

UBIQUIDADE

Revista de estudos sobre as Tecnologia de Informação e
Comunicação (TIC) Centro Universitário Padre Anchieta.

ISSN 2236-9031

A hand is shown interacting with a glowing, abstract digital structure. The structure is composed of flowing, translucent lines in shades of blue and purple, resembling a complex network or data flow. The hand is positioned on the left side, with fingers reaching towards the structure. The background is dark, making the glowing lines stand out prominently.

UNIAnchieta

FICHA TÉCNICA

Revista Ubiquidade, ISSN 2236-9031, v. 2, n. 1, jan./jun. 2019

Capa: Larissa Conelheiro Kovelis

Editoração e Diagramação: Prof. Dr. Juliano Schimiguel

Revisão: Isabel Cristina Alvares de Souza

Editora: UNIANCHIETA

Prof. Me. João Vasconcelos

Diretor Acadêmico

Prof. Me. Anselmo de Araujo Couto

**Coordenador dos Cursos de Análise e Desenvolvimento de Sistemas,
Ciência da Computação e Sistemas de Informação**

Prof. Dr. Juliano Schimiguel

Coordenador/Editor da Revista Ubiquidade

Todos os direitos reservados e protegidos pela Lei 9.610 de 19/02/1998. É permitida a reprodução e distribuição desta obra, desde que para fins educacionais e integralmente mantidas as informações autorais. É vedado seu uso comercial, sem prévia autorização, por escrito, dos autores e da Editora.

EXPEDIENTE

A revista Ubiquidade é uma publicação semestral vinculada ao Curso de Bacharelado em Ciência da Computação do UniAnchieta, exclusivamente eletrônica, que pretende divulgar contribuições originais, teóricas ou empíricas, relacionadas às áreas de Tecnologia de Informação e Comunicação (TIC) e está aberta para trabalhos científicos de pesquisadores nacionais ou internacionais.

O envio de trabalhos para apreciação, assim como o pedido de informações, pode ser feito por meio do endereço: ubiquidade@anchieta.br

EDITOR

Prof. Dr. Juliano Schimiguel (UniAnchieta, Unicsul)

CONSELHO EDITORIAL

Prof.a Dra. Aline Brum Loreto, Universidade Federal de Santa Maria-Campus Cachoeira do Sul (UFSM-CS)/RS

Prof. Dr. Carlos Adriano Martins, Unacid - Universidade Cidade de São Paulo, São Paulo/SP

Prof. Dr. Carlos Eduardo Câmara, Centro Universitário Padre Anchieta, Jundiaí/SP

Prof.a Dra. Cecília Sosa Arias Peixoto, Instituto Paulista de Ensino e Pesquisa, Campinas/SP

Prof. Dr. Cristiano Monteiro da Silva, Centro Universitário Anchieta, Jundiaí/SP

Prof. Dr. Hélio Rosetti Júnior, Instituto Federal do Espírito Santo, Vitória/ES

Prof. Dra. Jane Garcia de Carvalho, Unacid - Universidade Cidade de São Paulo, São Paulo/SP

Prof. Dr. Josney Freitas Silva, UEMG - Universidade do Estado de Minas Gerais - UEMG, Frutal/MG

Prof. Dr. Juliano Schimiguel, UniAnchieta, Unicsul

Prof. Ms. Juliano Silva Marçal, Centro Universitário Padre Anchieta, Jundiaí/SP

Prof. Dr. Luciano Soares Pedroso, Universidade Federal dos Vales do Jequitinhonha e Mucuri - UFRVJM, Teófilo Otoni/MG

Prof.a Dra. Lucy Mirian Campos Tavares Nascimento, Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Formosa/GO

Prof. Dr. Marcelo Eloy Fernandes, Universidade Nove de Julho, São Paulo/SP

Prof. Ms. Nadia Vilela Pereira, IFTO — Instituto Federal do Tocantins, Campus Palmas

Prof. Dr. Vivaldo José Breternitz, Universidade Prestibariana Mackenzie, São Paulo/SP

Prof.a Dra. Viviane Sartori, Universidad Europea del Atlántico (Uneatlatico), Salamanca, Espanha

SUMÁRIO

APRESENTAÇÃO	4
PROFESSOR.COM: OS DESAFIOS DA DOCÊNCIA NO SÉCULO XXI <i>(Viviane Lima Martins e Carlos Adriano Martins (Orientador)</i>	<i>5</i>
BOAS PRÁTICAS EM DATA CENTER PARA USO DE TECNOLOGIA VERDE <i>(Josiane Martins Silva Guirardi e Edilson Gomes Feitoza)</i>	<i>22</i>
CLOUDERA: UMA ABORDAGEM PARA ANÁLISE DE LOGS DO PROXY SQUID-CACHE USANDO HADOOP, FLUME, MAPREDUCE E IMPALA <i>(Volnei Cervi Puttini e Juliano Schimiguel)</i>	<i>45</i>
ANÁLISE DO PENSAMENTO COMPUTACIONAL NA EDUCAÇÃO VOLTADO À SERIOUS GAMES <i>(Fernanda Moreira Pouza e Carlos</i> <i>Eduardo Câmara)</i>	<i>63</i>
ANÁLISE DE DESEMPENHO DE UM ALGORITMO DESENVOLVIDO PARA SOLUÇÃO DE DEEP LEARNING UTILIZANDO REDES NEURAIS CONVOLUCIONAIS PARA ANÁLISE DE CONTRASTE DE IMAGENS <i>(Patrick Francisco Oliveira e Carlos Eduardo Câmara).....</i>	<i>84</i>

APRESENTAÇÃO

Neste número da Revista Ubiquidade (v.2. n.1) encontram-se um conjunto de artigos de diferentes áreas da Ciência da Computação, onde cada um aponta para múltiplas possibilidades de aplicação e pesquisa.

No primeiro artigo, Martins & Martins discutem os desafios da docência no Século XXI. Na sequência, Guirardi & Feitoza apresentam as boas práticas em *data center* para uso de tecnologia da informação 'verde'. Puttini & Schimiguel apresentam uma abordagem para análise de *logs* do *proxy squid-cache* usando *Hadoop*, *Flume*, *Mapreduce* e *Impala* (*Cloudera*).

O quarto artigo, de Pouza & Câmara, realiza uma análise do conceito de pensamento computacional na educação, voltado ao contexto e cenário de *serious games*. O objetivo é propor uma nova metodologia baseada em *serious games* e apresentar, por meio de um protótipo, uma melhor alternativa para um desenvolvimento cognitivo mais favorável para os discentes em alguns anos.

Oliveira & Câmara realizam uma análise de desempenho de um algoritmo desenvolvido para solução de *deep learning* utilizando redes neurais convolucionais para análise de contraste de imagens.

Boa leitura!

Prof. Dr. Juliano Schimiguel

Coordenador da Revista Ubiquidade

PROFESSOR.COM: OS DESAFIOS DA DOCÊNCIA NO SÉCULO XXI

PROFESSOR.COM: THE CHALLENGES OF TEACHING IN THE 21ST CENTURY

Viviane Lima MARTINS

ead.adriano@gmail.com

Pós-Graduação *Lato Sensu* em Tecnologias Digitais e Inovação na Educação
Universidade Cidade de São Paulo (UNICID), São Paulo-SP, Brasil

Carlos Adriano MARTINS (Orientador)

ead.adriano@gmail.com

Pós-Graduação *Lato Sensu* em Tecnologias Digitais e Inovação na Educação
Universidade Cidade de São Paulo (UNICID), São Paulo-SP, Brasil

RESUMO

O propósito da pesquisa é discutir a formação docente e a atuação dos professores para o uso da informática na sala de aula. Para tanto, foi realizada inicialmente uma pesquisa bibliográfica, tendo como aportes principais Valente e Tardiff, com o intuito de se aproximar das demais pesquisas a respeito do tema e com isso nortear a pesquisa com os professores da educação básica, na cidade de Guarujá/SP. Os resultados revelaram a dificuldade dos docentes, pois a formação inicial dos mesmos não foi suficiente. Além disso, percebeu-se que os professores entendem e acreditam que o uso da informática pode ajudar no processo ensino aprendizagem e que eles se preocupam com a infraestrutura que por sua vez pode interferir em suas práticas. Todos estes aspectos são relevantes e devem ser repensados com o objetivo de buscar novas alternativas que promovam ainda mais vantagens para todos os envolvidos.

PALAVRAS CHAVE: tecnologia; formação docente; prática pedagógica.

ABSTRACT

The purpose of the research is to discuss teacher education and the role of teachers in the use of information technology in the classroom. In order to do so, a bibliographical research was initially carried out, with the main contributions of Valente and Tardiff, in order to get closer to the other researches on the subject and with that to guide the research with the primary and secondary teachers of the city of Guarujá. The results revealed the difficulty of the teachers, since the initial training was not enough. In addition, it has been realized that teachers understand and believe that the use of information technology can help in the learning teaching process and that they care about the infrastructure which in turn can interfere in their practices. All of these aspects are relevant and should be rethought in order to seek new alternatives that will offer even more advantages to all those involved.

KEY WORDS: technology; teacher training; pedagogical practice

INTRODUÇÃO

Embora as novas tecnologias tenham se tornado importantes aliadas no processo educativo, nem todas as escolas, assim como seus profissionais, estão preparadas para esta nova realidade. Isto ocorre por diversos motivos, como: falta de equipamentos, local apropriado para instalações dos equipamentos, profissionais não qualificados, entre outros. Consequentemente, tudo isso vai influenciar na entrada de tecnologia na sala de aula bem como no cotidiano do professor.

Nesta pesquisa vamos nos ater a estudar a relação do professor com o universo tecnológico buscando compreender a seguinte questão-problema: Os professores conseguem acompanhar a evolução da informática disponível no contexto da educação formal?

Partimos do pressuposto que as dificuldades dos professores no uso das tecnologias começam no próprio processo de formação inicial, em que o futuro professor tem pouco acesso a estas mesmas tecnologias, posteriormente se sentem desestimulados a utilizá-las em sala de aula. Moran (*apud* MASETTO, 2006) afirma que em geral os professores têm dificuldades no domínio das tecnologias e, tentam fazer o máximo que podem, diante deste hábito mantêm uma estrutura repressiva, controladora, repetidora. Muitos tentam mudar, mas não sabem bem como fazê-lo e não sentem preparados para experimentar com segurança.

A definição do tema de pesquisa baseia-se no entendimento de que se trata de um tema pertinente e atual, portanto, partimos da premissa de que o estudo se torna significativo, pois pesquisar a relação entre professor e novas tecnologias implica em ampliar as discussões em torno não só do tema propriamente dito, mas também da organização do trabalho do docente, das propostas educacionais e as relações que se estabelecem neste contexto tecnológico.

Quanto mais a informática se faz presente na educação, mais desafios se apresentam e na opinião de Valente (1997) estes desafios estão relacionados à: entender o computador como uma nova maneira de representar o conhecimento, em segundo lugar, analisar com cuidado o que significa ensinar e aprender bem como demanda rever o papel do professor e em terceiro lugar, a formação do professor que envolve muito mais do que prover o professor com conhecimentos sobre computadores.

Outro desafio é aliar o uso da informática a formação intelectual do aluno e a esse respeito, a educadora argentina Emília Ferreiro, citada por Garcia (2001, p.163) chama a atenção e afirma:

trata-se de permanecer alerta às mudanças que estão acontecendo que podem envolver profundas mudanças na relação dos já letrados com os textos, e no modo com que as novas gerações são alfabetizadas - e de reconhecer que as mudanças necessárias em nível educativo são muito drásticas, de fato, porque já agora, a escola pública está imensamente desatualizada. Se a tecnologia da informática chegar a

servir, como se fosse um gatilho para repensar o que acontece na escola, que ela seja bem-vinda.

Analisando as palavras de Ferreira, é possível compreender quanto é importante à adequação do ambiente escolar informatizado e essa adequação passa pelo entendimento de quais devem ser os objetivos da utilização da informática e todas as tecnologias que a envolve alcançando assim seu maior interessado, ou seja, o aluno.

A discussão sobre o professor no contexto do letramento digital passa necessariamente pelo entendimento do significado do termo letramento. Como afirma Soares (2002, p.15)

Letramento é palavra e conceito recentes, introduzidos na linguagem da educação e das ciências linguísticas há pouco mais de duas décadas. Seu surgimento pode ser interpretado como decorrência da necessidade de configurar e nomear comportamentos e práticas sociais na área da leitura e da escrita que ultrapassem o domínio do sistema alfabético e ortográfico, nível de aprendizagem da língua escrita perseguido, tradicionalmente, pelo processo de alfabetização. Esses comportamentos e práticas sociais de leitura e de escrita foram adquirindo visibilidade e importância à medida que a vida social e as atividades profissionais tornaram--se cada vez mais centradas na e dependentes da língua escrita, revelando a insuficiência de apenas alfabetizar – no sentido tradicional – a criança ou o adulto.

O conceito de letramento acabou sendo transportado para a área da informática e isso se deu porque da mesma forma que a leitura e a escrita adquiriram visibilidade e importância no contexto das práticas sociais e com a informática não foi diferente, pois a introdução das novas tecnologias na sociedade suscita o desenvolvimento de novas práticas.

Para Soares (2002) o espaço da escrita passa a ser a tela do computador e não mais o papel, com isso mudam-se as relações entre autor e texto, leitor e autor e leitor e texto. Dessa forma entende-se que neste contexto emerge um novo tipo de letramento, ou seja, letramento digital. Nas palavras de Xavier (s/d, p.2)

O letramento digital implica realizar práticas de leitura e escrita diferentes das formas tradicionais de letramento e alfabetização. Ser letrado digital pressupõe assumir mudanças nos modos de ler e escrever os códigos e sinais verbais e não verbais, como imagens e desenhos, se compararmos às formas de leitura e escrita feitas no livro, até porque o suporte sobre o qual estão os textos digitais é a tela, também digital.

Dito de outra forma, o letramento digital é a forma de se utilizar os recursos de informática para aplicá-lo no seu cotidiano em benefício do próprio usuário e como afirma Xavier quando o indivíduo desenvolve competências para utilizar os equipamentos digitais se permite a reinventar seu cotidiano bem como estabelecer novas formas de ação, que se revelam em práticas sociais específicas e em modos diferentes de utilização da linguagem verbal e não verbal.

Concordando com as palavras de Xavier, Soares (2012) afirma que, além disso, o letramento digital também permite ao usuário se relacionar com seus pares, aprender

constantemente, construir, transformar, reconstruir, exercer autoria, compartilhar conhecimento etc., entretanto ainda é grande o número de pessoas que não domina as tecnologias disponíveis na sociedade do conhecimento.

Prado e Valente (2003) afirmam que a formação inicial deve fornecer a base para que na prática os docentes dominem o uso de tecnologias em sala de aula, principalmente o computador, por outro lado, Masseto (2004, p.135) afirma:

Nos próprios cursos de ensino superior, o uso de tecnologia adequada ao processo de aprendizagem variada para motivar o aluno não é tão comum, o que faz com que os novos professores do ensino fundamental e médio, ao ministrarem suas aulas, praticamente copiem o modo de fazê-lo e o próprio comportamento de alguns de seus professores de faculdade, dando aula expositiva e, às vezes, sugerindo algum trabalho em grupo com pouca ou nenhuma orientação.

Esta questão é explicada por Tardif (2002) da seguinte maneira: para ele, os saberes profissionais são caracterizados como temporais aqueles que são adquiridos com o tempo, dessa forma, a bagagem de conhecimento é construída ao longo da vida escolar por aproximadamente dezesseis anos de imersão em seu espaço de trabalho antes de começar a trabalhar formalmente como professor e quando estes entram no curso de formação não modificam suas crenças anteriores sobre o ensino e quando iniciam sua vida profissional são principalmente essas crenças que eles reativam para solucionar seus problemas. É provável que alguns professores ainda resistam ao uso de tecnologias na escola, por privilegiar antigas práticas, pois na Universidade ainda pode-se encontrar professores, com pouca habilidade em utilizar qualquer tipo de tecnologia. Segundo Rocha (2009, p. 71):

Para que se inclua tecnologia no processo de formação de professores, tanto inicial como continuada, é importante que uma disciplina nessa formação, com a denominação de tecnologia da informação e da comunicação na educação. Dada a sua constituição essa disciplina será importante na formação para possibilitar momentos de reflexão para que alunos e professores de um curso que tem a disciplina percebam o alcance que o desenvolvimento tecnológico tem em suas vidas. Tal disciplina poderá mostrar as possibilidades que apresenta para auxiliar na compreensão de outros temas trazidos por esse desenvolvimento à sociedade e a educação.

As considerações de autor são legítimas do ponto de vista da teoria, mas o que ocorre é que as universidades não disponibilizam, pelo menos de forma eficaz, tal formação, pois os cursos de graduação não preparam nem professores, nem gestores para trabalhar com tecnologias na escola.

OBJETIVOS

O objetivo geral desta pesquisa é:

- verificar a relevância que as novas tecnologias trazem à educação escolar e para a prática docente e identificar os principais desafios presentes na formação e na prática dos professores.

Quanto aos objetivos específicos, destacamos:

- analisar qual é o envolvimento dos professores com as novas tecnologias no interior da escola;
- discutir a formação inicial dos professores e seus aprendizados em relação às tecnologias, buscando compreender como isso se refletiu em sua prática docente.

Tais objetivos se justificam, pois, o computador é um grande aliado do professor no processo ensino aprendizagem, haja vista que possibilita aos alunos descobrir as próprias potencialidades.

No entendimento de Maio e Stezer (2011) a informatização representa uma “nova cultura no mundo do ensino”, e pressupõe mudança no comportamento didático, uma vez que, de forma gradativa e irreversível, a informática praticamente permeia todas as atividades humanas, porém é preciso lembrar que a informática não surgiu para melhorar os processos educativos, a educação é quem a adotou como uma ferramenta. Considerando este aspecto mencionado por Maio e Stezer de que a informática permeia todas as atividades humanas, fica evidente que a educação não poderia ficar de fora desta cultura.

Por outro lado, Gouvêa (1999) lembra que o papel do professor é mais importante ainda, pois da mesma forma que o professor um dia introduziu o livro e teve de começar a lidar de modo diferente com o conhecimento sem deixar as outras tecnologias de comunicação de lado, ele continuará a ensinar e aprender pela palavra, pelo gesto, pela emoção, pela afetividade, pelos textos lidos e escritos, pela televisão, mas agora também pelo computador e pela informação em tempo real.

