Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 7600754, 12 pages
https://doi.org/10.1155/2022/7600754

Research Article

Identifying Key Risk Factors in Air Traffic Controller Workload by SEIR Model

Yuming Heng,[1,2] Minggong Wu,[1,2] and Xiangxi Wen [1,2]

[1]*Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China*
[2]*National Key Laboratory of Air Traffic Collision Prevention, Xi'an 710051, China*

Correspondence should be addressed to Xiangxi Wen; wxxajy@163.com

Received 6 September 2022; Revised 26 September 2022; Accepted 5 October 2022; Published 19 October 2022

Academic Editor: Chiara Bedon

To address the controller workload with the forecast, the capacity of the air traffic management system is effectively enhanced. It should be based on a specific analysis of the controller workload. In the current controller workload studies, there is no clear means to analyze the process of controller workload development propagation. In this paper, we propose a new method for analyzing the factors influencing the controller workload. This method takes into account the influence of various situations in the actual work of controllers and objectively quantifies the complexity of work conditions. A complex network is constructed by treating various factors as nodes and the complexity relationships between these nodes as edges. The complexity network was then tested using the contagion model. The sum of the number of times of infecting other nodes and being infected in the detection result was defined as the infection capacity of the nodes, and the point with the strongest infection capacity was controlled and analyzed. The results show that the point with the strongest infection capacity is the key factor for the development of controller workload generation. In addition, the analysis of the key factors using a backpropagation neural network shows that the prediction of the controller workload can be made by the key factors. It will provide a new effective method to control controller workload and improve air traffic control capability.

1. Introduction

According to the International Air Transport Association (IATA) forecast made in 2018, China will overtake the United States as the world's largest aviation market by the mid-2020s, with the Chinese civil aviation market reaching 1.6 billion passengers per year by 2037 [1]. China's air traffic management system will then face a large impact, and the rapid increase in the number of flights will likely lead to a serious overload on controllers. The likelihood of air traffic safety incidents is also likely to increase. In addition to this, excessive workload can lead to a reduction in the capacity the controllers can manage, leading to flight delays and affecting air traffic [2]. Therefore, this paper intends to analyze and study the controller workload and preventive measures to cope with the development factors of controller workload.

Controller workload refers to the physical and psychological stress experienced by controllers during their work to maintain the safe, efficient, and rational operation of guaranteed air traffic. As an important support for the air traffic management system, the workload of controllers is an important factor affecting flight safety. To this end, Stein E of the Federal Aeronautics Administration (FAA) was the first to propose the air traffic workload input technique (ATWIT) in 1985 [3]. It was followed by the task load index method (NASA-TLX) in 1988 by Hart at the National Aeronautics and Space Administration Ames Research Centre (NASA AMES) [4]. Reid also proposed the subjective workload assessment technique (SWAT) method in 1988 [5]. In addition, the International Civil Aviation Organization (ICAO) document 9426 recommends two assessment methods: the "DORATASK" method proposed by the British Operations and Analysis Council [6] and the "MBB" method proposed by the German scientist Messerschmidt [7]. These models have been endorsed by ICAO since their presentation. In particular, the "DORATASK" method has

been documented by ICAO and recommended to civil aviation authorities worldwide. The "DORATASK" method measures the controller's workload through the controller's working hours. It divides the controller's working time into "visible working time, invisible working time, and response time." Visible work is defined as checking process sheets, land and air calls, etc. Invisible work is defined as controller thinking, etc. Recovery time is another part of the equation. It has no clear characteristics; however, it is essential for the controller's status maintenance. The "MBB" method is similar; however, the visible part of the work is refined, and the tasks are clearly categorized. Other methods, such as "FAA," "ATWIT," and "NASA-TLX," are different. They do so by means of questionnaires. They classify the load of controllers into several classes. The controller's workload is confirmed by filling out a questionnaire at the end of the work day or at regular intervals. These methods do not specify the magnitude of the controller's load and rely entirely on the controller's own perceptions. Therefore, they are not as applicable as the "DORATASK" method. It indirectly proves that methods that do not have criteria and rely exclusively on the controller's own perceptions have certain drawbacks. In other words, it is more subjective and more influenced by the controller's human factors. This drawback can be remedied to a certain extent by having clear criteria for the completion of the questionnaire.

With the development of China's civil aviation industry, scientists have improved the above-mentioned methods in combination with China's actual situation, and they have achieved better results [8–10]. This work further demonstrates the universality and soundness of the approach described above.

In terms of objective evaluation methods, the electroencephalogram (EEG) is a direct response to changes in brain activity. It directly reflects changes in brain activity and is generally considered the most effective method of detecting human fatigue, and it is naturally a hot topic in the study of controller fatigue at home and abroad. It is also a hot spot for the study of fatigue in controllers at home and abroad. Lal and Craig [11] and Eoh et al. [12] found that δ and θ wave activities increased significantly in the fatigue state by means of statistical analysis, and the ratios of β and $(\alpha + \theta)/\beta$ were significantly different. ARICO [13] proposed a controller brain fatigue coefficient based on the EEG signals collected by school control trainees in a simulator and established a related controller workload model. DASARI and KRISHNA [14, 15] continuously monitored the controllers' EEG signals through simulation experiments, and the experimental results showed that the controllers began to show fatigue at 70 min, and their judgement reaction ability gradually deteriorated. A team of scientists has demonstrated that relevant indicators can be extracted from physiological indicators, such as controller brain wave signals, skin electrical signals, and electrocardiograph (ECG) signals, for controller workload assessment [16–18]. Similarly, scientists from various countries have proposed controller workload evaluation methods based on facial features, voice features, and so on [19, 20].

