

Uma Revisão Sistemática das Abordagens de Previsão da Disponibilidade de Recursos Computacionais em Sistemas Distribuídos: O Uso da Carta de Controle I-MR e Técnicas de Controle Estatístico de Processos

1 Introdução

A disponibilidade de recursos computacionais é um aspecto crucial em sistemas distribuídos, uma vez que afeta diretamente o desempenho, a confiabilidade e a eficiência desses sistemas. A capacidade de prever a disponibilidade desses recursos de forma precisa e confiável desempenha um papel fundamental no gerenciamento eficaz desses ambientes computacionais. A previsão da disponibilidade de recursos possibilita a tomada de decisões informadas sobre a alocação, escalabilidade e otimização dos recursos, além de permitir a detecção precoce de possíveis falhas ou gargalos.

Nesse contexto, a aplicação de técnicas de controle estatístico de processos tem sido amplamente explorada como uma abordagem promissora para o monitoramento de recursos dos mais diversos tipos. As cartas de controle é uma ferramenta estatística que permite monitorar e controlar a variabilidade dos dados coletados ao longo do tempo, facilitando a identificação de desvios e anomalias nos recursos computacionais. O controle estatístico de processos, por sua vez, abrange uma série de técnicas e métodos estatísticos, incluindo-se as cartas de controle, para analisar e melhorar a estabilidade e o desempenho dos processos.

Diante desse cenário, esta revisão sistemática tem como objetivo identificar as abordagens mais eficazes e inovadoras para a previsão da disponibilidade de recursos computacionais em sistemas distribuídos, com foco no uso da carta de controle I-MR e técnicas de controle estatístico de processos. Além disso, pretende-se avaliar a qualidade e a confiabilidade dos estudos que investigaram o uso dessas abordagens, identificar as técnicas de controle estatístico de processos mais aplicadas e investigar as práticas e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais nesse contexto.

Por meio dessa revisão sistemática, busca-se fornecer uma análise abrangente e atualizada sobre as abordagens de previsão da disponibilidade de recursos computacionais em sistemas distribuídos, bem como propor recomendações e diretrizes para o aprimoramento do gerenciamento de recursos nesses ambientes. Com base na análise crítica dos estudos selecionados, espera-se contribuir para o avanço do conhecimento nessa área e para o aprimoramento das práticas de previsão da disponibilidade de recursos computacionais.

2 Metodologia

2.1 Definição dos critérios de busca

Para realizar a busca dos estudos relevantes, foram definidos critérios específicos para garantir a inclusão de estudos pertinentes à temática da previsão da disponibilidade de recursos computacionais em sistemas distribuídos. Os critérios de busca incluíram palavras-chave relacionadas ao tema, bem como filtros de idioma e período de publicação.

Aqui é definida uma lista de palavras-chave relevantes para o estudo, incluindo sinônimos e variações. Em seguida, é selecionada as bases de dados que serão utilizadas para fazer a pesquisa bibliográfica.

Termos de pesquisa:

- Disponibilidade de recursos computacionais
- Previsão de recursos
- Carta de controle I-AM (Individual com Amplitudes Móveis) e I-MR(Individual com Moving Range)
- Controle estatístico de processos
- Gerenciamento de recursos computacionais
- Análise de desempenho
- Sistemas distribuídos

Sinônimos e variações:

- Recursos de computação
- Capacidade computacional
- Controle estatístico de qualidade
- Gráfico de controle I-AM (Individual com Amplitudes Móveis)
- Gerenciamento de desempenho
- Sistema distribuído em nuvem
- Cloud computing

Objetivo da revisão sistemática

- Identificar as abordagens mais eficazes e inovadoras para previsão da disponibilidade de recursos computacionais em sistemas distribuídos;
- Avaliar a qualidade e a confiabilidade dos estudos que investigaram o uso da carta de controle I-AM para previsão da disponibilidade de recursos computacionais em sistemas distribuídos;
- Identificar as técnicas de controle estatístico de processos mais aplicadas em sistemas distribuídos;
- Investigar as práticas e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais em sistemas distribuídos;
- Propor recomendações e diretrizes para a melhoria do gerenciamento de recursos computacionais em sistemas distribuídos por meio da previsão da disponibilidade de recursos com o uso da carta de controle I-AM e técnicas de controle estatístico de processos.

