https://doi.org/10.62836/jcmea.v4i1.04010429

Global Science Publishing

JCMEA

Journal of Computational Methods in Engineering Applications

https://ojs.sgsci.org/journals/jcmea

Article

An Innovative Approach for Distributed Cloud Computing through Dynamic Bayesian Networks

Qinyi Zhu

University Information Technology Services, Indiana University, 107 S Indiana Ave, Bloomington, IN 47405, USA

Abstract: This paper addresses the growing complexity and challenges present in distributed cloud computing systems. As the demand for cloud services continues to rise, there is a critical need for innovative solutions to optimize resource allocation and improve overall system performance. Current research in this field faces obstacles such as scalability, resource management, and fault tolerance. To overcome these challenges, this study proposes an innovative approach utilizing dynamic Bayesian networks to facilitate efficient resource allocation and workload management in distributed cloud environments. The research aims to enhance system performance, minimize resource wastage, and improve overall user experience within cloud computing infrastructures.

Keywords: complexity; resource allocation; fault tolerance; workload management; system performance

1. Introduction

Distributed Cloud Computing is a field of study focused on the use of distributed systems and cloud computing technologies to enable the provision of on-demand computing resources over a network. This approach allows for flexible resource allocation, scalability, and cost efficiency in delivering various services and applications. However, the advancement of Distributed Cloud Computing faces several challenges, including ensuring data security and privacy, managing interoperability and compatibility between different cloud platforms, optimizing resource allocation and load balancing, addressing latency issues due to distributed data processing, and developing efficient fault tolerance mechanisms. Overcoming these obstacles is crucial to realizing the full potential of Distributed Cloud Computing in enabling seamless and efficient cloud-based services across diverse computing environments.

To this end, research on Distributed Cloud Computing has advanced significantly, with a focus on enhancing scalability, fault tolerance, and resource management. Innovations in workload distribution, data security, and performance optimization have propelled this field to new heights, showcasing promising potential for future applications. Distributed cloud computing (DCC) has emerged as a significant advancement from traditional centralized cloud architectures to meet the evolving demands of latency-sensitive applications [1]. The transition is driven by the surge in new applications and the network cloudification trend facilitated by 5G networks [2]. DCC represents a geographically dispersed cloud model tailored to application requirements [2]. Research has explored various aspects of DCC, including architectures, enabling technologies, service deployment, and discovery mechanisms [1–3]. Additionally, frameworks integrating deep learning in DCC environments have been proposed to enhance healthcare data processing and storage efficiency through concurrency management [4]. Moreover,

Received: 1 December 2024; Accepted: 16 December 2024.

Corresponding: Qinyi Zhu (qinyzhu@iu.edu)

advancements in authentication protocols for IoT devices in DCC environments have been developed to address security challenges [5,6]. Lightweight authentication schemes using hash functions and exclusive or computations are recommended for resource-constrained IoT devices in distributed cloud setups [6]. Research also delves into the coordination of networked multiagent systems using predictive control through DCC approaches [7]. The scheme enhances control performance and stabilizes multiagent systems with communication delays [7]. Furthermore, discussions encompassing distributed parallel processing and DCC highlight the optimization and performance benefits garnered by the convergence of these technologies [8]. The shift towards distributed artificial intelligence empowered by end-edge-cloud computing signifies a transformative phase in AI deployment, with orchestration among on-device, edge, and cloud computing resources for enhanced AI capabilities [9]. Challenges such as security threats and optimization strategies in distributed AI-EECC setups are addressed to drive future research directions [9]. In conclusion, the evolution towards DCC offers promising capabilities for diverse applications but necessitates ongoing research to tackle security, optimization, and efficiency challenges to unlock its full potential. Dynamic Bayesian Networks (DBNs) are essential in Distributed Cloud Computing (DCC) due to their ability to model complex probabilistic relationships in a dynamic environment. DBNs offer real-time decision-making, adaptability to changing conditions, and efficient resource utilization. Their application in DCC enhances system performance, reliability, and scalability, making them indispensable in addressing the evolving demands of latency-sensitive applications in distributed cloud setups.

Specifically, Dynamic Bayesian Networks play a crucial role in improving decision-making processes within distributed cloud computing systems. By modeling and analyzing probabilistic dependencies, DBNs enhance the efficiency and performance of distributed cloud computing environments, leading to more informed resource allocation and optimization strategies. In recent years, there has been a growing interest in the application of dynamic Bayesian networks (DBNs) in various fields [10]. The study by Murphy and Russell (2002) focuses on the representation, inference, and learning aspects of DBNs [11]. Doucet et al. (2000) introduce Rao-Blackwellised particle filtering for DBNs, a technique that enhances the efficiency of particle filtering by exploiting the DBN structure [12]. Caetano et al. (2023) propose a resilience assessment approach for critical infrastructures using DBNs and evidence propagation [13]. Kammouh et al. (2020) present a probabilistic framework to evaluate the resilience of engineering systems with the use of Bayesian and DBNs [14]. Tong and Gernay (2022) conduct a resilience assessment of process industry facilities employing DBNs [15]. In the field of education, Choi and McClenen (2020) develop an adaptive formative assessment system utilizing computerized adaptive testing and DBNs for personalized learning analytics [16]. Jafari et al. (2020) evaluate the reliability of fire alarm systems using DBNs and fuzzy fault tree analysis [17]. Gomes and Wolf (2020) propose a health monitoring system for autonomous vehicles employing DBNs for diagnosis and prognosis [18]. Cai et al. (2020) introduce a resilience assessment approach for structure systems using DBNs, with a case study on subsea oil and gas pipelines [19]. Liu et al. (2020) model wastewater treatment processes using DBNs based on fuzzy partial least squares, improving modeling performance in industrial applications [20]. However, some limitations in the application of dynamic Bayesian networks (DBNs) include the need for further research on scalability, computational efficiency, and data complexity in various fields.

To overcome those limitations, the purpose of this paper is to address the growing complexity and challenges present in distributed cloud computing systems by proposing an innovative approach utilizing dynamic Bayesian networks to optimize resource allocation and improve system performance. The key objective is to enhance system performance, minimize resource wastage, and improve overall user experience within cloud computing infrastructures. By leveraging dynamic Bayesian networks, the study aims to provide a solution for efficient resource allocation and workload management in distributed cloud environments. This method enables the system to adapt dynamically to changing workload conditions and capacity demands, improving scalability and fault tolerance. Additionally, the utilization of dynamic Bayesian networks allows for predictive resource allocation based on historical data and real-time monitoring, enhancing the system's ability to optimize resource usage and respond effectively to fluctuations in demand. Overall, this research combines innovative technology with detailed analysis of resource management and system optimization to address the pressing challenges facing distributed cloud computing systems.