On this basis, scientists in various countries began to focus on the factors affecting the control process and tried to prevent and adjust them. Edwards first developed the principle of a specific system interface for "people" working safely in 1972, known as the software, hardware, environment, liveware (SHEL) model. Hawkins then described the model in diagrams [21]. The model can be used to analyze human errors that occur because of mismatches between the elements of the interface. Errors tend to occur at the central point of contact between people and hardware, software, environment, and other people. The model is a direct guide to safety as it graphically depicts the vulnerable aspects of modern production, and the interfaces described are not only found on the front line but at all levels of the production organization. Hence, the model is universally relevant. For this reason, ICAO has included it as a safety management analysis method [22]. Shorrock derived the causes of control errors based on psychological principles by analyzing dangerous approach events that occurred during 1995–1997 [23]. The Reason model was developed in 1990 by Reason, a professor at the University of Manchester, in his book "Human Error," [24] which argues that accidents are not usually caused by isolated factors but by a combination of system defects. Defects at all levels of the organization do not necessarily lead to an accident. When defects occur at all levels simultaneously (like light penetrating cheese), the system loses its multilevel defense, and an accident occurs. However, the Reason model is an abstract theory. It does not indicate what exactly the "holes" in the different levels of defect cheese are, let alone how these "holes" are to be found in an accident investigation. Therefore, since 1997, Wiegmann and Shappell have been building a human factors analysis and classification system (HFACS) based on the Reason model [25]. HFACS has achieved certain results. However, it is tailored to flight (pilot-centric) and is not fully applicable to controller-centered air traffic control activities. Then, in 2021, Di Mascio et al. have conducted research on the impact of work on controller load. By simulating a busy three-runway international airport, he concluded that the maximum traffic manageable from the airside would generate an unacceptable workload for tower controllers [2]. Bedon and Mattei have conducted a study on the effects of the environment on the human body [26, 27]. In their study, they found that the psychological and mental state of a person is influenced by the structure when the environment in which he or she is placed is a glass structure. Since the working environment of the tower in which the controllers work is almost an all-glass structure, the working environment can impose a severe psycho-physical load on the controllers.

In conjunction with the above studies, we know that while many current studies have identified some of the factors influencing controller workload, no studies have been able to thoroughly consider all factors together. In particular, there are gaps in the understanding of the propagation of controller workload. Current research addressing how multiple influences on controller workload work together for controllers remains flawed. The propagation of workload needs to be studied more extensively and

networked in the context of the work situation. With the development of the times, controllers are facing increasingly diverse work. They, as individuals, also face a variety of events before and after work. However, there is a lack of clear evaluation of the many factors that affect the controller workload, and it is not clear exactly which of these factors will have a larger impact and which will not. The number of factors that may affect the controller workload is also increasing. From a network perspective, the various influencing factors work together toward the ultimate goal of controller workload. Therefore, from this perspective, the propagation of influencing factors in the network can be understood to the maximum extent only by introducing the complex network theory. Complex networks allow us to analyze the interrelationships between factors and the degree of influence they have on each other. In the field of civil aviation, complex networks were first applied to the field of air traffic control by Laudeman et al. to assess the control difficulty and controller workload [28]. Wu et al. used the identification of air traffic complexity to evaluate flight safety and quality [29]. Zhang et al.'s team used complex networks to perform a risk analysis of the flight network, which helped to reduce risk propagation and delays [30]. Therefore, it is feasible to use the complex network theory to analyze civil aviation safety issues. The previous studies mainly focused on the problems of civil aviation flight safety and delays using the complex network theory. We then want to use complex networks to make an identification of the key factors that affect the controller workload.

From the above, we concluded that the controller workload is not solely determined by one factor because of the interaction of multiple factors that affect the controller workload. The complex network theory can provide a new way of solving the problem. In this paper, we propose a new approach based on the complex network theory and contagion model to analyze the key factors affecting the controller workload. Usually, the contagion models are many, and typically, they include the susceptible-infected-recovered (SIR) model, susceptible-infected-exposed-recovered (SIRS) model, and so on. In practice, the propagation mechanism of the SEIR model and the mechanism of action of factors affecting the controller workload are most similar. In this paper, the mechanism of action of factors affecting the controller workload is explained using the SEIR model.

This paper proposes the use of actual data from China's aviation operations, the use of correlation coefficients to establish complex networks, the introduction of the SEIR propagation model to analyze the propagation effects of each node in the network, and the identification and control of key nodes through dynamics principles. The purpose of this paper is to discuss the causal factors and propagation paths of controller workload in a more comprehensive and thorough manner and to provide support for the controller workload control. It is found that the maximum information coefficient (MIC) can effectively reflect the nonlinear relationship between features. Complex networks can clearly show the relationship between features. The identification of key nodes using propagation model analysis becomes very simple as the features and their relationships are determined in advance. The network is tested using the SEIR model, where the more infectious nodes are more important in the network. The identified nodes are also controlled by means of improving the probability of the corresponding nodes in the model to prove that the identified nodes are critical nodes. Finally, the nodes were verified again using the backpropagation (BP) neural network. If the identified node is a critical node, then it should correspond to a prediction result that is better than the other nodes. In summary, this method can accurately identify the key factors affecting the controller workload and provide support for controlling the controller workload.

2. Network Construction

There is a complex relationship between the factors, through a complex network composed of factors (nodes) and correlations between factors (connected edges), denoted as $G = (V, E, W)$, where $V = \{v_1, v_2, v_3, \ldots, v_n\}$ is the set of nodes in the network corresponding to each influencing factor. $E = \{e_1, e_2, e_3, \ldots, e_n\}$ represents the set of connected edges, which reflect the interrelationship between factors. $W = \{w_1, w_2, w_3, \ldots, w_n\}$ is the set of edge weights of the network, and the weights of the connected edges reflect the degree of correlation between the factors in the network. In the construction of the network, it is necessary to determine the nodes and connected edges.

2.1. Feature Selection. As air traffic control is a "human-in-the-loop" system, there are many factors that have an impact on controller workload and are closely related. Therefore, a reasonable desection of the influencing factors is the key to the analysis of controller workload. This paper classifies the various influencing factors involved in controller workload with reference to the relevant requirements and research of civil aviation on controller workload. All the influencing factors are classified into four categories: human factors, operational environment factors, airspace posture factors, and organizational implementation factors. In total, there are 30 influencing factors. Data were collected from civil aviation controllers in the form of a questionnaire survey on the above 30 items and controller workload. The risk was taken in the range of [1, 10], with higher values indicating a greater impact on control compliance. Some of the data are shown in Table 1.

