Predator-Prey Model Based Resource Allocation in Edge Computing

bingjie liu (✉ 446297409@qq.com)
Henan Agricultural University

Yali Bai
University of Science and Technology Beijing

Li Ding
Henan Agricultural University

Research Article

Keywords: Edge Computing, Resource allocation, Stochastic differential equation, Stochastic perturbation, Stability analysis

Posted Date: March 1st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-1436493/v1

License: ☺️ ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Predator-Prey Model Based Resource Allocation in Edge Computing

Bingjie Liu, Yali Bai, Li Ding∗

Abstract—Edge computing layer fills the gap between cloud computing layer and terminal devices layer, having the characteristics of fast processing speed and short application response time. Because the edge nodes are highly dynamic, the first problem we need to consider is how to allocate these resources reasonably so that the system can achieve a continuous stable state. In this paper, we study the resource allocation problem of network services in edge computing environment, and proposing a resource allocation model based on stochastic differential equation. We use the predator-prey model to analyze the resource strategy for the system to reach a stable state. We also quantitatively analyze the relationship between cloud computing resources and edge computing resources providing services to users, and prove the existence, uniqueness as well as stability of the model solution. Numerical simulation results show the correctness and effectiveness of the model.

Index Terms—Edge Computing; Resource allocation; Stochastic differential equation; Stochastic perturbation; Stability analysis.

1 Introduction

Edge computing [1] provides computing, storage and network services near the data source through the intelligent terminal devices at the edge of the network. As the extension of cloud computing, edge computing has the characteristics of relieving network bandwidth pressure, enhancing service response ability, protecting data privacy and so on [2]. Because the edge computing is closer to users, it can decrease network latency [3], reduce the bandwidth consumption of the backhaul link [4, 5] and mobile energy consumption [6]. With the rapid development of mobile communication technology, edge computing has been widely used in many scenarios, such as 5g, AR / VR, real-time traffic monitoring, Internet of things, Internet of vehicles, smart home, smart city, etc. [7–11]. The architecture of edge computing is shown in Figure 1. From the network center to the edge of the network, we divide edge computing into different layers: cloud computing layer, edge computing layer and terminal layer.

Compared with traditional cloud computing technology, edge computing requires a large number of deployments, and the scale as well as amount of resources for a single edge computing node are relatively limited. With the rapid development of edge computing, it is of great significance to strengthen the cooperation of edge computing nodes, edge nodes and core cloud computing nodes. By means of the cooperation mechanism, not only can the utilization rate of computing, storage and network resources be effectively improved, but also the network service quality and user experience can be greatly improved. Therefore, edge computing resource allocation for collaboration mechanism is an important research direction. However, a lot of works have been done from the perspective of time delay and energy efficiency, which mainly focus on the resource allocation and optimization of edge computing , lacking the consideration of overall stability of the system.

Edge computing possesses fast processing speed and short application response time [12]. Accordingly, all users hope to provide resource services through edge computing to obtain a better experience. Edge computing is usually composed of lightweight edge servers and network devices, which has relatively limited resources. However, due to the randomness and suddenness of the network and the mobility of users, when the number of local users and business volume suddenly increase explosively, it will cause problems such as unstable network services and poor user experience. Therefore, facing of user resource requests, whether the edge computing environment is capable of continuing to work in a balanced state, and how to allocate these resource requests reasonably so that the system can achieve a persistent and stable state is a problem we need to consider. The biological population model is an effective method to analyze the dynamic behavior of species and make the system reach the equilibrium state.

In our work, we propose a resource allocation model for network services in the edge computing environment. The resource request and service process in the edge computing environment is constantly changing, and it is highly dynamic and complex. Edge computing and cloud computing obtain benefits by providing resource services, which is similar to the interaction between predators and prey in biological system. Moreover, the resource state of each layer is inevitably interfered by external stochastic factors, such as hacker attack, occupation of resources, battery power and network jitter. In order to make the model closer to the actual situation, we propose an edge computing resource allocation model based on stochastic differential equations, which can quantitatively analyze the relationship between cloud computing resources and edge computing resources to providing services for users to guarantee the stability and persistence of the system. The model can obtain different equilibrium solutions according to different parameter settings, thereby identifying the time dynamics and behaviors of each layer.

The main contributions of our paper are summarized as follows:

Bingjie Liu, and Li Ding are with the College of Information and Management Science, Henan Agricultural University, Zhengzhou, China, 450046. (E-mail: liubingjie0802@163.com; dingli@henau.edu.cn; Corresponding author: Li Ding).
Yali Bai is with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, China, 100083, (E-mail: m18701295508@163.com).
- We propose a resource allocation model for network services in an edge computing environment based on the predator-prey model in the biological population.
- An edge computing resource allocation model based on stochastic differential equations is proposed. Edge computing and cloud computing try to find the resource allocation strategy that makes the system reach a balanced state, while users try to find the resource request strategy that makes the system reach a balanced state.
- Lyapunov analysis method is used to prove the uniqueness and stability of the solution, and the relationship between cloud computing resources and edge computing resources providing services for users is quantitatively analyzed.

The remainder of the paper is organized as follows. Section 2 mainly discusses some related researches on resource allocation in edge computing. Section 3 introduces the system model and problem formulation. Section 4 provides the existence and uniqueness theorem, and stability analysis for the proposed stochastic differential equation model is given in 5. Numerical simulations are given in Section 6. Finally, the work is concluded in Section 7.