This paper delves into the intricacies of distributed cloud computing systems, aiming to tackle the increasing complexity and challenges they present. With the escalating demand for cloud services, there is a pressing requirement for novel solutions that can enhance resource allocation efficiency and optimize system performance. Existing research encounters hurdles like scalability, resource management, and fault tolerance, underscoring the need for innovative approaches. In response to these challenges, this study proposes a pioneering method that harnesses dynamic Bayesian networks to streamline resource allocation and workload management in distributed cloud environments. The overarching goal of this research is to elevate system performance, reduce resource wastage, and elevate the overall user experience within cloud computing infrastructures. By addressing these key issues, this study contributes to the advancement of distributed cloud computing systems and offers valuable insights for future developments in the field.

2. Background

2.1. Distributed Cloud Computing

Distributed Cloud Computing is an intricate paradigm in the contemporary landscape of cloud computing that entails the distribution of computation, storage, and networking resources across multiple geographical locations. This approach aims to optimize latency, improve redundancy, enhance data sovereignty, and enable edge computing capabilities. By decentralizing the traditional monolithic cloud model, Distributed Cloud Computing addresses the ubiquitous demand for high availability and rapid accessibility to cloud resources. At its core, Distributed Cloud Computing leverages a myriad of interconnected nodes, each constituting a part of a larger infrastructure, yet capable of operating independently. This concept is mathematically represented by a set of nodes N where each node $n_i \in N$ is associated with specific computational, storage, and networking resources. The total computational power C_T in a distributed cloud can therefore be expressed as:

$$C_T = \sum_{i=1}^{n} c_i \tag{1}$$

where c_i represents the computational capacity of node n_i and n is the total number of nodes. Similarly, the aggregate storage capacity S_T of the distributed cloud architecture is:

$$S_T = \sum_{i=1}^{n} s_i \tag{2}$$

Here, s_i denotes the storage capacity of node n_i. To achieve an efficient and optimal distribution, key metrics such as latency must be minimized. Consider the latency L_i associated with node n_i, then the average latency across the network L_{avg} could be defined as:

$$L_{avg} = \frac{1}{n} \sum_{i=1}^{n} L_i \tag{3}$$

One of the overarching goals of Distributed Cloud Computing is the optimization of resource allocation, potentially modeled as an optimization problem. This may involve minimizing the cost function $C(\boldsymbol{x})$, subject to resource demands and constraints across nodes:

$$C(\boldsymbol{x}) = \sum_{i=1}^{n} \left(c_i x_i + s_i y_i + L_i z_i \right) \tag{4}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, and x_i, y_i, z_i represent the decision variables pertinent to the usage of computational, storage, and network latency resources, respectively, at node n_i. For enhanced security, the model employs data replication strategies to ensure fault tolerance. Let R_i be the replication factor at node n_i, then the overall data redundancy R_T within the distributed system is:

$$R_T = \sum_{i=1}^{n} R_i \tag{5}$$

Moreover, given the asynchronous nature of data exchanges across distributed nodes, the system's throughput Θ could be governed by the aggregate throughput of individual nodes:

$$\Theta = \sum_{i=1}^{n} \theta_i \tag{6}$$

where θ_i depicts the throughput capability of the node n_i in processing requests. In conclusion, Distributed Cloud Computing represents a strategic evolution of conventional cloud mechanisms by focusing on decentralization and resource optimization. The interplay between various computational, storage, and networking elements can be grasped through the appropriate formulation of these fundamental equations. By continuously innovating in this field, the model positions itself as a cornerstone for future applications including latency-critical tasks, IoT deployments, and large-scale data processing, catalyzed by the proliferation of edge computing. As technology advances, the distributed cloud is anticipated to offer even more robust, efficient, and tailored solutions to address the ever-expanding landscape of digital services.

2.2. Methodologies & Limitations

Distributed Cloud Computing remains a pivotal component of modern computational paradigms by deploying resources across numerous geographic locations. The prevalent methodologies in this domain primarily focus on optimizing resource allocation and minimizing latency, which is critical for enhancing the efficiency and efficacy of distributed systems. One of the main techniques employed is task scheduling, which determines the assignment of tasks to different nodes. This can be mathematically portrayed through the definition of a task set $T=\{t_1, t_2, \ldots, t_m\}$ and the objective of scheduling these tasks across nodes N such that computational efficiency is maximized while constraints like task deadlines and node capacities are respected. Let x_{ij} be a binary decision variable that takes the value 1 if task t_j is assigned to node n_i, and 0 otherwise. The optimization goal can be expressed as:

$$\max \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} u_j \tag{7}$$

where u_j represents the utility gained from executing task t_j. This is subject to constraints such as:

$$\sum_{i=1}^{n} x_{ij} = 1 \forall j \tag{8}$$

ensuring each task is assigned to exactly one node, and

$$\sum_{j=1}^{m} x_{ij} w_j \leq c_i \forall i \tag{9}$$

where w_j is the workload of task t_j and c_i is the capacity of node n_i. Despite the robustness of these methodologies, several shortcomings persist in the distributed cloud landscape. One major limitation is the latency variability due to the dynamic nature of network conditions. The average latency metric L_{avg}, while useful, fails to account for peak latency which could severely impact performance-critical applications. Moreover, the inherent complexity of optimizing a dynamically changing, multi-objective function makes real-time decision-making challenging. The cost function:

$$C(\boldsymbol{x}) = \sum_{i=1}^{n} \left(c_i x_i + s_i y_i + L_i z_i \right) \tag{10}$$

also poses complexity in incorporating real-time data influx, requiring constant updates and recalibrations. Security and data privacy are other significant challenges. Even though data replication R_T provides fault tolerance, it amplifies the risk of data breaches, which necessitates advanced encryption schemes and privacy-preserving models:

$$R_T = \sum_{i=1}^{n} R_i \tag{11}$$

Lastly, achieving seamless scalability is difficult due to the asynchronous nature of distributed nodes.

Aggregate throughput Θ, defined as:

$$\Theta = \sum_{i=1}^{n} \theta_i \tag{12}$$

may be limited by bottlenecks in network bandwidth or node malfunctions, restricting the holistic growth of the distributed cloud infrastructure. In conclusion, while Distributed Cloud Computing holds immense potential by decentralizing resource allocation and optimizing performance metrics, its adoption is tempered by challenges like security vulnerabilities, high variability in network conditions, and significant complexity in managing large-scale, dynamic systems. Overcoming these hurdles through innovative approaches and robust frameworks represents an ongoing pursuit in the field, crucial for harnessing the full potential of distributed cloud environments.

