%.GraphSAGE has better behavioral pattern graph dimensionality reduction effect, which provides higher efficiency for its accurate detection of cyber threats in the real-time economic prediction environment and discovery of hidden cyber threat paths.

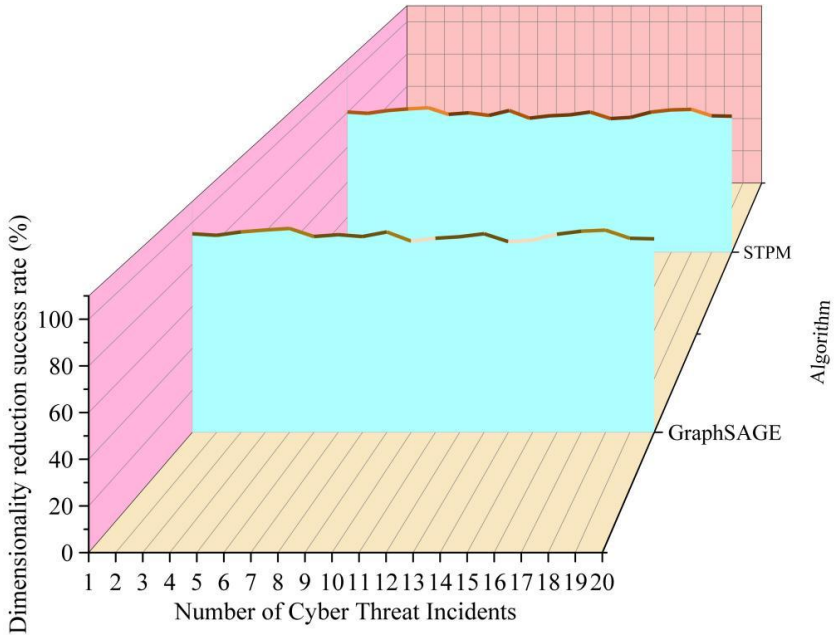

Figure 4: Threat behavior dimension reduction effects of 2 types of algorithms

3.2.3 Performance Analysis of Algorithms for Network Threat Detection

Using the 2 algorithms for 5 repetitions of the dimensionality reduced cyber threat behavior data, the statistical algorithm successfully detects the time required for 12 cyber threat paths. Table 3 shows the results of the comparison of the network threat path detection time between GraphSAGE and the state transfer probability metric algorithm. As the length of the threat path increases, the detection time of both algorithms increases. In the overall comparison, GraphSAGE successfully detects 12 network threat paths with an average time of 1.82ms-9.18ms, while the average time of the comparison algorithms reaches 7.07ms-15.94ms, which is much more time-consuming than that of GraphSAGE.GraphSAGE is more effective at dimensionality reduction of behavioral pattern graphs, which results in the ability of this algorithm to quickly detect hidden cyber threat paths and improve the security of real-time economic prediction environments.

Table 3: Comparison of the detection duration of algorithmic network threat paths

Object of detection	GraphSAGE			State transition probability measurement		
	Shortest Time Taken (ms)	Longest duration (ms)	Average value (ms)	Shortest Time Taken (ms)	Longest duration (ms)	Average value (ms)
P1	1.54	2.09	1.82	5.57	8.57	7.07
P2	1.67	2.35	2.01	5.76	8.82	7.29
P3	1.81	2.72	2.27	5.84	9.09	7.47
P4	3.75	5.48	4.62	7.78	11.95	9.87
P5	3.92	5.73	4.83	7.95	12.62	10.29
P6	4.01	5.82	4.92	8.44	12.19	10.32
P7	4.17	5.99	5.08	8.72	12.46	10.59
P8	7.34	9.24	8.29	10.37	15.78	13.08
P9	7.58	9.31	8.45	10.61	16.78	13.70
P10	7.93	9.47	8.70	11.96	17.54	14.75
P11	8.30	9.83	9.07	12.73	17.93	15.33
P12	8.41	9.95	9.18	13.45	18.42	15.94

4 Conclusion

In this paper, we use GraphSAGE, a behavioral pattern graph embedding algorithm, to downscale user economic access behavior data to quickly and accurately detect hidden cyber threats in real-time economic forecasting environments. Compared with the 76.47%-82.65% success rate of similar algorithms in downscaling, GraphSAGE downscales behavioral pattern graphs with a high success rate of 90.84%-97.12%. Supported by the high success rate of dimensionality reduction, the algorithm takes only 1.82ms-9.18ms on average for five successful detections of hidden network threat paths. Good cyber threat detection is of high value for the security of real-time economic forecasting environments. In the future, it can also be added to extend the quarterly GDP growth rate real-time prediction to annual GDP growth rate real-time prediction, further expanding the scope of application of the behavioral pattern graph embedding algorithm GraphSAGE.

About the Author

Mingming Cao was born in Harbin, Heilongjiang, P.R. China, in 1989.I am currently a lecturer at Harbin Finance University and a doctoral student at the School of Economics and Business Administration, Heilongjiang University. With years of experience working in the banking industry, she integrates theoretical knowledge with practical work experience in teaching related courses. Key courses taught include *Finance*, *Microeconomics*, and *Macroeconomics*. She has led or participated in multiple educational reform research projects and philosophy and social sciences planning projects.