2 Related Work

The issue of resource allocation in the edge computing environment has received extensive attention, and scholars have made a lot of contributions in resource allocation. In order to improve the efficiency of computing resources and optimize performance indicators, edge cloud collaborative network computing architecture can be used [13], taking advantage of edge computing and cloud computing. Huang et al. [14] proposed a collaborative resource allocation algorithm. By sharing resources between the cloud center and the adjacent edge servers, the alternating direction method of multipliers is used to minimize the network cost. Game theory is also widely used in resource allocation in the edge computing environment to optimize various performance indicators. A method based on game theory can solve the joint optimization problem between edge terminal user experience and network cost. On this basis, Guan et al. [15] proposed an edge computing resource scheduling mechanism based on Stackelberg dynamic game, in order to minimize costs while maintaining user experience. Under this mechanism, when the game model tends to balance, it proves that the resource allocation strategy is optimal. However, the data demand of cloud data centers is often very large, and this will lead to a reduction in energy consumption. For this problem, Silva et al. [16] Proposed a Gaussian process regression mechanism for resource allocation in collaborative architecture, which reduces energy consumption by reducing the demand for cloud data centers. The development of the above-mentioned technologies in edge computing is constantly maturing, especially the game theory approach. In recent years, non-cooperative games, cooperative games, big alliances, and intermediate alliance games have been increasingly used in edge computing. Considering the selfishness of nodes in edge computing, and each node cannot fully obtain the information of each other’s nodes, non-cooperative and incomplete information game models are used more in edge computing.

Time delay is an important index of resource allocation in edge computing, which has great influence on QoS of edge users. Literature [17] used a game theory method to study the dynamics of user interaction, and proposed a distributed equilibrium algorithm to calculate the equilibrium strategy for each user. Experimental results show that the response time delay is reduced by more than 20% compared with other methods. However, there are still some problems in resource optimization problems based on game theory models. For example, when the number of nodes is very large, it is more complicated to solve the equilibrium solution in the actual edge computing scenario. At present, some scholars solve this problem through the reinforcement learning method. Based on the perceptual environment ability of reinforcement learning, as well as the predictive ability and feature selection of deep learning, they can make reasonable predictions on the behavior of other nodes and provide environmental information by actively learning environmental information. However, the current development of this technology is still not very mature, so further research and development are needed. The game theory model given to deep reinforcement learning has important research value and a lot of research space in the future. Machine learning tools are undoubtedly the main carrier of artificial intelligence. The correct set of machine learning tools can provide a large number of intelligent services required for edge computing resource management, thereby releasing the true potential of edge computing. Based on machine learning, Yang et al. [18] proposed a deep reinforcement learning method to allocate resources in the edge computing network to reduce latency. This method uses the independent learning ability of deep reinforcement learning to allocate appropriate resources to each edge node, and makes good use of the advantages of reinforcement learning. The independent adaptation strategy also more clearly highlights the advantages of edge computing. You et al. [19] studied the resource allocation of mobile edge computing offloading system with delay constraint, which assigns priorities to users based on energy efficiency and reduces the complexity of the optimization problem. This method is a typical representative of resource allocation based on energy efficiency and is of great practical value. It is a very important way to model and analyze wireless resource allocation with the help of random geometric models. Use the random space model to model the edge nodes, and finally find the best resource allocation strategy for the edge computing network. Ko et al. [20] proposed a method based on spatial network model to analyze the communication delay and computing delay of mobile edge computing network by applying stochastic geometry, queuing and parallel computing theory. However, although the model based on random theory can consider the relationship between network connectivity and stability in edge computing scenarios, it still ignores the trade-off between network traffic and load.

Resource allocation in the process of transmitting data from terminal devices to edge computing processors is also important and challenging. The continuous development of edge computing scenarios has led to rapid growth in data traffic and rapid expansion of network infrastructure, which inevitably triggers the increase in energy consumption of edge computing networks. Moreover, this situation will continue to aggravate with the increase of edge devices. Therefore, industry and academia are committed to improving network energy efficiency and spectrum efficiency. Since spectrum efficiency and energy efficiency are important indicators of resource management in the edge computing environment, it is necessary to transmit data in the process of edge computing by reasonably allocating the transmission spectrum and energy resources [21][22]. Wang et al. [24] proposed a resource allocation scheme for a wireless multi-user MEC system, which considered the calculation of waiting time limit, and used the Lagrange duality
method to minimize the total energy consumption of the AP. This method is very helpful to improve the fairness of resource allocation, and can achieve the purpose of optimizing energy efficiency. Considering the same problem, Dinh et al. [25] proposed a computational offloading framework, which took into account the CPU frequency range of mobile devices and the data and service rates of AP, and adopted a semidefinite relaxation method to reduce energy consumption and task execution delay. This computational offloading framework has great advantages in terms of energy efficiency optimization and user QoS requirements.

Promoting fairness among users is another important issue in resource allocation in the edge computing environment. The traditional fairness problem is related to group scheduling between users, where each user should receive a considerable amount of resources. In the edge computing scenario, fairness issues not only appear between edge nodes, but also in resource allocation decisions between edge nodes at different levels. However, most of the resource allocation in edge computing focuses on the factors such as delay and energy consumption, but ignores the fairness in the process of resource allocation. Considering the fairness of resource allocation, scholars use biological methods to quantitatively analyze the time dynamics and behaviors of each layer in [26]. Some works have been done to study the multiple levels dynamic change of heterogeneous systems using biological methods. To meet the QoS requirements of different users, network virtualization technology can dynamically allocate available wireless resources among these virtual networks. On the basis of network virtualization technology, Banez et al. [27] proposed a network virtualization resource allocation and economic method based on a predator-prey food chain model. Considering the limited spectrum resources in the edge computing environment, some scholars also use the population model to dynamically perceive spectrum resources. Literature [28] and [29] studied the dynamics of cognitive radio networks using evolutionary game model and predator-prey model respectively. By analyzing the population dynamics in spectrum sharing ecosystem, the transient behavior of cognitive radio is analyzed. On this basis, Hui et al. [30] proposed a secure resource allocation mechanism in a mobile edge computing environment based on the biological population equation. This method can provide reliable support for the edge computing intrusion prevention system, possessing great research value. However, this model only considers the security resource allocation of the edge computing layer, and does not consider the equilibrium and stable state of the edge computing system. In this paper, the predator-prey model is used to model the resource allocation problem in the edge computing system. Through the predator-prey model, the internal law among the layers of the edge computing system is simulated, and to predict the trend of resource demand changes, which can be directed management adjustment system.