As the data obtained are of a fixed order type, they should be analyzed before the complex network is constructed. Cronbach's reliability coefficient was first proposed by the American educator Lee Cronbach in 1951 [31]. Cronbach's reliability coefficient α usually takes a value between [0, 1]. A value of 0.6 or below is generally considered to have insufficient confidence. A value of 0.7-0.8 is considered to have good confidence. A value of more than 0.8 indicates very good confidence. Equation (1) is used for the calculation of Cronbach's reliability coefficient α.

$$\alpha = \frac{m}{m-1}\left(1 - \frac{\sum_{i=1}^{m} \sigma_i^2}{\sigma_N^2}\right)\sqrt{a^2 + b^2}, \tag{1}$$

Table 1: Controller workload impact index.

Category	Item	No.	Selected values		
Human factors	Physiological (diurnal) rhythms	1	8	8	7
	Alcohol/drug use status	2	8	9	6
	Dietary status	3	2	3	3
	Sleep status	4	8	8	9
	Work proficiency	5	7	6	8
	Family social activities	6	9	8	8
	Control shifts	7	8	8	6
	Pilot and crew	8	8	8	9
	Length of duty	9	9	10	9
	Stress reaction ability	10	8	7	8
	Dietary status	11	9	8	9
	Psychological qualities	12	1	2	1
Operational environmental factors	Human-machine interface friendliness	13	2	1	3
	Equipment ergonomics	14	8	7	9
	Equipment condition	15	8	8	7
	Control room environment	16	9	10	9
	Rest room environment	17	9	9	8
	Geographical environment	18	3	4	3
Airspace situational factors	Weather conditions	19	5	2	2
	Military aviation activities	20	8	7	8
	Flight flow	21	8	8	9
	Route structure	22	8	6	8
	Emergencies	23	2	3	5
	Sector structure	24	6	7	6
Organizational implementation factors	Duty system	25	8	9	8
	Culture	26	4	5	3
	Operational training	27	8	8	7
	Site management	28	9	7	7
	Operating procedures	29	8	7	7
	Staffing	30	4	3	4
	Controller workload	31	9	8	8

where m is the number of items in the questionnaire, σ_i is the variance of the score of item i, σ_N is the variance of the scores of all questions, and α is the reliability coefficient of the questionnaire.

Upon substituting the relevant data into (1), the credibility coefficient of this questionnaire can be derived as $\alpha = 0.79$. It proves that this questionnaire has a high degree of credibility and can be used as a basis for complex network construction.

2.2. Complex Network Construction. Complex networks are built with the influences as nodes and the interrelationships between the influences as connected edges. A complex network built on this basis is more objective. Common network building methods include the empirical network building method, time series network building method, and correlation coefficient network building method.

2.2.1. Empirical Network Construction Method. The strategy of the empirical network-building method is to have experts score the indicators of the influencing factors in Table 1 through questionnaires and other forms of research and judgement. Judgements are made on the basis of the experts' experience in relation to the influencing factors. If the experts believe that they are related, then they are connected, and if not, then they are not.

The empirical method is prone to errors and omissions in the process of using it. It is also too subjective and can lead to poorly constructed edges and poorly connected nodes. The network model constructed is different from the actual one and does not reflect the actual situation. In practice, experts' opinions are not always correct. When experts make judgements based on experience, they are likely to make misjudgments for their own reasons. In this case, the method would be highly problematic. It would have an impact on the overall study.

2.2.2. Time Series Network Construction Method. The strategy of the time series netting method is to analyze and reconstruct the data using the time series phase space reconstruction method in the first place. The time series correlation coefficients are then calculated. Next, the adjacency matrix is constructed from the time series correlation coefficients. Finally, the adjacency matrix is used to construct the connected edges between the corresponding nodes, and the weights of the connected edges are usually taken as the correlation coefficients to build a complex network.

The time series correlation coefficient method can lead to a large number of isolated nodes and independent networks in the process of network building. It, in turn, affects the analysis of the interrelationships between the nodes. There is a large discrepancy between the results of the network building and the actual operating conditions.

2.2.3. Correlation Coefficient Method. The correlation coefficient method is used to calculate the correlation coefficients between the influencing factors and use them to construct a continuous edge. The common methods are the Spearman correlation coefficient method and Pearson correlation coefficient method.

The Pearson correlation method has been found to be suitable for fixed percentage changes in indicators and tends to underestimate the correlation between data. The Spearman correlation method is difficult to distinguish between direct and indirect correlations and often results in "false correlations" in the network.

In summary, this paper uses the maximum information coefficient to analyze the correlation between the characteristic variables.

MIC is computed using mutual information and grid partitioning methods as follows:

Assuming a finite set of ordered pairs $D\{(x_i, y_i), \quad i = 1, 2, \ldots, n\}$, the scatterplot consisting of x_i and y_i in D is gridded $x \times y$ and the mutual information $I(X,Y)$ in each grid is calculated. Mutual information $I(X,Y)$ is calculated as follows: for two random variables $X = \{x_i, \quad i = 1, 2, \ldots, n\}$ and $Y = \{y_i, \quad i = 1, 2, \ldots, n\}$, n being the number of samples. Equation (2) is used for the calculation of $I(X,Y)$.

$$I(X,Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \tag{2}$$

where $p(x, y)$ is the joint probability density of X and Y and $p(x)$ and $p(y)$ are the edge probability densities of X and Y. The maximum value $\max(I(X,Y))$ of $I(X,Y)$ under different division methods is selected as the mutual information value for dividing $x \times y$ grids. Then, the maximum information coefficient $\mathrm{mic}(I(X\colon Y))$ is the result obtained after its normalization. Equation (2) is used for the calculation of $\mathrm{mic}(I(X\colon Y))$.

$$\mathrm{mic}(X\colon Y) = \max_{|X||Y|<B} \frac{\max(I(X,Y))}{\log(\min(|X|, |Y|))}, \tag{3}$$

where B is the upper limit of the grid division $x \times y$, which is a function of the number of samples n. Experience gives the best results when B is taken as $n^{0.6}$. The same strategy is used in this paper.

$\mathrm{mic}(I(X\colon Y)) \in [0, 1]$. When the maximum mutual information coefficient is 0, it means that the two features are not correlated and are independent of each other. When it is 1, it means that the two features are fully correlated and can be substituted for each other. The larger the maximum mutual information coefficient, the stronger the correlation between features. The maximum information coefficient is used here to evaluate the correlation between features and to determine whether two nodes i and j (between features) have connected edges.