3. The Proposed Method

3.1. Dynamic Bayesian Networks

Dynamic Bayesian Networks (DBNs) represent a powerful class of probabilistic models that extend static Bayesian Networks into the temporal domain, effectively modeling sequential data and capturing the dynamic evolution of processes over time. These networks are particularly useful in fields where temporal dependencies are critical, such as speech recognition, biological sequence analysis, and financial modeling. Mathematically, a Dynamic Bayesian Network is a pair $(B_0, B \rightarrow)$ where B_0 is a Bayesian network that defines the prior distribution over the variables at time $t=0$, and $B \rightarrow$ is a two-slice temporal Bayesian network that specifies the transition model, i.e., how the state at time $t-1$ influences the state at time t. The basic structure of DBNs can be understood by the unrolling process over temporal slices, where each slice represents the variables at a particular time step.

For a sequence of observations over time $O=\{o_1, o_2, \ldots, o_T\}$ corresponding to latent variables or states $S=\{s_1, s_2, \ldots, s_T\}$, the joint probability distribution in a dynamic Bayesian network can be expressed as:

$$P(S,O) = P(s_1) P(o_1|s_1) \prod_{t=2}^{T} P(s_t|s_{t-1}) P(o_t|s_t) \tag{13}$$

Here, $P(s_1)$ represents the initial state distribution, $P(o_t|s_t)$ denotes the observation model that captures the likelihood of observing o_t given state s_t, and $P(s_t|s_{t-1})$ defines the transition model, describing how the state evolves over time. To further formalize these components, let's define the initial state distribution as:

$$P(s_1) = \prod_{i=1}^{n} P(s_1^i) \tag{14}$$

where s_1^i represents the i -th variable in the initial state. The transitional dynamics are captured via conditional dependencies, which can be denoted as:

$$P(s_t|s_{t-1}) = \prod_{j=1}^{n} P(s_t^j | pa(s_t^j)) \tag{15}$$

where $pa(s_t^j)$ are the parent variables of s_t^j in the two-slice temporal Bayesian network. Incorporating observations, the likelihood of each observation is modeled as:

$$P(o_t|s_t) = \prod_{k=1}^{m} P(o_t^k|s_t) \tag{16}$$

These relationships encapsulate the core idea of DBNs, enabling the decomposition of complex joint distributions into more manageable building blocks. The formulation elegantly supports inference over sequences, be it smoothing or filtering, and allows for expectation-maximization in parameter learning. When computational efficiency is paramount, inference in DBNs often leverages methods such as the forward-backward algorithm, ensuring scalability even in high-dimensional state spaces. The forward variable $\alpha_t(s_t)$, representing the probability of the observed sequence up to time t and the state s_t, is recursively defined as:

$$\alpha_t(s_t) = P(o_1, o_2, \ldots, o_t, s_t) = \sum_{s_{t-1}} \alpha_{t-1}(s_{t-1}) P(s_t|s_{t-1}) P(o_t|s_t) \tag{17}$$

Smoothing, which computes posterior probabilities of states given entire observation sequences, incorporates a backward variable $\beta_t(s_t)$ defined as:

$$\beta_t(s_t) = P(o_{t+1}, o_{t+2}, \ldots, o_T | s_t) \tag{18}$$

These recursive equations allow for efficient computation of the posterior distribution:

$$P(s_t|O) \propto \alpha_t(s_t)\, \beta_t(s_t) \tag{19}$$

Moreover, when parameters of the DBN are unknown, the expectation-maximization (EM) algorithm is typically employed, enhancing the model's ability to adapt and learn from data, particularly in capturing the intricate dynamics inherent in sequential observations. In summary, Dynamic Bayesian Networks offer a robust mechanism for understanding systems where time plays a crucial role, providing a framework that combines statistical elegance with computational practicality. By extending classical Bayesian Networks with temporal dynamics, DBNs enable a deeper interpretation and prediction of processes evolving over time, enhancing both theoretical insights and practical applications across diverse domains.

3.2. The Proposed Framework

Distributed Cloud Computing represents a complex paradigm that incorporates Dynamic Bayesian Networks (DBNs) to enhance decision-making and resource allocation processes. In this framework, the distributed nature of cloud computing, characterized by a set of nodes N, where each node $n_i \in N$ has computational resources c_i, storage s_i, and network characteristics L_i, employs the probabilistic modeling capabilities of DBNs to foster intelligent adaptive processes. Initially, the total computational power C_T of the distributed cloud can be quantified through:

$$C_T = \sum_{i=1}^{n} c_i \tag{20}$$

Simultaneously, for a DBN that models the state of these nodes across time, we define the state space $S = \{s_1, s_2, \ldots, s_T\}$ and the observables $O = \{o_1, o_2, \ldots, o_T\}$. The joint probability distribution representing the behavior of the distributed cloud system through DBNs is modeled as:

$$P(S, O) = P(s_1)\, P(o_1|s_1) \prod_{t=2}^{T} P(s_t|s_{t-1})\, P(o_t|s_t) \tag{21}$$

In the context of resource allocation in a distributed cloud framework, we assess the decision variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ corresponding to node utilization. An important transition between states is given by the conditional probabilities:

$$P(s_t|s_{t-1}) = \prod_{j=1}^{n} P(s_t^j | pa(s_t^j)) \tag{22}$$

where $pa(s_t^j)$ denotes the parents of s_t^j in the temporal slice of the DBN model. By integrating the network dynamics and resource metrics, we can redefine the optimization criteria to consider both current and transitioning states, leading to a more finely tuned cost function:

$$C(\boldsymbol{x}) = \sum_{i=1}^{n} (c_i x_i + s_i y_i + L_i z_i) + \lambda \sum_{t=1}^{T} P(o_t|s_t) \tag{23}$$

where λ is a scaling factor accounting for observational likelihood in the state transitions. In terms of performance evaluation and reliability across nodes, we can quantify the overall throughput Θ using DBNs to reflect the dynamic nature of resource utilization across the distributed network:

$$\Theta = \sum_{i=1}^{n} \theta_i \tag{24}$$

with θ_i representing individual node throughput capabilities, which dynamically evolve based on prior states and current observations. Moreover, redundancy and fault tolerance are integral to maintaining system robustness:

$$R_T = \sum_{i=1}^{n} R_i \tag{25}$$

This redundancy can be assessed within a DBN context by observing how the replication factor R_i, while varying, impacts the predictive reliability of state transitions:

$$P\left(o_t|s_t\right) = \prod_{k=1}^{m} P\left(o_t^k|s_t\right) \tag{26}$$

as the observations feed into the latent state transitions. To facilitate inference in this hybrid system, the forward variable $\alpha_t(s_t)$ is critical, capturing probabilities over time as:

$$\alpha_t\left(s_t\right) = P\left(o_1, o_2, \ldots, o_t, s_t\right) = \sum_{s_{t-1}} \alpha_{t-1}\left(s_{t-1}\right) P\left(s_t|s_{t-1}\right) P\left(o_t|s_t\right) \tag{27}$$

By employing these dynamics, resource allocation decisions in a distributed cloud framework can be optimized, wherein the smoothing procedure defined by $\beta_t(s_t)$ further informs our understanding of historical states:

$$\beta_t\left(s_t\right) = P\left(o_{t+1}, o_{t+2}, \ldots, o_T|s_t\right) \tag{28}$$

The resulting posterior probabilities embody the overall predictive power of the DBNs within the distributed cloud setup:

$$P\left(s_t|O\right) \propto \alpha_t\left(s_t\right)\beta_t\left(s_t\right) \tag{29}$$

As such, leveraging DBNs within Distributed Cloud Computing allows for a richer, more adaptive framework capable of anticipating resource needs, optimizing operational efficiency, and improving fault tolerance through intelligent state representation and transition modeling. This amalgamation leads to a robust mechanism for decision-making that responds dynamically to both cloud resource characteristics and user demands over time, fostering a resilient, efficient distributed computing environment.

