

Caracterização de Pesquisa

Marisangila Alves

¹Programa de Pós Graduação de Computação Aplicada
Universidade do Estado de Santa Catarina
Centro de Ciências Tecnológicas (UDESC) Joinville – SC – Brasil

marisangila.alves@outlook.com

Resumo. *Esse trabalho objetiva caraterizar a pesquisa da dissertação de mes-trado quanto as classificações definidas por diferentes autores. Serão apre-sentadas classificações quanto ao tipo de ciência, objetivo geral, nível de ma-turidade, raciocínio lógico e procedimento. Além disso, serão apresentadas as variáveis presentes bem como relacionamento entre seus tipos, abordagem, forma e representação de dados.*

1. Introdução

Serviços e aplicações fornecidos por serviços de computação em nuvem têm gerado um aumento significativo de tráfego na rede, além disso, redes móveis e dispositivos móveis estão em tendência de crescimento e se tornam a principal forma de acesso a rede para usuários finais [Costa et al. 2019]. A quinta geração de redes móveis – 5G, define que alguns requisitos sejam alcançados, como eficiência energética, alta densidade de dispositivos e, ultra baixa latência.

A arquitetura distribuída na rede móvel 5G permite reduzir a latência da rede, priorizando hospedeiros próximos aos dispositivos finais. Portanto, a utilização de sistemas de *cache* descentralizados e distribuídos na borda da Internet pode proporcionar aos usuários finais *Quality-of-Service* (QoS) e *Quality-of-Experience* (QoE). Para garantir QoS e QoE para o usuário final é interessante obter informações de suas conexões para proporcionar escalonamento adequado de serviços de *cache*.

1.1. Hipótese

Para isso é necessário identificar a relação entre variáveis e, a correlação entre elas e, em quais efeitos podem ser causados por interferências nessas variáveis. Identificado o problema central da pesquisa que resume-se em como melhorar QoS e QoE percebida pelo usuário em redes móveis 5G. De acordo com [Parvez et al. 2018] assume-se a premissa de que servidores *cache* distribuídos seja uma estratégia que contribui para a melhora de QoS e QoE. Então parte-se de uma pergunta de base: Como utilizar informações da rede para que servidores *cache* sejam alocados estrategicamente com o objetivo de melhorar a QoS e QoE? Para tal, existe uma primeira hipótese como ponto de partida que resume se em : (i) minimizar a latência *End-to-end* (E2E) e (ii) maximizar o número de usuários alocados no servidor *cache*. Para validar a hipótese é necessário realizar experimentos através de formulação linear ou uma heurística que prove que a relação entre *Round Trip Time* (RTT) e tamanho total do serviço solicitado resultam em uma vazão e, e com base na vazão tomar decisões de localização do servidor *cache* e, consequentemente minimizar a latência e maximizar o número de usuário servidos.

Na seção Classificação das Ciências, a pesquisa será classificada quanto ao tipo de ciência. A seção de Caracterização da Pesquisa, apresenta classificações quanto ao objetivo geral, nível de maturidade, raciocínio lógico e procedimento. Por fim, na seção Variáveis, variáveis presentes serão classificadas quanto a abordagem, tipo e forma e a representação quanto aos dados.

2. Classificação das Ciências

De acordo com [Wazlawick 2010] a ciência pode ser subdividida em dois macro tipos sendo eles, a ciência formal e a ciência empírica. A ciência formal trata-se de conclusões formais e lógicas, observa a relação entre ideias e formula uma teoria. Por outro lado, a ciência empírica resumidamente é a ciência do observável e testável. É possível realizar experimentos para evidenciar e provar uma hipótese e, com base dos dados e resultados das observações validar essa hipótese que fundamenta uma teoria.

Outra classificação é apresentada por [Wazlawick 2010] é a que divide a ciência em natural e social, ele argumenta que a ciência natural é voltada para tudo aquilo que é natural e, independe da ação humana, em computação por exemplo, estudo do *hardware*, que são componentes físicos. A ciência social estuda elementos da relação humana, por exemplo influência da informática na educação.

[Wazlawick 2010] classifica a ciência pura ou a ciência aplicada, a ciência pura ou ciência de base não objetiva aplicação tecnológica imediata. Conceitos são provados mas não necessariamente apresentam aplicações imediatas, como por exemplo, a Cosmologia que estuda a formação do universo e não tem como objetivo a aplicação prática.