3 System Model and Problem Formulation

In this chapter, in order to ensure the long-term stable operation of the edge computing system and reasonable resource management, the use and demand situation of the three-tier resources in the edge computing system architecture are modeled. The process of terminal user resource request, cloud computing and edge computing to provide resource services is shown in Figure 2. The process of resource service provided by cloud computing and edge computing is similar to the predator of biological population. The more the prey is, the more service resources will be provided. The process of resource request by users is similar to the prey of biological population. The more predators, the fewer remaining resource requests will be. This paper does not consider the interaction between edge computing and cloud computing. According to the user’s resource requirements, resource strategies for edge computing and cloud computing are given to ensure the balance of the system.

In the edge computing environment, users can make resource requests for edge computing and cloud computing. Let $C(t)$ denotes the amount of resources that cloud computing provides computing resource services to users at time $t$, $E(t)$ represents the amount of resources that edge computing provides computing resource services to users at time $t$, and $U(t)$ represents the resource requests of users at time $t$. Accompanied by more and more resource requests from users, cloud computing and edge computing provide more and more resource services for users. Let $a_1(t)$ denote the resource growth rate of cloud computing to provide computing resource services at time $t$, and the growth rate is directly proportional to the change of resource services provided per unit time. Mathematically, it can be expressed by the Malthus equation

$$a_1(t) = \frac{1}{C(t)} \frac{dC(t)}{dt}$$  \hspace{1cm} (1)

where $a_1(t)$ is number belonging to $(0, 1)$, and the solution of formula (1) is $C(t) = C(0)e^{a_1(t)}$. Therefore, $\lim_{t \to \infty} C(t) = \infty$, where $C(0)$ is the initial value of $C(t)$.

In fact, the growth of resources will be limited by the system capacity, so we use the classical logistic equation [31], which is recognized as the optimal mathematical model to describe the population growth law when resources is limited

$$\frac{1}{C(t)} \frac{dC(t)}{dt} = a_1(t) \left(1 - \frac{C(t)}{N_1} \right)$$  \hspace{1cm} (2)

where $N_1$ is the maximum capacity of cloud computing resources. The solution of formula (2) is $C(t) = \frac{N_1}{1 + \frac{N_1 - C(0)}{C(0)}e^{-a_1(t)}}$, and $\lim_{t \to \infty} C(t) = N_1$. The dynamic change of $C(t)$ can be described by the following differential equation

$$\frac{dC(t)}{dt} = C(t) \left[ a_1(t) \left(1 - \frac{C(t)}{N_1} \right) + b_13(t) U(t) \right]$$  \hspace{1cm} (3)
where \( b_{13} (t) \) represents the arrival rate of cloud computing resources requested by users at time \( t \).

Assuming that the maximum capacity of edge computing resources is \( N_2 \), \( a_2 (t) \) represents the growth rate of computing resource services provided by edge computing at time \( t \), and similarly, \( a_2 (t) \) can be expressed by the classic logistic equation as

\[
\frac{1}{E(t)} \frac{dE(t)}{dt} = a_2 (t) \left( 1 - \frac{E(t)}{N_2} \right)
\]

(4)

The solution of formula (4) is \( E(t) = \frac{N_2}{1 + e^{-a_2 (t) t}} \), where \( E(0) \) is the initial value of \( E(t) \). The dynamic change of \( E(t) \) can be described by the following differential equation

\[
\frac{dE(t)}{dt} = E(t) \left[ a_2 (t) \left( 1 - \frac{E(t)}{N_2} \right) + b_{23} (t) U(t) \right]
\]

(5)

where \( b_{23} (t) \) represents the arrival rate of edge computing resources requested by users at time \( t \).

With the increase of resource services provided by cloud computing and edge computing, the remaining user requests will decrease, and the resource requests \( U(t) \) of users will decrease with the increase of \( C(t) \) and \( E(t) \). Assuming that \( N_3 \) represents the maximum threshold of user resource requests in the system, represents the growth rate of user resource requests at time \( t \), and user resource requests will not increase unlimited, so

\[
\frac{1}{U(t)} \frac{dU(t)}{dt} = a_3 (t) \left( 1 - \frac{U(t)}{N_3} \right)
\]

(6)

The solution of formula (6) is \( U(t) = \frac{N_3}{1 + e^{-a_3 (t) t}} \), where \( U(0) \) is the initial value of \( U(t) \). The dynamic changes of user resource requests can be described by the following differential equation

\[
\frac{dU(t)}{dt} = U(t) \left[ a_3 (t) \left( 1 - \frac{U(t)}{N_3} \right) - b_{31} (t) C(t) - b_{32} (t) E(t) \right]
\]

(7)

where \( b_{31} (t) \) and \( b_{32} (t) \) respectively represent the occupancy rate of cloud computing resources and edge computing resources by users at time \( t \).

Based on the above analysis, we establish the resource allocation model of edge computing system as follows:

\[
\begin{align*}
\frac{dC(t)}{dt} &= C(t) \left[ a_1 (t) \left( 1 - \frac{C(t)}{N_1} \right) + b_{13} (t) U(t) \right] \\
\frac{dE(t)}{dt} &= E(t) \left[ a_2 (t) \left( 1 - \frac{E(t)}{N_2} \right) + b_{23} (t) U(t) \right] \\
\frac{dU(t)}{dt} &= U(t) \left[ a_3 (t) \left( 1 - \frac{U(t)}{N_3} \right) - b_{31} (t) C(t) - b_{32} (t) E(t) \right]
\end{align*}
\]

(8)