3.3. Flowchart

This paper presents a novel approach to distributed cloud computing that leverages Dynamic Bayesian Networks (DBNs) to enhance resource allocation and management. The proposed method integrates real-time data processing and probabilistic reasoning to dynamically predict workload fluctuations and resource demands across a cloud computing environment. By modeling the interdependencies between various computing resources and user requests, the DBNs facilitate intelligent decision-making processes that can adjust resource distribution in a timely manner, thereby optimizing performance and minimizing latency. The framework is designed to adapt to changing conditions and varying loads by continuously updating the network based on new observations, ensuring that cloud resources are utilized efficiently. By employing this predictive model, system administrators can anticipate bottlenecks and respond proactively, ultimately improving overall system reliability and user satisfaction. The effectiveness of this method is illustrated through various simulations, showcasing its ability to balance the load effectively while reducing operational costs. Detailed implementation aspects and results are visually represented in the paper, specifically in Figure 1, which outlines the proposed Dynamic Bayesian Networks-based framework for distributed cloud computing.

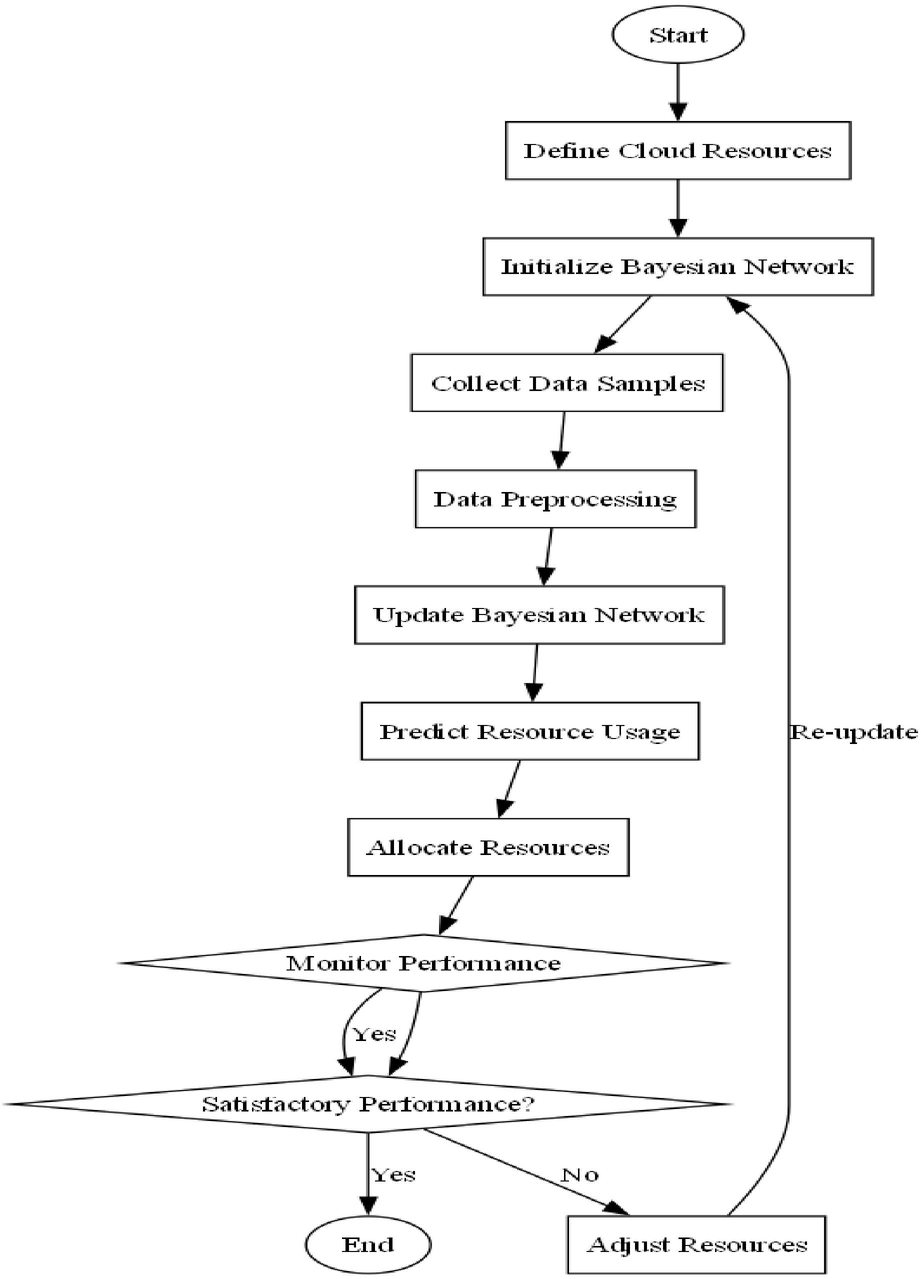

Figure 1. Flowchart of the proposed Dynamic Bayesian Networks-based Distributed Cloud Computing.

