

Planejamento de Pesquisa Bibliográfica sistemática

Wilton Jaciel Loch ¹

¹Programa de Pós-graduação em Computação Aplicada(PPGCA)
Universidade do Estado de Santa Catarina(UDESC)

Resumo. *Este trabalho, desenvolvido para a disciplina de Metodologia da Pesquisa, apresenta a classificação da pesquisa sendo desenvolvida pelo autor. Esta é classificada quanto sua abordagem, objetivo, grau de amadurecimento e outros aspectos. Conclui-se de maneira geral, que a pesquisa é explicativa, detutiva, quantitativa e enquadra-se no terceiro grau de amadurecimento.*

1. Classificação da pesquisa

Inicialmente, para facilitar a posterior avaliação do trabalho, é apresentado um pequeno resumo do objetivo de pesquisa. O projeto proposto pelo autor busca desenvolver um escalonador distribuído para tarefas comunicantes com alto paralelismo de dados, utilizando informações da rede de comunicação como um dos critérios fundamentais para o processo de tomada de decisão. Esta proposta baseia-se na necessidade de transferência de altos volumes de dados entre diferentes tarefas após a conclusão do processamento[Wang et al. 2018] e no impacto negativo desta no tempo de conclusão dos trabalhos[Chowdhury et al. 2011].

1.1. Quanto ao objetivo geral

A primeira classificação a ser realizada é quanto ao tipo de objetivo da pesquisa, que segundo [Gil 2010] pode ser exploratória, descritiva ou explicativa. Até o presente momento o estado da pesquisa tem sido puramente exploratório, já que de acordo com [Gil 2010] uma pesquisa exploratória visa a construção de maior familiaridade com o problema sendo abordado, tornando-o mais explícito e propiciando a elaboração de hipóteses acerca do tema. Assim, todo o processo de levantamento de trabalhos relacionados, pergunta e objetivos de pesquisa, delimitação parcial do escopo e do contexto tratado que foram apresentados em trabalhos anteriores do autor são fruto direto de uma pesquisa exploratória preliminar.

Tratando-se da proposta em si e do objetivo buscado com conclusão do trabalho, a pesquisa assume um caráter explicativo. [Gil 2010] define que pesquisas explicativas buscam elucidar os fatores que contribuem para a ocorrência de um fenômeno, bem como a razão ou o porquê de tais acontecimentos, tratando a interferência entre diferentes variáveis. Nesse sentido, o trabalho desenvolvido busca – através da construção da proposta – avaliar o impacto que pode ser obtido no tempo de conclusão dos trabalhos decorrente da utilização da rede como critério de escalonamento, sendo que a associação desse e de outros fatores com o objetivo de melhora foi previamente explorado e definido por pesquisas descritivas realizadas por outros autores.

Sendo esta então uma pesquisa explicativa, há existência de uma hipótese que guie o desenvolvimento. Esta não será formalmente definida neste momento, porém fundamenta-se no princípio de que uma vez identificado em trabalhos anteriores que a

rede possui influência negativa sobre o tempo de conclusão dos trabalhos, supõe-se que utilizar formas de escalonamento que sejam atentas à informações da rede pode contribuir para a redução de tal impacto. Assim, objetiva-se mensurar a interferência do fator referente à rede no tempo de conclusão, através do escalonamento.

1.2. Quanto ao grau de amadurecimento

O segundo fator a ser considerado na classificação da pesquisa é o grau de amadurecimento. Nesse sentido, identifica-se que pesquisa em desenvolvimento situa-se no grau três de amadurecimento, apresentando algo presumivelmente melhor que o já proposto na literatura. Segundo [Wazlawick 2008], pesquisas pertencentes ao terceiro grau geralmente são relativas à áreas específicas que demandam que novas propostas sejam comparadas quantitativamente com outras abordagens já existentes porém, não há *benchmarks* ou bases de dados reconhecidas internacionalmente para a realização de testes comparativos, fazendo com que recaia sobre o autor o dever de propor e realizar os experimentos necessários. O contexto do desenvolvimento de escalonadores para tarefas comunicantes explorado até o momento segue esta mesma forma de funcionamento. Cada autor realiza experimentos de modo a quantificar as melhorias decorrentes de sua proposta em relação às demais, entretanto não há padronização e cada autor seleciona arbitrariamente – de acordo com seu julgamento do que é mais propício, sensato e passível de realização – as diversas características dos experimentos como tamanho dos *clusters*, configuração dos nós de computação, tamanho e estrutura das cargas a serem processadas, entre outros.