In the edge computing environment, the resource status of each layer is inevitably interfered by external stochastic factors, such as hacker attacks occupying resources, the impact of battery power, and major emergencies. In order to make the model closer to the actual situation, stochastic disturbances are added to the main parameters of the above model. Assuming that \( (\Omega, F, \{F_t\}_{t \geq 0}, P) \) is a complete probability space, the filtration \( \{F_t\}_{t \geq 0} \) is monotonically increasing and contains all null sets. Assume that the perturbation in the environment is mainly manifested as the perturbation of the parameters \( a_1 (t) \), \( a_2 (t) \) and \( a_3 (t) \) : \( a_1 (t) \to a_1 (t) + \sigma_1 (t) dB_1 (t) \), \( a_2 (t) \to a_2 (t) + \sigma_2 (t) dB_2 (t) \), \( a_3 (t) \to a_3 (t) + \sigma_3 (t) dB_3 (t) \), where \( B_i (t) (i = 1, 2, 3) \) are the independent one-dimensional standard Brownian motions defined in a complete probability space, and satisfies \( B_i (0) = 0 \), \( \sigma_i > 0 (i = 1, 2, 3) \) denotes the intensities of stochastic perturbations. Therefore, the stochastic resource allocation model of the edge computing system is as follows

\[
\begin{align*}
\frac{dC(t)}{dt} &= C(t) \left[ a_1 (t) \left( 1 - \frac{C(t)}{N_1} \right) + b_{13} (t) U(t) \right] dt + \sigma_1 (t) C(t) dB_1 (t) \\
\frac{dE(t)}{dt} &= E(t) \left[ a_2 (t) \left( 1 - \frac{E(t)}{N_2} \right) + b_{23} (t) U(t) \right] dt + \sigma_2 (t) E(t) dB_2 (t) \\
\frac{dU(t)}{dt} &= U(t) \left[ a_3 (t) \left( 1 - \frac{U(t)}{N_3} \right) - b_{31} (t) C(t) - b_{32} (t) E(t) \right] dt + \sigma_3 (t) U(t) dB_3 (t)
\end{align*}
\]

(9)

## 4. Existence and uniqueness of Solutions

In this section, we will discuss the solution of the stochastic resource allocation model in the edge computing environment established in the previous section. Before studying the dynamic behavior characteristics, what we first consider is the existence of global positive solution of model (9), and we further prove the existence and uniqueness of the solution of model (9).

**Theorem 1** For any initial value \( (C(0), E(0), U(0)) \in \mathbb{R}_+^3 \), the solution \( (C(t), E(t), U(t)) \) of model (9) on \( t \geq 0 \) is unique, and it will keep in area \( \mathbb{R}_+^3 \) with probability 1.

**Proof** Because in the edge computing environment, the coefficients of the model (9) are locally Lipschitz continuous, for any given initial value \( (C(0), E(0), U(0)) \in \mathbb{R}_+^3 \), the local solution \( (C(t), E(t), U(t)) \) on \( t \in [0, \tau_c) \) is of uniqueness, where \( \tau_c \) denotes the explosion time. In order to prove that the solution is global, we need to prove \( \tau_c = \infty \).

For the first equation of model (9), let \( x = \frac{1}{C} \), apply Itô’s formula, we can get the following results

\[
dx = \left[ \frac{a_1 (t)}{N_1} + \left( \sigma_1^2 (t) - a_1 (t) \right) x - b_{13} (t) Ux \right] dt - \sigma_1 (t) dB_1 (t)
\]

(10)

It can be seen that formula (10) is a nonhomogeneous stochastic differential equation about variable \( x \). We give the corresponding homogeneous stochastic differential equations

\[
dx = \left[ \left( \sigma_1^2 (t) - a_1 (t) \right) x - b_{13} (t) Ux \right] dt - \sigma_1 (t) dB_1 (t)
\]

(11)

Let \( V = \ln x \), apply Itô’s formula, we can obtain

\[dV = d\ln x = \left[ \frac{1}{2} \sigma_1^2 (t) - a_1 (t) \right] dt - \sigma_1 (t) dB_1 (t)
\]

(12)

Integrating formula (12), we can get

\[x(t) = x_0 \exp \int_0^t \left[ \frac{1}{2} \sigma_1^2 (s) - a_1 (s) \right] ds - \int_0^t \sigma_1 (s) dB_1 (s)
\]

(13)

\[ - \int_0^t b_{13} (s) U(s) ds]

Let \( \varphi (t) = \int_0^t \sigma_1 (s) dB_1 (s) + \int_0^t \left[ a_1 (s) - \frac{1}{2} \sigma_1^2 (s) \right] ds \), then \( x(t) = x_0 \exp \left[ - \int_0^t \varphi (s) dt \right] \left[ 1 + \int_0^t \sigma_1 (s) dB_1 (s) \right] \). The explicit solution of formula (10) is obtained by using the constant variation formula

\[C(t) = \frac{1}{C(0)} \exp \left[ \int_0^t \frac{a_1 (s) U(s) ds + \varphi (t)}{C_0} \right]
\]

(14)
In the same way, we can get

$$E(t) = \frac{\exp\left[\int_0^t b_{23}(s)U(s)\,ds + \phi(t)\right]}{\int_0^t + \int_0^t \frac{a_1(s)}{N_1} \exp\left[\int_0^t b_{23}(\tau)U(\tau)\,d\tau + \phi(s)\right]ds}$$  \hspace{1cm} (15)$$

and

$$U(t) = \frac{\exp\left[\phi(t) - \int_0^t b_{31}(s)C(s)\,ds - \int_0^t b_{32}(s)E(s)\,ds\right]}{\frac{1}{\int_0^t + \int_0^t \frac{a_1(s)}{N_1} \exp\left[\phi(s) - \int_0^t (b_{31}(\tau)C(\tau) + b_{32}(\tau)E(\tau))d\tau\right]ds}$$  \hspace{1cm} (16)$$

where

$$\phi(t) = \int_0^t \sigma_2(s)dB_2(s) + \int_0^t \left[2a_1(s) - \frac{1}{2}\sigma_2^2(s)\right]ds$$  \hspace{1cm} (17)$$

$$\psi(t) = \int_0^t \sigma_3(s)dB_3(s) + \int_0^t \left[2a_1(s) - \frac{1}{2}\sigma_3^2(s)\right]ds$$  \hspace{1cm} (18)$$