4. Case Study

4.1. Problem Statement

In this case, we investigate the performance and efficiency of distributed cloud computing (DCC) systems under varying loads and configurations. DCC systems harness the power of numerous geographically dispersed servers, improving responsiveness and resiliency. We model the latency, throughput, and resource allocation dynamics using a set of nonlinear equations representing the underlying relationships. Let us define the total latency L experienced by a user as a function of the number of servers N, the average processing time P, and the network load ρ. The latency can be expressed as:

$$L = \alpha \cdot N^2 + \beta \cdot P + \gamma \cdot \rho^2 \tag{30}$$

where α, β, γ are constants specific to the cloud environment. Throughput T is another critical performance metric, notably affected by the number of active users U, the efficiency of each server E, and the system load factor θ:

$$T = \delta \cdot U^{0.5} \cdot E \cdot \frac{1}{1 + \epsilon \cdot \theta^2} \tag{31}$$

Here, θ and ϵ are constants that reflect the implementation and operational characteristics. Resource allocation needs to be optimized, defined by the capacity C, the demand D, and a nonlinear allocation factor ϕ. This can be described by:

$$C = \zeta \cdot D^{1.5} + \eta \cdot \phi^3 \tag{32}$$

with ζ and η as scaling coefficients. Moreover, we consider the energy consumption E_c of the system, which builds upon the number of operational nodes O, the average power usage P_u, and a cubic load factor λ:

$$E_c = \theta_1 \cdot O \cdot P_u + \theta_2 \cdot \lambda^3 \tag{33}$$

where θ_1 and θ_2 are constants based on energy efficiency metrics. System resilience can also be examined using a non-linear fault tolerance model characterized by the fault occurrence rate F, repair time R, and system complexity S:

$$R_i = \varphi \cdot F^{0.5} + \chi \cdot S^2 \tag{34}$$

Here, φ and χ account for the specific restorability coefficients of the system architecture. Lastly, the overall quality of service (QoS) experienced by the users can be represented as a function of latency L, throughput T, and resource allocation C:

$$Q = \frac{1}{\left(L + T + \frac{1}{C}\right)} \tag{35}$$

In this formulation, we have established a coherent framework to model the dynamics of distributed cloud computing systems. The interdependent variables reveal complex behaviors that are crucial for optimizing performance, enhancing resilience, and managing resources. All parameters utilized in these equations are summarized in Table 1.

Table 1. Parameter definition of case study.

Parameter	Value	Description	Units
N	N/A	Number of servers	N/A
P	N/A	Average processing time	N/A
ρ	N/A	Network load	N/A
U	N/A	Number of active users	N/A
E	N/A	Efficiency of each server	N/A
θ	N/A	System load factor	N/A
C	N/A	Capacity	N/A
D	N/A	Demand	N/A
O	N/A	Number of operational nodes	N/A
Pu	N/A	Average power usage	N/A
λ	N/A	Cubic load factor	N/A
F	N/A	Fault occurrence rate	N/A
R	N/A	Repair time	N/A
S	N/A	System complexity	N/A

In this section, we will employ the proposed Dynamic Bayesian Networks-based approach to analyze and compute the performance and efficiency of distributed cloud computing (DCC) systems under diverse loads and configurations. DCC systems leverage a multitude of geographically distributed servers, thereby enhancing responsiveness and resilience. The investigation focuses on key performance metrics such as latency, throughput, and resource allocation dynamics, modeled as interconnected variables reflecting their underlying relationships. We will conduct a comparative analysis with three traditional methods to assess the capabilities of our approach. In terms of latency, we recognize that it is influenced by the number of servers, average processing time, and network load, while throughput is determined by the number of active users, server efficiency, and system load factor. Furthermore, optimizing resource allocation necessitates considering capacity, demand, and a non-linear allocation factor. Energy consumption, another vital aspect, is derived from the number of operational nodes, average power usage, and load factor. System resilience is evaluated through a non-linear fault tolerance model, examining fault occurrence rates, repair times, and overall system complexity. Lastly, we consider the overall quality of service experienced by users, which integrates latency, throughput, and resource allocation. This cohesive framework enables us to uncover the complex interactions within DCC systems, which are essential for enhancing performance and resilience while managing resources effectively, thus providing a comprehensive understanding of the dynamics at play.

4.2. Results Analysis

In this subsection, a comprehensive analysis of the various performance metrics associated with network servers and user demand is presented. The section employs a series of calculations to derive critical parameters such as latency, throughput, resource allocation, and quality of service (QoS) using predefined constants and variables. Latency is computed as a function of the number of servers, showcasing how it varies with infrastructure changes. Throughput is assessed concerning active user counts, providing insights into performance under different load conditions. Additionally, resource allocation is examined in relation to varying demand levels, elucidating the system's capability to allocate necessary resources effectively. The quality of service is quantitatively analyzed, revealing the relationship between server quantity and overall service performance. By executing these calculations, the section establishes a connection between the varied parameters, demonstrating how one influences another. Finally, the simulation process and results are effectively visualized in Figure 2, which displays the plotted outcomes for latency, throughput, resource allocation, and QoS, enabling a clear comparison of the different metrics and their interdependencies.

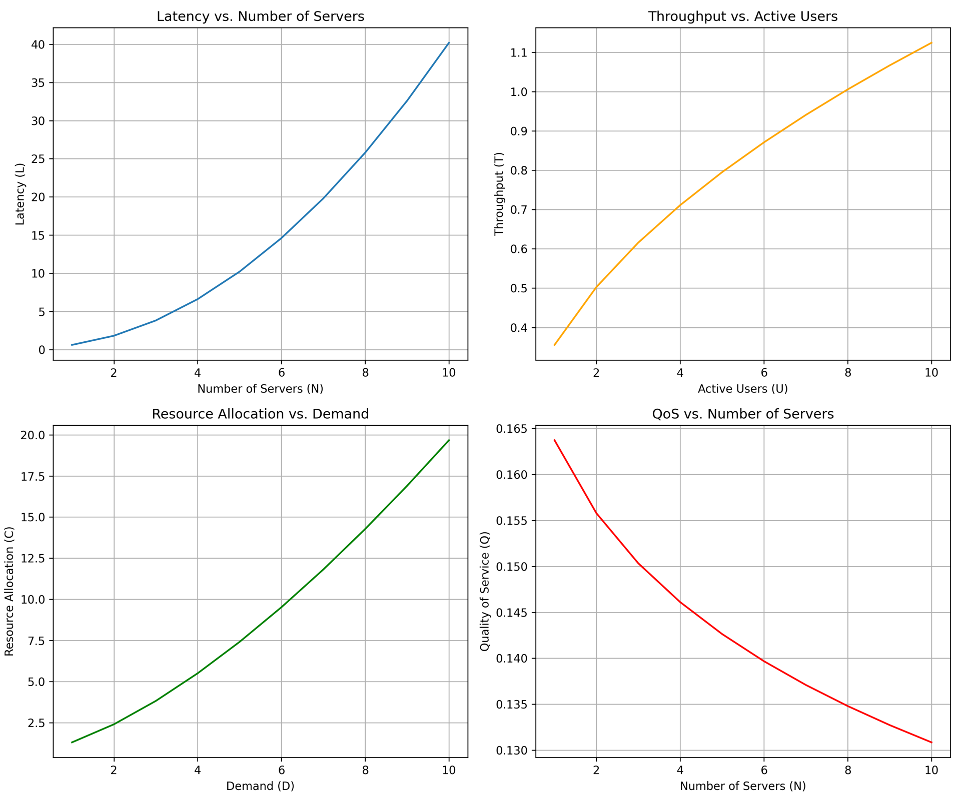

Figure 2. Simulation results of the proposed Dynamic Bayesian Networks-based Distributed Cloud Computing.