Perguntas de pesquisa

- Quais são os métodos mais utilizados para previsão da disponibilidade de recursos computacionais em sistemas distribuídos?
- Quais são os principais modelos estatísticos de processos aplicados?
- Quais foram as medidas comuns de avaliação da qualidade de ajuste em modelos de regressão ou de previsão?
- Quais são as tendências e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais em sistemas distribuídos?

2.2 Fontes de informação

Foram selecionadas diversas bases de dados acadêmicas e industriais para realizar a busca dos estudos relevantes. Exemplos de fontes de informação incluem a IEEE Xplore,

ACM Digital Library, SpringerLink, ScienceDirect, Scopus, Web of Science e Wiley. Além disso, foram consultadas conferências e periódicos relevantes na área de sistemas distribuídos e computação.

2.3 Estratégia de busca

A estratégia de busca envolveu a combinação de termos de busca relacionados à previsão da disponibilidade de recursos computacionais, sistemas distribuídos, computação em nuvem e técnicas de controle estatístico de processos. As palavras-chave foram adaptadas conforme a sintaxe e os recursos de cada base de dados.

Strings de busca:

- "Disponibilidade de recursos computacionais" AND "Carta de controle I-AM"
- "Previsão de recursos" AND "Controle estatístico de processos"
- "Gerenciamento de recursos computacionais" AND "Análise de desempenho"
- "Sistemas distribuídos" AND "Carta de controle I-AM"
- "Computação em nuvem" AND "Controle estatístico de processos"
- ("Disponibilidade de recursos computacionais" OR "Recursos de computação" OR "Capacidade computacional") AND ("Carta de controle I-AM" OR "Gráfico de controle I-AM" OR "Controle estatístico de qualidade")
- ("Previsão de recursos" OR "Gerenciamento de recursos computacionais" OR "Gerenciamento de desempenho") AND ("Controle estatístico de processos" OR "Carta de controle I-AM" OR "Gráfico de controle I-AM")
- ("Análise de desempenho" OR "Sistemas distribuídos" OR "Sistema distribuído em nuvem" OR "Cloud computing") AND ("Carta de controle I-AM" OR "Gráfico de controle I-AM" OR "Controle estatístico de qualidade")

String de busca em inglês

- "Availability of computational resources" AND "I-MR control chart"
- "Resource Forecasting" AND "Statistical Process Control"
- "Computing resource management" AND "Performance analysis"
- "Distributed Systems" AND "I-MR Control Chart"
- "Cloud Computing" AND "Statistical Process Control"
- ("Availability of Computing Resources" OR "Computing Resources" OR "Computing Capacity") AND ("I-MR Control Chart" OR "I-MR Control Chart" OR "Statistical Quality Control")
- ("Resource Forecasting" OR "Computing Resource Management" OR "Performance Management") AND ("Statistical Process Control" OR "I-MR Control Chart" OR "I-MR Control Chart")
- ("Performance analysis" OR "Distributed systems" OR "Distributed cloud system" OR "Cloud computing") AND ("I-MR control chart" OR "I-MR control chart" OR "Statistical quality control")

2.4 Processo de seleção dos estudos

Após a realização da busca inicial, os estudos foram selecionados em duas etapas: triagem com base nos títulos e resumos, seguida pela análise completa dos artigos selecionados. Durante a triagem, os critérios de inclusão foram aplicados para identificar os estudos potencialmente relevantes. Na análise completa, os estudos foram avaliados em relação aos critérios de inclusão e exclusão para garantir a seleção dos estudos mais adequados ao escopo da revisão.

Critérios de inclusão:

- Tema: O trabalho deve abordar o tema de monitoramento ou previsão de recursos computacionais.
- Método: O trabalho deve utilizar a carta de controle I-MR ou alguma variação de carta de controle ou alguma ferramenta ou método de controle estatístico de processo.;
- Ano: Artigos publicados entre 2018 e 2023;
- Idioma: Artigos publicados em português ou inglês;

Critérios de exclusão:

- Tema: Estudos que não abordem a temática central da pesquisa;
- Disponibilidade: o trabalho deve estar disponível integralmente;
- Integridade: Artigos que não estejam disponíveis integralmente;
- Duplicidade: Artigos duplicados.