ASPECTOS METODOLÓGICOS

No que tange à metodologia, para análise da problemática apresentada, estudamos o universo dos docentes do ensino fundamental e do ensino médio de escolas públicas e privadas da cidade do Guarujá, através de pesquisa de campo. Além disso, foi feita revisão bibliográfica de livros e artigos de estudiosos que tratam do assunto.

Segundo Silva e Menezes (2001) para realizar uma pesquisa científica é necessário que o autor escolha um tema e defina um problema a ser investigado e após a execução operacional da pesquisa, escreva-se um relatório desse estudo, e este seja apresentado de forma planejada, ordenada, lógica e conclusiva.

Do ponto de vista da abordagem do problema esta pesquisa se caracteriza como quantitativa, pois, neste tipo de abordagem considera-se que tudo pode ser quantificável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las e requer uso de recursos e técnicas estatísticas.

Para a abordagem do tema e considerando que do ponto de vista de seus objetivos que é exploratória foi realizado um levantamento de trabalhos acadêmicos anteriormente realizados por outros autores, estes levantamentos promoveram maior familiaridade com as atuais discussões a respeito do tema. Em seguida foi elaborado um questionário para professores do ensino fundamental.

Informações sobre a pesquisa de campo

A coleta dos dados foi realizada através de um questionário, definida por Silva e Menezes como:

Uma série ordenada de perguntas que devem ser respondidas por escrito pelo informante. O questionário deve ser objetivo, limitado em extensão e estar acompanhado de instruções. As instruções deve esclarecer o propósito de sua aplicação, ressaltar a importância da colaboração do informante e facilitar o preenchimento. (SILVA E MENEZES, 2001, p.40)

No caso desta pesquisa, as perguntas foram divididas em perguntas abertas e fechadas além de serem feitas em linguagem compreensível que pudesse facilitar a compreensão dos informantes e evitar a possibilidade de interpretação dúbia, sugerir ou induzir a resposta.

A pesquisa foi direcionada a professores do ensino fundamental e médio com idade entre 25 e 65 anos de ambos os gêneros, os questionários foram respondidos por professores da rede pública e privada de ensino, da cidade de Guarujá, SP. A escolha dos informantes foi realizada de forma aleatória sem a preocupação de estabelecer o local onde este atua. Parte destes questionários foi entregue pessoalmente aos professores enquanto a outra parte foi enviada por e-mail.

Com relação à análise, nesta pesquisa os dados foram analisados considerando seu aspecto exploratório e quantitativo, descritos à medida que ocorreu a pesquisa.

DISCUSSÃO E ANÁLISE DOS RESULTADOS

Os resultados, bem como as discussões a seguir, representam a parcela mais importante da pesquisa e estão agrupados da seguinte forma: perfil do entrevistado, percepções sobre informática, a educação e prática docente com uso de informática na sala de aula. Na medida em que se apresentam os dados analisados, inserimos alguns depoimentos dos professores, com o objetivo de identificar e comparar a realidade destes professores com outros descritos em pesquisas publicadas.

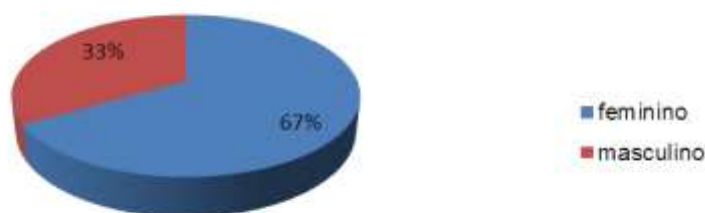
Outro ponto importante a ser destacado é que em nossas entrevistas os professores não foram obrigados a citar seus nomes, por isso nas justificativas apresentadas não aparecem seus nomes e identificamos apenas como depoimento dos professores entrevistados.

O instrumento de coleta foi um questionário semiestruturado, com oito perguntas, as quais são analisadas a seguir, nos gráficos.

Análise quantitativa

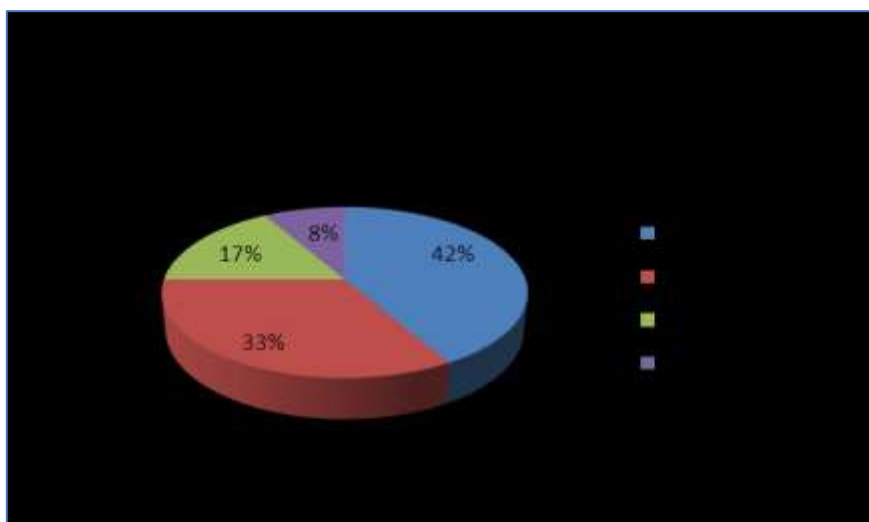
Gráfico 1. Gênero dos entrevistados

Gráfico 1: Gênero dos entrevistados



Para esta pesquisa foram entrevistados professores por meio de questionários com perguntas abertas e fechadas, do total entrevistado 67% são do sexo feminino e 33% do sexo masculino, entre 25 e 68 anos como demonstra o gráfico 1.

Gráfico 2. Faixa etária dos entrevistados



Em relação à faixa etária dos docentes entrevistados, temos uma variação de idade entre 25 e 68 anos. Sendo que a faixa entre 25 e 35 anos é predominante, seguindo por docentes entre 36 e 46 anos de acordo com o gráfico 2.

Gráfico 3. Formação inicial



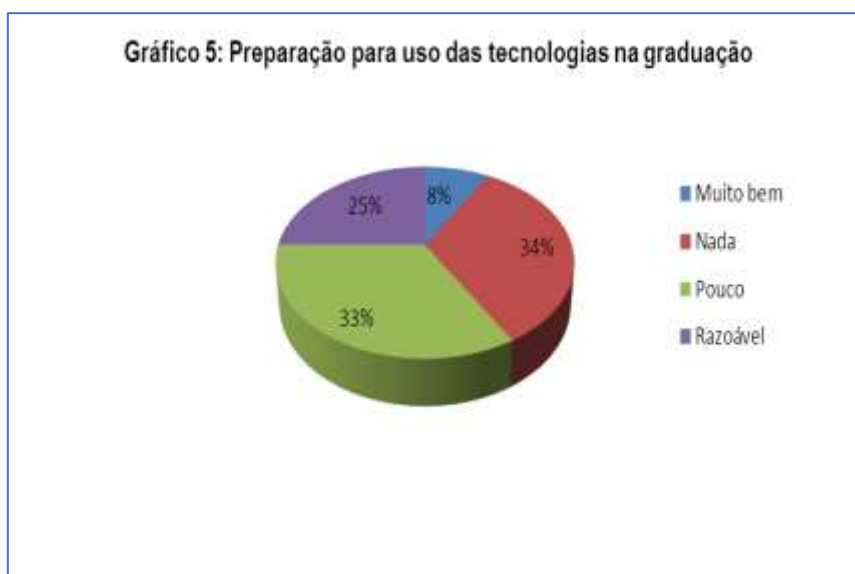
Como se pode verificar pelo gráfico 3, 67% dos entrevistados são formados em Pedagogia, enquanto 33% são de outras licenciaturas, a maior parte destes professores possui pós-graduação e entre elas: em psicopedagogia, gestão escolar, supervisão, entre outras.

Gráfico 4. Tempo de formação



Entre os entrevistados, a maior parte, ou seja 67% está formados há mais de vinte anos, o que significa dizer que estes professores foram, pelo menos teoricamente, contemplados com a inserção da informática na educação, como podemos ver no gráfico 4. Diz-se teoricamente, porque como se pode ver no gráfico 5, somente 8% afirma que foi muito bem preparado para o uso de informática durante seu processo de formação.

Gráfico 5. Preparação para uso das tecnologias na graduação



O percentual é bem pequeno se comparado ao percentual que afirma que não foi preparado em nada. Isso demonstra que, uma vez que esta formação é insuficiente, a saída é contar com um especialista em informática, geralmente o profissional responsável pelo laboratório de informática da escola. Por outro lado, a impressão que se tem é que alguns professores se utilizam desta realidade para justificar suas dificuldades em utilizar este recurso, pois como também mencionado anteriormente trata-se em alguns casos de resistência a adequação aos novos recursos pedagógicos.

Gráfico 6. Rede de ensino que atua



Conforme o gráfico 6, a maior parte dos entrevistados é docente de escola pública, e embora esses professores tenham dificuldade em relação a equipamentos ou mesmo utilização destas tecnologias, usam computadores, eles fazem uso constante do computador, inclusive para planejamento de suas aulas.

Percepções sobre informática na educação

Nesta pesquisa, perguntou-se o que os professores entendem por informática na educação e neste sentido destacam-se as seguintes respostas:

A - Há muita confusão quando se trata de definir informática na educação, os termos se confundem com sua função, para simplificar entendo que se trata de um recurso pedagógico que auxilia no processo ensino aprendizagem e como um recurso é preciso identificar como e quando deve ser utilizado.

B - A informática é um recurso pedagógico fabuloso, entenda RECURSO, colabora e muito na ampliação de horizontes, mas ainda não consegue substituir um professor, uma que vez as relações humanas vêm se perdendo por conta do uso inadequado da informática, tecnologia.

C - A informática que chamaremos de tecnologia da informação faz parte do cotidiano de qualquer pessoa atualmente e principalmente jovens, que são em sua maioria os estudantes, portanto informática na educação é uso desta tecnologia como ferramenta de suporte ao processo de ensino e aprendizagem, mas deixando bem explícito, que ela só não faz milagres, se não houver o desejo de aprender, a curiosidade, a busca pelo conhecimento.

As demais respostas a esta pergunta, resume-se em definir a informática como uma ferramenta ou suporte a ação educativa, as respostas acima foram destacadas por trazerem alguns pontos interessantes de se discutir. Na primeira fala a professora ressalta que mais do que entender o que é informática é saber fazer uso correto do recurso e neste ponto retomamos as palavras de Pinheiro (2007) já discutidas anteriormente, quando a autora diz que o uso da informática e das tecnologias disponíveis tanto pode facilitar como também pode prejudicar os relacionamentos entre as pessoas e que tudo vai depender de como se educa as pessoas para seu uso.

A segunda entrevistada reforçando o que disse a primeira lembra que embora seja um recurso importante não substitui o professor. Além disso, no próprio curso de formação, é necessário que seja discutida a relação das tecnologias com os interesses do capital globalizado e possam desenvolver o censo crítico comprometido com os valores éticos, morais e políticos e humanos, com isso compreender e perceber a responsabilidade das tecnologias no complexo das relações do ser humano com o ambiente, dele com sua própria subjetividade e dele com os outros socialmente. Concordamos com o último professor, que destaca que o uso da informática só vai fazer sentido, se este estiver relacionado com o desejo de aprender e de construir novos conhecimentos, do contrário de nada servirá o acesso a este recurso.

A pesquisa quis saber se a informática pode melhorar o processo ensino aprendizagem. Os entrevistados foram unânimes em afirmar que sim e as justificativas se baseiam no fato de que a informática inova, aproxima o aluno do conhecimento, para efeito de discussão destacamos as seguintes justificativas:

A - O fato que a informática chegou e não há como ignorá-la, podemos até fazer um comparativo em antes da Informática e Depois da Informática. É notório que ela só colaborou no ensino aprendizagem; porém a visão e experiências vividas por um profissional da Educação nenhuma máquina pode transmitir. Enfim, melhorou muito,

B - Com certeza. Pois a maioria dos jovens nasceu, cresceram e vivem dentro do contexto da informática e ignorar essa ferramenta no processo de ensino e aprendizagem, é ignorar o mundo deles, sem o uso desta ferramenta não estaremos falando o “dialeto” deles e seremos ignorados, portanto, se usado corretamente, melhora com certeza essa relação ensino e aprendizagem.

C - Sim. As escolas não aproveitaram o crescimento tecnológico, utilizando-se ainda da metodologia GLS – giz, lousa e saliva. Os alunos são interligados, aproveitam e absorvem o que a tecnologia digital tem a oferecer e quando chegam ao ambiente escolar não encontram uma forma diferenciada no processo de aprendizagem que apresente a mesma motivação para os estudos como os recursos que utilizam diariamente fora da escola.

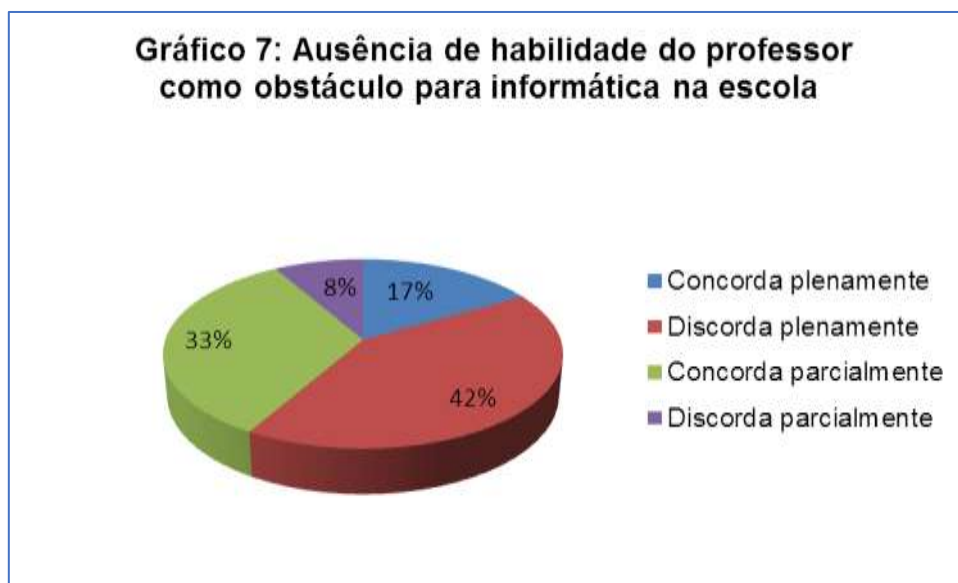
D - Sim. Com uso de ferramenta como a informática muda-se o modo de ensinar, a informática nesse caso, faz mediação entre professor aluno e conhecimento. Também devemos ter em mente que o aprendizado mediado pela informática deve ir além, ou seja, sair da escola para entrar no mundo fora dela.

Quando nos reportamos a essa questão de processo ensino aprendizagem, não podemos nos ater aos conceitos simples considerando a complexidade do processo, por isso, não é tão simples responder a essa questão. As respostas destes docentes certamente vão de acordo com a visão que cada um tem sobre como se dá o processo ensino aprendizagem, bem como relação sendo mediada pela informática.

Dois dos professores acima consideram que se não utilizarem ferramentas como a informática poderá ser ignorada pelo aluno e que estes perderam a motivação de estudar, a nosso ver, a situação não passa necessariamente por aí, o uso de informática não pode ser visto pelo professor dessa forma, ou então estariam reduzindo sua própria prática e supervalorizando outras tecnologias em detrimento da sua. Lembrando que a partir dessa nova realidade o professor deve sim repensar sua prática, se adequar aos novos tempos, mas sem esquecer quem é mais importante neste contexto, pois como já discutido anteriormente, a informática ajuda, mas não faz milagres e nem resolve os problemas da educação.

No percurso dessa pesquisa, verificamos que são vários os obstáculos para a inserção do computador nas escolas e que dentre eles destaca-se o despreparo do professor. Esta pergunta também foi feita aos nossos entrevistados e as respostas são as que seguem:

Gráfico 7. Ausência de habilidade do professor como obstáculo para informática na escola



No gráfico 7, o percentual de 33% que concordam parcialmente que o despreparo do professor é o principal obstáculo, inicialmente a resposta é sim, mas fazem ressalva que isto ocorre em parte, pois não se pode a responsabilidade somente ao professor. Aqueles 8% que discordam parcialmente, afirmam inicialmente que o despreparo do professor é o principal obstáculo se a situação for pensada no contexto da escola quem possui toda estrutura e o professor não sabe fazer uso da ferramenta e nesse caso, ele, o professor precisa buscar sua própria capacitação e como exemplo disso, destacamos a seguinte resposta:

A - Entendo que um profissional da Educação não deve esperar que alguém o prepare para tal, o mundo está inteiramente globalizado, não dá para acreditar que, em pleno século XXI, existam professores esperando uma preparação para assim ministrar suas aulas utilizando este recurso. Se vier preparação, ótimo!!! Se não vier, deve-se se atualizar de qualquer maneira, é um preparo tanto para sua vida social quanto profissional.

Os 17% que concordam plenamente, ou seja, o despreparo do professor é um dos principais obstáculos e representando esta resposta destacamos uma das justificativas:

B - O uso da informática depende da capacidade e do conhecimento do uso desses recursos. Sem uma preparação adequada, não tem como utilizar as ferramentas de maneira consistente. O que percebo é que tem muita gente que ainda tem dificuldades em usar o computador. Na necessidade de usar programas e sistemas de informação, esta dificuldade acaba prejudicando bastante. Como usar softwares educacionais sem saber usar o básico do computador? E preparar aulas sem ter facilidade para usar o computador. Salvar arquivos em pen drive é uma atividade simples, mas para muitos é um parto. Logo, antes de inserir este tópico na escola, é necessário qualificar os profissionais da educação para o seu uso.