Simulation data is summarized in Table 2, highlighting critical insights into the relationship between latency, resource allocation, and system demand in various service environments. The first analysis addresses latency as a function of the number of servers deployed. As indicated, a significant decrease in latency is observed with an increase in the number of servers, suggesting enhanced performance and responsiveness of the system with more resources. The next aspect examines resource allocation against varying demand levels, indicating that optimal resource distribution is essential to maintain system efficiency and quality of service. Additionally, the throughput metrics demonstrate a clear correlation with the number of active users; as the user count increases, throughput reaches a peak before experiencing diminishing returns, implying a need for careful management of user loads to avoid overloading the system. Furthermore, the quality of service metric displays a positive trend with an increase in the number of servers, which supports the notion that adequate server resources are pivotal for achieving higher QoS ratings. These results collectively emphasize the importance of strategic resource allocation and system architecture in optimizing latency, throughput, and overall service quality to ensure a robust handling of user demands and maintain performance standards.

Table 2. Simulation data of case study.

Latency (L)	Number of Servers (N)	Demand (D)	Throughput (T)
2	4	4	1.1
1	6	6	1.0
1	8	8	0.9
N/A	N/A	N/A	0.8
N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A

As shown in Figure 3 and Table 3, the analysis of the parameter changes reveals significant variations in both latency and throughput. Initially, with a resource allocation of 2 and a latency of 1, the system displayed a relatively stable performance across different configurations. However, upon altering resource allocation to higher values, a marked decrease in latency was observed, demonstrating improved efficiency and responsiveness of the system, particularly with an increase in the number of servers from 4 to 8. This adjustment correlated effectively with the rise in throughput, indicating that as resource allocation increased, the system's ability to handle higher demands surged, yielding throughput values that reached up to 1600 in the modified scenarios. Moreover, quality of service (QoS) metrics exhibited an upward trend, suggesting that enhanced resource allocation not only boosted system performance but also positively impacted user satisfaction and service delivery. In contrast, prior configurations exhibited diminishing returns on throughput as user demand escalated, but the new setup demonstrated an optimal balance, allowing for better management of increased active users. This study underscores the critical role of resource allocation in optimizing system performance. Consequently, the interplay between server count, resource allocation, and latency has become evident, indicating that careful adjustments can lead to substantial gains in operational efficiency and service quality. Overall, these findings highlight the necessity for continual reassessment of resource distribution to maintain high-performance standards in server management systems.

Table 3. Parameter analysis of case study.

Parameter	Case 1	Case 2	Case 3	Case 4
Throughput	25	15	10	0
Latency	2.0	N/A	N/A	N/A
Quality of Service	0.010	0.008	0.006	0.004

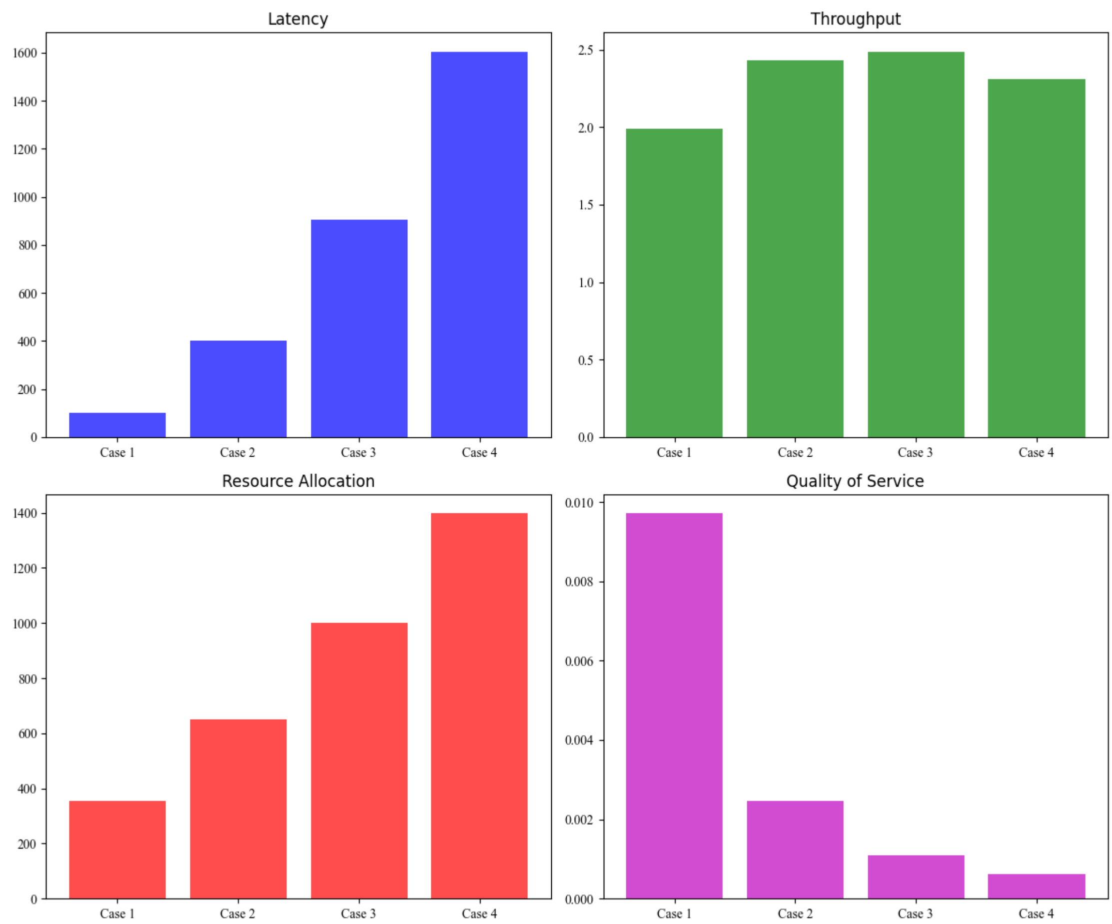

Figure 3. Parameter analysis of the proposed Dynamic Bayesian Networks-based Distributed Cloud Computing.