Consider the comparison equation

$$d\bar{C}(t) = \bar{C}(t)\left[a_1(t)\left(1 - \frac{\bar{C}(t)}{N_1}\right) + b_{13}(t)\bar{U}(t)\right]dt + \sigma_1(t)\bar{C}(t)dB_1(t)$$

$$d\bar{E}(t) = \bar{E}(t)\left[a_2(t)\left(1 - \frac{\bar{E}(t)}{N_2}\right) + b_{23}(t)\bar{U}(t)\right]dt + \sigma_2(t)\bar{E}(t)dB_2(t)$$

$$d\bar{U}(t) = a_3(t)\bar{U}(t)\left(1 - \frac{\bar{U}(t)}{N_3}\right)dt + \sigma_3(t)\bar{U}(t)dB_3(t)$$  \hspace{1cm} (19)$$

Similarly, the explicit solution of equation (19) is obtained

$$\bar{C}(t) = \frac{\exp\left[\int_0^t b_{13}(s)U(s)\,ds + \psi(t)\right]}{\int_0^t + \int_0^t \frac{a_1(s)}{N_1} \exp\left[\int_0^t b_{13}(\tau)U(\tau)\,d\tau + \phi(s)\right]ds}$$  \hspace{1cm} (20)$$

$$\bar{E}(t) = \frac{\exp\left[\int_0^t b_{23}(s)U(s)\,ds + \phi(t)\right]}{\int_0^t + \int_0^t \frac{a_1(s)}{N_1} \exp\left[\int_0^t b_{23}(\tau)U(\tau)\,d\tau + \phi(s)\right]ds}$$  \hspace{1cm} (21)$$

$$\bar{U}(t) = \frac{\exp[\psi(t)]}{\int_0^t + \int_0^t \frac{a_1(s)}{N_1} \exp[\phi(s)]ds}$$  \hspace{1cm} (22)$$

According to the comparison theorem of stochastic differential equations, it can be easily get: $C(t) \leq \bar{C}(t)$, $E(t) \leq \bar{E}(t)$ and $U(t) \leq \bar{U}(t)$. Because the parameters $a_i(t)$ and $b_{ij}(t)$ ($i, j = 1, 2, 3$) are positive $\omega$-periodic functions, so $C(t)$, $E(t)$, and $U(t)$ will not explode in finite time, that is $\tau_e = \infty$.

Hence, Theorem 1 follows.

Due to the randomness, suddenness of the network and the mobility of users, the resource request and service process in the edge computing environment are highly dynamic and complex. It can be seen from the proof process that the model can quickly identify the time dynamics and behaviors of each layer under different parameter settings. We can set appropriate parameters according to different user requirements.

5 Stability analysis of resource service

In previous section, we have obtained the existence and uniqueness of the solution of model (9). However, once the user’s resource request is not served in time or the service quality is poor, the user request rate will decrease until it tends to zero. We hope that under certain conditions, no matter how the external environment changes, it will not make effects on the user’s continuous resource request and the process of continuous resource service provided by edge computing and cloud computing. We define this continuous process as model sustainability. For the convenience of subsequent theoretical analysis, we first give some definitions and lemmas to understand the established stochastic resource allocation model.

**Definition 1** Suppose that $f(t)$ is a continuous bounded function on $R_+$, then define $f^u = \sup_{t \in R_+} f(t)$, $f^l = \inf_{t \in R_+} f(t)$.

**Definition 2** (1) If $\lim_{t \to +\infty} x(t) = 0$, then $x(t)$ is unservicable.

(2) If $\limsup_{t \to +\infty} \int_0^t x(s)\,ds = 0$, then $x(t)$ is unservicable.

(3) If $\liminf_{t \to +\infty} \int_0^t x(s)\,ds > 0$, then $x(t)$ is unservicable.

**Lemma 1** Suppose that $x(t) \in [\Omega \times R_+, R_+]$, where $R^n_+ := \{a | a > 0, a \in R \}$.

(1) If there are normal numbers $\lambda_0$, $T$ and lambda $\geq 0$, such that for any $t \leq T$ there are $\ln x(t) \leq \lambda_0 \int_0^t x(s)\,ds + \sum_{i=1}^n \beta_i B_i(t)$, where $\beta_i (1 \leq i \leq n)$ is a constant, then under $\tau_e \to +\infty$, $\limsup_{t \to +\infty} \frac{1}{T} \int_0^T x(s)\,ds \leq \frac{\lambda_0}{\lambda_0}$.

(2) If there are normal numbers $\lambda_0$, $T$ and lambda $\geq 0$, such that for any $t \leq T$ there are $\ln x(t) \leq \lambda_0 \int_0^t x(s)\,ds + \sum_{i=1}^n \beta_i B_i(t)$, where $\beta_i (1 \leq i \leq n)$ is a constant, then under $\tau_e \to +\infty$, $\liminf_{t \to +\infty} \frac{1}{T} \int_0^T x(s)\,ds \geq \frac{\lambda_0}{\lambda_0}$.

According to the definition and lemma given above, we use Itô’s formula and the strong law of large numbers of local martingale to analyze the unservicable and sustainable conditions of $C(t)$, $E(t)$ and $U(t)$, respectively.

**Theorem 2** (1) If $\limsup_{t \to +\infty} \int_0^t \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds < 0$, then $U(t)$ is said to be unservicable.

(2) If $\limlim_{t \to +\infty} \int_0^t \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds = 0$, then $U(t)$ is said to be unservicable.