ID	Autor	Título	Ano da publicação
A01	Kansal, P. and Kumar, M. and Verma, O.P.	Classification of resource management approaches in fog/edge paradigm and future research prospects: a systematic review	2022
A02	Djigal, H. and Xu, J. and Liu, L. and Zhang, Y.	Machine and Deep Learning for Resource Allocation in Multi-Access Edge Computing: A Survey	2022
A03	Jiao, Y. and Wang, P. and Niyato, D. and Suankaewmanee, K.	Auction mechanisms in cloud/fog computing resource allocation for public blockchain networks	2019
A04	Guo, S. and Qi, Y. and Yu, P. and Shao, S. and Qiu, X.	Edge Network Resource Synergy for Mobile Blockchain in Smart City	2020
A05	Gao, Z. and Lin, B. and Xiao, K. and Wang, Q. and Mo, Z. and Yang, Y.	A dynamic resource allocation algorithm based on auction model in mobile blockchain network	2019
A06	Xu, Y. and Zhu, K. and Li, S.	Hierarchical Combinatorial Auction in Computing Resource Allocation for Mobile Blockchain	2020
A07	Wang, H. and Xu, H. and Huang, H. and Chen, M. and Chen, S.	Robust Task Offloading in Dynamic Edge Computing	2023

A08	Fan, Y. and Jin, Z. and Shen, G. and Hu, D. and Shi, L. and Yuan, X.	Three-stage Stackelberg game based edge computing resource management for mobile blockchain	2021
A09	Chen, X. and Lin, J. and Lin, B. and Xiang, T. and Zhang, Y. and Huang, G.	Self-learning and self-adaptive resource allocation for cloud-based software services	2019
A10	Abdellatif, K. and Abdelmouttalib, C. and EL-Azouzi R. and Sadik M. and Sabir E. and Tmiri A. and Altman E.	Graph-Based Computing Resource Allocation for Mobile Blockchain	2019
A11	Xiong, Z. and Feng, S. and Niyato, D. and Wang, P. and Han, Z.	Optimal pricing-based edge computing resource management in mobile blockchain	2018
A12	Jiao, Y. and Wang, P. and Niyato, D. and Xiong, Z.	Social Welfare Maximization Auction in Edge Computing Resource Allocation for Mobile Blockchain	2018
A13	Hanzelik, P.P. and Kummer, A. and Abonyi, J.	Edge-Computing and Machine-Learning-Based Framework for Software Sensor Development	2022
A14	Li, Y. and Wang, L. and Lee, D. and Jin, R.	Monitoring Runtime Metrics of Fog Manufacturing via a Qualitative and Quantitative (QQ) Control Chart	2022
A15	Ibidunmoye, O. and Rezaie, A.-R. and Elmroth, E.	Adaptive anomaly detection in performance metric streams	2018
A16	Kuang, Y. and Das, D. and Sir, M. and Pasupathy, K.	Likelihood ratio-based CUSUM charts for real-time monitoring the quality of service in a network of queues	2023
A17	Ahmad, H. and Ahmadi Nadi, A. and Amini, M. and Sadeghpour Gildeh, B.	Monitoring processes with multiple dependent production lines using time between events control charts	2023
A18	Yashchin, E.	Gradient analysis of Markov-type control schemes and its applications	2022
A19	Belay, M.A. and Blakseth, S.S. and Rasheed, A. and Salvo Rossi, P.	Unsupervised Anomaly Detection for IoT-Based Multivariate Time Series: Existing Solutions, Performance Analysis and Future Directions	2023
A20	Liao, M.-Y. and Wu, C.-W.	Process capability monitoring and change-point analysis for S-type quality characteristic	2023
A21	Weiβ, C.H. and Testik, M.C.	Nonparametric Control Charts for Monitoring Serial Dependence based on Ordinal Patterns	2023
A22	Pourmajidi, W. and Erwin, T. and Steinbacher, J. and Miransky, A. and	On challenges of cloud monitoring	2020