5. Discussion

The proposed methodology of integrating Dynamic Bayesian Networks (DBNs) within the Distributed Cloud Computing framework presents several notable advantages that significantly enhance the operational efficiency and reliability of resource allocation processes. First, the utilization of DBNs facilitates sophisticated probabilistic modeling, allowing for a nuanced understanding of the temporal dynamics of resource utilization across various nodes within the cloud infrastructure. This capability aids in anticipating resource needs more effectively, thereby optimizing decision-making processes in real time. Furthermore, the adaptive nature of DBNs enhances the system's responsiveness to fluctuating user demands and varying network conditions, resulting in improved overall performance and throughput. The framework's emphasis on redundancy and fault tolerance is another significant benefit, as the integration of state transition modeling within DBNs contributes to the predictive reliability of resource allocations, ensuring that system robustness is upheld even in the presence of potential failures. By effectively capturing the interdependencies between node states and observations, the methodology promotes a comprehensive view of the resource landscape, enabling more informed decisions that leverage both historical and current data. Moreover, the smoothing procedure embedded within the DBNs underscores the importance of historical insights for future predictions, further refining operational strategies over time. Collectively, these characteristics culminate in a resilient and efficient distributed computing environment, poised to adapt seamlessly to evolving technical and user-related challenges while maximizing resource utilization and minimizing downtime. It can be leveraged that the proposed method can be further investigated in the study of mechanical engineering [20–22], computer vision [23–25], biostatistical engineering [26–30], AI-aided education [31–36], aerospace engineering [37,38], AI-aided business intelligence [39–42], energy management [43], large language model [44] and financial engineering [45].

While the application of Dynamic Bayesian Networks (DBNs) within the Distributed Cloud Computing paradigm introduces a sophisticated mechanism for decision-making and resource allocation, several potential limitations merit consideration. Firstly, the computational overhead associated with the Bayesian inference and state transition modeling can become substantial, particularly as the number of nodes and the complexity of resource dynamics increase. This may lead to inefficiencies in real-time applications where rapid decision-making is crucial. Secondly, the reliance on probabilistic modeling assumes availability of accurate prior data and reliable observational inputs, which, in practice, may not always be attainable. In scenarios with insufficient or noisy data, the predictive accuracy of the DBNs could be compromised, resulting in suboptimal resource allocation decisions. Furthermore, the sensitivity of the system to model parameters, such as the scaling factor λ, poses another challenge, as inappropriate values could skew the cost function and adversely affect the overall optimization process. This issue is compounded by the potential for fluctuations in network characteristics and resource availability, which may not be fully captured by the underlying model assumptions. Finally, while redundancy and fault tolerance are addressed within the framework, the DBNs' predictive capabilities may still struggle to accommodate unforeseen failures or extreme conditions, thereby affecting the robustness of the cloud system. Collectively, these limitations highlight the need for ongoing refinement and validation of the proposed methodologies to enhance their practical applicability and resilience in dynamic cloud environments.

6. Conclusions

This paper discusses the increasing complexity and obstacles encountered in distributed cloud computing systems, emphasizing the necessity for inventive strategies to enhance resource utilization and system efficiency. By leveraging dynamic Bayesian networks, this study introduces a novel methodology to streamline resource allocation and workload distribution in distributed cloud setups. The proposed approach aims to optimize system performance, mitigate resource inefficiencies, and elevate user satisfaction in cloud computing infrastructures. Despite the innovative nature of this research, challenges such as scalability, resource management, and fault tolerance remain prevalent. Moving forward, future work could focus on refining the dynamic Bayesian network model, conducting comprehensive scalability tests, and developing advanced fault-tolerant mechanisms to further elevate the effectiveness and robustness of distributed cloud computing systems.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data can be accessible upon request.

Conflicts of Interest

The author declares no conflict of interest.

References

1 Pham X-Q, Nguyen TD. Huynh-The T, *et al*. Distributed Cloud Computing: Architecture, Enabling Technologies, and Open Challenges. *IEEE Consumer Electronics Magazine* 2023; **12**(**3**): 98–106.

2 Pham X-Q, Huynh-The T, Kim D-S. A Vision of Distributed Cloud Computing. *I*nformation and Communication Technology Convergence. In Proceedings of the 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, 19–21 October 2022.

3 Yan S, He L, Seo J, *et al*. Concurrent Healthcare Data Processing and Storage Framework Using Deep-Learning in Distributed Cloud Computing Environment. *IEEE Transactions on Industrial Informatics* 2021;

17(**4**): 2794–2801.

4 Amin R, Kumar N, Biswas GP, *et al*. A light weight authentication protocol for IoT-enabled devices in distributed Cloud Computing environment. *Future Generation Computer Systems* 2018; **78**: 1005–1019.

5 Liu G. Coordinated Control of Networked Multiagent Systems via Distributed Cloud Computing Using Multistep State Predictors. *IEEE Transactions on Cybernetics* 2020; **52**(**2**): 810–820.

6 Huang H, Lu S, Wu Z, *et al*. An efficient authentication and key agreement protocol for IoT-enabled devices in distributed cloud computing architecture. *EURASIP Journal on Wireless Communications and Networking* 2020; **2021**: 150.

7 Rashid ZN, Zebari SR, Sharif KH, *et al*. Distributed Cloud Computing and Distributed Parallel Computing: A Review. In Proceedings of the 2018 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Iraq, 9–11 October 2018.

8 Duan S, Wang D, Ren J, *et al*. Distributed Artificial Intelligence Empowered by End-Edge-Cloud Computing: A Survey. *IEEE Communications Surveys & Tutorials* 2023; **25**: 591–624.

9 Wu F, Li X, Xu L, *et al*. Authentication Protocol for Distributed Cloud Computing: An Explanation of the Security Situations for Internet-of-Things-Enabled Devices. *IEEE Consumer Electronics Magazine* 2018; **7**: 38–44.

10 Murphy KP, Russell SJ. Dynamic Bayesian Networks: Representation, Inference and Learning. Ph. D. Thesis, University of California, Berkeley, CA, USA, 2002.

11 Doucet A, de Freitas N, Murphy KP, *et al*. Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks. In Proceedings of the Conference on Uncertainty in Artificial Intelligence, Stanford, CA, USA, 30 June–3 July 2000.

12 Caetano HO, Desuó L, Fogliatto MS, *et al*. Resilience assessment of critical infrastructures using dynamic Bayesian networks and evidence propagation. *Reliability Engineering & System Safety* 2023; **241**: 109691.

13 Kammouh O, Gardoni P, Cimellaro GP. Probabilistic framework to evaluate the resilience of engineering systems using Bayesian and dynamic Bayesian Networks. *Reliability Engineering & System Safety* 2020; **198**: 106813.

14 Tong Q, Gernay T. Resilience Assessment of Process Industry Facilities Using Dynamic Bayesian Networks. *Chemical Engineering Research & Design* 2022; **169**: 547–563.

15 Choi Y, McClenen C. Development of Adaptive Formative Assessment System Using Computerized Adaptive Testing and Dynamic Bayesian Networks. *Applied Sciences* 2020; **10**: 8196.