1.3. Quanto ao raciocínio lógico

A pesquisa deve também enquadrar-se em um tipo de raciocínio lógico, que compreende as seguintes possibilidades: dialético, indutivo, dedutivo e hipotético-dedutivo. A partir de tais classificações e avaliando a pesquisa em desenvolvimento conclui-se que é empregado o método dedutivo. Este trata-se de um método científico que utiliza como base a existência de premissas supostamente verdadeiras para a confecção de uma conclusão também verdadeira, cujas informações já eram previamente contidas, mesmo que de forma implícita, nas premissas [Lakatos and Marconi 2005]. Naturalmente, se alguma das premissas mostrar-se falsa, a conclusão é invalidada. Com base em tais informações o trabalho enquadra-se nesse método por partir de uma série de premissas a respeito das características dos trabalhos de tarefas comunicantes (como a inerente necessidade de envio de dados entre diferentes tarefas [Dean and Ghemawat 2008]) e da influência de fatores alternativos sobre tais trabalhos (*e.g.*, impacto da rede e localidade dos dados no tempo de conclusão [Chowdhury et al. 2011]) para concluir que a utilização de tais critérios no processo de escalonamento deve gerar melhorias na característica selecionada como objetivo.

1.4. Quanto ao procedimento de coleta

A próxima classificação a ser realizada é quanto ao procedimento utilizado para a coleta de dados, que pode residir em uma das três formas de pesquisa: bibliográfica, de campo ou de laboratório. Para o trabalho em desenvolvimento são empregadas duas formas de coleta de dados, sendo a primeira uma pesquisa bibliográfica e a segunda uma pesquisa de laboratório. [Lakatos and Marconi 2005] definem que uma pesquisa bibliográfica é a que abrange a investigação de todo o corpo de trabalhos já publicados a respeito de certa

área específica, cujo objetivo é aproximar o pesquisador de todo o conteúdo produzido nesse contexto. Essa forma de coleta de dados é empregada no presente trabalho para que seja possível entender de forma mais completa os métodos, técnicas, experimentos e outros fatores pertinentes que são utilizados em propostas relacionadas, de modo a guiar o desenvolvimento do produto final da pesquisa.

[Lakatos and Marconi 2005] definem ainda, que um procedimento de coleta de dados de laboratório busca analisar o que ocorrerá em situações controladas, utilizando instrumental preciso, específico e ambientes controlados. Nessa forma de coleta o pesquisador executa experimentos tendo em consideração o objeto, objetivo, instrumental e técnicas, podendo chegar a resultados esperados ou inesperados. Essa forma será utilizada no processo de validação da proposta, de modo que serão executados experimentos em ambientes controlados com o objetivo de comparar a efetividade do produto desenvolvido com as demais abordagens pertinentes disponíveis na literatura.

1.5. Quanto à influência das variáveis

As variáveis tratadas na pesquisa podem receber classificações de acordo com seu tipo, divididos de acordo com a forma de influência da variável: dependentes, independentes e intervenientes. As variáveis independentes são relativas às causas ou fatores determinantes de certo fenômeno, sendo condicionantes diretas do resultado final e podendo geralmente ser controladas. Variáveis dependentes correspondem aos resultados medidos ou observados, sendo condicionadas pela ocorrência do fenômeno e afetadas tanto por variáveis independentes quanto intervenientes. Por fim, variáveis intervenientes correspondem a fatores externos que influenciam as variáveis dependentes e que podem eventualmente ser observados ou medidos, entretanto não podem ser controlados [Cervo et al. 2006].