Outra professora entrevistada afirma o seguinte: “Sim. Não há cursos de reciclagem para os docentes com formação superior a 10 anos e isso de fato prejudica o próprio docente”. Os demais entrevistados que totalizam 42% dos entrevistados, discordam deste dado e afirmam que não existe o principal obstáculo, mas sim um conjunto deles, o que acontece é o professor está na linha de frente e quando as coisas começam a não funcionar a culpa recai no professor. Um dos entrevistados baseado em sua própria experiência afirma o seguinte:

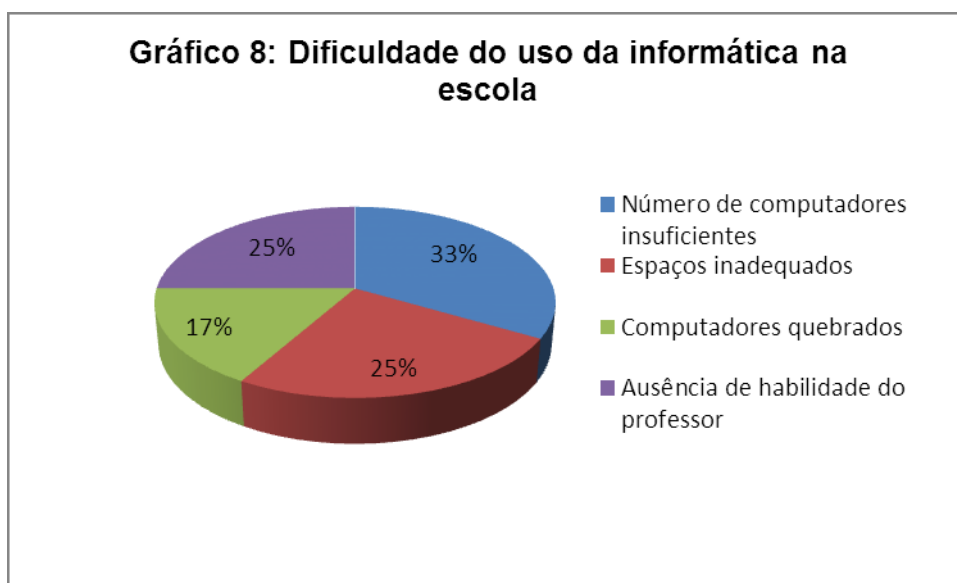
C - Não posso afirmar se as pesquisas estão certas ou erradas. Mas posso afirmar e com base no contexto profissional que transito, que as dificuldades estão mais relacionadas à falta de estruturas e suporte dentro da rede educacional, como também a falta de seriedade e compromisso de gestores públicos com relação à implantação de uma estrutura sólida, séria e que sirva para sedimentação e transmissão do conhecimento para outras gerações de profissionais.

O depoimento do professor se aproxima de outras pesquisas que também revela situações como essa e revelam que não é o só o professor, mas como já mencionado é o conjunto, aliás deve-se se lembrar de que junto a esses problemas existem outros que certamente compromete a prática do professor.

Neste contexto, a pesquisa buscou verificar o grau de habilidade dos professores em relação ao uso de informática e o resultado foi igual para duas situações, ou seja, 50% consideram que seu nível é alto, enquanto os demais 50% considerara razoável. Quem considerou sua habilidade alta, informou como se utiliza dos recursos em sala de aula, informando inclusive, programas que utilizam os demais foram vagos em suas respostas, e, neste caso, não especificaram suas formas de utilização com a informática.

Outra questão aberta sobre as principais dificuldades no uso pedagógico de computadores na escola que trabalham revelou que fatores de infraestrutura, como número de computadores insuficiente, ausência de preparo dos professores, computadores quebrados e conseqüentemente falta de manutenção, além de espaços inadequados como se pode verificar no gráfico 8.

Gráfico 8. Dificuldade do uso da informática na escola



De acordo com os professores entrevistados estas dificuldades estão presentes em suas escolas de origem, e analisando estes dados, há de se compreender o que disse alguns professores na questão anterior, sobre o a ausência de qualificação de o professor ser um dos principais obstáculos e conforme estes mesmos professores:

A - Sem em dúvida a maior dificuldade é a ausência de equipamentos que acarreta também a falta de domínio da tecnologia, outra dificuldade e a disponibilidade, com a implantação de uma rotina estabelecida no plano mensal, bimestral ou mesmo anual. A somatória de todos esses eventos torna o domínio desta tecnologia mais distante para os profissionais da educação.

B - Hoje, o aluno entende mais de informática do que o professor. Acredito que uma grande dificuldade seja a questão física, computadores de fácil acesso a todos os alunos, com manutenção das máquinas em dia. Já passei por situações em que havia atividades específicas para serem executadas em computadores e não havia máquina disponível a todos. Dividir a turma em grupos é interessante quando há quem fique com o grupo restante. Enfim, de acordo com experiências vividas, é a grande dificuldade que costumo encontrar.

C - Primeiramente a falta de recursos. Mesmo onde eles existem, usá-los é burocrático. Trabalho na rede municipal de Guarujá e em todas as salas de aulas tem lousas interativas. Acontece que às vezes (e são muitas) elas não funcionam. Em outras, existem equipamentos, mas para usá-los tem que reservar, montar, perde-se muito tempo. Laboratórios de informática são raros, onde existem são difíceis de acessá-los. E pensando em informática, hoje em dia, o acesso à informação é fácil, num clique no celular os alunos buscam o que querem. Então, temos que estar preparados para não só usar, expor, mas também para mediar à aprendizagem por meio destes recursos. Eis a grande dificuldade. Ou seja, recursos humanos qualificados, recursos materiais disponíveis e acessíveis e os softwares também

acessíveis. Com tudo isso, dá para planejar bem e usar a informática na educação de maneira mais fácil.

Os depoimentos dos professores evidenciam a realidade que cada um deles vivencia nas escolas que atuam e que não pode ser ignorado ou simplesmente se buscar culpados pelas falhas que ocorrem, as pesquisas estão aí para provar que a inserção a informática na educação foi um ganho tanto para professores como para os alunos e para a sociedade como o todo, por estas razões é que o uso da informática, assim como qualquer tecnologia (inclusive a do professor) deva de fato, promover o acesso a informações indispensáveis para entender o mundo, sem esquecer que isso deva fazer diferença no processo ensino aprendizagem de forma positiva.

CONSIDERAÇÕES FINAIS

Para a elaboração deste trabalho, além de leituras e revisão bibliográfica de autores como Valente, Soares e Tardiff, também foi realizada uma pesquisa de campo para saber qual é a realidade de professores que atuam em escolas públicas e privadas na cidade do Guarujá-SP. Dentre do que foi possível analisar, percebemos que os avanços nesta área são positivos, porém existem problemas comuns no uso da informática por partes dos professores em sala de aula.

Para os professores entrevistados a formação não é percebida como suficiente e adequada, para o uso de informática centrada no processo ensino aprendizagem. Os docentes acreditam que há obstáculos para o uso da informática na escola, mas também não discordam que este é sim um obstáculo entre tantos que se tem dentro da escola.

A habilidade do professor em relação ao uso pedagógico é classificada entre alto e razoável, o que indica que mesmo não tendo tido formação adequadas estes profissionais buscaram outras formas de se aperfeiçoar. Quanto aos recursos e a infraestrutura, estes, segundo os professores não são suficientes e isto acaba interferindo no trabalho do professor e na qualidade do ensino ofertado aos alunos.

A questão norteadora desta pesquisa foi verificar se os professores conseguem acompanhar a evolução da informática na educação e a resposta foi positiva, pois como foi possível verificar, mesmo com todas as dificuldades e limitações enfrentadas pelos professores, eles buscam seu aperfeiçoamento porque, certamente, entendem que se eles mesmos não se preocuparem com isso, ficaram para trás. Essa disposição do professor é positiva até mesmo porque melhora sua prática, pois o professor ganha novas dinâmicas e se aproxima de seu aluno que já possuem muita intimidade com as tecnologias existentes, assim, a relação com o professor fica menos autoritária e mais colaborativa na construção do conhecimento.

Por outro lado, a responsabilidade deve ser compartilhada por aqueles que fazem a política educacional e dos recursos educacionais para que todo o processo seja realizado com compromisso e não apenas para dizer que a escola é informatizada como se tem visto em muitas escolas no Brasil, como por exemplo, proporcionar aos professores atualização das práticas docentes, de forma que não só a informática, mas toda tecnologia disponível a educação seja incorporada e utilizada de forma plena e democrática.

Além destas considerações, são interessantes também refletir sobre as concepções de ensino aprendizagem que permeia as práticas educativas que se utilizam da informática e as novas tecnologias. Com toda evolução dos últimos tempos na educação vai se perceber que os conceitos sobre ensino e aprendizagem ganharam novas abordagens e as práticas pedagógicas não ficaram atrás. Como afirmou Gouvêa (1999), nos tempos atuais o papel do professor é mais importante, o que vai mudar são os instrumentos que hoje se utiliza nesse processo e por isso também cabe ao professor dos tempos atuais refletirem sobre seus conceitos de aprendizagem e ensino integrado às novas tecnologias, para tanto é necessário um conhecimento específico da necessidade de mudança no fazer do professor. Neste sentido, Richter et al afirma (2003, p.12)

Para realizar um trabalho incorporando o uso das tecnologias no processo ensino-aprendizagem, devemos refletir sobre as seguintes questões: a) o que se entende hoje por processo ensino-aprendizagem; b) quais as possibilidades de uso das tecnologias no desenvolvimento de projetos de trabalho interdisciplinares; c) as vantagens e desvantagens do uso da informática na educação.

Neste contexto, os autores lembram que ao integrar o computador às práticas escolares é preciso repensar o processo ensino aprendizagem para que o conhecimento seja construído e contextualizado. A construção do conhecimento ocorre com a realização de uma ação, e a contextualização é necessária para assegurar o significado da ação, tendo em vista a realidade do aluno.

Apesar destas dificuldades, reconhecemos que do ponto de vista das frentes de informatização das escolas, o Brasil tem feito um esforço importante na área da educação pública. Por outro lado, percebemos que ainda se faz necessário maior investimento nesta área, para a tecnologia seja, definitivamente um apoio ao professor e a inclusão digital seja uma realidade, tanto para alunos como para professores.

REFERÊNCIAS

DINIZ, Melissa. O grande desafio de quem ensina. In: Revista Nova Escola- edição especial. Disponível em <<http://www.revistaescola.abril.com.br>> Acesso em 20/11/2018.

GARCIA, Joaquin Ramos García (org). **Ensinar ou Aprender a Ler e a Escrever**. Artmed Editora Porto Alegre 2001.

GOUVÊA, Sylvia Figueiredo. **Os caminhos do professor na Era da Tecnologia** - Acesso Revista de Educação e Informática, Ano 9 - número 13 - abril 1999.

HEYDENREICH, Dietmar MICHEL, Daniela RAMOS, José Sérgio et al. **Escolas Públicas: A Informática como instrumento pedagógico**. Relato de experiência. São Paulo -2005. Disponível em < <http://www.fbh.org.br/>> Acesso em 03/09/2018.

MAIO, Angélica Carvalho Di. SETZER. Alberto. **Educação, Geografia e o desafio de novas tecnologias**. Disponível em <[http:// www.scielo.gpeari.mctes.pt/scielo](http://www.scielo.gpeari.mctes.pt/scielo)> Acesso em 10/12/201

MASETTO, M. T. **Novas tecnologias e mediação pedagógica**. 8. ed. Campinas, SP: Papirus, 2006.

PINHEIRO, Patrícia Peck. **Boas Práticas legais no uso da tecnologia dentro e fora da sala de aula**: Guia rápido para as instituições educacionais. Disponível em <<http://www.criancamaissegura.com.br>> Acesso em 20/10/2018.

PRADO, ME. B.B. VALENTE, J.A. **Formação de Educadores a Distância Via Internet**. São Paulo. Avercamp, 2003.

ROCHA, Carlos Alves. **Mediações tecnológicas na Educação superior**. IBPEX Curitiba. 2009.

RICHTER, Denis. BRAGA, Flávia Spineli. et al. **Informática no processo Ensino-Aprendizagem**: contribuindo para uma nova escola. Disponível em: <<http://wwwrevista.fct.unesp.br>> Acesso em 20/12/2018.

SILVA, L. S.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. Manual de orientação. Florianópolis, 2001. Disponível em: <<http://www.scribd.com/doc/2367267/DA-SILVA-MENEZES-2001-Metodologia-da-pesquisa-e-elaboracao-de-dissertacao>> Acesso em: 14/10/ 2018.

SOARES, M. **Novas práticas de leitura e escrita**: letramento na cibercultura. Educação e Sociedade: Campinas, 2002.

TARDIF, Maurice. **Saberes docentes e formação profissional**. Petrópolis, RJ: Vozes, 2002.

VALENTE, José Armando. Visão analítica da informática na Educação no Brasil. A questão da formação do professor. **In**: Revista Brasileira de Informática na Educação – Número 1 – 1997. Disponível em<<http://www.geogebra.im-uff.mat.b>> Acesso em 03/11/2018.

VALENTE, José Armando. **O computador na sociedade do conhecimento**. Campinas, SP: UNICAMP/NIED, 1999.

XAVIER. Antônio Carlos dos Santos. **Letramento digital e ensino**. Disponível em <<http://www.ufpe.br>> Acesso em 30/09/2018.

BOAS PRÁTICAS EM DATA CENTER PARA USO DE TECNOLOGIA VERDE

GOOD PRACTICES IN DATA CENTER FOR USE OF GREEN TECHNOLOGY

Josiane Martins Silva, GUIRARDI

josianemartinssilva@hotmail.com

Sistemas de Informação, Centro Universitário Padre Anchieta

Edilson Gomes, FEITOZA

edilson.feitoza@anchieta.br

Professor do Centro Universitário Anchieta (Jundiaí/SP), Especialista em Rede de Computadores pela UNIP/SP.

Resumo

A eficiência energética em data centers é uma das principais preocupações dos gestores de tecnologia da informação. Uma das razões é que o consumo de energia elétrica se tornou responsável pelas maiores despesas relacionadas com o ambiente. O artigo tem a finalidade de estimular a preservação ambiental e explorar a aplicabilidade de boas práticas de tecnologia da informação verde para obter melhor eficiência energética em data centers. As práticas apresentadas no artigo consistem-se em virtualização, e consolidação de servidores, desativação de servidores inativos, aquisição de servidores energeticamente eficientes, layout da área mais eficiente, confinamento de corredores quentes e frios, dispositivos de gerenciamento de fluxo de ar, controle de temperatura e umidade, economizadores a ar e a água e aparelhos de iluminação com eficiência energética.

Palavras-chave: TI verde; data center; eficiência energética; práticas.

Abstract

Energy efficiency in data centers is a major concern of information technology managers. One of the reasons is that electricity consumption has become responsible for the higher expenses related to the environment. The article aims to stimulate environmental preservation and explore the applicability of green information technology best practices for better energy efficiency in data centers. The practices presented in the article consist of virtualization and server consolidation, deactivation of inactive servers, acquisition of energy efficient servers, more efficient layout of the area, confinement of hot and cold corridors, air flow management devices, control of temperature and humidity, air and water economizers and energy efficient lighting appliances.

Keywords: Green IT; data center; energy efficiency; practices.

INTRODUÇÃO

Para demonstrar responsabilidade com o meio ambiente as empresas têm adotado o termo "data center verde" ou "data center sustentável" que abrange o uso consciente da tecnologia a favor da sustentabilidade. Com o uso de alternativas e boas práticas que podem reduzir custos, minimizar efeitos causado ao ambiente, diminuir o consumo excessivo de energia elétrica, reutilizar recursos naturais e inovação tecnológica.

Possuir data center sem planejamento sustentável tem provocado grande consequência nas organizações como o aumento excessivo do consumo de energia elétrica, o descarte incorreto de equipamento não utilizados, e danos ambientais.

Segundo (CARRERO, 2015, p.688) “Os custos de energia dos *datacenters* são uma grande preocupação da TI Verde, porque mais da metade de todos os custos de energia elétrica relacionados à TI são pelos *datacenters*.”

O funcionamento do data center verde, tem como propósito, estimular a preservação ambiental. Visa discutir o impacto causado ao meio ambiente e o consumo excessivo de energia. Com a aplicação de técnicas e ferramentas disponíveis no mercado, escolher alternativas qualificadas, para minimizar essas decorrências.

Concentrar no gerenciamento de datacenter, serviços e dispositivos de TI, climatização, energia, estrutura física e monitoramentos. Implementar algumas alternativas como: virtualização e consolidação de servidores, desativar servidores sem uso, aquisição de servidores energeticamente eficientes, layout e confinamento de corredores quentes e frios, dispositivos de gerenciamento do fluxo de ar, ajustes de temperatura e umidade de entrada, economizadores a ar e a água.

A disposição estrutural do artigo é composta por seções com conceitos de sustentabilidade, Tecnologia da informação verde, Data center, Computação em nuvem, Eficiência energética, e análise de resultados. A sessão sustentabilidade mostra em geral como preservar o ambiente. Tecnologia da informação verde (TI verde) destaca o uso de técnicas conscientes da área de TI voltada para o meio ambiente.

A sessão data center ressalta sobre o seu conceito, uso e importância para as organizações. Computação em nuvem sobre os recursos disponibilizados em nuvens vantagens e desvantagens e eficiência energética descreve formas de economizar energia no data center. Para análise de resultado foi realizada a verificação dos requisitos estudados, e se essas ferramentas e técnicas são realmente eficientes.

Para o desenvolvimento deste artigo foi examinado artigos, e publicações em sites relacionados a sustentabilidade e tecnologia da informação verde, com a finalidade de autenticidade das informações sobre o assunto abordado.

Portanto o projeto discorrerá a respeito de técnicas que podem ser implementadas aos data centers, buscando minimizar desperdícios e tornar processos relacionados a tecnologia da Informação mais eficazes.

SUSTENTABILIDADE

De acordo com os autores Vita e Engelmann a sustentabilidade “visa estabelecer um equilíbrio entre o que a natureza nos oferece e o limite de consumo dos recursos naturais, de maneira que a qualidade de vida do homem esteja em harmonia com a capacidade do planeta.” (VITA; ENGELMANN, 2016, p.130).

São ações que utilizam os recursos naturais de maneira eficiente para que não comprometa o amanhã. Fica notável que as ações sustentáveis vividas hoje, pode garantir condições suficiente para manter a qualidade de vida da população no futuro.

A sustentabilidade expõe o modo que as pessoas devem agir, buscando qualidade de vida, porém respeitando a natureza. Para (AQUINO, et al, 2015, p.107), existem três fatores importantes dentro da sustentabilidade, “o desenvolvimento sustentável está pautado nos seus três pilares: social, ambiental, e econômico [...]”. Para que ela realmente ocorra, esses princípios necessitam estar relacionados entre si.