When $\mathrm{mic}(i, j) < 0.2$ is chosen, the correlation between the two features is considered very small, and there are no connected edges between the two feature nodes. When $\mathrm{mic}(i, j) \geq 0.2$, the correlation between the two nodes is considered connected, and there are connected edges. The edge weight between nodes is set to $\mathrm{mic}(i, j)$. Based on this, we can construct a weighted undirected feature network. The obtained feature network adjacency matrix is A, and the weighting matrix is B. Equations (4) and (5) are used to calculate the adjacency matrix (A) and the weighting matrix (B).

$$A = \begin{vmatrix} 0 & a_{12} & \cdots & a_{1N} \\ a_{21} & 0 & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & 0 \end{vmatrix}, \begin{cases} a_{ij} = 0 \text{ if } \mathrm{mic}_{ij} < 0.2, \\ a_{ij} = 1 \text{ if } \mathrm{mic}_{ij} \geq 0.2, \end{cases} \tag{4}$$

$$B = \begin{vmatrix} 0 & b_{12} & \cdots & b_{1N} \\ b_{21} & 0 & \cdots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \cdots & 0 \end{vmatrix}, \begin{cases} b_{ij} = 0 \text{ if } \mathrm{mic}_{ij} < 0.2, \\ b_{ij} = 1 \text{ if } \mathrm{mic}_{ij} \geq 0.2. \end{cases} \tag{5}$$

Through computational analysis, we obtain the network as shown in Figure 1.

As shown in Figure 1, the complex network takes on a shape similar to concentric circles. With node 31 as the center, the nodes are distributed from the inside out according to their degree values. In the whole network, node 31 has the highest degree value and the best connectivity. It proves that the 30 factors selected to influence controller compliance all have an impact on controller workload, and the selected metrics are reasonable.

While reflecting the relationships between nodes, complex networks are also capable of reflecting different properties of different nodes in different ways:

(1) According to the position of the nodes in the network, the impact value of each node gradually decreases from the inside to the outside. The outermost node, node 13, has a mean impact value of 2.7, while the inner node, node 9, has a mean impact value of 8.3. It means that the outer nodes have less impact on the controller workload than the inner nodes, and the closer the node is to the center, the more impact it has on the controller workload.

(2) Nodes in the inner layer of the network have larger mean values and larger standard deviations of impact values in the inner layer compared to those in the outer layer. It indicates that the nodes in the inner layer represent characteristics that vary more across conditions and have a greater impact on the controller workload. In contrast, the nodes in the outer layer have smaller mean values and smaller standard deviations, and their variability is smaller. For example, in node 13, "Human-machine interface

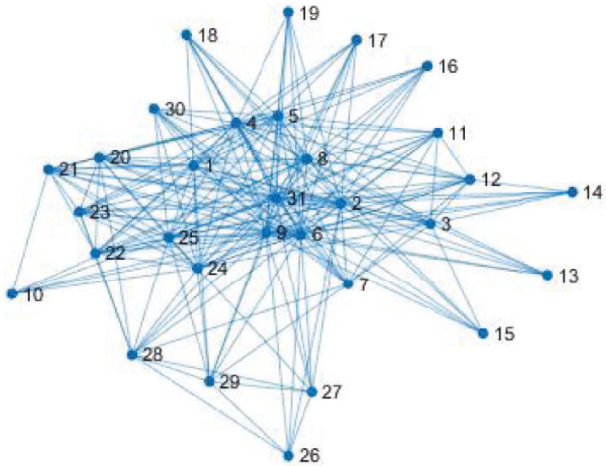

Figure 1: Complex network based on MIC.

friendliness," if the interface design of the operating system is not smooth and appropriate, it will lead to a certain amount of controller workload. However, it does not become a factor that causes a surge in controller workload.

3. Improved SEIR Model Application

The SEIR model is an improvement on the traditional SIR model. It is traditionally used to analyze infectious diseases that are latent, infectious, and immune after cure, such as SARS, avian influenza, and COVID-19. It has been generalized and combined with the complex network theory to analyze the transmission capacity of networks and evaluate node performance.

3.1. Improving the SEIR Model. To adapt the SEIR model to the actual civil aviation operation control, adjustments are made in conjunction with the actual civil aviation operation control. The specific improvement measures are as follows:

3.1.1. Susceptible Node (S). A susceptible node is defined in the model as a node that is connected to an infected node and is susceptible to infection by the infected node, becoming a latent node. In practice, it means that it is temporarily unaffected by other adverse factors. However, there is a risk of interference from other factors, which may further interfere with controller workload and other factors, leading to a vicious circle and overloading of controllers.

The probability of infection between nodes is reflected in the model by the weights of the edges attached to the nodes. The higher the edge weight, the higher the probability of infection. (6) is used for calculating the probability of the infection of node P_{SE}.

$$P_{SE} = \rho, \tag{6}$$

where P_{SE} is the probability of a node changing from a susceptible state to a latent state.

3.1.2. Exposed Nodes (E). An exposed node is the one that has a potential risk of transmission. In practice, it means that it has been affected by a factor but has not actually placed an additional burden on the controller's work, nor has it had a negative outcome on other factors.

At this stage, the node has the potential to recover, enter the contagious phase, and remain latent at the same time. Set the probability of recovery P_{ER} at 10%. Equations (7) and (8) are used to calculate the probability of entering the contagious phase P_{EI} and the probability of remaining latent P_{EE}.

$$P_{EI} + P_{EE} + P_{ER} = 100\%, \tag{7}$$

$$P_{EI} = \exp\left(S_i^2\right), \tag{8}$$

where P_{ER} is the probability of the node changing from the latent state to the immune state. P_{EE} is the probability of maintaining the latent state. P_{EI} is the probability of the node changing from the latent state to the infectious state. S_i^2 is the variance of the node.

3.1.3. Infected Nodes (I). An infected node is the one that has been affected and will affect other nodes. Infected nodes include two categories in the model, namely initially infected nodes and infected nodes. The initially infected nodes refer to the earliest nodes that are in the infected state.

The presence of an initially infected node should be inversely proportional to the standard deviation of the node. The larger the standard deviation, the greater its impact on the controller workload. Equation (9) is used for the calculation of the initial probability of infection P_I of a node.

$$P_I = \exp\left(-S_i\right), \tag{9}$$

where P_I denotes the probability of a node becoming initially infected and S_i is the standard deviation of the node.