	Featherstone J. and Lyons K. and Wigglesworth J. and Saraf D.		
A23	Monshizadeh Naeen, H. and Zeinali, E. and Toroghi Haghighat, A.	Adaptive Markov-based approach for dynamic virtual machine consolidation in cloud data centers with quality-of-service constraints	2020
A24	Monshizadeh Naeen, H. and Zeinali, E. and Toroghi Haghighat, A.	A stochastic process-based server consolidation approach for dynamic workloads in cloud data centers	2020
A25	Alam, M.R. and Gerostathopoulos, I. and Prehofer, C. and Attanasi, A. and Bures, T.	A framework for tunable anomaly detection	2019
A26	Jeske, D.R. and Stevens, N.T. and Tartakovsky, A.G. and Wilson, J.D.	Statistical methods for network surveillance	2018
A27	P. A. Bommanavar and R. Krishnan and S. Shahedha	Optimal Resource Estimation Policy Selection for Ecommerce Applications in Cloud	2019
A28	J. Li and W. Liang and Y. Ma	Robust Service Provisioning With Service Function Chain Requirements in Mobile Edge Computing	2021
A29	L. Wang and Y. Zhang and R. Jin	A Monitoring System for Anomaly Detection in Fog Manufacturing	2020
A30	Mishra, Shashank and Gupta, Aman and Jairam Naik, K. and Mandal, Jyotsna Kumar and De, Debashis	Efficient Computing Resource Sharing for Mobile Edge-Cloud Computing Networks	2023
A31	Benmakrelouf, Souhila and St-Onge, Cédric and Kara, Nadjia and Tout, Hanine and Edstrom, Claes and Lemieux, Yves	Abnormal behavior detection using resource level to service level metrics mapping in virtualized systems	2020
A32	Benmakrelouf, Souhila and Kara, Nadjia and Tout, Hanine and Rabipour, Rafi and Edstrom, Claes	Resource needs prediction in virtualized systems: Generic proactive and self-adaptive solution	2019
A33	Toumi, Hajer and Brahmi, Zaki and Gammoudi, Mohhamed Mohsen	RTSLPS: Real time server load prediction system for the ever-changing cloud computing environment	2022

A34	Teoh, J.W. and Teoh, W.L. and Khoo, Michael B.C. and Castagliola, Philippe and Moy, W.H.	On designing an optimal SPRT control chart with estimated process parameters under guaranteed in-control performance	2022
A35	You, Huay Woon	Optimal estimated process parameters side sensitive group runs chart based on expected average run length	2018

2.5 Avaliação da qualidade dos estudos

Os estudos selecionados foram submetidos a uma avaliação da qualidade e confiabilidade. Foram considerados critérios como a rigorosidade metodológica, a validade dos resultados, a clareza na apresentação dos métodos utilizados e a contribuição para o campo de estudo. Estudos de alta qualidade foram priorizados para análise e discussão na revisão sistemática.

2.6 Extração e análise dos dados

Após a seleção dos estudos, os dados relevantes foram extraídos e organizados em uma planilha para permitir uma análise sistemática. As informações extraídas incluíram detalhes sobre os estudos (autor, ano, contexto de aplicação, métodos utilizados), as abordagens de previsão de disponibilidade de recursos, as técnicas de controle estatístico de processos empregadas, as métricas avaliadas e os principais resultados obtidos.

Perguntas de pesquisa

- Quais são os métodos mais utilizados para previsão da disponibilidade de recursos computacionais em sistemas distribuídos?
 - (Aprendizado de máquina) **Aprendizado profundo (deep learning): é uma subárea do aprendizado de máquina que utiliza redes neurais artificiais com múltiplas camadas para aprender representações abstratas dos dados e realizar tarefas complexas, como classificação, regressão, geração e reconhecimento de padrões. Alguns dos artigos que usam ou propõem modelos de aprendizado profundo para previsão da disponibilidade de recursos são** [A02], [A09], [A13], [A15], [A19], [A20], [A25], [A31] e [A33].
 - Mecanismos de leilão: são modelos baseados em teoria dos jogos e microeconomia que permitem a alocação eficiente de recursos escassos entre agentes racionais que competem por eles. Alguns dos artigos que usam ou propõem mecanismos de leilão para previsão da disponibilidade de recursos são [A3], [A04] [A05], [A06], [A08], e [A12].
 - Modelos estatísticos: são uma representação matemática de um fenômeno real que usa conceitos e métodos estatísticos para descrever e explicar a relação

entre variáveis. Esses modelos são baseados em dados observados e tentam capturar as características essenciais do processo subjacente ao fenômeno estudado. Alguns dos artigos que usam ou propõem esses modelos para previsão da disponibilidade de recursos são [A07], [A11], [A014], [A016], [A017], [A18], [A20], [A21], [A23], [A24], [A26], [A27], [A28], [A29], [A34] e [A35].