(3) If $A > 0$, then $U(t)$ is said to be sustainable, where

$$A = \liminf_{t \to +\infty} \frac{1}{T} \int_0^T \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds - \frac{b_{31}^2}{(a_1/N_1)^2} \frac{1}{T} \int_0^T \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds$$  \hspace{1cm} (23)$$

$$+ \frac{b_{32}^2}{(a_2/N_2)^2} \frac{1}{T} \int_0^T \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds.$$  \hspace{1cm} (24)$$

Proof. See Appendix.

**Theorem 3** (1) If $\limsup_{t \to +\infty} \int_0^t \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds < 0$, then $C(t)$ is said to be unservicable.

(2) If $\limsup_{t \to +\infty} \int_0^t \left[\sigma_3(s) - \frac{1}{2}\sigma_3^2(s)\right]ds = 0$, then $C(t)$ is said to be unservicable.
(3) If \( \liminf_{t \rightarrow +\infty} \int_{0}^{t} \left[ a_{1}(s) - \frac{1}{2} \sigma_{1}^{2}(s) \right] ds > 0 \), then \( C(t) \) is said to be sustainable.

Proof. See Appendix.

**Theorem 4** (1) If \( \limsup_{t \rightarrow +\infty} \left( \int_{0}^{t} \left[ a_{2}(s) - \frac{1}{2} \sigma_{2}^{2}(s) \right] ds \right) < 0 \), then \( E(t) \) is said to be unserviceable.

(2) If \( \limsup_{t \rightarrow +\infty} \left( \int_{0}^{t} \left[ a_{3}(s) - \frac{1}{2} \sigma_{3}^{2}(s) \right] ds \right) = 0 \), then \( E(t) \) is said to be unsustainable.

(3) If \( \liminf_{t \rightarrow +\infty} \left( \int_{0}^{t} \left[ a_{2}(s) - \frac{1}{2} \sigma_{2}^{2}(s) \right] ds \right) > 0 \), then \( E(t) \) is said to be sustainable.

The proof of Theorem 4 is similar to the proof of Theorem 3, and is omitted here. From Theorem 3 and Theorem 4, we found the conditions of \( C(t) \) and \( E(t) \) satisfying sustainability are related to stochastic perturbation. When the stochastic perturbation exceeds a certain threshold, the system does not satisfy the sustainability. From Theorem 2, we can see that the sustainability of \( u \) is related to \( C(t) \) and \( E(t) \). Only when cloud computing and edge computing continue to provide services, users’ resource requests will continue.

If the system is still able to return to its original equilibrium state after a transitional process under the influence of external factors, we call the system stable, otherwise the system is unstable. Next, we analyze the stability of the model. For the convenience of subsequent theoretical analysis, we first give some definitions and lemmas to understand the established stochastic resource allocation model.

**Definition 3** (32) The model (9) is globally attractive, if the condition is satisfied: \( \lim_{t \rightarrow +\infty} |C_{1}(t) - C_{2}(t)| = 0 \), \( \lim_{t \rightarrow +\infty} |E_{1}(t) - E_{2}(t)| = 0 \), \( \lim_{t \rightarrow +\infty} |U_{1}(t) - U_{2}(t)| = 0 \), where \((C_{1}(t), E_{1}(t), U_{1}(t))\) and \((C_{2}(t), E_{2}(t), U_{2}(t))\) are two arbitrary positive solutions of the model (9) with initial values \((C_{0}(0), E_{0}(0), U_{0}(0))\) and \((C_{0}(0), E_{0}(0), U_{0}(0))\).

**Lemma 2** (32) Let \((C(t), E(t), U(t))\) be the solution of model (9) on \( t \geq 0 \), and satisfies the initial value \((C_{0}(0), E_{0}(0), U_{0}(0)) \in \mathbb{R}^{3}_{+}\), then \((C(t), E(t), U(t))\) is uniformly continuous for almost every sample path.

**Lemma 3** (34) Let \( f \) be a nonnegative, integrable and uniformly continuous function defined on \( \mathbb{R}_{+} \), then \( \lim_{t \rightarrow +\infty} f(t) = 0 \).

**Theorem 5** If the conditions \( a_{1}/N_{1}^2 - b_{13}^u > 0 \), \( a_{2}/N_{2}^2 - b_{23}^u > 0 \) and \( (a_{3}/N_{3}^2 - b_{13}^u)^{2} - b_{23}^u > 0 \) are satisfied, then model (9) is globally attractive.

Proof. See Appendix.

Theorem 5 gives the global attractiveness of model (9). Namely, the solution of model (9) will tend to a stable normal state. According to the above theorems, we can conclude that in an open external environment with stochastic perturbation, this allocation model can continue operating in a feasible space as long as the resource allocation in the edge computing environment follows the global macro rules.

### 6 Numerical Simulations

In this section, to verify the effectiveness of the stochastic resource allocation model proposed in this paper, we use MATLAB to perform numerical simulations on the deterministic model and stochastic model of the edge computing system under the original state and the case of increasing stochastic perturbation. The selection of parameters in reference [32-35], the simulation parameter settings are as follows:

- \( a_{1}(t) = 0.5 + h \sin(2\pi t/24) \), \( a_{2}(t) = 0.8 + h \sin(2\pi t/24) \), \( a_{3}(t) = 0.9 + h \sin(2\pi t/24) \),
- \( b_{13}(t) = 0.7 + h \sin(2\pi t/24) \), \( b_{23}(t) = 0.85 + h \sin(2\pi t/24) \), \( b_{31}(t) = 0.2 + h \sin(2\pi t/24) \), \( b_{32}(t) = 0.1 + h \sin(2\pi t/24) \),
- where \( h \) is a positive constant. In order to better show the periodicity, we choose the parameters of the model to be a function of the period of 24 hours.

![Fig. 2: The solution of model (8).](image)

![Fig. 3: The solution of stochastic model (9).](image)
computing and edge computing makes the resource requests of a large number of users satisfied, and the remaining resource requests decrease. Conversely, when \( C(t) \) and \( E(t) \) decrease, \( U(t) \) increases, indicating that cloud computing and edge computing reduce the amount of resources provided, which makes users’ remaining resource requests increase.