A sustentabilidade social está relacionada ao indivíduo, seu bem-estar e suas condições de vida. “Entende-se como a criação de um processo de desenvolvimento sustentável pela visão de uma sociedade equilibrada, que busca um novo estilo de vida adequado ao momento presente e ao futuro.” (FROEHLICH, 2014, p.157).

A sustentabilidade ambiental refere-se à conservação do meio ambiente, seus recursos naturais e o modo que são utilizados. “[...] é definida como a desmaterialização da atividade econômica, pois uma diminuição do processamento de material pode reduzir a pressão sobre os sistemas naturais e ampliar a prestação de serviços ambientais para a economia.” (Sartori, et al, 2014, p.4).

A sustentabilidade econômica baseia na produção, organização de bens e serviços. “Do ponto de vista econômico, a sustentabilidade prevê que as empresas têm que ser economicamente viáveis.” (DIAS; MARQUES, 2013, p.72) É manter o crescimento econômico, sem prejudicar o meio ambiente.

Cada um dos princípios acima mencionado deve ser analisado dentro de cada caso para que todos trabalhem em harmonia, alinhado ao planejamento estratégico da organização, para atingir um resultado satisfatório, dentro das empresas e também na vida de toda a sociedade.

Além dos elementos ambientais, sociais e econômicos a sustentabilidade empresarial abrange ainda outros aspectos importantes. De acordo com os autores (DOS SANTOS; DA SILVA, 2017, p.84) “Há também requisitos de suma importância como a ética, transparência, a prática da governança corporativa de maneira efetiva, o bem-estar dos funcionários e a seleção de fornecedores engajados com a sustentabilidade[...]”.

Além disso inserir o desenvolvimento sustentável nas empresas é uma oportunidade de inovação. Utilizar tecnologias que tornam os processos de forma mais limpa, pode reduzir custos e obter benefícios e, ainda, preservar o meio ambiente.

Entretanto a crescente inovação tecnológica é responsável pela comodidade, acessibilidade, conforto e vantagens econômicas. No entanto, tem causado consequências para a sociedade, o consumismo exagerado, o desperdício de energia, de água e de diversos outros recursos.

Com isso as reservas de recursos naturais não renováveis estão cada vez mais escassas, os resíduos sólidos aumentam a cada dia, e os espaços para armazenamento do lixo

eletrônico estão diminuindo. As pessoas e as empresas estão muito preocupadas com os benefícios oferecidos pela natureza, esquecendo os danos que vêm causando ao meio ambiente ao longo dos anos.

O intuito do desenvolvimento sustentável é a preservação do planeta. Agir de modo sustentável é responsabilidade de toda a sociedade, em suas casas, nas empresas e em todos os lugares. Algumas ações são importantes para garantir melhoria de vida da população:

- Garantir o replantio de áreas exploradas
- Preservar de áreas verdes
- Uso de fontes de energia renováveis
- Reciclar
- Descartar corretamente o lixo eletrônico
- Controlar a exploração de recursos naturais
- Reduzir do consumo de água e energia elétrica
- Evitar o consumismo desnecessário
- Promover e aplicar os conceitos de sustentabilidade.

Desse modo para reduzir os efeitos causados ao meio ambiente deve ser colocado em prática as ações citadas acima. Deve evitar o consumo exagerado de tecnologias sem necessidade, aplicar práticas de redução de energia elétrica e água, controlar a exploração de recursos naturais não renováveis, reduzir o lixo eletrônico, descartando-o em locais específicos para o descarte, doar equipamentos eletrônicos que ainda funcionem, devolver ao fabricante, vender e comprar aparelhos usados.

TECNOLOGIA DA INFORMAÇÃO VERDE

Nas palavras de (LUNARDI, et al, 2012, p.3) “TI Verde possibilita identificar o que as organizações precisam para se tornarem mais verdes e a extensão com que têm progredido em direção à TI Verde.”

É a aplicabilidade inteligente de práticas constante do uso de técnicas e recursos de computação da forma mais eficiente possível, economia de energia e consciência ambiental nas áreas relacionadas à tecnologia.

Para (DA SILVA, et al, 2015, p.60). “A redução do consumo de energia é uma ação prioritária da TI Verde”. Os autores acima afirmam que o objetivo principal da TI verde é utilizar técnicas adequadas para minimizar o consumo de energia elétrica nas empresas.

A dependência crescente das tecnologias nas corporações, as áreas de tecnologia da informação têm se transformado em grandes consumidores de energia, e uso de processos de gestão e manutenção inadequados e antigos tem deixado visível que o efeito sobre o meio ambiente é gigantesco.

A TI Verde consiste na criação, produção, uso e descarte conscientes e voltados para o meio ambiente. Podem ser consideradas boas práticas de acordo com Soldati (2017) todas aquelas que dentro da área de tecnologia, envolvem:

Controlar e otimizar o consumo excessivo de energia elétrica gerado pelos servidores e dispositivos de rede, principalmente quando os equipamentos são antigos, e que não possui controle eficiente de consumo de energia.

- Reciclar componentes eletrônicos;
- Gerenciar o descarte de lixos eletrônicos e resíduos;
- Diminuir o uso de metais pesados e outros componentes tóxicos na fabricação de componentes e produtos eletrônicos;
- Usar soluções de software para problemas de hardware; e
- Adotar uma forma de auditoria constante.

Os equipamentos com mais tecnologia consomem menos energia, pode possuir partes de seus componentes reciclados e com maior durabilidade de uso, isso faz com que o descarte desse equipamento seja o mais demorado possível.

O autor (PEREIRA, 2009, p.104) acredita que as empresas que adota as práticas de TI verde, além de cuidar do ambiente está também reduzindo custos.

Entre as questões afirmativas sobre as práticas da TI Verde na empresa e a importância da TI Verde para o desenvolvimento sustentável. A empresa constata no que se refere às práticas da TI Verde adotada pela empresa, dispõe da racionalização do uso de energia, racionalização no uso de insumos (papel, equipamentos e materiais de uso), utilização de certificações ambientais[...]; e existe a conscientização social a limitação de recursos através da evidência com as certificações ambientais obtidas e do conscientização dos próprios funcionários.

Apesar de a economia de energia vir em primeiro lugar, também outros elementos precisam ser considerados, como por exemplo diminuir o uso de insumos, utilizar produtos que não oferecem risco a natureza, descartar lixo de maneira correta e também conscientizar os funcionários sobre a importância da implantação de técnicas de TI verde na empresa.

As práticas de TI verde pode ser aplicada em três etapas: grupo tático, grupo estratégico e grupo *deep it*. Nas palavras das autoras Pinto e Savoine (2011, p.7).

TI Verde de incrementação tática: As práticas contidas nesse nível não afetam a infraestrutura de TI, e também não modificam as políticas internas da empresa, apenas agregam algumas medidas de controle do uso excessivo de energia elétrica. [...]TI Verde Estratégico: As práticas relacionadas a esse nível exigem mudança na infraestrutura de TI da empresa. [...] TI Verde a Fundo: Este nível é a integração dos dois níveis anteriores, porém com uma diferença, ele requer maiores gastos, pois visa uma mudança total de instalações, desempenho de equipamentos e padronização de processos[...].

Cada um dos níveis acima citado apresenta suas próprias particularidades e funções. E nas subseções abaixo será especificado, as características e aplicabilidade de cada grupo.

Grupo tático

De acordo com Abreu “Nesse nível é proposto medidas de redução de gastos elétricos quando excessivos”. (ABREU, 2012, p.5) Esse grupo inclui medidas que controlam o consumo excessivo de energia elétrica. São medidas simples de ser implementadas e não geram custos.

Esse grupo não faz alteração nem de infraestrutura e nem de políticas internas. “Não modifica a infraestrutura de TI, nem as políticas internas.” (NUNES, et al, 2011, p.2) A implementação da TI Verde neste grupo pode ocasionar resistência por parte dos

colaboradores. Para suavizar essa resistência é viável investir em campanhas de conscientização de sustentabilidade.

Grupo estratégico

Gianelli afirma que nesse grupo “É necessária a realização de uma auditoria sobre a infraestrutura de TI e seu uso relacionado ao meio ambiente, desenvolvendo e implementando novos meios viáveis de produção de bens ou serviços de forma ecológica.” (GIANELLI, 2016, p.24).

Refere-se a um nível mais profundo que o grupo tático, pois abrange o planejamento de práticas voltadas para infraestrutura.

Grupo *deep it*

Este grupo é a integração dos dois grupos anteriores. “Incorpora o projeto e implementação estrutural de um porque tecnológico visando a maximização do desenvolvimento com o mínimo gasto elétrico, incluindo projetos de sistemas de refrigeração, iluminação e disposição de equipamento no local [...]” (NUNES, et al, 2011, p.2). A implementação das práticas relativas a este grupo demanda um custo muito maior se comparado aos custos dos outros dois grupos.

Esses grupos especificados, podem ser indicadores da evolução do compromisso da TI com o meio ambiente e mostrar que ao adotar qualquer grupo, mesmo que seja o mais simples, já traz um retorno não só para o meio ambiente, mas também mostra e traz uma redução de custos para a empresa que, afinal de contas, é o que mais interessa.

DATA CENTER

É considerado um espaço físico planejado para abrigar servidores, e outros componentes como sistemas de processamento e armazenamento de dados e ativos de rede. Surgiu da necessidade de manter todos esses recursos em um mesmo ambiente. Nele estão concentrados os sistemas computacionais de uma empresa ou organização onde necessita de uma atenção específica. (Bandeira, et al, 2017)

Os equipamentos são montados em racks com o propósito de abrigar servidores e bancos de dados, e processar grandes quantidades de informação. Possuem sistemas de resfriamento dos racks, para manter a temperatura estável e proteção contra incêndios. Geralmente, por questões de segurança ficam localizados em locais com acesso restrito de funcionários.

Zunchi e Amâncio (2013) consideram que o data center é o sucessor dos centros de processamento de dados dos anos 70 e 80, com capacidade de processamento e de armazenamento superior à do ambiente legado. Está relacionado a um ambiente com convivência de equipamentos de diversos fornecedores.

Conforme apontado é possível notar que o data center moderno tem uma maior capacidade de armazenamento e processamento de dados. Com a adesão de novas tecnologias e substituição de equipamentos ultrapassado, fica notável a mudança quando comparada ao antigo centro de processamento de dados.

Da Silva e Junior consideram que “A construção de um Data Center leva em consideração vários fatores” (DA SILVA; JUNIOR, 2013, p.35). Para Zunchi e Amâncio. “Antes de tudo, um data center precisa ser bem planejado, ou melhor, precisa ser projetado” (ZUNCHI; AMÂNCIO, 2013, p.45).

De acordo com as afirmações dos autores citados no parágrafo acima, ambos descrevem que uma construção de um data center não é tão simples, e que antes de tudo necessita de um projeto, ou seja, um planejamento. Conseqüentemente requer uma análise inicial e demanda de diversos aspectos que são extremamente importantes.

Para Da Silva e Junior (2013), em empresas de alto caráter sustentável, tudo é metodicamente projetado. Favorecendo a questão ambiental quando a mesma investe em alta tecnologia, nos sistemas de refrigeração e no fornecimento de energia para os recursos de TI. Tudo é feito de acordo com a regulamentação de segurança de dados exigidas pelo Banco Central e órgãos internacionais do segmento.

Zunchi e Amâncio (2013), esclarecem que sem um planejamento adequado as conexões dos novos equipamentos passam a ser feitas sem seguir um padrão necessário. E para que um data center seja planejado de forma eficiente necessita da estabilidade de três fatores: tendências tecnológicas, padronização e modularidade.

De acordo com os autores, o grande obstáculo em planejar um novo projeto de data center, é conseguir maior eficiência em relação aos elementos de construção e instalação como:

- Energia consumida nos equipamentos de rede e servidores;
- Ar condicionado;
- Fornecimento e manutenção da energia;
- Elementos que não são prejudiciais ao meio ambiente;
- Segurança e controle de acesso.

Os autores Zunchi e Amâncio destacam os fatores responsáveis pela dificuldade em planejar um novo data center:

O desafio dos novos projetos de data center é conseguir maior eficiência dos diversos elementos construtivos e de instalação, o que se traduz em menor consumo. Entre os vários elementos a serem considerados, pode-se citar: energia para alimentação dos equipamentos de rede e servidores; ar condicionado; geração e manutenção da energia; elementos de detecção e combate a incêndio que não agridam a natureza e que não coloquem em risco a continuidade das atividades; segurança e controle de acesso uma vez que os dados armazenados podem ser de extrema importância para quem os armazena” (ZUNCHI; AMÂNCIO, 2013, p.50).

Em conformidade ao que foi escrito pode-se observar que vários elementos são levados em consideração quando se trata de um data center sustentável. Nota-se responsabilidade desde o consumo de energia até o controle de acesso. Já que estão armazenadas informações de extrema importância para a organização.

As oportunidades para o aumento da eficiência energética de um data center vão desde o projeto de estruturas, até medidas baratas e simples. Para que isso ocorra, os departamentos de TI e de instalações da empresa precisam trabalhar muitas vezes em conjunto com ajuda externa para buscar novas formas de enfrentar desafios ambientais.

COMPUTAÇÃO EM NUVEM

A definição de computação em nuvem nas palavras de (ALVES, et al, 2013, p.3) “A computação em nuvem é a expressão usada para descrever sistemas que permitem que usuários, desenvolvedores e empresas usem a Internet para acessar programas e dados armazenados em centrais de processamento de dados”.

Pois com essa implantação, as organizações necessitam de uma pequena infraestrutura composta pelos dispositivos utilizados e por uma robusta rede de acesso, deixando a responsabilidade de manter a infraestrutura a plataforma e aplicativos atualizados para os provedores.

Nas palavras de Souza “A computação em nuvem pode ser definida, de forma simplificada, como um paradigma de infraestrutura que permite o estabelecimento do SaaS sendo um grande conjunto de serviços baseados na web com o objetivo de fornecer funcionalidades[...]”. (Souza, 2009) Como o armazenamento é feito através da internet, fica possível acessar informações da empresa sendo indispensável a conexão com a internet.

Computação em nuvem, se refere à ideia de utilizar, em qualquer lugar e independente de plataforma, as mais variadas aplicações por meio da internet com a mesma facilidade de tê-las instaladas em nossos próprios computadores, com a evolução constante da tecnologia computacional e das telecomunicações, o acesso à internet vem se tornando cada vez mais amplo e rápido e esse cenário está criando uma condição perfeita para a popularização da computação em nuvem (ALECRIM, 2015).

De acordo com os autores (ALVES, et al, 2013, p.4). Existem quatro tipos de implantação de computação em nuvem que estão divididos em Nuvem Privada, Nuvem Pública, Nuvem Comunitária e Nuvem Híbrida. Abaixo uma breve ideia sobre cada modelo:

“No modelo de implantação de nuvem privada, a infraestrutura de nuvem é provisionada para uso exclusivo por uma única organização compreendendo vários consumidores internos (ALVES, et al, 2013, p.4)”. Quando a aplicação ou máquina virtual é alojada dentro da empresa.

“No modelo de implantação de nuvem pública, a infraestrutura de nuvem é provisionada para uso aberto ao público em geral, sendo acessada por qualquer usuário que conheça a localização do serviço” (ALVES, et al, 2013 p.4). Quando a aplicação ou máquina virtual é alojada em um fornecedor que atende diversos clientes. Os benefícios desse modelo é rápido provisionamento do que será consumido, escalabilidade, e padronização de processos.

“Na implantação do modelo de nuvem comunidade, a infraestrutura de nuvem é provisionada para uso exclusivo por uma comunidade específica de consumidores de organizações que têm preocupações comuns (ALVES, et al, 2013, p.4). Esse modelo permite ser administrado pelas próprias empresas ou por terceiros.

Já no modelo de implantação de nuvem híbrida nos dizeres dos autores (ALVES, et al, 2013, p.4), “[...] a infraestrutura de nuvem existe em uma composição de duas ou mais infraestruturas de nuvem distintas, que permanecem como entidades únicas ligadas por uma tecnologia padronizada[...] que permite a portabilidade de dados e aplicações”. É considerado nuvem híbrida quando a empresa faz uso de ambos os locais de armazenamento.

Ao falar dos ganhos da computação em nuvem Souza descreve que: “As características essenciais são vantagens que as soluções de computação em nuvem oferecem. Algumas destas características, em conjunto, definem exclusivamente a computação em nuvem e faz a distinção com outros paradigmas”. (SOUZA, 2009).

Seguindo o raciocínio do autor, a computação em nuvem traz benefícios para as empresas como: escalabilidade, maior agilidade, baixo custo, acesso móvel, colaboração, segurança, diminuição de risco, alta disponibilidade, gerenciamento do uso de recursos, praticidade e atualização. É muito prático fazer uma atualização no servidor já instalado em nuvem, pois todas as configurações do sistema podem ser realizadas remotamente.

Mesmo com todas as vantagens descritas pelo autor acima, Cunha (2012) diz que a computação em nuvem também possui desvantagens: a possibilidade de vazamento de informações, o risco de perda de dados e/ou informações sensíveis pela empresa contratante, necessidade de conexão com a internet.

EFICIÊNCIA ENERGÉTICA

“A eficiência resume-se na busca de menor quantidade de energia para o fornecimento da mesma quantidade de valor energético.” (BUSSE, 2010, p.2) Baseia-se no modo com que uma tarefa é executada com o mínimo de gasto energético e utilizar de modo eficiente a energia para atingir um resultado esperado.

Simple atitudes que muitas vezes são esquecidas, ou até mesmo deixada de lado pelas empresas na hora de projetar um data center podem contribuir para a redução do consumo de energia elétrica.

“A economia dos recursos de energia é umas principais vantagens que as empresas obtêm com a TI Verde.” (Nunes, et al, 2011, p.5) A energia está presente no cotidiano das empresas, porém existem formas de amenizar o consumo de energia elétrica nos data center e serão citadas nos tópicos abaixo.

Virtualização de servidores

Segundo Soldati “Um servidor virtual é uma instância de alguma plataforma de sistema operacional que roda em uma configuração física qualquer de servidor.” (SOLDATI, 2011, p.101).

Com o passar do tempo sua tecnologia foi aprimorada e utilizada de forma que traz benefícios em sua utilização em data centers pois além de reduzir custos ela facilita a manutenção e utilização de recursos.

O autor Soldati (2011, p.101) acrescenta que “Com a utilização de servidores virtuais em um Data Center, a primeira melhoria observada é na capacidade de processamento dos servidores, que são menos sujeitos à redundância [...]” Por existir vários servidores virtuais em um servidor físico, o processamento, é sempre direcionado para um servidor e de modo algum fica parado.