- Quais são os principais modelos estatísticos de processos aplicados?
 - Programação linear: envolve maximizar ou minimizar uma função linear sujeita a restrições lineares, sendo amplamente aplicada na otimização de recursos e tomada de decisões eficientes.. Alguns dos artigos que usam ou propõem Programação linear para previsão da disponibilidade de recursos são [A10], [A11] e [A12], [24] e [28].
 - Cartas de controle: são ferramentas que permitem monitorar a qualidade de um processo por meio da análise de variáveis que indicam o seu desempenho. Eles consistem em gráficos que mostram os limites superior e inferior de controle, que representam os valores aceitáveis para a variável monitorada, e os pontos que correspondem às observações da variável ao longo do tempo. Alguns dos artigos que usam ou propõem gráficos de controle para previsão da disponibilidade de recursos são [A09], [A14], [A15], [A16], [A20], [A21], [A26] [A29], [A31], [A32], [A34] e [A35].
 - Testes de hipóteses: são métodos que permitem testar se uma afirmação sobre uma população é verdadeira ou falsa, com base em uma amostra. Eles envolvem a formulação de duas hipóteses opostas, a nula e a alternativa, e a escolha de um nível de significância que determina a probabilidade máxima de rejeitar a hipótese nula quando ela é verdadeira. Alguns dos artigos que usam ou propõem testes de hipóteses para previsão da disponibilidade de recursos são [A02], [A03], [A06], [A07], [A08], [A18].

2.7 Síntese dos resultados

Os resultados dos estudos foram sintetizados e analisados com base nos objetivos da revisão sistemática. As principais descobertas foram agrupadas e organizadas de acordo com os tópicos relevantes para responder às perguntas de pesquisa. A análise dos resultados levou em consideração a efetividade das abordagens, as técnicas de controle estatístico de processos mais aplicadas, os desafios identificados e as tendências atuais na previsão. Aqui é apresentada a síntese narrativa dos artigos selecionados organizada pelos seguintes temas:

- T1: Métodos baseados em leilão
- T2: Métodos baseados em aprendizado de máquina
- T3: Métodos baseados em monitoramento e detecção de anomalias
- T4: Métodos baseados em otimização e controle estatístico

T1: Métodos baseados em leilão

Os métodos baseados em leilão são aqueles que usam mecanismos de mercado para alocar recursos computacionais entre os provedores e os consumidores de serviços em ambientes de computação em nuvem e névoa. Esses métodos visam maximizar o bem-estar social, a eficiência alocativa e a justiça dos participantes do leilão. Alguns exemplos de artigos que usam esse tipo de método são:

- Em [A03] é proposto um mecanismo de leilão para a alocação de recursos de computação em nuvem/nevoeiro para redes blockchain públicas. Os autores consideram dois cenários: um em que os usuários finais podem escolher entre a nuvem e o nevoeiro, e outro em que os provedores de nevoeiro podem subcontratar recursos da nuvem. Eles formulam os problemas de alocação de recursos como jogos de leilão e analisam as propriedades dos equilíbrios de Nash e de Bayes-Nash. Eles também propõem algoritmos distribuídos para encontrar os equilíbrios e avaliam o desempenho dos mecanismos de leilão por meio de simulações. Eles mostram que os mecanismos de leilão podem melhorar a eficiência, a justiça e a utilidade dos participantes em comparação com outros esquemas de alocação.
- Em [A12] é proposto um mecanismo de leilão de maximização de bem-estar social para a alocação de recursos de computação de borda para o blockchain móvel. Os autores consideram um cenário em que os usuários móveis podem comprar recursos de computação de borda para executar suas tarefas de mineração de blockchain. Eles modelam o problema como um jogo de leilão com externalidades alocativas, ou seja, o efeito que a alocação de recursos tem sobre a utilidade dos outros usuários. Eles mostram que o mecanismo de leilão proposto é verdadeiro, individualmente racional e computacionalmente eficiente. Eles também definem uma função de poder de hash que caracteriza a probabilidade de minerar com sucesso um bloco. Eles avaliam o desempenho do mecanismo de leilão por meio de simulações e mostram que ele pode melhorar o bem-estar social, a utilidade dos usuários e a receita do provedor de computação de borda.
- Em [A05] é proposto um algoritmo de alocação dinâmica de recursos baseado em um modelo de leilão na rede blockchain móvel. Os autores consideram um cenário em que os usuários móveis podem comprar recursos de computação de borda para executar suas tarefas de mineração de blockchain. Eles modelam o problema como um jogo de leilão baseado em Vickrey, onde os usuários enviam seus lances para os provedores de borda e o provedor escolhe o lance mais alto. Eles propõem um algoritmo distribuído para encontrar o equilíbrio de Nash do jogo e ajustar os lances dos usuários de acordo com a demanda e a oferta dos recursos. Eles avaliam o desempenho do algoritmo por meio de simulações e mostram que ele pode melhorar a utilidade dos usuários, a receita do provedor e a eficiência da alocação de recursos.