![Fig. 4: Comparison of deterministic and stochastic models with periodicity and aperiodicity.](image1)

\[ C(t) = \begin{cases} C(t)_{\text{deterministic model with periodicity}} & \text{if } C(t) \text{ is periodic} \\ C(t)_{\text{deterministic model without periodicity}} & \text{if } C(t) \text{ is aperiodic} \\ C(t)_{\text{stochastic model with periodicity}} & \text{if } C(t) \text{ is periodic and affected by stochastic perturbation} \\ C(t)_{\text{stochastic model without periodicity}} & \text{if } C(t) \text{ is aperiodic and affected by stochastic perturbation} \end{cases} \]

We set the initial values of \( C(0) = 0.5 \). Figure 4 shows a comparison of periodicity and aperiodicity between the deterministic model and the stochastic model. The red and black lines represent the deterministic models with periodicity and aperiodicity respectively, which show how it change at different periods of the day. The green and blue lines represent the stochastic models with periodicity and aperiodicity respectively. Whether the stochastic models are affected by periodic factors or not, they will show more complex dynamic behaviors.

Given different stochastic perturbation intensities \( \sigma_1(t) = 0.04 \), \( \sigma_2(t) = 0.05 \), \( \sigma_3(t) = 0.03 \) and \( \sigma_1(t) = 0.18 \), \( \sigma_2(t) = 0.15 \), \( \sigma_3(t) = 0.2 \). According to theorem 2-4, it can be concluded that

\[ C(t) , E(t) \text{ and } U(t) \text{ are sustainable}. \]

![Fig. 5: Comparison between stochastic model with different stochastic perturbation and deterministic model in \( C(t) \).](image2)

![Fig. 6: Comparison between stochastic model with different stochastic perturbation and deterministic model in \( E(t) \).](image3)

![Fig. 7: Comparison between stochastic model with different stochastic perturbation and deterministic model in \( U(t) \).](image4)

C(t), E(t) and U(t) are sustainable. Figure 5, Figure 6 and Figure 7 compare the dynamic differences between the stochastic model with different stochastic perturbation and the deterministic model. It can be seen that the amplitude of the stochastic model is closely related to the intensity of the stochastic perturbation. The larger the stochastic perturbation, the more obvious the oscillation.

### 7 Conclusions

This paper proposes a resource allocation model in edge computing environment based on stochastic differential equations. This model quantitatively analyzes the relationship between cloud computing resources and edge computing resources to provide services for users. Through the solution and analysis of the model, it promotes the coordinated development of resource supply and demand in the edge computing environment. And the feasibility of the model is analyzed from the sustainability as well as stability. In an open external environment with stochastic perturbation, this allocation mode can run stably and persistently, as long as the
resource allocation in the edge computing environment follows the requirements of the global macro rules.

**APPENDIX**

**Proof of Theorem 2.**

Proof (1) Applying Itô’s formula to formula (9), we can get the following results

\[ d\ln C = \left[ a_1(t) - \frac{1}{2} \sigma_1(t)^2 + \frac{a_1(t)}{N_1} U - b_{13}(t) U \right] dt + \sigma_1(t) dB_1(t) \]

\[ d\ln E = \left[ a_2(t) - \frac{1}{2} \sigma_2(t)^2 - \frac{a_2(t)}{N_2} E + b_{23}(t) U \right] dt + \sigma_2(t) dB_2(t) \]

\[ d\ln U = \left[ a_3(t) - \frac{1}{2} \sigma_3(t)^2 - \frac{a_3(t)}{N_3} U - b_{31}(t) C - b_{32}(t) E \right] dt + \sigma_3(t) dB_3(t) \]

According to the property of the superior limit, for any \( \varepsilon > 0 \), there exists a \( T > 0 \) such that for any \( t > T \) there is

\[ \frac{1}{t} \int_0^t \left( a_3(s) - \frac{1}{2} \sigma_3^2(s) \right) ds \leq \limsup_{t \to +\infty} \frac{1}{t} \int_0^t \left( a_3(s) - \frac{1}{2} \sigma_3^2(s) \right) ds + \frac{\varepsilon}{2} \]

\[ \frac{1}{t} \int_0^t \sigma_3(s) dB_3(s) \leq \frac{\varepsilon}{2} \]

Substituting inequalities (31) and (32) into (30), we can get

\[ \frac{1}{t} \int_0^t \left( a_3(s) - \frac{1}{2} \sigma_3^2(s) \right) ds \leq \limsup_{t \to +\infty} \frac{1}{t} \int_0^t \left( a_3(s) - \frac{1}{2} \sigma_3^2(s) \right) ds + \varepsilon \]

\[ \frac{1}{t} \int_0^t \sigma_3(s) dB_3(s) \leq \frac{\varepsilon}{2} \]

When \( \limsup_{t \to +\infty} \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds = 0 \), there is \( \ln \frac{\bar{U}(t)}{\bar{U}(0)} \leq et - \left( a_3/N_3 \right)^{t} \int_0^t U(s) ds \). And because of \( (a_3/N_3)^{t} > 0 \), it can be obtained from Lemma 1 and the arbitrariness of \( \varepsilon \)

\[ \limsup_{t \to +\infty} \frac{1}{t} \int_0^t U(s) ds \leq \frac{e}{(a_3/N_3)^{t}} \leq 0 \]

Since the solution of model (9) is nonnegative, it is easy to get

\[ \limsup_{t \to +\infty} \frac{1}{t} \int_0^t U(s) ds = 0 \]