An infected node represents a node that has been infected by another node and has moved from the latent to the infected stage, and it is infectious and self-healing. The probability of recovery P_{IR} is set at 20% in the model. In practice, it is shown as being affected by a link, but relevant measures are taken to adjust it and terminate the miscommunication without causing additional disruption to the controller. The probability of a node remaining infected P_{II} is 80%.

3.1.4. Recovered Nodes (R). A recovered node refers to a node that is in an immune state and no longer has the ability to be infected, and it is also protected from infection by the spreading node I. In practice, it is shown to have taken relevant measures for factor control, and the impact has been contained. Its interference with the controller workload has been circumvented.

3.2. Improving the SEIR Model. We performed 500 propagation experiments following the propagation rules

described in 2.1. The results obtained after averaging the propagation results are shown in Figure 2.

The trend of the curve smoothed out as the transmission time progressed. With susceptibility, the node curve gradually stabilized and reached immunity at a later stage. The peak value of the transmission node curve is 12.05, and the time taken to reach the peak is 11.

4. Nodes Identification and Control

4.1. Key Nodes Identification. In previous research, the identification of key nodes was mainly done by the characteristics of the nodes in a complex network. For example, the degree value of the node, point strength, and other indicators are judged. Although this method is relatively simple, the importance of the key nodes does not actually correspond exactly to the degree value and other indicators. This leads to incomplete identification of key nodes when this method is used.

Therefore, infectivity was chosen as the metric for evaluating key nodes. We define infectivity as the ability of a node to infect other nodes and be infected by other nodes throughout the propagation cycle. Infectivity can be used as an important indicator to evaluate the role of a node in the propagation cycle and is calculated as shown. Equation (10) is used for the calculation of infectivity I_i.

$$I_i = \frac{1}{T}\sum_{t=1}^{T}\left[m_i(t) + n_i(t)\right], \tag{10}$$

where I_i denotes the infection capacity of a node, $m_i(t)$ denotes the number of times a node is infected by other nodes during the i^{th} propagation, $n_i(t)$ denotes the number of times a node infects other nodes during the i^{th} propagation, and T denotes the propagation period.

We ran 500 propagation simulations of the model followed by data statistics. Plot the results of the propagation simulations as a scatter plot.

It is important to note that when performing the infectivity calculations in this paper, nodes that have already been infected or have achieved immunity will not be reinfected as they are already infected. However, the propagation process may still be carried out on them. It is only because they are already infected or immune that no further changes occur and the transmission process actually occurs. This "number of infections" should, in fact, also be taken into account. Therefore, in this paper, the number of infections of a node is calculated by including the number of times it was infected for the first time and the number of times it was "infected" when it was immune or infected, and the results are shown in Figure 3.

Figure 3plots the number of times a node is infected by other nodes as the horizontal axis and the number of times a node infects other nodes as the vertical axis. The node is positioned near the top right corner of the graph and is more infectious when both the number of times it has been infected by other nodes and the number of times it has infected other nodes are larger. Conversely, when the node's position is near the origin and the value of both

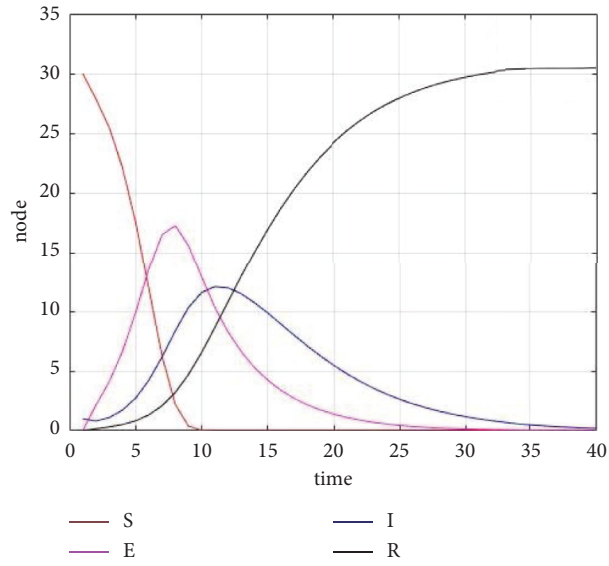

Figure 2: SEIR model propagation process.

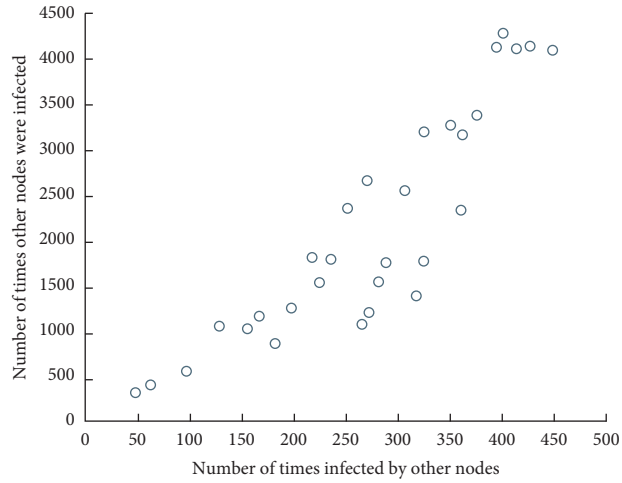

Figure 3: Scatter diagram of propagation results.

indicators is small, it is less infectious. The data of some of the nodes with higher infectious capacity are shown in Table 2.

From the calculation of infection capacity and the data in the table, it can be found that the number of infections and the number of infections of other five nodes, i.e., NO. 1, NO. 9, NO. 21, NO. 4, and NO. 25, are relatively high. Their infection capacity is also at a high value. It proves that they will play a key role in the propagation process and have a high influence capacity. Therefore, they are defined as critical nodes.

In previous studies, sleep quality, circadian rhythm, and environmental factors were identified as factors affecting controller workload [32]. Compared with the traditional results of controller workload influencing factor identification, the new item of duty system (NO. 25) is added to the key factors identified in this paper. In practice, the reasonableness of the duty system will largely affect the

Table 2: Nodes with strong infection ability.