Segundo Da Silva (2017, p.5) A virtualização de servidores surgiu da necessidade de consolidar servidores em uma única máquina física, utilizar de forma mais eficiente os recursos computacionais e diminuir gastos.

De acordo com a afirmação do autor, o fato de que vários sistemas executados em apenas uma máquina física, diminui gastos tanto energético quanto gastos operacionais. Já que servidores virtuais carregam mais rápido, reduz a indisponibilidade, e também diminui o tempo de recuperação de falhas.

Fica notável na virtualização a redução do uso de servidores físicos:

A virtualização de servidores proporciona uma redução drástica das necessidades de alimentação de TI. A virtualização quase sempre reduz consideravelmente o número de servidores instalados. A eliminação de um servidor é uma eliminação estrutural de consumo de aproximadamente 200-400 W, dependendo da tecnologia. Portanto, o consumo de energia elétrica evitado é de aproximadamente US\$380 por ano por servidor eliminado, e o TCO total em 10 anos poupado por essa eliminação estrutural é de aproximadamente US\$7.680 por servidor eliminado. Essa economia é substancialmente maior do que o custo do servidor em si. (RASMUSSEN, 2012, p.7).

Conforme Rasmussen afirma a virtualização reduz o número de servidores instalado. De acordo da tecnologia aplicada, a retirada de um servidor é uma eliminação de consumo em torno de 200- 400 W. A economia gerada no uso da virtualização pode ser maior do que o custo do servidor.

A virtualização pode ser indicada a qualquer tipo de empresa que tem necessidade de mais de um servidor. Seu custo de implantação não é alto, já que seu investimento com servidores e aquisição de equipamentos e software traz um retorno com a redução de gastos com manutenção, e economia de energia. Algumas vantagens da virtualização segundo Dos Santos, (2014):

- Menor Aquecimento / Economia de Energia
- Redução de Custos de Aquisição e Redução do Espaço Físico
- Menor tempo de parada em manutenções programadas
- Rápida recuperação de falhas, backup otimizado e Recuperação de Desastres fácil
- Ambientes de Testes
- Provisionamento rápido de novos aplicativos e servidores
- Independência de fornecedor
- Isolamento de Serviços e Aplicações
- Manter sistema legado
- Automatização de processos e contabilização de recursos
- Migração para a Computação em Nuvem facilitada

Mesmo possuindo vários benefícios a virtualização de servidores também possui desvantagens. Algumas desvantagens segundo (GHANNOUM; RODRIGUES, 2018):

- Grande uso do espaço em disco
- Sem acesso direto ao hardware
- Desempenho
- Segurança
- Acúmulo de máquinas virtuais
- Necessidade de um servidor físico
- Performance de softwares ou sistemas

Consolidação de servidores

Para (DA SILVA, 2017, P.6). “A consolidação de servidores busca diminuir a quantidade de servidores físicos utilizados para um mesmo serviço sem que estes fiquem ociosos ou com baixa relação de utilização/energia gasta.” Já para Schulz e Silva (2012, p.127), “A consolidação é uma forma de associar os processos em um menor número possível de máquinas”.

A consolidação de servidores, procura reduzir o número de servidores físicos. Consolidar é uma excelente estratégia, pois diminui gastos com energia e refrigeração por integrar diversos servidores em uma única máquina física. Da mesma forma diminui a necessidade de espaço físicos em racks, reduzindo também gastos com manutenção.

Diante das definições citadas pelos autores acima compreende que a consolidação de servidores visa facilidade de gerenciamento dos servidores, já que torna o sistema mais simples. Reduz custos de operação, economiza energia elétrica, já que com os equipamentos centralizados é possível planejar melhor a refrigeração e outros aspectos de alto consumo energético. Faz uma utilização mais eficiente dos servidores, sendo que é possível utilizar um mesmo servidor para serviços distintos.

Existe várias formas de consolidar servidores, segundo Da Silva (2017, p.7):

A consolidação de servidores pode ser realizada de diversas formas, como combinando aplicações em um único servidor com uma instância de sistema operacional rodando; criando grupos (também conhecidos como *clusters*) de servidores, que atuem como uma única aplicação e necessitem apenas de um único servidor de *backup*; virtualizando servidores, conforme visto no item 1 e diminuindo o portfólio de aplicações, removendo serviços redundantes.

As vantagens de consolidar servidores são as mesmas encontradas na virtualização e também na desativação de servidores. Outra questão para completar o contexto da citação é simplificar a complexidade dos servidores, possibilitando a facilidade da padronização dos servidores, diminuindo gastos com manutenções e aumentando o tempo total de disponibilidade dos serviços.

Desativação de servidores inativos

Normalmente o equipamento de TI em relação a sua capacidade é muito pouco utilizado. Mesmo com o dispositivo inativo o equipamento ainda consome energia. Para que isso não ocorra, deve-se verificar e remover servidores sem finalidade, que não estão ativos.

Para Da Silva (2017, p.6) esta observação “envolve geralmente a desativação destes servidores, enquanto se aguarda alguma indisponibilidade referente a ele. Outra possibilidade é a migração da carga de trabalho para outro servidor, e posterior análise de como o serviço se comporta”.

Dessa forma permite a remoção dos servidores antigos e sem uso, diminuindo o consumo de energia elétrica e o desperdício de calor. Pode também ser feita uma migração para outro servidor.

Aquisição de servidores energeticamente eficientes

A obtenção de hardwares que possuem inovações tecnológicas, e que são energeticamente eficientes possibilita o mesmo desempenho do equipamento utilizando uma quantidade inferior de energia.

Da Silva acredita que “No caso de servidores, a utilização de fontes de energia mais eficientes, melhores reguladores de voltagem, processadores que consomem menos energia e *coolers* mais eficientes são as funcionalidades desejadas.” (DA SILVA, 2017, p.8)

De acordo com a citação acima esses servidores usam tecnologia apropriadas nos termos como, processadores que gastam menor energia, melhores reguladores de tensão, fontes de alimentação mais eficientes e ventiladores de resfriamento mais ativo.

Layout de área mais eficiente

Para que calor seja eliminado da forma mais eficiente nos servidores, é necessário a implantação do procedimento de posicionar os racks de maneira que produza corredores quentes e frios. “Nessa estratégia, organizam-se as fileiras de *racks* de tal modo que a parte da frente dos servidores fique virada umas para as outras, deixando a parte de trás dos servidores confinada” (DA SILVA, 2017, p.9). Essa disposição pode reduzir perdas de energia e prolongar a vida dos servidores.

Esse esquema de distribuição controla de modo eficiente o fluxo de ar minimizando gastos com climatização. Sendo assim para data center que possuir uma distribuição de modo diferente necessita de reajustes que pode tornar o projeto um pouco mais trabalhoso. Terão que ser feitas adequações nos cabos para os racks reposicionados, ajustes no sistema de climatização, custos para nova configuração na distribuição de energia elétrica e custos com mão de obra.

Confinamento de corredores quentes e frios

Os corredores frios enclausurados proporcionam a separação física do ar frio do ar quente por meio do confinamento do corredor frio. Já o corredor quente é separado fisicamente dos demais ambientes de produção.

“O confinamento de corredores, sejam eles quentes ou frios, refere-se a barreiras físicas usadas junto aos corredores quentes e frios para eliminar a mistura do ar frio insuflado com o ar quente” (DA SILVA, 2017, p.10). Desse modo as temperaturas são gerenciadas facilmente, e tem uma necessidade menor de climatização.

A finalidade do enclausuramento do corredor frio, e do corredor quente é a mesma. Otimizar a temperatura do ar de entrada para os servidores com redução de custos operacional e aumento na capacidade de refrigeração.

Segundo Da Silva (2017, p.10), “Em relação à *datacenters* que utilizam o *layout* de corredores quentes e frios, o confinamento pode resultar em uma economia de 5% a 10% em energia, e reduzir a energia de ventiladores e *chillers* em 20% a 25%.”

No entanto essa alteração pode ser executada tanto em corredores quentes como também em corredores frios, sem existir entre um e outro nenhuma diferença substancial. Ambos

manifestam benefícios e restrições, então melhor escolha é aproveitar ao máximo as vantagens dos dois e eliminar as desvantagens.

Dispositivos de gerenciamento de fluxo de ar

Aperfeiçoar o fluxo de eliminação de ar quente e a entrada de ar fresco nos ambientes embaixo dos pisos elevados do data center, mantendo a organização do espaço para que a estruturação do cabeamento não atrapalhe o funcionamento do sistema de refrigeração.

Segundo Villela (2011, p.5), gerenciamento do fluxo de ar pode ser feito da seguinte forma:

Desfazer bloqueios sob o piso para aproveitar da capacidade atual e implantar um gerenciamento efetivo de cabeamento. Assegurar que as aberturas no piso sejam compatíveis com a carga térmica do equipamento, adicionando ou removendo azulejos perfurados nas entradas de ar dos equipamentos. Considerar a inclusão de retornos por dutos.

Os itens acima mostram que utilizar racks para melhor controle do fluxo de ar, dispõem de corredores independentes de ar quente e frio e que possuem entradas de ar em diferentes direções.

Controle de temperatura e umidade

Deve ser instalado em todo data center uma série de sensores de temperatura, umidade, potência e pressão. Colocado para controlar a regulação de temperatura e umidade. Verificar se a unidade HVAC (aquecimento, ventilação e ar-condicionado) usa a temperatura de entrada do servidor, e não a temperatura do ar de retorno, como a temperatura do ponto de ajuste.

Nas palavras de Da Silva “Modificar os níveis atuais pelos quais um *datacenter* trabalha requer um controle rigoroso da temperatura de todos os pontos passíveis de *hot spots*. É necessário distribuir sensores por racks, piso elevado e na própria unidade CRAC” (DA SILVA, 2017, p.11).

Os CRAC (*Computer Room Air Conditioner*) possuem sensores de temperatura e umidade próximo a entrada de ar. Cada unidade do CRAC tem capacidade de esfriar e esquentar, e de umidificar e desumidificar. Onde a umidade é muito alta a água pode condensar dentro dos servidores causando curtos circuitos e onde a umidade é muito baixa pode gerar falhas de descarga eletrostática.

A verificação de indicadores temperatura e umidade na proporção a temperatura geral vai sendo elevada gradativamente. “A *Energy Star* recomenda que a diferença entre o servidor mais baixo no *rack* e o mais alto esteja entre 21°C e 25°C (ou seja, 4°C de diferença entre o ponto mais baixo e o mais alto).” (DA SILVA, 2017, p.11).

Geralmente as unidades do CRAC de data center utilizam muita energia para manter os níveis necessário de umidade então expandir o limite de umidade permite economizar energia.

Economizadores a ar e a água

O economizador a ar aproveita o ar externo do local e o distribui diretamente aos servidores. “[...]consiste no arrefecimento de um edifício por meio de ventilação pela introdução de ar exterior, utilizando a energia livre desse ar, desde que este apresenta uma temperatura inferior à do ar interior [...]” (TAVARES, 2012, p.1)

Integrado com o sistema central de climatização, o economizador a ar, filtra partículas contaminantes antes de remeter ao data center o ar externo. Mesmo em ambiente com climas mais quentes, essa solução pode ser eficiente pois consegue aproveitar as temperaturas do inverno ou noites mais frias.

O economizador integrado a água usa a capacidade de resfriamento evaporativo e podem pré-resfriar a água antes de chegar ao resfriador. Eles podem ser integrados ou não com o resfriador. Ao contrário do economizador de ar o economizador a água pode ser adaptado.

Para Da Silva (2017, p.12), “Economizadores a água podem ser uma ótima opção para *datacenters* com plantas de resfriamento por água ou ar. Sua torre de resfriamento gera água gelada, que servirá como substituta do *chiller* durante os meses de inverno”.

Assim disponibiliza, uma redundância ao *chiller*, reduzindo o risco de indisponibilidade no data center por consequência de ter ocorrido falha ou erro inesperado.

Aparelhos de iluminação com eficiência energética

Usar uma tecnologia de iluminação mais eficiente, luzes com baixo consumo de energia elétrica, desligar as luzes quando não houver ninguém no local ou de acordo com o horário, e para diminuir a necessidade de iluminação é recomendado o uso de cores claras. Nas palavras de Da Silva “[...] o uso de luzes com baixo uso de energia, como LED, o desligamento das luzes quando não houver ninguém na área do prédio, o uso de cores claras no ambiente para diminuir a necessidade de iluminação [...]” Da Silva (2017, p.17).

Sensores integrados e switches identificam movimento, temperatura e níveis de luz, notificando esses dados através de um software de fácil utilização, modificando os dados de consumo de energia em conhecimentos práticos, que constituem, através da sua API, outros sistemas de construção, reduzindo custos.

ANÁLISE DE RESULTADOS

Através dos resultados obtidos foi feita duas tabelas para mostrar de forma clara ao leitor sobre as práticas de TI verde aplicada ao data center. A tabela 1 mostra algumas práticas apresentadas no decorrer do texto. Mostrando a porcentagem de economia de energia elétrica ao ser implantada. Esses métodos são eficazes para data centers novos, mas outras já podem ser implantadas em data center já existente.

Tabela 1. Estratégias práticas para reduzir o consumo de energia nos data centers, indicando a porcentagem de economia de energia elétrica.

	Economia	Orientação	Limitações
Dimensionamento correto do DCPI	10 - 30%	Uso de uma arquitetura modular e dimensionável de alimentação e resfriamento A economia é maior para sistemas redundantes	Para novos projetos e algumas expansões
Virtualize os servidores	10 – 40%	Não é tecnicamente uma solução de infraestrutura física, mas tem um impacto radical Envolve a consolidação de aplicativos em menos servidores, normalmente servidores <i>blade</i> Também libera a capacidade de alimentação e refrigeração para a expansão	Requer grandes mudanças nos processos de TI Para conseguir economia em uma instalação existente, pode ser necessário desligar alguns equipamentos de alimentação e refrigeração
Arquitetura mais eficiente de sistemas de ar condicionado	7 – 15%	O resfriamento por fileira é mais eficiente em alta densidade (White Paper 130) Caminhos de ar mais curtos exigem menos potência de ventilador As temperaturas de alimentação e retorno do CRAC são mais elevadas, aumentando a eficiência, a capacidade e evitando a desumidificação, reduzindo, assim, os custos de umidificação.	Para novos projetos Os benefícios são limitados aos projetos de alta densidade
Modos de economia dos condicionadores de ar	4 –15%	Muitos condicionadores de ar oferecem opções de economizador Isso pode proporcionar uma economia substancial de	Para novos projetos Difícil de readaptar

		<p>energia, dependendo da localização geográfica</p> <p>Alguns data centers têm sistemas de ar condicionado com modos de economizador, mas a operação com o economizador fica desativada</p>	
Layout de área mais eficiente	5 –12%	<p>O layout da área afeta muito a eficiência do sistema de ar condicionado</p> <p>Envolve a distribuição corredor quente / corredor frio através de um posicionamento adequado do sistema de ar condicionado.</p>	<p>Para novos projetos</p> <p>Difícil de readaptar</p>
Equipamentos de energia mais eficientes	4–10%	<p>Os novos sistemas de <i>nobreak</i> de alta qualidade apresentam 70% menos perdas do que os nobreaks antigos com cargas típicas</p> <p>A eficiência com cargas leves é o principal parâmetro, NÃO a eficiência com carga plena</p> <p>Não se esqueça de que as perdas dos nobreaks devem ser resfriadas, o que dobra seus custos</p>	<p>Para novos projetos ou readaptações</p>
Posicione corretamente as placas de piso ventiladas		<p>Muitas placas ventiladas estão localizadas incorretamente ou em quantidade errada em grande número de data centers comuns</p> <p>Os locais corretos NÃO são intuitivamente óbvios</p> <p>Uma avaliação profissional pode garantir um resultado ideal</p> <p>Benefício adicional - redução de pontos de concentração de calor.</p>	<p>Somente para data centers que usam piso elevado</p> <p>Fácil, mas exige orientação especializada para alcançar o melhor resultado</p>

Coordenação de aparelhos de ar condicionado	0–10%	<p>Muitos data centers têm vários condicionadores de ar que realmente lutam entre si</p> <p>Um pode aquecer enquanto o outro resfria</p> <p>Um pode desumidificar enquanto o outro umidifica</p> <p>O resultado é um desperdício gigantesco</p> <p>O diagnóstico pode exigir uma avaliação profissional</p>	Em qualquer data center com mais de um condicionador de ar
Instale aparelhos de iluminação com eficiência energética	1 – 3%	<p>Desligue algumas ou todas as luzes de acordo com a hora do dia ou o movimento</p> <p>Use uma tecnologia de iluminação mais eficiente</p> <p>Não se esqueça que a potência de iluminação também deve ser resfriada, duplicando o custo</p> <p>O benefício é maior em data centers de baixa densidade ou parcialmente preenchidos</p>	A maioria dos data centers pode se beneficiar
Instale painéis de isolamento	1 – 2%	<p>Diminua a temperatura de entrada dos servidores</p> <p>Também economiza energia através do aumento da temperatura do ar de retorno do CRAC.</p> <p>Barato e fácil com os novos espelhos cegos encaixáveis, como os da Schneider Electric</p>	Para qualquer data center, antigo ou novo

Fonte: RAUSMUSSEN (2012, p.11)

A tabela acima se refere a algumas técnicas para redução de energia elétrica em data center. A economia referida na tabela está associada com o projeto e instalação, outras citadas podem ser incorporadas aos equipamentos fornecidos pelos fabricantes.

A análise de resultados foi obtida de forma hipotética, com base em informações de soluções já consolidadas e largamente utilizadas por várias empresas, métodos que podem ser implantados em novos datacenters ou podem ser adaptados a datacenters já existentes.

Virtualização de Servidores: Solução que tem como objetivo consolidar servidores, isto é, ser capaz de unir 5 ou 10 servidores físicos em um só com um ganho imediato de espaço físico, redução de energia elétrica e menos uso de ar condicionado. Ainda garantindo o benefício de ter um gerenciamento técnico centralizado e um custo de licenciamento e suporte técnico também reduzido.

Layout eficiente: Com a consolidação de servidores, é possível gerenciar melhor o ambiente de data center, criando corredores adequados de correntes de ar quente e frio melhorando a distribuição física e a acessibilidade aos servidores físicos.

Deve-se verificar e remover servidores sem finalidade, que não estão ativos. Pois desativar servidores sem uso diminui o desperdício de calor.