References

[1] Liu, W. L., & McKibbin, W. J. (2025). Long-Term Projections of the World Economy. CAMA Working Papers, (2025-31).

[2] Garnitz, J., Lehmann, R., & Wohlrabe, K. (2019). Forecasting GDP all over the world using leading indicators based on comprehensive survey data. Applied Economics, 51(54), 5802-5816.

[3] JIANG, T., Yan-Jun, W. A. N. G., Jia-Shuang, Y. U. A. N., Ying, C. H. E. N., Xiang, G. A. O., Cheng, J. I. N. G., ... & Cheng-Yi, Z. H. A. O. (2018). Projection of population and economy in the Belt and Road countries (2020–2060). Advances in Climate Change Research, 14(2), 155.

[4] Wang, S. (2021). An interview with Shouyang Wang: research frontier of big data-driven economic and financial forecasting. Data Science and Management, 1(1), 10-12.

[5] Okoh, O. F., & Grace, I. (2022). Mathematical modeling and machine learning for economic forecasting: A hybrid approach to predicting market trends. Acta Electronica Malaysia, 6(1), 07-15.

[6] Shoja, M., & Soofi, E. S. (2017). Uncertainty, information, and disagreement of economic forecasters. Econometric Reviews, 36(6-9), 796-817.

[7] Claveria, O., Monte, E., & Torra, S. (2019). Economic uncertainty: a geometric indicator of discrepancy among experts' expectations. Social Indicators Research, 143(1), 95-114.

[8] Musaev, A., & Grigoriev, D. (2025). Ensemble Multi-Expert Forecasting: Robust Decision-Making in Chaotic Financial Markets. Journal of Risk and Financial Management, 18(6), 296.

[9] An, J., & Dorofeev, M. (2019). Short-term foreign exchange forecasting: decision making based on expert polls. Investment Management & Financial Innovations, 16(4), 215.

[10] O'Mahony, T., Luukkanen, J., Vehmas, J., & Kaivo-oja, J. R. L. (2024). Time to build a new practice of foresight for national economies? Ireland, and uncertain futures in forecasts and scenarios. foresight, 26(1), 18-34.

[11] Channe, P. S. (2024). The Impact of AI on Economic Forecasting and Policy-Making: Opportunities and Challenges for Future Economic Stability and Growth. York University.

[12] Wang, L., & Zhao, L. (2022). Digital economy meets artificial intelligence: forecasting economic conditions based on big data analytics. Mobile information systems, 2022(1), 7014874.

[13] Tang, Y. M., Chau, K. Y., Li, W., & Wan, T. W. (2020). Forecasting economic recession through share price in the logistics industry with artificial intelligence (AI). Computation, 8(3), 70.

[14] Umronov, E., Kadirov, A., Abdujabborov, A., Muydinov, I., Karimova, M., & Sobirjonov, T. (2024, May). Economic levels forecasting system by evaluating with more accuracy using ml, dl and ai systems. In 2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 1027-1031). IEEE.

[15] Shawon, R. E. R., Rahman, A., Islam, M. R., Debnath, P., Sumon, M. F. I., Khan, M. A., & Miah, M. N. I. (2024). AI-driven predictive modeling of us economic trends: insights and innovations. Journal of Humanities and Social Sciences Studies, 6(10), 01-15.

[16] Challoumis, C. (2024). HOW CAN AI PREDICT ECONOMIC TRENDS IN THE MONEY CYCLE?. evolution.

[17] Khan, A. A., Jamshed, H. U. M. A., Ahmed, S., Iqbal, S., Mansoor, Y., & Waheed, U. (2024). Smart Growth Predictions: Deep Learning Applications in Economic Forecasting. Technical Journal, 29(02), 61-68.

[18] Deliparaschos, K. M., Michail, K., & Zolotas, A. C. (2020). Facilitating autonomous systems with AI-based fault tolerance and computational resource economy. Electronics, 9(5), 788.

[19] Dixit, R., Chinnam, R. B., & Singh, H. (2020, March). Artificial intelligence and machine learning in sparse/inaccurate data situations. In 2020 IEEE Aerospace Conference (pp. 1-8). IEEE.

[20] Wang, H., Hao, L., Sharma, A., & Kukkar, A. (2022). Automatic control of computer application data processing system based on artificial intelligence. Journal of Intelligent Systems, 31(1), 177-192.

[21] Gadde, H. (2022). AI in Dynamic Data Sharding for Optimized Performance in Large Databases. International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence, 13(1), 413-440.

[22] Kokogho, E., Odio, P. E., Ogunsola, O. Y., & Nwaozomudoh, M. O. (2024). AI-powered economic forecasting: Challenges and opportunities in a data-driven world. International Journal Management and Organizational Research, 3(6), 74-83. https://doi.org/10.54660/IJMOR.2024.3.6.74-83.

[23] Fu, J., Zhou, X., & Mei, G. (2022). Internet Digital Economy Development Forecast Based on Artificial Intelligence and SVM-KNN Network Detection. Computational Intelligence and Neuroscience, 2022(1), 5792694.