A tabela 1 apresenta um levantamento das variáveis mais relevantes para a realização dos experimentos relacionados à pesquisa. As duas primeiras variáveis correspondem às características fundamentais dos trabalhos e portanto, assumem um aspecto independente e definem condições iniciais sobre a carga que deve ser processada. A terceira e quarta variáveis representam características intervenientes que vão inevitavelmente existir no ambiente de testes e que devem ser controladas para que sua influência seja mínima no resultado final. Por fim, as últimas cinco variáveis dizem respeito às métricas da qualidade das soluções encontradas e naturalmente dependem tanto das variáveis independentes que condicionam a carga a ser processada, quanto das variáveis intervenientes relativas à utilização adicional do sistema.

1.6. Quanto a abordagem de pesquisa

[Gil 2010] afirma que uma pesquisa pode ainda ser classificada de acordo com a sua forma de abordagem, que pode ser quantitativa ou qualitativa. Assim, uma pesquisa é considerada quantitativa quando todos os seus elementos podem ser quantificados numericamente e conclusões podem ser feitas a partir de análises matemáticas ou estatísticas sobre os conjuntos de dados oriundos do processo de coleta. Esta forma de pesquisa implica na existência de bases amplas e estruturadas de dados, levando à conclusões objetivas. Em contrapartida, uma pesquisa é considerada qualitativa quando seus elementos são considerados invariavelmente subjetivos e dessa forma não é possível quantificar as informações colhidas de forma numérica. Esta forma de abordagem depende da interpretação e atribuição de significado por parte do pesquisador e tem foco não somente

nos resultados mas também no processo originário. Como a pesquisa desenvolvida pelo autor tem um caráter quase completamente numérico, no qual todas as variáveis independentemente de seus tipos são numericamente quantificáveis e conclusões a respeito da qualidade das soluções obtidas são estatisticamente analisadas e comparadas, conclui-se que esta trata-se de uma pesquisa quantitativa.

1.7. Quanto à forma das variáveis

As variáveis de natureza quantitativa, como as empregadas no projeto desenvolvido, podem ser também classificadas de acordo com sua forma, divididas em contínuas ou discretas e cardinais ou intervalares. De acordo com [Barbetta 2008] variáveis discretas são as que assumem valores enumeráveis, de forma que tomando dois elementos consecutivos pertencentes a um conjunto, não há nenhum terceiro elemento existente entre eles. Variáveis contínuas entretanto, são as que assumem valores que não são enumeráveis, de forma que tomando dois elementos consecutivos, há uma infinidade de valores existentes entre estes. Adicionalmente, as variáveis consideradas cardinais são as que assumem valores bem definidos do conjunto, enquanto variáveis intervalares compreendem múltiplos possíveis valores dentro do conjunto como um todo ou de faixas específicas deste. Novamente na tabela 1 são apresentadas as classificações das variáveis mais relevantes quanto suas formas. A primeira variável trata-se de uma distribuição de tempo, e portanto é contínua e intervalar já que pode assumir uma miríade de valores entre o mínimo e máximo. A segunda variável dita quantas tarefas cada trabalho terá e portanto é discreta e cardinal. As duas próximas variáveis são medidas de carga no sistema e são aqui classificadas como discretas e cardinais, já que a primeira pode ser medida em *bits/s* e a segunda em ciclos de processamento utilizados, no caso do processador, ou *bytes*, no caso de memória. As cinco últimas variáveis correspondentes às métricas do resultado são todas definidas como contínuas e cardinais, já que a primeira trata-se de um valor de porcentagem, enquanto as quatro últimas são medições temporais.

| Nome | Tipo | Forma |
|---------------------------------------|---------------|----------------------|
| Dist. tempo entre chegadas | Independente | Contínua, Intervalar |
| Tamanho dos trabalhos | Independente | Discreto, Cardinal |
| Carga paralela na rede | Interveniente | Discreta, Cardinal |
| Carga paralela nas máquinas | Interveniente | Discreta, Cardinal |
| Localidade das escolhas | Dependente | Contínua, Cardinal |
| Tempo de conclusão médio (Trabalhos) | Dependente | Contínua, Cardinal |
| Tempo de conclusão máximo (Trabalhos) | Dependente | Contínua, Cardinal |
| Tempo de conclusão médio (Tarefas) | Dependente | Contínua, Cardinal |
| Tempo de conclusão máximo (Tarefas) | Dependente | Contínua, Cardinal |

Table 1. Variáveis e suas características.