Adquirir hardwares que possuem novas tecnologias, e que são energeticamente eficientes possibilita o mesmo desempenho do equipamento, utilizando uma quantidade inferior de energia.

Além do layout eficiente é importante o confinamento de corredor quente e frio, já que ele proporciona a separação do ar quente do ar frio trazendo economia com ar condicionado. Dispositivos de gerenciamento do fluxo do ar, controle de temperatura e umidade, e também os economizadores a ar e a água fazem parte da economia de energia com refrigeração e climatização.

Os economizadores a ar aproveitam o ar de fora e o distribui diretamente sobre os servidores enquanto o economizador a água usa o resfriamento evaporativo que pré resfria a água antes dela chegar ao resfriador. Por fim aparelhos de iluminação eficiente usando lâmpadas de LED, e manter as luzes apagadas quando não houver movimento.

A tabela 2 faz uma comparação entre data center convencional e computação em nuvem. Nela mostra as principais diferenças: a tecnologia usada, o modo de funcionamento, segurança, serviço, desvantagem e custo.

Tabela 2. Comparando data center e computação em nuvem

	Data center	Computação em nuvem
Tecnologia	Segurança, disponibilidade, desempenho, gerenciamento e eficiência	Dinâmica, escalabilidade, flexibilidade, acessibilidade e maior distribuição dos arquivos.
Modo de funcionamento	Funciona como um servidor comum, ligado para que os dados estejam armazenados durante todo	Não necessita de um local físico. Os dados estão armazenados na nuvem, em diversos servidores espalhados.

	o dia e sejam acessados a qualquer momento.	
Segurança	<p>Pode ser considerado mais seguro, pois é mais simples escolher e dar acesso a somente algumas pessoas para chegar nos arquivos armazenados.</p> <p>A empresa é responsável pelos seus dados.</p>	<p>Apesar de ser mais acessível, permite que, por meio de conexões com a internet, as informações sejam localizadas de qualquer lugar.</p> <p>Os dados ficam em provedores de terceiros.</p>
Serviços	Não há qualquer tipo de compartilhamento de espaço com outra empresa — o servidor é dedicado (ou seja, exclusivo)	É mais flexível e possui poder de processamento maior, além de muito menos burocracias. Se for preciso alterar a memória para salvar mais informações, tudo é feito de maneira simplificada.
Desvantagem	A empresa tem capacidade de armazenamento limitada, não sendo possível alterar a quantidade de espaço sem que seja feita a compra e instalação de novos equipamentos.	Os recursos de armazenamento estarão sendo compartilhados na nuvem do provedor.
Custos	É necessário investir certas quantias para iniciar o processo de uso do serviço, pois serão instalados equipamentos novos, um espaço deve ser reservado para essas máquinas, além da equipe que vai administrar. Essa opção acaba por ficar mais cara devido à quantidade de fatores envolvidos, como softwares e gastos de energia.	Os provedores de <i>Cloud Computing</i> oferecerem diversos planos mais acessíveis, com funções que o empresário pode determinar como mais necessárias ou não conforme a realidade atual do negócio. Outro ponto é que esse serviço não depende da montagem de uma infraestrutura, novos equipamentos, ficando disponível para uso em espaços de tempo mais curtos.

Fonte: DIAS (2017)

A decisão de escolher por um data center local ou em nuvem vai depender do porte da empresa. Ter um data center interno pode representar gastos com gestões de acesso físico, energia, climatização e equipamentos.

A computação em nuvem pode ser paga mensalmente ou por hora de acordo com o uso. Qualquer empresa que deseja ter alto performance computacional sem investir em equipamentos e ambiente físico pode optar por computação em nuvem. Ambos possuem vantagens para as empresas então as organizações precisam levar em consideração o valor disponível para o investimento.

CONCLUSÕES

Este artigo proporcionou um estudo da aplicabilidade de práticas de TI Verde em data center, com o foco na economia de energia elétrica. Buscado desde o descarte correto de lixo eletrônico, utilização da virtualização de servidores e computação em nuvem. Deixando notável que as práticas de TI verde utilizadas em data centers, traz benefícios não somente para o meio ambiente, mas também para as empresas que as adotam.

Para eficiência energética não deve considerar apenas o fato de se produzir gastando menos. Mas que essa produção tenha um menor impacto ambiental. A principal contribuição dessa conduta é trazer um equilíbrio ambiental e também qualidade de vida para a sociedade.

A virtualização facilita muitas operações, estão entre as principais vantagens da virtualização a redução do espaço físico e redução dos gastos com energia e refrigeração, no entanto é necessário fazer uma avaliação se essa virtualização não trará contratempo em situação de falhas, como por exemplo a queima de um host físico fará todas as máquinas virtuais parar ou então se as máquinas virtuais terão desempenho necessário para operações críticas.

Tanto os corredores de ar frio quanto o quente apresentam benefícios e restrições. A melhor opção a ser adotada é apostar em eliminar as desvantagens e aproveitar ao máximo das vantagens de ambos.

As organizações que aderem recursos de TI verde terão economia desde a compra de equipamentos, na infraestrutura, de energia, manutenção e uma boa imagem. No entanto torna se necessário que os profissionais da área de TI se conscientizem sobre os impactos ambientais de seus data centers. Inserindo a questão de sustentabilidade na lista dos parâmetros.

Portanto o tempo todo surgem novidades que devem ser avaliadas e analisadas a sua aplicabilidade num cenário verde. As contribuições deste artigo não se restringem ao que foi escrito e fica aberto para que sejam analisadas pesquisas futuras. Sugere em trabalhos futuros o estudo de práticas de TI verde que não foram apresentadas neste artigo.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, A. F. *Ti Verde – Implementação de Práticas Sustentáveis em Empresa de Tecnologia da Informação.* 2012. Disponível em:

<<https://www.aedb.br/seget/arquivos/artigos12/25916208.pdf>>. Acesso em: 17 dez. 2018.

ALECRIM, Emerson. *O que é cloud computing (computação nas nuvens)?*, 2015. Disponível em: <<https://www.infowester.com/cloudcomputing.php>>. Acesso em: 11 set. 2018.

ALVES, D. M.; COSTA, M.; FURTADO, M. R. S.; MORAVIA, R. V. *Computação em nuvem: um estudo sobre seus conceitos, tecnologia e aplicação*. Disponível em: <http://revistapensar.com.br/tecnologia/pasta_upload/artigos/a57.pdf>. Acesso em: 01 ago. 2018.

AQUINO, A. R.; PALETTA, F. C.; CAMELLO, T. C. F.; MARTINS, T. P.; ALMEIDA, J. R.; SILVA, A. C. M.; GALVÃO, V.; BORDON, I. C. A.; RIBEIRO, L. S. S.; STECHER, L.; DE OLIVEIRA, M. J. A.; DA SILVA, M. V.; DOS SANTOS, R. M.; MATTIOLO, S. R.; SCAGLIUSI, S. R.; GARCIA, V. S. G. *Sustentabilidade Ambiental*. 1. ed. Rio de Janeiro: Rede Sirius; OUERJ, 2015. Disponível em: <http://www.rsirius.uerj.br/pdfs/sustentabilidade_ambiental.pdf>. Acesso em: 10 set. 2018.

BANDEIRA, A. K. R.; MONTEIRO, A. J. F.; ALVES, J. B.; DE OLIVEIRA, J. M. S.; GUIMARÃES, O. M. *Fundamentos de data center*. 2017. Disponível em: <<https://pt.slideshare.net/Mauryliosilva/fundamentos-de-data-center-79294047>>. Acesso em 17 dez.

BUSSE, B. N. *Textos acadêmicos sobre eficiência energética: uma amostra quantitativa dos últimos 40 anos de pesquisa*. 2010. Disponível em: <<https://www.ipog.edu.br/download-arquivo-site.sp?arquivo=textos-academicos-sobre-eficiencia-energetica-1644617.pdf>>. Acesso em: 17 dez.2018.

CARRERO, M. P.; LOPES, N. S. *Boas práticas da ti verde adotadas pelas empresas como forma de uso eficiente dos recursos energéticos*. Programa de Apoio à Iniciação Científica - PAIC 2015-2016. Disponível em: <<https://www.google.com/search?q=praticas+verde+em+data+center&ei=O3saXOmCEcuAwgTnppeQDQ&start=40&sa=N&ved=0ahUKEwipjPWWsazfAhVLgJAKHWfTBdI4HhDy0wMItQE&biw=1385&bih=652#>>>. Acesso em: 17 dez. 2018.

CUNHA, J. F. S. R. *Computação em nuvem e considerações contratuais*. 2012

Disponível em: <<http://www.publicadireito.com.br/artigos/?cod=e6c46d6ccf72dec9>>. Acesso em: 10 set. 2018.

DA SILVA, L. B.; BOMFIM, L. H. da S.; JUNIOR, G. L.; DE OLIVEIRA, M. N. *TI verde: uma proposta de economia energética usando virtualização*. *Interfaces Científicas - Exatas e Tecnológicas* • Aracaju • V.1 • N.2 • p. 57 - 74 • Jun. 2015. Disponível em: <<https://periodicos.set.edu.br/index.php/exatas/article/download/2255/1294>>. Acesso em: 11 ago. 2018.

DA SILVA, M.K. *Soluções de TI verde e eficiência energética para datacenters*. Disponível em :< <https://www.riuni.unisul.br/handle/12345/3022> >. Acesso em: 01 ago. 2018.

DA SILVA, Newton Rocha; JUNIOR, Flavio Hourneaux. *Ti verde: sustentabilidade por meio da virtualização de servidores*. IPTEC - Revista Inovação, Projetos e Tecnologias, São Paulo, v. 1, n. 1, p. 30-42, jan./dez. 2013. Disponível em:

<<http://www6.uninove.br/ojs/journaliji/index.php/iptec/article/view/3/pdf>>. Acesso em: 18 set. 2018.

DIAS, L. S.; MARQUES, M. D. *Responsabilidade social das organizações empresariais diante do desenvolvimento sustentável*. IX Fórum Ambiental da Alta Paulista, v. 9, n. 10, 2013, p. 67-75. Disponível em: <http://www.amigosdanatureza.org.br/publicacoes/index.php/forum_ambiental/article/viewFile/564/732>. Acesso em: 07 set. 2018.

DIAS, Luiz. *Data Center ou Computação em Nuvem? Entenda cada tipo*. 2017 Disponível em: <<http://www.administradores.com.br/artigos/empreendedorismo/data-center-ou-computacao-em-nuvem-entenda-cada-tipo/104706/>> Acesso em: 18 set. 2018

DOS SANTOS; Élcio Henrique; DA SILVA, Mirela Auxiliadora. *Sustentabilidade empresarial: um novo modelo de negócio*. Revista Ciência Contemporânea jun./dez. 2017, v.2, n.1, p. 75 – 94. Disponível em: <<http://uniesp.edu.br/sites/biblioteca/revistas/20180301124814.pdf>>. Acesso em: 18 set. 2018.

DOS SANTOS, F. U. *Benefícios da virtualização de servidores*. 2014. Disponível em: <<https://www.bluesolutions.com.br/2014/07/os-12-beneficios-da-virtualizacao-de-servidores-no-datacenter/>>. Acesso em: 10 set. 2018.

FROEHLICH, Cristiane. *Sustentabilidade: Dimensões e métodos de mensuração de resultados*. DESENVOLVE: Revista de Gestão do Unilasalle (ISSN 2316-5537), Canoas, v. 3, n. 2, set. 2014. Disponível em: <<https://revistas.unilasalle.edu.br/index.php/desenvolve/article/viewFile/1316/1182>>. Acesso em: 10 set. 2018.

GIANELLI, Álvaro. *Práticas sustentáveis em ti verde no instituto federal de São Paulo: limites e possibilidades*. 2016. Disponível em:

<<http://www.fae.br/mestrado/dissertacoes/2016/PR%C3%81TICAS%20SUSTENT%C3%81VEIS%20EM%20TI%20VERDE%20NO%20INSTITUTO%20FEDERAL%20DE%20S%C3%83O%20PAULO%20limites%20e%20possibilidades.pdf>>. Acesso em 17 dez. 2018

LUNARDI, G. L.; ALVES, A. P. F.; SALLES, A. C. *TI Verde e seu Impacto na Sustentabilidade Ambiental*. Disponível em:

<http://www.repositorio.furg.br/bitstream/handle/1/5381/2012_ADI1891.pdf?sequence=1>. Acesso em 07 set. 2018.

[NUNES, A. C. P.; DAS CHAGAS, A. C.; CAMILO, A.; DOS SANTOS, N. M. A TI Verde na Sociedade Atual](https://docplayer.com.br/5306016-A-ti-verde-na-sociedade-atual.html). 2011. Disponível em: <<https://docplayer.com.br/5306016-A-ti-verde-na-sociedade-atual.html>>. Acesso em: 10 set. 2018.

PEREIRA, Glauber Ruan Barbosa. *Práticas da ti verde que contribuem para o desenvolvimento sustentável: um estudo de caso em indústrias do RN*. 2009. 126 f. Dissertação (Mestrado em Políticas e Gestão Públicas; Gestão Organizacional) - Universidade Federal do Rio Grande do Norte, Natal, 2009. Disponível em: <https://repositorio.ufrn.br/jspui/bitstream/123456789/12129/1/Pr%C3%A1ticasTIVerde_Pereira_2009.pdf>. Acesso em 07 set. 2018.

PINTO, Thays Mayara da Costa; SAVOINE, Márcia Maria. *Estudo sobre ti verde e sua aplicabilidade em Araguaína*. Revista científica do ITPAC. Volume 4. Número 2. Abril de 2011. Publicação 3. Disponível em: <[HTTPS://ASSETS.ITPAC.BR/ARQUIVOS/REVISTA/42/3.PDF](https://assets.itpac.br/arquivos/revista/42/3.pdf)>

RASMUSSEN, NEIL. *Implantação de data centers com eficiência energética*. 2012. disponível em: <http://www.smartups.com.br/white_papers/NRAN-6LXSHX_R1_BR.pdf>. acesso em 28 set. 2018.

SARTORI, S.; LATRÔNICO, F.; CAMPOS, L. M.S. *Sustentabilidade e desenvolvimento sustentável: uma taxonomia no campo da literatura*. Ambiente & Sociedade. São Paulo v. XVII, n. 1 n p. 1-22 n jan.-mar. 2014. Disponível em: <<http://www.scielo.br/pdf/asoc/v17n1/v17n1a02.pdf>>. Acesso em: 10 set. 2018.

SCHULZ, M. A.; SILVA, T. N. *Ti verde e eficiência energética em data centers*. Revista de gestão social e ambiental – RGAA, SÃO PAULO, V.6, N.2, P.121 – 133, MAIO/AGO.2012.

SOBRAL, F. A.; JUNIOR, F. H. *Data center verde: implantação de um sistema de refrigeração em uma instituição de ensino superior*. Disponível em:

<<http://repositorio.uninove.br/xmlui/bitstream/handle/123456789/963>>. acesso em: 01 ago. 2018.

SOLDATI, M. K. *TI verde – virtualização de servidores*. Fasci-Tech – Periódico Eletrônico da FATEC-São Caetano do Sul, São Caetano do Sul, v. 1, n. 5, Out/Dez 2011, p. 97 a 104.

SOUSA, Flávio. *Computação em Nuvem: Conceitos, Tecnologias, Aplicações e*

Desafios. III Escola Regional de Computação Ceará, Maranhão e Piauí –

ERCEMAPI. 1 ed. Piauí: SBC, 2009, v. 1, p. 150-175.

TAVARES, C. M. *Impacto da Influência do Ar Novo na Climatização de Edifícios*. 2012. Disponível em: <http://recipp.ipp.pt/bitstream/10400.22/2619/1/DM_CristovaoTavares_2012_MES.pdf>. Acesso em: > Acesso em 17 dez. 2018.

VILLELA, Vitor Souza. *Eficiência Energética em data centers*. Disponível em: <<https://dcomp.sor.ufscar.br/verdi/topicosCloud/EficienciaEnergetica.pdf>>. Acesso em 10 set. 2018.

VITA, J. B.; ENGELMANN, W. *Direito, economia e desenvolvimento sustentável II*. Disponível em:

<<https://www.conpedi.org.br/publicacoes/02q8agmu/y9agq5n5/PAS72Up0fy364A49.pdf>>.

Acesso em 16 dez. 2018.

ZUCCHI, Wagner Luiz; AMÂNCIO, Anderson Barreto. *Construindo um Datacenter*. Revista USP, São Paulo, n. 97 • P. 43-58 • Março/Abril/Maio 2013. Disponível em: <<http://www.revistas.usp.br/revusp/article/view/61684/64573>>. Acesso em: 10 set. 2018.

CLOUDERA: UMA ABORDAGEM PARA ANÁLISE DE LOGS DO PROXY SQUID-CACHE USANDO HADOOP, FLUME, MAPREDUCE E IMPALA

*CLOUDERA: AN APPROACH FOR PROXY SQUID-CACHE LOGS ANALYSIS
USING HADOOP, FLUME, MAPREDUCE AND IMPALA*

Volnei Cervi PUTTINI

vcputtini@gmail.com, Pós-Graduação Lato Sensu em BI em Bigdata, UniAnchieta

Juliano SCHIMIGUEL

Professor do Centro Universitário Padre Anchieta, Jundiaí/SP, schimiguel@gmail.com

Resumo

Nosso objetivo é demonstrar uma abordagem no contexto do Big Data a qual permita a coleta, armazenagem persistente, tratamento e geração de resultados para análises dos dados mediante o uso das ferramentas disponíveis na plataforma Cloudera. Os dados utilizados são os registros de log gerados em tempo real pelo servidor proxy Squid-cache como resultado dos acessos à Internet feitos a partir da rede local. Usando o Apache Hadoop para o armazenamento de massa e processamento MapReduce quando necessário, conversão dos dados armazenados no HDFS para tabelas Impala para criação de rotinas de manipulação dos dados usando a linguagem SQL, e o Apache Flume para coleta de transmissão dos dados entre a fonte e o Apache Hadoop.

Palavras-chaves: Big Data; Apache Hadoop; Apache Flume; Cloudera. Log.

Abstract

Our goal is to demonstrate an approach in the context of Big Data that enables collection, persistent storage, processing and results generation for data analysis using the tools available on the Cloudera platform. The data used are real-time logs generated by the Squid-cache proxy server, as a result of Internet access from the local network. Using Apache Hadoop for mass storage and MapReduce processing when necessary, converting HDFS-stored data to Impala tables for creating data manipulation routines using the SQL language and Apache Flume to collect data transmission between source and Apache Hadoop.