Os principais critérios e indicadores utilizados para avaliar esses métodos são:

- Bem-estar social: a soma dos utilitários ou lucros dos participantes do leilão
- Eficiência alocativa: a razão entre o bem-estar social obtido pelo leilão e o bem-estar social ótimo
- Justiça: a distribuição equitativa dos utilitários ou lucros entre os participantes do leilão

- Veracidade: a propriedade que garante que os participantes do leilão não têm incentivo para mentir sobre suas valorações ou custos
- Segurança: a propriedade que garante que o leilão é resistente a ataques maliciosos ou fraudes

Os principais desafios e oportunidades para melhorar esses métodos são:

- Os desafios são: Falta de padronização e interoperabilidade entre diferentes plataformas de blockchain móvel, a falta de regulamentação e segurança jurídica, a escalabilidade e o alto consumo de energia.
- Já as oportunidades incluem a descentralização e transparência, a redução de custos e intermediários, a eficiência e a segurança.

T2: Métodos baseados em aprendizado de máquina

Os métodos baseados em aprendizado de máquina são aqueles que usam técnicas de inteligência artificial para aprender padrões, prever demandas, otimizar decisões e adaptar-se às mudanças nos ambientes de computação em nuvem e névoa. Esses métodos visam melhorar o desempenho, a eficiência e a qualidade dos serviços prestados. Alguns exemplos de artigos que usam esse tipo de método são:

- Em [A15] é apresentado um método adaptativo de detecção de anomalias em fluxos de métricas de desempenho de sistemas distribuídos. O método usa uma abordagem baseada em agrupamento para identificar padrões normais e anormais nas métricas, e um mecanismo de atualização online para ajustar os modelos de agrupamento às mudanças nos dados. O método também usa uma técnica de seleção de recursos para reduzir a dimensionalidade dos dados e eliminar os recursos irrelevantes ou redundantes. Os autores avaliam o método em dois conjuntos de dados reais e comparam os resultados com outros métodos existentes. Eles mostram que o método proposto pode detectar anomalias com alta precisão e baixo custo computacional, e que pode se adaptar às variações nos fluxos de dados. O artigo também discute as limitações e as possíveis extensões do método.
- Em [A31] é proposto um algoritmo para a detecção dinâmica e em tempo real de variações anormais e mapeamento de métricas de nível de serviço para métricas de nível de recurso em sistemas virtualizados. O algoritmo combina técnicas de divergência e detecção de outliers para uma análise de carga de trabalho em tempo real e uma detecção de comportamento anormal do sistema, bem como uma detecção de periodicidade. O algoritmo é capaz de detectar e mapear mudanças incomuns em ambas as métricas de nível de serviço e de recurso (por exemplo, vazão e uso de CPU) em um sistema virtualizado, sem nenhum conhecimento prévio ou suposições sobre seu comportamento interno. O algoritmo também fornece uma análise efetiva do estado do sistema e gera notificações ou alertas para adaptação de recursos. Os autores avaliam o algoritmo usando vários tipos de cargas de trabalho e mostram que ele pode detectar variações anormais com uma precisão que varia de 92% a 100%. O artigo também discute os desafios e as vantagens do algoritmo e sugere possíveis extensões.
- Em [A33] é proposto um sistema de predição de carga de servidor em tempo real para o ambiente de computação em nuvem em constante mudança. O sistema se baseia na

classificação de tarefas recebidas e na detecção de interferência de máquinas virtuais para capturar a tendência da carga de trabalho e a taxa de interferência. O sistema também usa uma versão melhorada da árvore adaptativa de Hoeffding, aumentada por detectores de mudança de conjunto. O sistema visa garantir um uso eficiente dos recursos e evitar violações de SLA. Os autores avaliam o sistema usando dados reais e sintéticos e mostram que ele pode entregar uma grande flexibilidade para lidar com mudanças e uma boa precisão com um tempo rápido de avaliação e uma pequena pegada de memória.