Node number	Number of times infected by other nodes	Number of times infected by other nodes	Infection capacity
1	403	4277	9.36
9	428	4126	9.108
21	451	4081	9.064
4	396	4117	9.026
25	416	4093	9.018
2	376	3367	7.486

controller workload. It is also more relevant to reality. It also indirectly proves the feasibility of the method.

At the same time, our study ranked all the factors affecting controllers' workload by the indicator of infectivity. It is possible to clearly analyze exactly which influencing factors are important and which ones can be ignored. It was missing in the previous study.

4.2. Critical Node Control. According to the Civil Aviation Administration of China (CAAC) management manual, the control measures for critical nodes include two main categories.

4.2.1. Preventive Measures. These include personnel capacity enhancement, team building, and staffing, as well as system and system construction and maintenance.

They mainly refer to improving controllers' business capabilities by organizing personnel to attend business training, improving the construction of hardware facilities to improve controllers' resting environment, increasing personnel reserves, and reasonably arranging the duty system to reduce the occurrence of long hours of continuous work because of unreasonable scheduling or insufficient personnel reserves. It is expressed in the model as reducing the infection probability of the nodes, setting the initial propagation probability of the control nodes to 0 and their infection probability to 10%.

4.2.2. Emergency Disposal Measures. They are the emergency measures to adjust the nodes to return to normal state after they have already been disturbed. In practice, it is expressed as a temporary increase of controllers, emergency arrangement of technical support, etc. In the model, the recovery probability of the node is increased and adjusted to 80%.

The five nodes with the top infectivity in Table 2 were artificially controlled, and the results of network propagation after control are shown in Figure 4.

As can be seen in Figure 4, after the implementation of control measures at key nodes, the peak propagation was reduced to 2.425, and the propagation cycle was delayed until the 15^{th} cycle. It proves that the control measures are effective in reducing the risk of propagation and have had a good effect.

In previous studies, the control of controller influence factors was not as effective as the results in this paper. In terms of convergence speed and convergence effect, this paper achieves the best results for the control of the identified nodes.

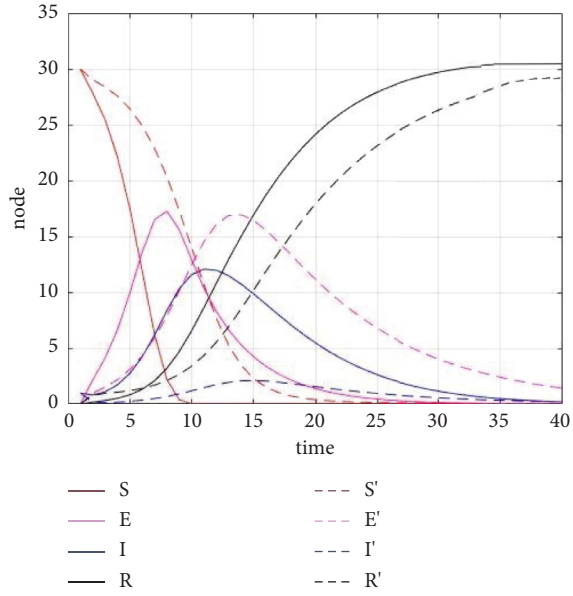

Figure 4: Comparison of SEIR model propagation process.

5. Neural Network-Based Analysis and Validation

To verify the impact of the identified nodes on the controller workload in actual operation, this paper chose to construct a BP neural network and select the identified nodes as the input layer for the controller workload prediction. A comparison experiment was set up to verify the validity of the identified nodes by selecting other nodes as the input layer.

5.1. Principle of BP Neural Networks. First proposed by Rumelhart in 1986, BP neural networks are multilayer feedforward neural nets trained according to the error backpropagation algorithm [33].

This paper uses MATLAB 2018b software to implement the design, training, and validation of the BP neural network.

5.1.1. Data Normalization Process. When using a BP neural network for prediction, the data should be normalized first. Equation (11) is used to calculate the data normalization result.

$$Z = Z_{\min} + \frac{Z_{\max} - Z_{\min}}{X_{\max} - X_{\min}} (X - X_{\min}), \tag{11}$$

where X is the initial data, $X_{\max}$ and $X_{\min}$ are the maximum and minimum values in the initial data, respectively, Z is the data after normalization, $Z_{\max}$ and $Z_{\min}$ denote the maximum and minimum values when normalized, respectively, and the values of both are 1 and -1, respectively.

5.1.2. Construction of the BP Neural Network. Hecht completed the proof that a 3-layer BP neural network can achieve the approximation of an arbitrary nonlinear function in 1987 [34].

Therefore, the BP neural network established in this paper is a 3-layer network structure of BP neural network. At the same time, since the number of nodes in the input layer L takes the number of evaluation indicators, the number of nodes in the output layer N takes the number of categories classified. For a specific item, L and N are fixed values. Hence, to obtain a better training effect and prediction accuracy, it is necessary to determine the number of nodes M of the hidden layer. Equation (12) is used for calculating the number of nodes M of the hidden layer.

$$M < \sqrt{L + N} + A, \tag{12}$$

where A is a constant that takes values between 1 and 10.

5.1.3. Training the BP Neural Network. The transfer function between the input layer and the hidden layer and between the hidden layer and the output layer uses a nonlinear transformation function, i.e., the sigmoid function (i.e., S function). At the same time, to make the number of iterations lower and improve the convergence accuracy, the adaptive learning gradient descent algorithm with an additional momentum factor is used in the training process in this paper.

5.2. Data Acquisition. To verify the validity of the model and the reasonableness of the assessment method, data from Xiamen Gaoqi International Airport was selected for modeling analysis in this paper. Xiamen Gaoqi International Airport is located in Xiamen, Fujian Province, and as a Class 4E civil international airport, it plays an important role as a regional aviation hub in the southeast coastal region of China. In the year 2020, Gaoqi Airport completed a total of 16,710,197 passengers, 278,336.4 tons of cargo and mail throughput, and 139,827 take-offs and landings, all of which are among the top in China.

In this paper, a total of 11,643 sets of valid data were collected, 90% of which (10,479 items) was used for BP neural network building and model training, and the remaining 10% (1,163 items) was used for model validation analysis.

5.3. Neural Network Prediction. After the normalization of data according to the above steps, all independent variables are selected for neural network prediction. The number of neurons in the input layer is 9, the number of neurons in the output layer is 1, and the number of neurons in the hidden layer can be found in the range [4, 13] according to equation (12).