Key-words: Big Data; Apache Hadoop; Apache Flume.; Cloudera; Log.

INTRODUÇÃO

Historicamente sempre foi um grande desafio efetuar análises consistentes em dados de log devido a fatores como: possuem formatos de dados complexos e proprietários, são comumente armazenados no mesmo host da instalação do sistema, estão em formato ASCII (texto puro), por vezes binários porém este formato está fora do nosso escopo, são semiestruturados, são gerados em tempo real ou muito próximo disso e seus arquivos podem conter milhões de eventos registrados, a segurança desses dados deve ser considerada sempre como uma prioridade. Dado este cenário analisar manualmente é tarefa impossível, e criar-se programas capazes de automatizar este processo também torna-se complexo pois algumas necessidades devem ser sanadas tais como a interpretação dos dados registrados por diferentes sistemas simultaneamente, o armazenamento deve ser centralizado e persistente a fim de podermos ter acesso a todo o histórico a qualquer momento, além de se criar rotinas padronizadas capazes de manipular os registros e deles extrair as informações necessárias de modo consistente e principalmente coletar os dados em tempo real e transferi-los a um sistema capaz de armazená-los permitindo processamento centralizado e ágil. A segurança destes dados também é preocupação importante pois se tratam de informações sensíveis podendo ser utilizadas inclusive como prova em ações judiciais.

Analisando tais características constatamos que nossas necessidades podem ser suportadas pelas tecnologias quem vem sendo consolidadas no âmbito da ciência do Big Data. O Big Data se estabeleceu devido a alguns fatores primários os quais remetem as variedades nos formatos dos dados e como eles devem ser armazenados, na velocidade na qual esses dados são gerados e seu imenso volume. Além destas variáveis (TAURION, 2013) complementa com "[...] veracidade dos dados (os dados tem significado ou são sujeira?) e valor para o negócio. Outra questão que começa a ser debatida é a privacidade, tema bastante complexo e controverso.", aqui ele nos proporciona um vislumbre das complexidades em se distinguir dados úteis dos não úteis além de toda a problemática envolvendo as questões sobre privacidade. Quando colocamos o gerenciamento de log sobre esta perspectiva, se torna importante que dediquemos nossos esforços na criação de soluções para este fim, conforme demonstrados nos trabalhos correlatos abaixo citados e comentados.

(TEOH et al, 2004) apresentam uma abordagem para inspeção de logs de mediante o desenvolvimento de uma aplicação própria. Apesar de não fazem uso de técnicas de Big Data a aplicação é capaz de análises estatísticas e visualizações para auxiliar na tarefa dos administradores, tal qual é um dos nossos objetivos a serem alcançados.

Conforme proposto por (JAKUB e BRANISOVÁ, 2015, tradução nossa) em "Detecção de anomalias de arquivos de log usando técnicas de mineração de dados", utilizam de técnicas avançadas "para detecção de anomalias em arquivos de log, baseado em técnicas de mineração de dados para criação de regras dinâmicas. Para suportar o processamento paralelo, utilizamos o framework Apache Hadoop, fornecendo armazenamento distribuído e processamento distribuído de dados.", o foco do trabalho é mais amplo quando pretende tratar de vários sistemas diferentes e não somente de um específico como no nosso caso, porém também fazem uso do Apache Hadoop e

MapReduce, mas não informando sobre os meios utilizados para coleta dos dados. Nós iremos fazer uso do Apache Flume para coletar os dados nos servidores proxy Squid e transmiti-los para o Apache Hadoop aonde eles serão armazenados.

Ao unir as técnicas de análise e processamento massivo de dados proporcionada pela ciência do Big Data aplicada ao campo do Gerenciamento de Log, poderemos desenvolver soluções capazes de garantir tanto as necessidades dos administradores de sistemas bem como dos requisitos de segurança da informação.

SEGURANÇA DA INFORMAÇÃO E O GERENCIAMENTO DOS LOGS

A Segurança da Informação conceitua as bases para podermos melhor entender o por que que devemos dar as informações a importância necessária.

A informação é um ativo que, como qualquer outro ativo importante, é essencial para os negócios de uma organização e conseqüentemente necessita ser adequadamente protegida (ABNT NBR ISO/EIC 27002, 2005).

Ainda há as seguintes definições que devem ser citadas:

1. Preservação da confidencialidade, integridade e disponibilidade da informação; adicionalmente, outras propriedades, tais como autenticidade, responsabilidade, não repúdio e confiabilidade, podem também estar envolvidas (ABNT NBR ISO/EIC 27002, 2005).
2. Podemos definir a Segurança da Informação como uma área do conhecimento dedicada à proteção de ativos da informação contra acessos não autorizados, alterações indevidas ou sua indisponibilidade. (SÊMOLA,2003,p.43).

Conforme (ABNT NBR ISO/EIC 27002, 2005), um ativo pode ser "qualquer coisa que tenha valor para organização". Desse modo dentro do contexto deste trabalho o registro dos acessos à Internet é um ativo a ser monitorado, controlado e mantido, possivelmente indefinidamente, a fim que as organizações possam através dessas informações gerir estas atividades que tem como potencial impactar negativamente nas atividades internas das empresas, causando prejuízos das mais diversas ordens.

Conforme explicado por Lyra (2008), "Quando falamos em segurança da informação, estamos nos referindo a tomar ações para garantir a confidencialidade, integridade, disponibilidade e demais aspectos da segurança das informações dentro das necessidades do cliente."

Dadas estas características das informações obtidas mediante as análises dos registros de log, devem ser geridas sob a mesma ótica, desse modo torna-se indispensável que o gerenciamento dos logs tenham estabelecidas políticas próprias.

Segundo (Kent e Souppaya, 2006, tradução nossa), "Um log é a gravação dos eventos que ocorrem dentro dos sistemas e das redes das organizações. Os logs são compostos pelas entradas de logs; cada entrada contém informações relacionadas a um

evento específico que ocorreu dentro de um sistema ou da rede. Muitos logs contém registros relacionados a segurança dos computadores.[...]”, definem perfeitamente o que é o log e a sua importância no desenvolvimento das políticas de segurança as quais impactam diretamente na qualidade, segurança e confiabilidade dos ativos das empresas, daí a necessidade em se auditar os logs constantemente.

Porém auditar logs é uma tarefa considerada tediosa e muito trabalhosa, mas de modo algum deve ser negligenciada pelos administradores de sistemas. Em trabalho correlato, (SIMON et al, 2008, tradução nossa) demonstram que "auditoria procura identificar e evitar ações suspeitas e fraudulentas por parte do usuário, coletando dados sobre suas atividades no banco de dados.", os autores nos fazem compreender a importância da auditoria quando falam que "As informações coletadas são analisadas a fim de descobrir problemas de segurança e sua origem. A necessidade de identificar quais foram as ações e quais os padrões suspeitos são importantes requisitos para a segurança do sistema.", os autores focam entretanto seu trabalho na auditoria visando a segurança dos dados armazenados nos Sistemas Gerenciadores de Banco de Dados (SGBD) propondo a utilização de Sistemas Gerenciadores de Stream de Dados (SGSD) para enviar os dados dos registros de log para o sistema de destino a fim de serem processados e analisados. É interessante notar que neste artigo não são usados termos comuns ou ferramentas do Big Data porém são claras as semelhanças quando descrevem que "[...] o processamento de fluxos de dados (*streams*) é realizado através de máquinas de processamento de *streams* [...]" a qual é usada para transmitir dos dados coletados para o destino final, e também que "A integração desses dois métodos, auditoria e processamento de *streams*, pode garantir que os registros de log sejam processados em tempo real [...]" corroborando nossa ideia sobre a importância em se criar soluções destinadas à este modelo de processamento.

Assim o registro dos acessos à Internet é um ativo a ser monitorado, controlado e mantido, possivelmente indefinidamente, a fim que as organizações possam através dessas informações gerir estas atividades que tem como potencial impactar negativamente em suas atividades, causando prejuízos das mais diversas ordens. Tais conceitos convergem sempre para a tríade denominada de CID, ou seja, Confidencialidade, Integridade e Disponibilidade, sendo este paradigma os pilares fundamentais da Segurança de Informação.

TECNOLOGIAS ENVOLVIDAS

PROXY SQUID: FONTE DOS DADOS

O servidor Squid atua como *proxy*, sendo um intermediário entre a rede local e a Internet, provendo diversas funcionalidades as quais nos permitem controlar os acessos, otimizar a utilização da banda e registrar em tempo real todas as requisições efetuadas pelos usuários da rede em um arquivo de log.

PLATAFORMA OPERACIONAL CLOUDERA

O Cloudera é uma plataforma de software open-source baseada em GNU-Linux, o qual é mantido pela empresa que leva o mesmo nome, conforme pode ser lido no site oficial, "Cloudera Inc., é uma empresa de software sediada nos EUA que fornece uma plataforma de software para engenharia de dados, armazenamento de dados, aprendizado de máquina e análise que é executada na nuvem ou no local. (CLOUDERA, 2018)."

Assim Cloudera tornou-se uma plataforma operacional de grande valia para aqueles que desejam iniciar estudos em Big Data ou aqueles que possuem pretensões mais profissionais. O Cloudera permite com sua distribuição denominada Quickstart, composta por um ambiente totalmente integrado baseado na distribuição Linux CentOS no modelo de VM, permite a instalação simples e rápida para testes e estudos, podendo também ser ampliada para ambientes mais complexos caso necessário. Porém dadas as necessidades de hardware poderoso e armazenamento de enormes capacidades, para o uso em escala realmente profissional, talvez seja mais conveniente por optar por instalações do Cloudera em sistemas de terceiros com o dimensionamento adequado para grandes escalas.

Apache Hadoop

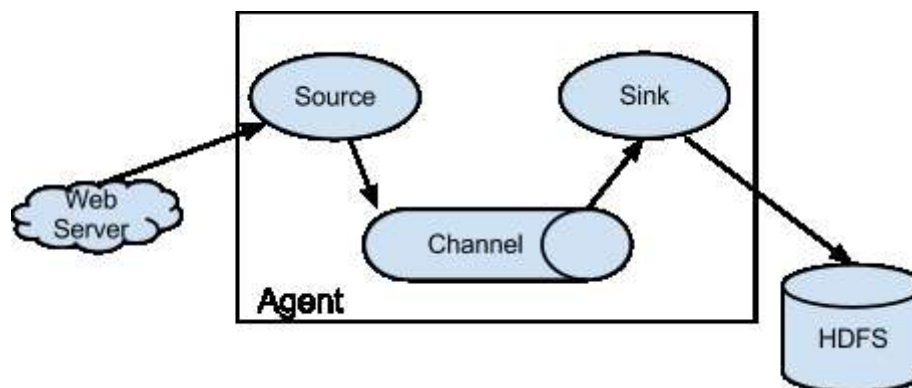
O Hadoop é uma plataforma open-source destinada a armazenamento massivo de dados de qualquer tipo, além de permitir o processamento desses dados em clusters. Suas principais características são: o sistema de arquivos HDFS - Hadoop File System que permite o armazenamento distribuído por milhares de sistemas interligados via rede, além da sua arquitetura MapReduce destinada ao processamento paralelo das aplicações, abstraindo toda a complexidade em se mapear, organizar e retornar os dados armazenados no cluster.

Apache Flume

Conforme descrito no site oficial do projeto "O Flume é um serviço distribuído, confiável e disponível para coletar, agregar e mover com eficiência grandes quantidades de dados de log. Tem uma arquitetura simples e flexível baseada em fluxos de dados de fluxo contínuo. Ele é robusto e tolerante a falhas, com mecanismos de confiabilidade ajustáveis e muitos mecanismos de failover e recuperação. Ele usa um modelo de dados extensível simples que permite a aplicação analítica on-line." Em suma o Flume é um coletor universal de dados, que por meio de suas configurações é capaz de ler os arquivos de log em tempo de execução e transferir estes dados pela rede até o nó central aonde ele deverá ser armazenado.

O Flume é capaz mediante sua estrutura de agentes de capturar em lotes ou tempo real os dados oriundos de registros de logs de sistemas dos mais variados tipos.

Figura 1: Modelo de processamento de *data stream*.



Fonte: <https://flume.apache.org/>

A Figura 1 mostra como é estruturado o software denominado agente, que é o responsável por ler, processar e transmitir os dados entre a origem e o seu destino.

- Sources: Recuperam os dados e enviam para os canais.
- Channels: Os canais mantêm as filas de dados e os transmitem entre fontes e coletores, atuando também na sincronização do fluxo.
- Sinks: Os sinks (coletores) processam os dados obtidos dos canais e os entregam a um destino, no nosso caso o HDFS.
-

Para que o agente possa operar ele deve possuir ao menos um grupo completo dos componentes. O agente é executado em sua própria instância Java VM.

Apache Impala

O Impala é um sistema capaz de efetuar consultas via instruções SQL aos dados armazenados no HDFS ou no Apache HBase. Sua arquitetura não utiliza os mecanismos do MapReduce, pois acessa diretamente os dados por meio de um mecanismo de consulta distribuído próprio, resultando em enorme ganho de desempenho, sendo inclusive superior ao Hive.

As tabelas são os principais depósitos dos dados do Impala. Apesar de seus esquemas de linha e colunas serem parecidos aos dos SGDB tradicionais, possui recursos diferenciados, tais como particionamento normalmente associados aos sistemas de data warehouse de ponta.

A estrutura lógica das tabelas se baseia nas definições das suas colunas, partições e demais características. Já a estrutura física é baseada na estrutura de diretórios do HDFS, sendo que cada tabela tem seus dados armazenados em arquivos dentro desta estrutura. (IMPALA, 2018).

HUE

Hue é uma interface visual entre as quantidades enormes de dados dos data warehouses e as ferramentas de BI/Machine Learning. Ele é destinado ao desenvolvedor de aplicativos de dados permitindo que ele possa iniciar seus projetos rapidamente.

Em resumo o Hue é um IDE, que permite tanto aos programadores bem como aos cientistas/analistas acessarem seus dados de forma simples, como por exemplo digitar no editor de textos instruções SQL para serem então executadas pelo Apache Impala, por exemplo. (HUE, 2018)

IMPLEMENTAÇÃO

Implementar localmente uma infraestrutura de hardware para suportar processamento massivo é algo que normalmente pode não ser viável para as empresas, pois como é sabido o Hadoop necessita a partir de dezenas de nós para que suas qualidades possam ser utilizadas plenamente. Não obstante, isso também é verdade quanto a capacidade de armazenamento dos dados, que como já sabemos é gigantesca. Sendo assim o ambiente implementado é suficiente apenas para efetuar os testes demonstrados nesse trabalho.

· Hardware Físico

– CPU: Intel Xeon E3-1240

– RAM: 32GB Totais

· Appliance Virtual17

– VM: Cloudera 5.13.0 QuickStart

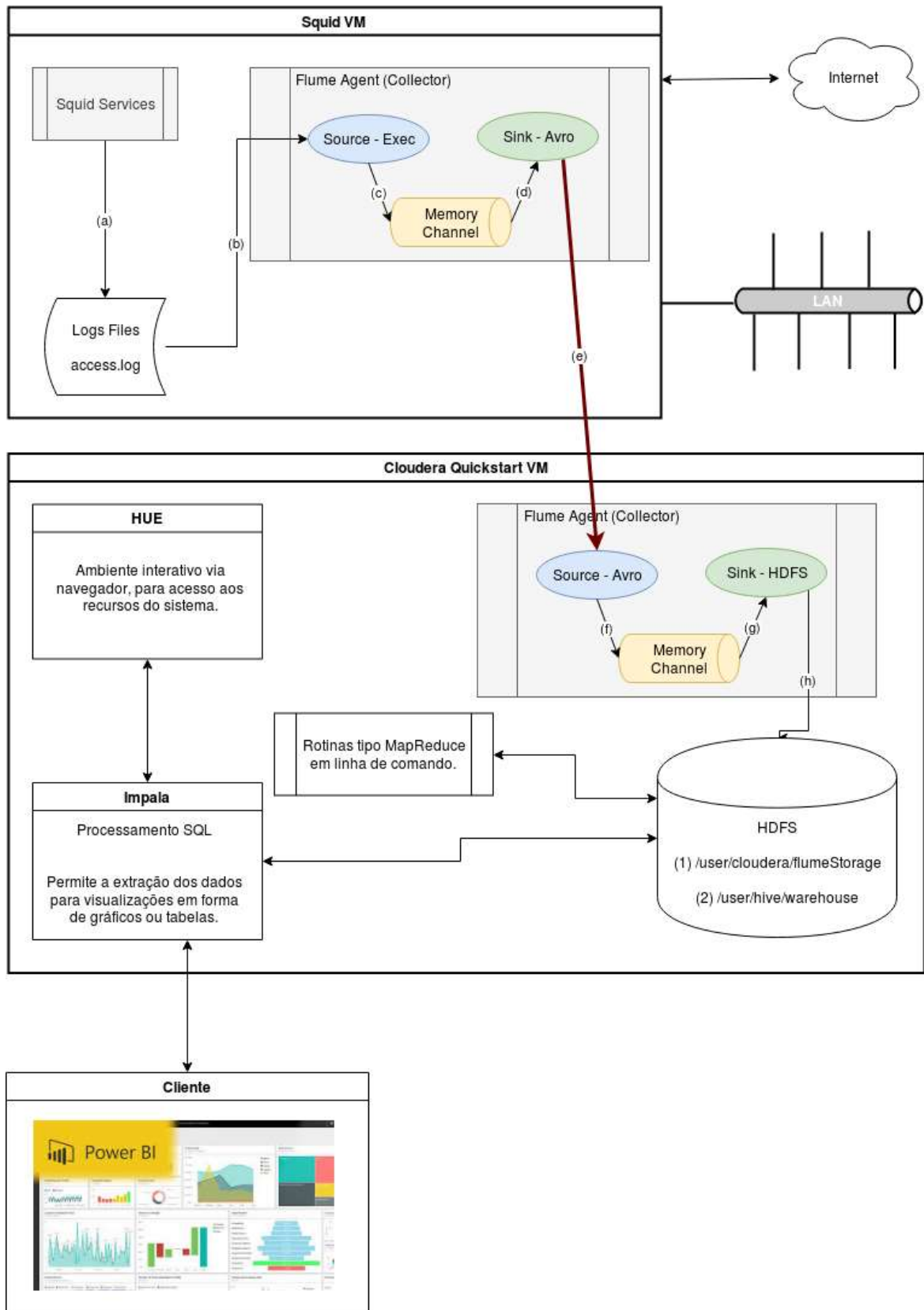
– Tamanho do VHD18: 55GB

– CPU Virtual: 04

– RAM Virtual: 10GB

– Hadoop Cluster Nodes: 01

Figura 2 - Fluxo dos dados pelos sistemas até ser armazenado para processamento.