1.8. Quanto ao tipo de ciência

Segundo [Wazlawick 2010] as ciências podem ser classificadas como: formais ou empíricas, puras ou aplicadas, exatas ou inexas, duras ou moles e nomotéticas ou idiógráficas. Ciências formais são as que estudam os processos lógicos e matemáticos por detrás das ideias, sem possuir foco direto em sua utilização no mundo real, enquanto

as ciências empíricas estudam os fenômenos do mundo real e utilizam tais observações para fundamentar hipóteses. Ciências puras estudam os conceitos fundamentais de uma área do conhecimento, sem foco direto no desenvolvimento de aplicações relacionadas aos resultados produzidos. Em contrapartida, as ciências aplicadas focam-se no desenvolvimento de conhecimento que possa ser diretamente aplicado na melhoria de um processo do mundo real, visando a obtenção de ganhos específicos. As ciências exatas são as que possuem resultados precisos e reproduzíveis bem como leis previsíveis, enquanto ciências inexatas podem prever apenas de forma geral o comportamento de um fenômeno estudado, porém os resultados não necessariamente são sempre previsíveis. Ciências duras são as que necessitam de elevado rigor científico nos processos e experimentos para que seus resultados possam ser confirmados, enquanto ciências moles geralmente baseiam seus resultados em estudos de caso. Ciências nomotéticas estudam fenômenos que tendem a se repetir e utilizam os conhecimentos gerados a partir de seu estudo para prever casos futuros e estipular leis, já as ciências idiográficas estudam fenômenos que não se repetem [Wazlawick 2010]. A partir de tais classificações, a pesquisa em desenvolvimento enquadra-se em: ciência formal, já que não estuda diretamente fenômenos do mundo real e faz uma proposta algorítmica e matemática para a solução de um problema; ciência aplicada, já que o objetivo do desenvolvimento é a resolução de um problema do mundo real, visando a melhora de um processo prático; ciência exata, já que os experimentos são reproduzíveis e possuem resultados precisos e de certa forma previsíveis; ciência dura, já que é necessário rigor matemático e estatístico para a validação dos resultados, através de experimentos extensivos; e ciência nomotética já que os processos em que a proposta se baseia repetem-se.

References

- Barbetta, P. A. (2008). *Estatística aplicada às ciências sociais*. Ed. UFSC.
- Cervo, A., Bervian, P., and da Silva, R. (2006). *Metodologia científica*. Pearson Prentice Hall.
- Chowdhury, M., Zaharia, M., Ma, J., Jordan, M. I., and Stoica, I. (2011). Managing data transfers in computer clusters with orchestra. *ACM SIGCOMM Computer Communication Review*, 41(4):98–109.
- Dean, J. and Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1):107–113.
- Gil, A. (2010). *Como elaborar projetos de pesquisa*. Atlas.
- Lakatos, E. M. and Marconi, M. d. A. (2005). *Fundamentos de metodologia científica*. Atlas.
- Wang, K., Zhou, Q., Guo, S., and Luo, J. (2018). Cluster Frameworks for Efficient Scheduling and Resource Allocation in Data Center Networks: A Survey. *IEEE Communications Surveys Tutorials*, 20(4):3560–3580. Conference Name: IEEE Communications Surveys Tutorials.
- Wazlawick, R. (2008). *METODOLOGIA DE PESQUISA PARA CIÊNCIA DA COMPUTAÇÃO*. Elsevier.

Wazlawick, R. S. (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.