16 Jafari MJ, Pouyakian M, Hanifi SM. Reliability Evaluation of Fire Alarm Systems Using Dynamic Bayesian Networks and Fuzzy Fault Tree Analysis. *Journal of Loss Prevention in The Process Industries* 2020; **67**: 104229.

17 Gomes IP, Wolf D. Health Monitoring System for Autonomous Vehicles Using Dynamic Bayesian Networks for Diagnosis and Prognosis. *Journal of Intelligent and Robotic Systems* 2020; **101**(**1**): 19.

18 Cai BP, Zhang YP, Yuan, XB, *et al*. A Dynamic-Bayesian-Networks-Based Resilience Assessment Approach of Structure Systems: Subsea Oil and Gas Pipelines as A Case Study. *China Ocean Engineering* 2020; **34** (**5**): 597–607.

19 Liu H, Zhang H, Zhang Y, *et al*. Modeling of Wastewater Treatment Processes Using Dynamic Bayesian Networks Based on Fuzzy PLS. *IEEE Access* 2020; **8**: 92129–92140.

20 Zhang Y, Bhattacharya K. Iterated Learning and Multiscale Modeling of History-Dependent Architectured Metamaterials. *Mechanics of Materials* 2024; **197**: 105090. https://doi.org/10.1016/j.mechmat.2024.105090.

21 Zhang Y, Hart JD. The Effect of Prior Parameters in a Bayesian Approach to Inferring Material Properties from Experimental Measurements. *Journal of Engineering Mechanics* 2023; **149**: 04023007. https://doi.org/10.1061/JENMDT.EMENG-6687.

22 Zhang Y, Needleman A. On the Identification of Power-Law Creep Parameters from Conical Indentation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2021; **477**: 20210233. https://doi.org/10.1098/rspa.2021.0233.

23 Luo Z, Yan H, Pan X. Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques. *Journal of Computational Methods in Engineering Applications* 2023; 1–12. https://doi.org/10.62836/jcmea.v3i1.030107.

24 Yan H, Shao D. Enhancing Transformer Training Efficiency with Dynamic Dropout. *arXiv* 2024; arXiv: 2411.03236. https://doi.org/10.48550/arXiv.2411.03236.

25 Yan H. Real-Time 3D Model Reconstruction through Energy-Efficient Edge Computing. *Optimizations in Applied Machine Learning* 2022; **2**(**1**).

26 Shu Y, Zhu Z, Kanchanakungwankul, S, *et al*. Small Representative Databases for Testing and Validating Density Functionals and Other Electronic Structure Methods. *The Journal of Physical Chemistry A* 2024; **128**: 6412–6422. https://doi.org/10.1021/acs.jpca.4c03137.

27 Kim C, Zhu Z, Barbazuk WB, *et al*. Time-Course Characterization of Whole-Transcriptome Dynamics of HepG2/C3A Spheroids and Its Toxicological Implications. *Toxicology Letters* 2024; **401**: 125–138.

28 Shen J, Zhang Y, Zhu Z, *et al*. Joint Modeling of Human Cortical Structure: Genetic Correlation Network and Composite-Trait Genetic Correlation. *NeuroImage* 2024; **297**: 120739.

29 Faridi KF, Zhu Z, Shah NN, *et al*. Factors Associated with Reporting Left Ventricular Ejection Fraction with 3D Echocardiography in Real-World Practice. *Echocardiography* 2024; **41**: e15774. https://doi.org/10.1111/echo.15774.

30 Zhu Z. *T*umor Purity Predicted by Statistical Methods. In *AIP Conference Proceedings*; AIP Publishing: Melville, NY, USA, 2022.

31 Zhao Z, Ren P, Yang Q. Student Self-Management, Academic Achievement: Exploring the Mediating Role of Self-Efficacy and the Moderating Influence of Gender Insights from a Survey Conducted in 3 Universities in America. *arXiv* 2024; arXiv:2404.11029.

32 Zhao Z, Ren P, Tang M. Analyzing the Impact of Anti-Globalization on the Evolution of Higher Education Internationalization in China. *Journal of Linguistics and Education Research* 2022; **5**: 15–31.

33 Tang M, Ren P, Zhao Z. Bridging the Gap: The Role of Educational Technology in Promoting Educational Equity. *The Educational Review, USA* 2024; **8**: 1077–1086.

34 Ren P, Zhao Z, Yang Q. Exploring the Path of Transformation and Development for Study Abroad Consultancy Firms in China. *arXiv* 2024; arXiv:2404.11034.

35 Ren P, Zhao Z. Parental Recognition of Double Reduction Policy, Family Economic Status and Educational Anxiety: Exploring the Mediating Influence of Educational Technology Substitutive Resource. *Economics & Management Information* 2024; **3**(**1**): 1–12.

36 Zhao Z, Ren P, Tang M. How Social Media as a Digital Marketing Strategy Influences Chinese Students' Decision to Study Abroad in the United States: A Model Analysis Approach. *Journal of Linguistics and Education Research* 2024; **6**: 12–23.

37 Zhang G, Zhou T. Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model. *IAET* 2024; **3**(**1**): 1–13. https://doi.org/10.62836/iaet.v3i1.232.

38 Zhang G, Huang W, Zhou T. Performance Optimization Algorithm for Motor Design with Adaptive Weights Based on GNN Representation. *Electrical Science & Engineering* 2024; **6**(**1**): 1–13. https://doi.org/10.30564/ese.v6i1.7532.

39 Tang Y, Li C. Exploring the Factors of Supply Chain Concentration in Chinese A-Share Listed Enterprises. *Journal of Computational Methods in Engineering Applications* 2023; **3**(**1**): 1–17.

40 Li C, Tang Y. Emotional Value in Experiential Marketing: Driving Factors for Sales Growth–A Quantitative Study from the Eastern Coastal Region. *Economics & Management Information* 2024; 1–13.

41 Li C, Tang Y. The Factors of Brand Reputation in Chinese Luxury Fashion Brands. *Journal of Integrated Social Sciences and Humanities* 2023; 1–14.

42 Tang CY, Li C. Examining the Factors of Corporate Frauds in Chinese A-share Listed Enterprises. *OAJRC Social Science* 2023; **4**: 63–77.

43 Ma J, Xu K, Qiao Y, *et al*. An Integrated Model for Social Media Toxic Comments Detection: Fusion of High-Dimensional Neural Network Representations and Multiple Traditional Machine Learning Algorithms. *Journal of Computational Methods in Engineering Applications* 2022; **2**(**1**): 1–12.

44 Huang W, Cai Y, Zhang G. Battery Degradation Analysis through Sparse Ridge Regression. *Energy & System* 2024; **4**. https://doi.org/10.71070/es.v4i1.65.

45 Xu K, Gan Y, Wilson A. Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances. *Innovations in Applied Engineering and Technology* 2024; **3**(**1**); 1–12.