Os principais critérios e indicadores utilizados para avaliar esses métodos são:

- Desempenho: a capacidade dos métodos de atingir os objetivos desejados, como precisão, velocidade, confiabilidade ou segurança
- Eficiência: a razão entre os benefícios obtidos pelos métodos e os custos incorridos pelos mesmos, como consumo energético, uso da rede ou complexidade computacional
- Qualidade: a satisfação dos usuários finais com os serviços prestados pelos métodos, como disponibilidade, latência ou experiência do usuário

Os principais desafios e oportunidades para melhorar esses métodos são:

- Lidar com a grande quantidade, variedade e veracidade dos dados gerados pelos ambientes de computação em nuvem e névoa
- Projetar técnicas de aprendizado de máquina que sejam adaptáveis, robustas e transparentes aos cenários dinâmicos e incertos
- Considerar aspectos éticos, legais e sociais na aplicação do aprendizado de máquina, como privacidade, segurança ou responsabilidade

T3: Métodos baseados em otimização e controle estatístico

Os métodos baseados em monitoramento e detecção de anomalias são aqueles que usam técnicas estatísticas ou computacionais para coletar, analisar e identificar comportamentos anormais nos ambientes de computação em nuvem e névoa. Esses métodos visam garantir a qualidade, a confiabilidade e a segurança dos serviços prestados. Alguns exemplos de artigos que usam esse tipo de método são:

- [A20] apresenta uma solução para o problema de monitoramento da capacidade do processo e análise de mudanças em características de qualidade do tipo S. O artigo propõe um algoritmo de monitoramento baseado em gráficos de controle de capacidade do processo para monitorar a capacidade do processo em tempo real. O algoritmo é avaliado por meio de simulações e os resultados mostram que ele é capaz de melhorar significativamente o desempenho do sistema em comparação com outros algoritmos existentes
- [A24] apresenta uma solução para o problema de monitoramento de cargas de trabalho dinâmicas em data centers em nuvem. O artigo propõe um algoritmo de consolidação de servidores baseado em processos estocásticos para reduzir o custo total dos data

centers enquanto satisfaz os requisitos de qualidade de serviço. O algoritmo é avaliado por meio de simulações e os resultados mostram que ele é capaz de melhorar significativamente o desempenho do sistema em comparação com outros algoritmos existentes.

- Em [A27] apresenta uma solução para o problema de monitoramento de recursos em tempo real em aplicativos de comércio eletrônico em nuvem. O artigo propõe um algoritmo de seleção de política de estimativa de recursos para escolher a melhor política de estimativa de recursos para uma carga de trabalho específica. O algoritmo é avaliado por meio de simulações e os resultados mostram que ele é capaz de melhorar significativamente o desempenho do sistema em comparação com outros algoritmos existentes.

Os principais critérios e indicadores utilizados para avaliar esses métodos são:

- Qualidade: a capacidade dos métodos de detectar anomalias com precisão, sensibilidade e especificidade
- Confiabilidade: a capacidade dos métodos de manter um desempenho consistente em diferentes condições e cenários
- Segurança: a capacidade dos métodos de proteger os dados e os serviços contra ataques maliciosos ou fraudes

Os principais desafios e oportunidades para melhorar esses métodos são:

- Lidar com a complexidade, a heterogeneidade e a incerteza dos dados provenientes dos ambientes de computação em nuvem e névoa
- Projetar técnicas de monitoramento e detecção de anomalias que sejam escaláveis, rápidas, adaptáveis e eficientes
- Considerar os requisitos e as expectativas dos usuários finais na definição e na avaliação das anomalias

3 Resultados

4 Análise dos Estudos

5 Discussão

6 Conclusão

Referências Bibliográficas