In the process of network training, the additional momentum factor was set to 0.9, the initial value of learning rate was set to 0.01, the training error convergence accuracy target was 10-5, and the maximum number of iterations was set to 1000.

After training, testing, and validating the network for different numbers of hidden layer neurons within the range of values, it was concluded that the training and testing errors of the BP neural network reached a small level when the number of hidden layer neurons was 10.

To assess the predictive accuracy of the model, the program was used to perform a goodness-of-fit analysis of the output results.

The prediction performance of the model is shown in Table 3 after selecting different nodes for multiple prediction validation.

The comparative experiments showed that when using neural networks for prediction, the training and test sets had the highest accuracy when the five identified nodes were chosen as the input layer, while the accuracy decreased with varying degrees when other nodes were chosen as the input layer.

The fitting results show that the identified nodes of the model constructed in this paper perform better in the fitting experiments, with better fitting results and higher confidence in the prediction results. It proves that the identified key factors affecting the controller workload are feasible.

6. Discussion

This paper focuses on the main factors that influence the controller workload in controller work. The results of the experiment show the five factors, namely physiological rhythm, duty length, sleep status, duty scheduling, shift system, and flight flow, as the main factors that significantly affect the controller workload. The relationship between the different factors is bidirectional. It indicates that they influence each other, interfere with each other, and ultimately act together to the controller workloads.

The first part is the construction of a complex network based on MIC, which well explains the closeness of the relationship between the different factors. It explains the degree of influence between different factors. The larger the value of MIC between different factors, the stronger the connection and the stronger the influence on each other.

The second part is the identification of key nodes based on the SEIR model. In the SEIR model, we define infection capacity. The stronger the infectivity, the greater the role of the node in the model. The final results show five factors, namely physiological rhythm, duty length, sleep status, duty scheduling, shift system, and flight flow, as the main factors that have a significant impact on the controller workload. In practical terms, these five factors will directly affect the

Table 3: Prediction performance of different models.

Prediction model	Input layer nodes	Training set accuracy	Test set accuracy
1	1, 4, 9, 21, 25	0.99	0.99
2	1, 2, 4, 9, 21	0.95	0.93
3	1, 2, 6, 4, 9	0.94	0.91
4	1, 2, 3, 4, 10	0.94	0.91
5	1, 2, 3, 10, 17	0.91	0.91
6	2, 6, 10, 17, 23	0.91	0.91

controller's state at work. They also have a strong influence on other factors. Therefore, the identified results are reliable.

Finally, the validation results based on the BP neural network further prove that these five factors are the most important factors affecting the controller workload.

In the study of this paper, there are still some problems that can be optimized. Firstly, the statistics on the factors affecting the controller load can still be optimized. The thirty indicators in this paper should be only some of the many factors that affect controller load. These indicators can still be enriched and improved. Secondly, the specific propagation rules can be improved in the process of using the SEIR model. In detail, the probabilities of mutation, infection, and propagation of various nodes in the model can be further optimized. These probabilities should be derived based on the actual situation. The third aspect is that when calculating the correlation coefficient between factors, this paper chooses to deal with the correlation negligible when the correlation coefficient is smaller than 0.2 as proposed by Buda Andrzej. However, this specific criterion should be open to further discussion in the course of specific analysis.

Moreover, this paper provides ideas for the next direction of research: controller workload assessment by analyzing the influencing factors. In other words, by quantifying the influencing factors, the controller workload can be evaluated and predicted. In the following study, the first research can be carried out on the link of using the influencing factors to achieve the prediction of controller workload. There is also further optimization of the SEIR model for more accurate identification of the interaction pattern between the influencing factors. The identified factors can be effectively controlled to reduce controller workload and improve control efficiency and safety. At the same time, we should optimize the construction of complex networks, i.e., we should study the third deficiency point presented above. The determination of the correlation coefficient criteria is carried out for this study. Determine what the correlation coefficient is when the factors are negligibly correlated with each other.

7. Conclusions

Considering the development of civil aviation in a comprehensive manner, the airport authorities want to increase the traffic as much as possible within the permissible limits of ensuring safety, low delays, and controller workload. For this purpose, a study of controller workload is needed. However, the key influences on controller workload have not been considered in the relevant literature. However, in the process of controller workload generation and development, a few key factors determine the overall generation and development process. To explore the process of controller workload generation and development and to solve the problem of controller workload evaluation, a controller workload influencing factor identification method based on the SEIR model is proposed. The method overcomes the subjectivity of traditional methods in a better way. The method has strong strategic and tactical significance. It can provide a basis for the controller workload generation, and development process and can better identify the key aspects of the process. The method consists of two main parts, complex network modeling based on actual operational data and critical node identification based on the SEIR model.

The structure of the complex network modeling based on actual operational data consists of nodes and connected edges. The nodes are the factors that may affect the controller workload analyzed by collecting the actual operational data. The edges are described by the correlation between the factors represented by the nodes. In analyzing the correlations, we compare various methods, such as the empirical method and the correlation coefficient method. The MIC was finally chosen as the indicator to evaluate the correlation. The calculated results were filtered to retain the stronger correlations and to remove the weaker ones.

In the identification of key nodes based on the SEIR model, we made adjustments to the SEIR model in conjunction with the actual civil aviation situation. It enables the SEIR model to be more relevant to the civil aviation reality. To analyze the model propagation results, we defined contagion capacity indicators. The indicators with high contagion capacity are controlled as key nodes, and the controlled network is again subjected to propagation experiments. We are able to identify the key factors affecting controller load by this method. It allows a comparison of the position of the identified nodes in the network. The experimental results show that our approach works and that the identified nodes have an important role in the network.

Finally, we use BP neural networks for the validation of the node identification results. By selecting independent data and constructing comparison experiments, we verified that the identified nodes have a strong effect on the controller workload and that the method is feasible and effective.

Data Availability

The data used to support the findings of this study have not been made available because the data are also a part of an ongoing study.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (no. 71801221).

References

[1] IATA, "IATA Forecast Predicts 8. 2 billion Air Travelers in 2037," pp. 10–24, 2018.