A popular classificação, exatas e inexatas é mencionada por [Wazlawick 2010], consiste em assumir que ciências exatas apresentam resultados reproduzíveis e precisos. Por outro lado, as ciências inexatas podem prever fenômenos, porém nem sempre com precisão, como a Meteorologia ou Economia, porque dependem de muitas variáveis que causam complexidade para avaliar seus dados e resultados.

Existem as ciências duras ou moles discutidas por [Wazlawick 2010]. Ciências duras são aquelas que requerem evidências de maior rigor científico, como as ciências formais ou as ciências naturais na área de medicina, em outras palavras exigem maior rigor na avaliação de experimentos empíricos. Ciência mole pode ser aquela que parte de evidências anedóticas pela razão na qual pode ser difícil realizar experimentos totalmente controlados ou sem definições e métricas bem estabelecidas em uma determinada área.

Finalmente, [Wazlawick 2010] discute as ciências nomotéticas e idiográficas, resumidamente pode se dizer que nomotética é uma ciência que permite que seus resultados sejam sempre precisos, portanto determinísticos, existe uma lei geral que permite anteciper um fenômeno. As ciências idiográficas analisam fenômenos únicos, em que é difícil identificar padrões determinísticos, um bom exemplo de ciência idiográfica é a História.

A partir das definições descritas por [Wazlawick 2010] a pesquisa desenvolvida na dissertação pode ser classificada como:

- **Ciência Empírica:** a hipótese proposta pode ser observada através de experimentos, considerando que trata-se de uma proposta lógica para manipular elementos físicos, como latência que é caracterizada pelo tempo ou o tamanho do serviço, caracterizado pelo número de bits em uso na memória do servidor *cache*. É possível

manipular, controlar e analisar os dados gerados pela execução de uma heurística. Portanto pode ser classificada entre as ciências empíricas.

- **Ciência Natural:** a pesquisa lida diretamente com elementos físicos, o RTT que é o tempo entre o envio de pacote e sua confirmação de entrega, o tamanho da *cache* delimitado por elementos físicos representados por bits presentes na *Random Access Memory* (RAM)
- **Ciência Aplicada:** a resolução do problema em questão pode ser aplicada diretamente, como trata-se de uma proposta lógica que pode ser programável e implementada em redes de acesso de provedores.
- **Ciência Exata:** executando uma heurística com exatamente os mesmos cenários, ou seja, mesmo parâmetros e valores de variáveis é possível prever o comportamento da rede.
- **Ciência Mole:** deve-se ao fato de que a avaliação das observações empíricas não requer alto rigor científico, em razão de não existir dados, padrões ou métricas definidas para realizar experimentos, além disso, estudos em redes móveis 5G são relativamente novos e emergentes.
- **Ciência Nomotética:** é possível reproduzir os experimentos e obter os mesmo resultados.

3. Caracterização de Pesquisa

Nessa seção será discutida a classificação da pesquisa, apresentadas as definições da classificação e a justificativa para tal classificação.

3.1. Objetivo Geral

[Gil 2010] define que a pesquisa científica pode ser classificada de acordo com seu objetivo geral. Uma delas é a pesquisa descritiva que pretende identificar uma relação de correlação entre variáveis e explicar a dependência entre elas. De acordo com essa definição a pesquisa em questão pode ser classificada como uma pesquisa **descritiva**, justificado pelo fato de que a pesquisa observa o efeito que a manipulação da localização da *cache* pode causar nas variáveis de vazão e latência E2E e, partir de causalidade busca provar a hipótese da influência da localização na QoS e QoE visando alcançar o objetivo da pesquisa.

Segundo [Volpato 2012] a pesquisa com hipótese que analisa a interferência entre variáveis, explica uma relação de causa e efeito. Portanto é evidente que a pesquisa possui uma **hipótese de interferência**, que é manipulação da localização da *cache* que determina a variação da vazão que por sua vez influencia na redução latência que é usada como métrica para alcançar melhor QoS e QoE. Além disso, o tamanho da *cache* é uma variável determinante para alocar um serviço em uma localização adequada que depende da vazão e também influencia na latência E2E.

3.2. Nível de Maturidade

[Wazlawick 2008] caracteriza as pesquisa quanto ao nível de maturidade, em quatro níveis. O nível Apresentação de algo diferente consiste em, conclusões geralmente baseadas em evidências anedóticas e casos de estudos que não podem ser generalizados. É possível definir a pesquisa como **apresentação de algo diferente (Nível 2)**, como se trata de uma

área emergente, sem concessão de técnicas para serem comparadas ou por estar englobada a área de redes de computadores, que em geral trata-se de uma área que não possui experimentos padronizados.