[24] Kuzior, A., Brożek, P., Kuzmenko, O., Yarovenko, H., & Vasilyeva, T. (2022). Countering cybercrime risks in financial institutions: Forecasting information trends. Journal of Risk

and Financial Management, 15(12), 613.

[25] Qasaimeh, M., Hammour, R. A., Yassein, M. B., Al-Qassas, R. S., Torralbo, J. A. L., & Lizcano, D. (2022). Advanced security testing using a cyber-attack forecasting model: A case study of financial institutions. Journal of Software: Evolution and Process, 34(11), e2489.

[26] Yusof, Z. B. (2024). Exploration of advanced persistent threats: techniques, mitigation strategies, and impacts on critical infrastructure. International Journal of Advanced Cybersecurity Systems, Technologies, and Applications, 8(12), 1-9.

[27] Lenel, L., Köster, R., & Fritsche, U. (2020). Introduction (Futures Past. Economic Forecasting in the 20th and 21st Century). Futures Past. Economic Forecasting in the 20th and 21st Century.

[28] He, Y., & Li, X. (2022). Feasibility of economic forecasting model based on intelligent algorithm of smart city. Mobile Information Systems, 2022(1), 9723190.

[29] Hopp, D. (2022). Economic nowcasting with long short-term memory artificial neural networks (LSTM). Journal of Official Statistics, 38(3), 847-873.

[30] Amini, A., & Kalantari, R. (2024). Gold price prediction by a CNN-Bi-LSTM model along with automatic parameter tuning. Plos one, 19(3), e0298426.

[31] Atif, D. (2025). Enhancing Long-Term GDP Forecasting with Advanced Hybrid Models: A Comparative Study of ARIMA-LSTM and ARIMA-TCN with Dense Regression. Computational Economics, 65(6).

[32] Al-Karkhi, M. I., & Rządkowski, G. (2025). Innovative machine learning approaches for complexity in economic forecasting and SME growth: A comprehensive review. Journal of Economy and Technology, 3, 109-122.

[33] Ivashchenko, A., & Ivashchenko, T. (2025). THE POTENTIAL OF HYBRID LSTM-GENERATIVE AI ECO-MODEL IN FORECASTING FINANCIAL AND ECONOMIC INDICATORS. COLLECTION OF PAPERS NEW ECONOMY, 171.

[34] Ravi, K., Bhuria, R., Sarasu, P., Khan, S., Kaur, B., & Bhoyar, M. (2025, February). AI-Driven Predictive Modeling for Real-Time Economic Forecasting. In 2025 International Conference on Technology Enabled Economic Changes (InTech) (pp. 1477-1481). IEEE.

[35] Soundenkar, S., Bhosale, K., Jakhete, M. D., Kadam, K., Chowdary, V. G. R., & Durga, H. K. (2024). AI Powered Risk Management: Addressing Cybersecurity Threats in Financial Systems. Library of Progress-Library Science, Information Technology & Computer, 44(3).

[36] Niu, W., Zhang, X., Yang, G., Chen, R., & Wang, D. (2017). Modeling attack process of advanced persistent threat using network evolution. IEICE TRANSACTIONS on Information and Systems, 100(10), 2275-2286.

[37] Lerums, J. E., La'Reshia, D. P., & Dietz, J. E. (2018, May). Simulation modeling cyber threats, risks, and prevention costs. In 2018 IEEE International Conference on

Electro/Information Technology (EIT) (pp. 0096-0101). IEEE.

[38] Moustafa, N., Misra, G., & Slay, J. (2018). Generalized outlier gaussian mixture technique based on automated association features for simulating and detecting web application attacks. IEEE Transactions on Sustainable Computing, 6(2), 245-256.

[39] Moskal, S., Yang, S. J., & Kuhl, M. E. (2018). Cyber threat assessment via attack scenario simulation using an integrated adversary and network modeling approach. The Journal of Defense Modeling and Simulation, 15(1), 13-29.

[40] Yeboah-Ofori, A., & Islam, S. (2019). Cyber security threat modeling for supply chain organizational environments. Future internet, 11(3), 63.

[41] Ajmal, A. B., Khan, S., Alam, M., Mehbodniya, A., Webber, J., & Waheed, A. (2023). Toward effective evaluation of cyber defense: threat based adversary emulation approach. IEEE Access, 11, 70443-70458.

[42] Rauf, H., Shah, S. I. H., Ali, T., Gul, H., & Soomro, M. (2025). USING GENERATIVE AI FOR SIMULATING CYBER SECURITY ATTACKS AND DEFENSE MECHANISMS: A NEW APPROACH TO AI-DRIVEN CYBER THREAT MODELING. Spectrum of Engineering Sciences, 3(3), 361-381.

[43] Choudhary, C., Chawla, J., Alkhayyat, A., Prabha, C., & Gulhane, M. (2025, August). AI-Driven Threat Simulation in Wireless Networks using Advanced Machine Learning and Hyperparameter Optimization. In 2025 5th International Conference on Soft Computing for Security Applications (ICSCSA) (pp. 197-202). IEEE.