(3) According to the formula (24) (25) and (26), we can get

\[ \ln \frac{C(t) - \ln C(0)}{t} \leq \frac{1}{t} \int_0^t \left[ a_1(s) - \frac{1}{2} \sigma_1^2(s) \right] ds \]

\[ - \left( a_1/N_1 \right) \frac{1}{t} \int_0^t \int_0^s C(s) ds + b_{13} \frac{1}{t} \int_0^t U(s) ds + \frac{1}{t} \int_0^t \sigma_1(s) dB_1(s) \]

\[ \ln \frac{E(t) - \ln E(0)}{t} \leq \frac{1}{t} \int_0^t \left[ a_2(s) - \frac{1}{2} \sigma_2^2(s) \right] ds \]

\[ - \left( a_2/N_2 \right) \frac{1}{t} \int_0^t \int_0^s E(s) ds + b_{23} \frac{1}{t} \int_0^t U(s) ds + \frac{1}{t} \int_0^t \sigma_2(s) dB_2(s) \]

\[ \ln \frac{U(t) - \ln U(0)}{t} \leq \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds \]

\[ - \left( a_3/N_3 \right) \frac{1}{t} \int_0^t \int_0^s U(s) ds + \frac{1}{t} \int_0^t \sigma_3(s) dB_3(s) \]

From Lemma 1, it is concluded that

\[ \limsup_{t \to +\infty} \frac{1}{t} \int_0^t C(s) ds \leq \frac{1}{t} \int_0^t \left[ a_1(s) - \frac{1}{2} \sigma_1^2(s) \right] ds \]

\[ \left( a_1/N_1 \right)^{t} \frac{1}{t} \int_0^t U(s) ds \]

\[ \limsup_{t \to +\infty} \frac{1}{t} \int_0^t E(s) ds \leq \frac{1}{t} \int_0^t \left[ a_2(s) - \frac{1}{2} \sigma_2^2(s) \right] ds \]

\[ \left( a_2/N_2 \right)^{t} \frac{1}{t} \int_0^t U(s) ds \]

\[ \limsup_{t \to +\infty} \frac{1}{t} \int_0^t U(s) ds \leq \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds \]

(38)

(39)

(40)

(41)

\[ \ln U(t) - \ln U(0) \]

\[ \geq \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds - \left( a_3/N_3 \right)^{t} \frac{1}{t} \int_0^t U(s) ds \]

\[ - b_{31} \frac{1}{t} \int_0^t \int_0^s U(s) ds - \frac{1}{t} \int_0^t \sigma_3(s) dB_3(s) \]

\[ \ln U(t) - \ln U(0) \geq \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds \geq \liminf_{t \to +\infty} \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds - \frac{\varepsilon}{2} \]

\[ \frac{1}{t} \int_0^t \sigma_3(s) dB_3(s) \geq - \frac{\varepsilon}{2} \]
It can be obtained from Lemma 1 and the arbitrariness of $\varepsilon$

$$\liminf_{t \to +\infty} \frac{1}{t} \int_0^t U(s)ds \geq \frac{A}{(a_1/N_1)^{1/2} + b_{13}^{1/2}} + (a_3/N_3)^{1/2} > 0 \quad (44)$$

where

$$A = \liminf_\frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds$$

Substituting the above formula into formula (48), we can get the following results

$$\liminf_{t \to +\infty} \frac{1}{t} \int_0^t U(s)ds \geq \frac{1}{\int_0^t U(s)ds} \geq C(s)ds \quad (53)$$

In other words, $U(t)$ is sustainable.

Hence, Theorem 2 follows.

**Proof of Theorem 3.**

Proof (1) From formulas (35) and (40), we can get

$$\ln C(t) - \ln C(0) \leq \frac{1}{t} \int_0^t \left[ a_1(s) - \frac{1}{2} \sigma_1^2(s) \right] ds + \frac{b_{13}^{1/2}}{(a_3/N_3)^{1/2}} \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds \leq 0$$

Take the superior limit on both sides of inequality to get

$$\limsup_{t \to +\infty} \frac{1}{t} \int_0^t \left[ a_1(s) - \frac{1}{2} \sigma_1^2(s) \right] ds + \limsup_{t \to +\infty} \frac{b_{13}^{1/2}}{(a_3/N_3)^{1/2}} \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds < 0$$

Therefore, $\lim C(t) = 0$, $C(t)$ is unserviceable.

(2) In the same way, it can be concluded that

$$\limsup_{t \to +\infty} \frac{1}{t} \int_0^t C(s)ds = 0$$

(3) It can be seen from formula (24)

$$\ln C(t) - \ln C(0) \leq \frac{1}{t} \int_0^t \left[ a_1(s) - \frac{1}{2} \sigma_1^2(s) \right] ds + \frac{b_{13}^{1/2}}{(a_3/N_3)^{1/2}} \frac{1}{t} \int_0^t \left[ a_3(s) - \frac{1}{2} \sigma_3^2(s) \right] ds$$

$$\leq \frac{a_3(t)}{N_3} \int_0^t [U(t) - U(t)] + b_{31} C(t)$$

Integrate from 0 to $t$ and divide by $t$ on both sides of the inequality to get

$$\frac{V(t) - V(0)}{t} \leq b_{31} C(t) - \frac{a_3(t)}{N_3} \int_0^t [U(t) - U(t)]$$

Integrate from 0 to $t$ and divide by $t$ on both sides of the inequality to get

$$\frac{V(t) - V(0)}{t} \leq b_{31} \frac{1}{t} \int_0^t C(s)ds$$

$$\leq \frac{a_3(t)}{N_3} \int_0^t [U(t) - U(t)]$$

Because $\frac{a_3(t)}{N_3} \geq 0$, then

$$\frac{1}{t} \int_0^t [U(t) - U(t)] ds \leq \frac{b_{31}^{1/2}}{(a_3/N_3)^{1/2}} \frac{1}{t} \int_0^t C(s)ds \quad (51)$$

**Acknowledgements** The authors are grateful to all study participants.

**Authors contribution** Bingjie Liu conceived the main idea and wrote the manuscript; all authors contributed to data analysis, simulations, and the writing of this paper.

**Funding** This research work was not funded by any organization/institute/agency.
mechanism for security of mobile edge computing system,” IEEE Access, 2019, no. 99, pp.1-1.