3.3. Raciocínio Lógico

[Marcone and Lakatos 2005] definem a pesquisa com método indutivo como, ideias particulares e específicas que levam a induzir uma ideia generalizada fundamentando-se em premissas. Analisando a pesquisa com base nessa definição, é possível classificá-la com de tipo lógico **indutivo**, considerando que existe uma hipótese de que a localização do servidor *cache* pode melhorar QoS e QoE em um cenário específico experimental, a prova dessa hipótese leva a regra de que podemos manipular a localização de servidores *cache* para prover melhor QoS e QoE para o usuário final.

3.4. Procedimento

Quanto ao procedimento, de acordo com [Gil 2010] a pesquisa experimental pode ser resumida de acordo com três principais elementos. A manipulação, que consiste a ação do pesquisador manipular variáveis de estudo. A possibilidade de controle exercido sobre uma variável. O grupo de controle em que o experimento é realizado em um grupo de dados que sofre manipulação e outro grupo não sofre manipulação.

Como a hipótese propõe observar o comportamento da rede e a relação das variáveis e, provar a influência da localização dos servidores *cache* na redução de congestionamento da rede, usando como métrica vazão e latência. Considerando a intervenção do pesquisador na manipulação de variáveis independentes que permitem o controle das variáveis dependentes. Além disso, o experimento pode ser realizado com um grupo de dados em que ocorre manipulação de localização e outro em que não ocorre essa manipulação. A pesquisa pode ser classificada como **experimental**.

4. Variáveis

É possível classificar a pesquisa quanto abordagem, tipo e forma de variável e dados segundo as definições encontradas na literatura.

4.1. Abordagem

[Gil 2010] estabelece que pesquisa com carácter quantitativo analisa uma grande quantidade de dados, considerando que os dados naturalmente são, ou, podem ser transformados em quantificáveis, resulta em dados amplos, estruturados e conclusões objetivas. Portanto, é possível classificar a pesquisa como naturalmente **quantitativa**. As variáveis são em geral quantificáveis, como medida de tempo ou tamanho de conteúdo em *bytes*. Além disso, serão analisados uma quantidade ampla de dados.

4.2. Tipo de Variável

Segundo [Cervo et al. 2007] as ciência objetiva identificar a relação entre fenômenos, ou seja, relacionamento entre variáveis. [Cervo et al. 2007] divide as variáveis em dependentes, independentes e intervenientes. Variáveis independentes são aquelas que podem ser a origem ou causa que reflete em uma determinada consequência e, podem ser manipuladas. Variáveis dependentes são aquelas que resultam da causa, além disso, pode existir

correlação entre variáveis dependentes e, podem ser controladas a partir da manipulação das variáveis independentes. Por fim, variáveis intervenientes são aquelas que não podem ser controladas mas possuem efeitos sobre as variáveis dependentes. Todavia, não são causadas pelas variáveis independentes. É recomendável fixar os valores das variáveis intervenientes para reduzir o ruído dos dados.

Assumindo a definição escrita, a pesquisa apresenta variáveis que possuem relacionamento que resultam em fenômenos observáveis (Figura 1). A localização da *cache* é uma variável em que será exercido o controle do experimento, com estimativas baseadas na vazão, que por sua vez depende do RTT. A localização interfere na latência. Se o servidor *cache* está localizado dentro da rede de acesso, esse fator pode fazer com que a latência seja reduzida, seria o contrário se o conteúdo da *cache* estiver localizado em sua origem ou rede externa, possivelmente a latência seria impactada pelo congestionamento resultante do compartilhamento de enlaces presente no núcleo da rede. Assim como vazão, se o usuário for direcionado para um servidor *cache* no qual, o enlace não esteja sofrendo com o congestionamento, consequentemente aumenta a vazão, conforme a largura de banda disponível. Ressalto que o relacionamento descrito, se for válido, pode ser usado para provar a hipótese.