[2] P. Di Mascio, R. Carrara, L. Frasacco, E. Luciano, A. Ponziani, and L. Moretti, "How the tower air traffic controller workload influences the capacity in a complex three-runway airport," *International Journal of Environmental Research and Public Health*, vol. 18, no. 6, pp. 2807–2814, 2021.

[3] M. Zhang, S. Han, and C. Pei, "Review of air traffic controller workload research," *Ergonomics*, vol. 14, no. 4, pp. 61–64, 2008.

[4] S. G. Hart and L. E. Staveland, "Development of NASA- TLX (task load index): results of empirical and theoretical research," *Advances in Psychology*, vol. 52, no. 6, pp. 139–183, 1988.

[5] G. B. Reid and T. E. Nygren, "The subjective workload assessment technique: a scaling procedure for measuring workload," *North - Holland: Advances in Psychology*, vol. 52, pp. 185–218, 1988.

[6] R. G. Stamp, "The DORATASK method of assessing ATC sector capacity an overview," *Dora communication 8924*, Civil Aviation Authority, London, UK, 1992.

[7] Icao, *Doc.9426: air traffic services planning manual*, ICAO, Montreal, 1992.

[8] S. Han, M. Hu, B. Jiang, J. Zhan, Y. Gao, and Y. Chen, "Research on the relationship between sector capacity and controller workload," *Air Traffic Management*, vol. 6, no. 1, pp. 42–45, 2000.

[9] L. Wan, *The research on the evaluation of controller's workload*, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2005.

[10] A. Zheng, *The research of atc workload based on terminal sector complexity factors*Civil Aviation Flight Academy, Guanghan, China, 2013.

[11] S. K. Lal and A. Craig, "Driver fatigue: electroencephalography and psychological assessment," *Psychophysiology*, vol. 39, no. 3, pp. S0048577201393095–S0048577201393321, 2002.

[12] H. J. Eoh, M. K. Chung, and S. H. Kim, "Electroencephalographic study of drowsiness in simulated driving with sleep deprivation," *International Journal of Industrial Ergonomics*, vol. 35, no. 4, pp. 307–320, 2005.

[13] P. Arico, G. Borghini, G. Di Flumeri et al., "Reliability over time of EEG-based mental workload evaluation during Air Traffic Management (ATM) tasks," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 7242–7245, Milan, Italy, August 2015.

[14] D. Dasari, C. Crowe, C. Ling, M. Zhu, and L. Ding, "EEG pattern analysis for physiological indicators of mental fatigue in simulated air traffic control tasks," *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting*, vol. 54, no. 3, pp. 205–209, 2010.

[15] V. K. Krishnan, D. Dasari, and L. Ding, "EEG correlates of fluctuation in cognitive performance in an air traffic control task," *Office of Aerospace Medicine, Federal Aviation Administration, Washington DC, USA*, 2014.

[16] L. Wan and F. Chen, "Study on relationship between controllers' cognitive behaviour and fatigue based on EEG," *China Safety Science Journal*, vol. 28, no. 7, pp. 1–6, 2018.

[17] X. Zhao, R. Fan, J. Rong, and W. W. Guo, "Research on travel time of basic road based on cellular automata," *Journal of Beijing University of Technology*, vol. 37, no. 10, pp. 1511–1523, 2011.

[18] F. Chen and L. Wan, "Analysis of shift fatigue of air traffic controllers based on EEG," *Journal of Civil Aviation Flight University of China*, vol. 31, no. 6, pp. 16–21, 2020.

[19] L. Wang and R. Sun, "Research on fatigue monitoring method of controller based on facial feature recognition," *Chinese Safety Science Journal*, vol. 22, no. 7, pp. 66–71, 2012.

[20] J. Krajewski, R. Wieland, and A. Batliner, "An acoustic framework for detecting fatigue in speech based human-computer-interaction," in *Computers Helping People with Special Needs, Proceedings*vol. 5105, p. 54, Springer, Berlin Heidelberg, 2008.

[21] F. Hawkins, *Human factors in flight*, Gower Technical Press, London, UK, 1990.

[22] International Civil Aviation Organization (Icao), *Safety management manual*, The UN Secretariat, New York, NY, USA, 2006.

[23] S. T. Shorrock, "Errors of perception in air traffic control," *Safety Science*, vol. 45, no. 8, pp. 890–904, 2007.

[24] J. Reason, *Human error*, pp. 1–19, Cambridge University Press, New York, NY, USA, 1990.

[25] D. A. Wiegmann and S. A. Shappell, *A human error approach to aviation accident analysis*, Ashgate Pub Co, London, UK, 2003.

[26] C. Bedon and S. Mattei, "Remote facial expression and heart rate measurements to assess human reactions in glass structures," *Advances in Civil Engineering*, vol. 202116 pages, Article ID 1978111, 2021.

[27] C. Bedon and S. Mattei, "Facial expression-based experimental analysis of human reactions and psychological comfort on glass structures in buildings," *Buildings*, vol. 11, no. 5, p. 204, 2021.

[28] I. V. Laudeman, S. G. Shelden, and R. Branstrom, *Dynamic density: an air traffic management metric*, pp. 1–4, National Aeronautics and Space Administration, California, CA, USA, 1998.

[29] M. Wu, Z. Ye, X. Wen, and X. Jiang, "Air traffic complexity recognition based on complex networks," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 46, no. 5, pp. 839–850, 2020.

[30] X. Zhang, S. Zhao, and H. Mei, "Analysis of airport risk propagation in Chinese air Transport network," *Journal of Advanced Transportation*, vol. 2022 Article ID 9958810, 13 pages, 2022.

[31] W. J. Conover, *Practical nonparametric statitics*, pp. 428–441, John Wiley & Sons, New York, NY, USA, 1999.

[32] M. A. Nealley and V. J. Gawron, "The effect of fatigue on air traffic controllers," *The International Journal of Aviation Psychology*, vol. 25, no. 1, pp. 14–47, 2015.

[33] X. Wen, X. Zhang, Y. Zhu, and X. Li, "Intelligent fault diagnosis technology: MATLAB application," *Beijing University of Aeronautics and Astronautics Press*, vol. 9, 2015.

[34] N. R. Hecht, "Kolmogorov's mapping neural network existence theorem," in *Proceedings of the International Conference on Neural Networks*, IEEE Press, New York, NY, USA, June 1987.