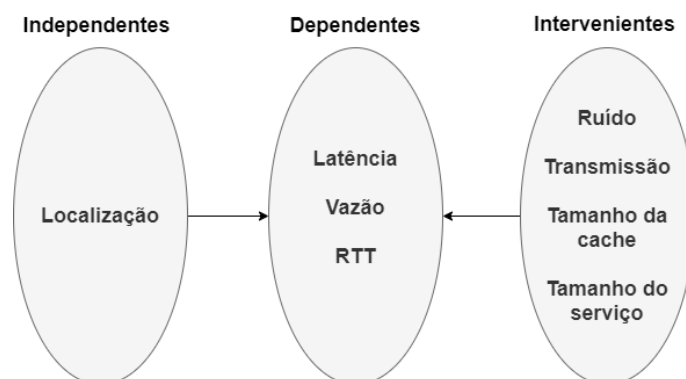


Figura 1. Relacionamento entre Variáveis

Por outro lado as variáveis intervenientes, que podem impactar nas variáveis diretamente como tamanho do espaço de armazenamento disponível no servidor *cache*, isso define se o serviço será alocado dependendo diretamente da capacidade de *hardware* e demanda requisitada para esse servidor. O tamanho do conteúdo solicitado é um fator limitante de depende expressamente da capacidade do servidor em servi-lo. Assim como, ruído e atrasos de transmissão causados pelo meio físico de transporte de dados.

4.3. Forma de Variáveis e Dados

De acordo [Barbetta 2006] variáveis numéricas podem ser classificadas como quantitativas e variáveis subjetivas como qualitativas. além disso, é possível transformar uma variável qualitativa em um dado quantitativo. Sendo assim, é possível classificar a seguintes variáveis:

- **Latência:** quantitativa cardinal, representada por um valor estimado em medidas de tempo, que representa o intervalo de tempo de requisição do conteúdo pelo usuário até a entrega do conteúdo ao usuário. Representada em dados numéricos.

- **RTT**: quantitativa cardinal, representado por um valor em medida de tempo, resultante do intervalo de envio de uma pacote e a sua confirmação de entrega. Representada em dados numéricos.
- **Vazão**: quantitativa cardinal, taxa de envio de dados representada por uma razão numérica, resultante divisão do tamanho serviço(*bytes*) requisitado pelo RTT(tempo). Representada em dados numéricos.
- **Localização**: qualitativa categórica, representada por interno ou externo, indica se a localização do conteúdo servido ao usuário final é interno ou externo a rede. Representada em dados quantitativos numéricos(0,1).

Referências

- Barbetta, P. A. (2006). *Estatística Aplicada às Ciências Sociais*.
- Behraves, R., Coronado, E., Harutyunyan, D., and Riggio, R. (2019). Joint user association and vnf placement for latency sensitive applications in 5g networks. In *2019 IEEE 8th International Conference on Cloud Networking (CloudNet)*, pages 1–7.
- Buchinger, D., Cavalcanti, G. A. S., and Hounsell, M. S. (2014). Mecanismos de busca acadêmica: uma análise quantitativa. *Revista Brasileira de Computação Aplicada*, 6:108–120.
- Cervo, A., Bervian, P. A., and Silva, R. (2007). *Metodologia Científica*.
- Costa, F. R., da Rosa Righi, R., da Costa, C. A., and Both, C. B. (2019). Nuoxus: A proactive caching model to manage multimedia content distribution on fog radio access networks. *Future Generation Computer Systems*, 93:143 – 155.
- Gil, A. C. (2010). *Como Elaborar Projetos de Pesquisa*. Atlas, São Paulo.
- Harutyunyan, D., Shahriar, N., Boutaba, R., and Riggio, R. (2019). Latency-aware service function chain placement in 5g mobile networks. In *2019 IEEE Conference on Network Softwarization (NetSoft)*, pages 133–141.
- Mach, P. and Becvar, Z. (2017). Mobile edge computing: A survey on architecture and computation offloading. *IEEE Communications Surveys Tutorials*, 19(3):1628–1656.
- Mao, Y., You, C., Zhang, J., Huang, K., and Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys Tutorials*, 19(4):2322–2358.
- Marcone, M. A. and Lakatos, E. M. (2005). *Fundamentos da Metodologia Científica*.
- Parvez, I., Rahmati, A., Guvenc, I., Sarwat, A. I., and Dai, H. (2018). A survey on low latency towards 5g: Ran, core network and caching solutions. *IEEE Communications Surveys Tutorials*, 20(4):3098–3130.
- Volpato, G. L. (2012). Tipos lógicos de pesquisa. aula 20. vídeoaula.
- Wazlawick, R. (2008). *Metodologia de Pesquisa para Ciência da Computação*. Elsevier, Rio de Janeiro.
- Wazlawick, R. (2010). Uma reflexão sobre a pesquisa em ciência da computação à luz da classificação das ciências e do método científico. *Revista de Sistemas de Informação da FSMA*, (6):3–10.