Descrição do fluxo dos dados

- (a) O proxy Squid intercepta o tráfego de Internet e grava no arquivo de log, por padrão access.log, todos os registros.
- (b) O agente Flume usa como fonte para ingestão os dados gravados no arquivo access.log por meio do source tipo exec, que irá executar o comando tail -F access.log e enviar a saída para o canal (c).
- (c) O canal, armazena os dados e os envia ao coletor (d).
- (d) O coletor (sink) irá formatar os dados para o protocolo escolhido, no caso Avro, e mediante o uso de uma porta de comunicação enviará os datagramas para o source do host de destino (e).
- (e) Este agente está sendo executado dentro do ambiente Cloudera. Seu source foi configurado com o tipo Avro e atribuídos o endereço IP deste host e a porta de comunicação.
- (f) O canal de memória foi configurado do mesmo modo que o usado no item (c), pois é necessário que ambos tenham capacidades semelhantes.
- (g) o coletor (sink) foi configurado como tipo HDFS, pois irá armazenar os dados recebidos do canal diretamente no sistema Hadoop HDFS..

Áreas para armazenamento - HDFS

Armazenamento do dados coletados:

Esta área é destinada à gravação dos dados coletados o servidor proxy Squid. Os dados recebidos são gravados conforme a figura abaixo:

Figura 3. Estrutura em árvore dos diretórios

```
/user/cloudera/flumeStorage
├── 2019 [ano atual]
│   ├── 01 [número do mês: 01 a 12]
│   │   ├── 01 [número do dia: 01 a 31]
│   │   │   ├── FlumeData.1547474603681.avro
│   │   │   └── FlumeData.1547478228760.avro
│   │   └── 12
│   │       ├── 01 [número do dia: 01 a 31]
│   │       │   ├── FlumeData.1547474603450.avro
│   │       │   └── FlumeData.1547478228597.avro
```

Esta organização tem o objetivo de tornar o acesso aos dados o mais simples e rápido possível. No topo da árvore está o diretório que recebe como nome o ano atual, no nível abaixo temos o diretório que recebe o nome do número do mês corrente e abaixo dele os dias do mês. Dentro do diretório dia do mês são gravados os dados coletados em arquivos que representam 01(uma) hora de logs, sendo assim teremos a cada dia 24(vinte e quatro) arquivos.

Podemos então notar o quanto esta estrutura simplifica a localização dos dados e como os arquivos gerados são menores sua leitura é muito rápida. Para simplificar os processos de identificação dos arquivos seu nome é assim composto:

- Prefixo que identifica o coletor: FlumeData.
- Data de criação: No formato Unix timestamp, a qual é o número de segundos que se passaram desde 00:00:00 de quinta-feira, 1 de janeiro de 1970 até a data da criação do arquivo.
- Sufixo: Identifica o Source, neste caso o Avro.

O conteúdo desses arquivos são as linhas de log separadas por espaços em branco.

Exemplo:

```
1546553031.980    30120    192.168.1.20    TCP    MISS/200    668    GET
https://tr2.terra.com/broadcast/sub/
ch=pt-BR.BreakingNewsm=529s=M/ch=pt-BR.NewStoriesm=251495s=
M0.0460431667643667051546552990823? - ORIGINAL DST/208.84.244.40 appli-
cation/json
```

Armazenamento definitivo dos dados:

Os dados coletados serão movidos para outro local, o qual será o banco de dados que irá permitir tanto a persistência quanto a possibilidade de acesso para execução das análises necessárias.

Figura 4. Local padrão de armazenagem no ambiente Cloudera

```
/user/hive/warehouse
├─ squidlogs.db [Nome do banco de dados]
│   └─ accesslog [Nome do tabela]
│       └─ logdate='yyyy-mm-dd' [Partições nomeadas com datas]
│           └─ 474571a55b9088db-edee7a0800000000_1636462698_data.0.parq
```

A parte mais relevante na estrutura acima são as partições, o Impala permite criar uma tabela particionada por um atributo definido pelo usuário, no nosso caso o atributo logdate, o qual nada mais é que a data dos logs registrados. Desse modo são criadas pastas e dentro delas são gravados apenas os registros referentes à aquela data. Isto permite uma busca mais otimizada pelos algoritmos do Impala.

Tratamento dos dados

Estrutura do Log

Os registros do log são gerados uma linha por vez e representam uma ação que ocorreu, por exemplo, a tentativa de acesso a um Web site que foi negada pelo proxy Squid. O formato do registro que estamos utilizando é o padrão do proxy Squid que é assim composto:

- Não possui cabeçalho informando os nomes dos campos ou quaisquer outras das suas características, ou seja não possui esquema.
- Os campos são separados por espaços em branco, ASCII (32₁₀).
- O registro possui 10(dez) campos distintos os quais são identificados apenas pela sua posição na linha.

Tabela 1. Descrição dos campos registrados pelo Squid no arquivo access.log.

Posição	Nome Atributo	Função
0	unixts	Data e hora da ocorrência da ação. Formato Unix timestamp.
1	timems	Tempo gasto nesta requisição, dado em milissegundos.
2	host	IP ou nome do host que fez a requisição ao cache.
3	action	A ação tomada pelo cache. Exemplo: TCP_HIT ou TCP_MISS.
4	lenreg	Tamanho em bytes do objeto requisitado.
5	method	O método utilizado, podendo ser: GET, PUT, POST, etc.
6	url	A URL (objeto) solicitada ao cache.
7	cliname	Pode ou não conter o nome usuário se houver autenticação.
8	direct	Informa se o proxy fez um requisição direta. Geralmente um tunel SSL.
9	mime	Informa o tipo do arquivo baixado: HTML, TXT, IMAGEM, etc.).

Nota: A coluna Nome Atributo, é apenas ilustrativa.

Tabelas Impala

- `accesslog_tmp`: Os dados são primeiro armazenados nesta tabela e depois inseridos na tabela principal. Tem por objetivo manter os dados em seu estado original até serem copiados para a tabela de final.
- `accesslog`: Contém os dados dos logs para armazenamento definitivo e processamento.

Ambas possuem esquemas semelhantes, conforme abaixo apresentado:

Os campos em destaque na Tabela 2, são utilizados para auxiliar a criação das consultas, facilitando a manipulação das seleções por datas e horas, diminuindo a necessidade de utilização de formulas muito complexas. É importante salientar que o campo `logdate` apesar de poder ser utilizado nas seleções, é um campo especial com a função primária de separar a tabela pela partição data de criação dos logs.

Tabela 2. Esquema da tabela Impala.

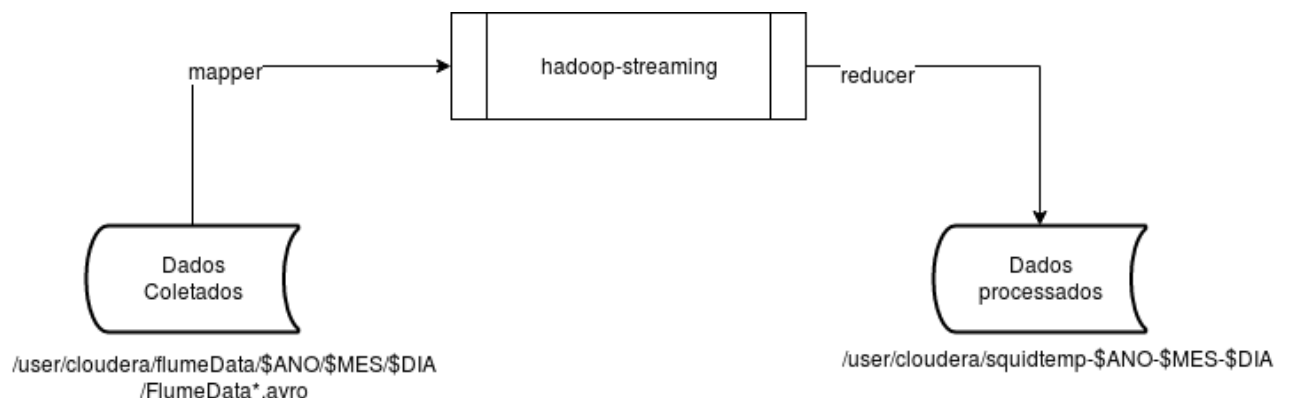
Campo	Tipo	Notas
Unixts	double	Data no formato UNIX timestamp.
timems	string	Duração da requisição em milissegundos.
Host	string	IP do Host.
Action	string	Ação tomada pelo proxy.
Lenreg	int	Tamanho em bytes da requisição.
method	string	Métodos utilizados.
url	string	A URL solicitada.
cliname	string	Contém o nome do cliente. Padrão '-'
Direct	string	Requisição direta pelo proxy ou via outro.
mime	string	Tipo de arquivo baixado.
sdatetime	string	Data e Hora no formato AAAA-MM-DD HH:MM:SS.
stime	string	Apenas a hora no formato HH:MM:SS.

iday	int	Dia.
imonth	int	Mês.
iyear	int	Ano.
logdate	string	Campo especial que identifica a partição: AAAA-MM-DD.

Parte 1: Pré-processamento MapReduce

Quando os dados são coletados pelo Flume eles são armazenados exatamente como são lidos, ou seja, o formato da linha de log, como já explicado, tem seus campos separados por espaços em branco. Para nossos fins é mais conveniente que os campos estejam usando como separador a vírgula. Para tanto foram criadas duas rotinas em linguagem Python que serão executadas pelo programa cliente hadoop no modo linha de comando.

Figura 3: Ilustração do fluxo do processamento do script.



Parte 2: Carga dos dados para o Impala

Dois procedimentos são necessários nessa fase sendo eles:

1. Carga dos dados do HDFS para a tabela temporária: A carga dos dados armazenados no HDFS para a tabela temporária é importante para que tenhamos uma cópia dos dados puros no caso da necessidade de reprocessá-los sem necessidade de reprocessamento.
2. Carga dos dados da tabela temporária para a tabela de produção com o processamento dos campos adicionais, preenchendo-os com os dados auxiliares para consultas por data e hora.

Listagem 1: O procedimento automatizado completo é feito pelo Bash script

```
#!/bin/bash

# Este script ira sempre processar os dados do DIA ANTERIOR sendo
# necessario que todos os arquivos de log estejam completamente
# gerados, completando o dia.

MAPPER="/home/cloudera/python/squid_map.py"
REDUCER="/home/cloudera/python/squid_reduce.py"
# O comando 'date' ira retorna a data do dia *anterior*
ONTEM=`date -d "1 days ago" +"%Y/%m/%d"`
ANO=${ONTEM:0:4}
MES=${ONTEM:5:2}
DIA=${ONTEM:8:2}
echo ">> Referente: $ANO-$MES-$DIA "

# Executa os script MapReduce
hadoop jar
/usr/lib/hadoop-0.20-mapreduce/contrib/streaming/hadoop-streaming-2.6.0-mr1-cdh
5.13.0.jar \
-D mapreduce.job.name="CvntLogDaily" \
-files ${MAPPER},${REDUCER} -mapper ${MAPPER} -reducer ${REDUCER} \
-input /user/cloudera/flumeStorage/${ANO}/${MES}/${DIA}/FlumeData*.avro \
-output /user/cloudera/squidtemp-${ANO}-${MES}-${DIA}

# Ajusta o modo de acesso para o diretorio para leitura
hadoop fs -chmod 0777 /user/cloudera/squidtemp-${ANO}-${MES}-${DIA}
# A particao da tabela deve ser criada *antes* da carga dos dados caso ela ja nao exista
impala-shell -u cloudera -d squidlogs --var=ANO=$ANO --var=MES=$MES
--var=DIA=$DIA -q 'alter table accesslog_tmp add partition
(logdate="${var:ANO}-${var:MES}-${var:DIA}"); load data inpath
"/user/cloudera/squidtemp-${var:ANO}-${var:MES}-${var:DIA}/part-00000"
```

```

overwrite into table squidlogs.accesslog_tmp
partition(logdate="${var:ANO}-${var:MES}-${var:DIA}");'
# Remove o diretorio apos os dados serem lidos para a tabela
hadoop fs -rm -r -f /user/cloudera/squidtemp-${ANO}-${MES}-${DIA}
# Insere os dados do log processado na particao correta da tabela temporaria.
# Alem de inserir os dados, popula os campos adicionais com a data da
# transacao do log, nos seguintes formatos: aaaa-mm-dd hh:mm:ss,
# dia, mes e ano em campos separados
impala-shell -u cloudera -d squidlogs --var=ANO=$ANO --var=MES=$MES
--var=DIA=$DIA -q 'insert overwrite table squidlogs.accesslog partition
(logdate) select unixts, timems, host, action, lenreg, method, url, cliname, direct, mime,
from_unixtime(cast(unixts as bigint), "yyyy-MM-dd HH:mm:ss") as sdatetime,
from_unixtime(cast(unixts as bigint), "HH:mm:ss") as stime,
cast(from_unixtime(cast(unixts as bigint), "dd") as int) as iday,
cast(from_unixtime(cast(unixts as bigint), "MM") as int) as imonth,
cast(from_unixtime(cast(unixts as bigint), "yyyy") as int) as iyear, logdate from
squidlogs.accesslog_tmp where logdate="${var:ANO}-${var:MES}-${var:DIA}";'
# eof

```

DEMONSTRAÇÃO

Esta seção demonstrará algumas maneiras de se extrair informações sobre os eventos dos logs. As possibilidades para extração dos dados são muitas, principalmente se usarmos a linguagem SQL e o ambiente interativo HUE. O banco de dados contém aproximadamente de 10 milhões de linhas de eventos gerados pelos acessos à Internet. Para nossos objetivos esse número é expressivo e suficiente, porém no contexto do Big Data é uma quantidade bem pequena.

Figura 4: Contagem simples mensal dos acessos aos domínios .com.br, agrupados por hosts.

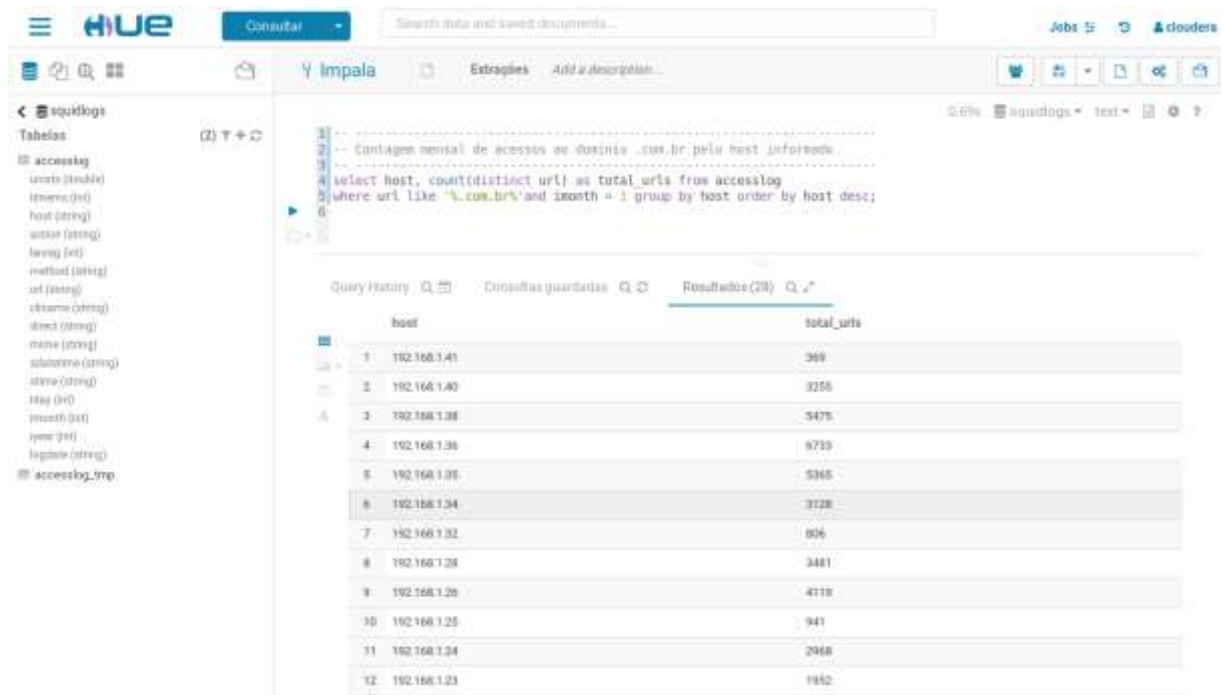


Figura 5: Gráfico representado a contagem mensal.

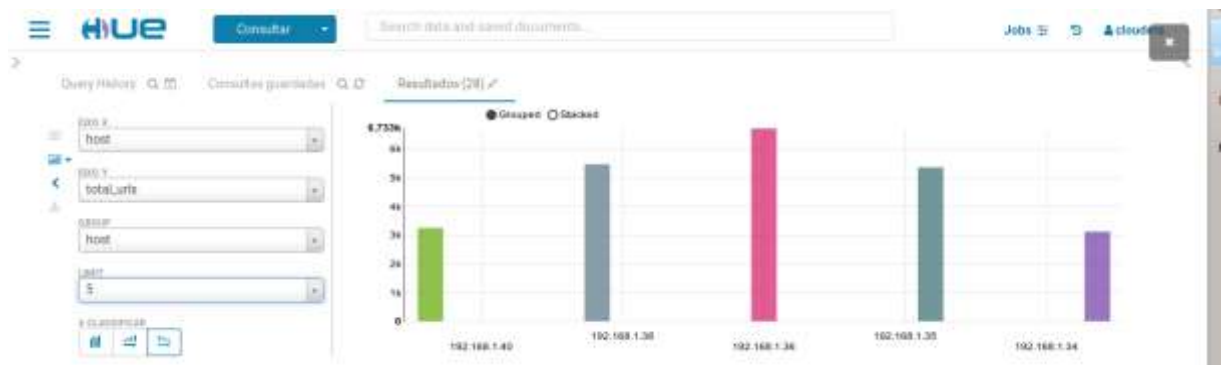
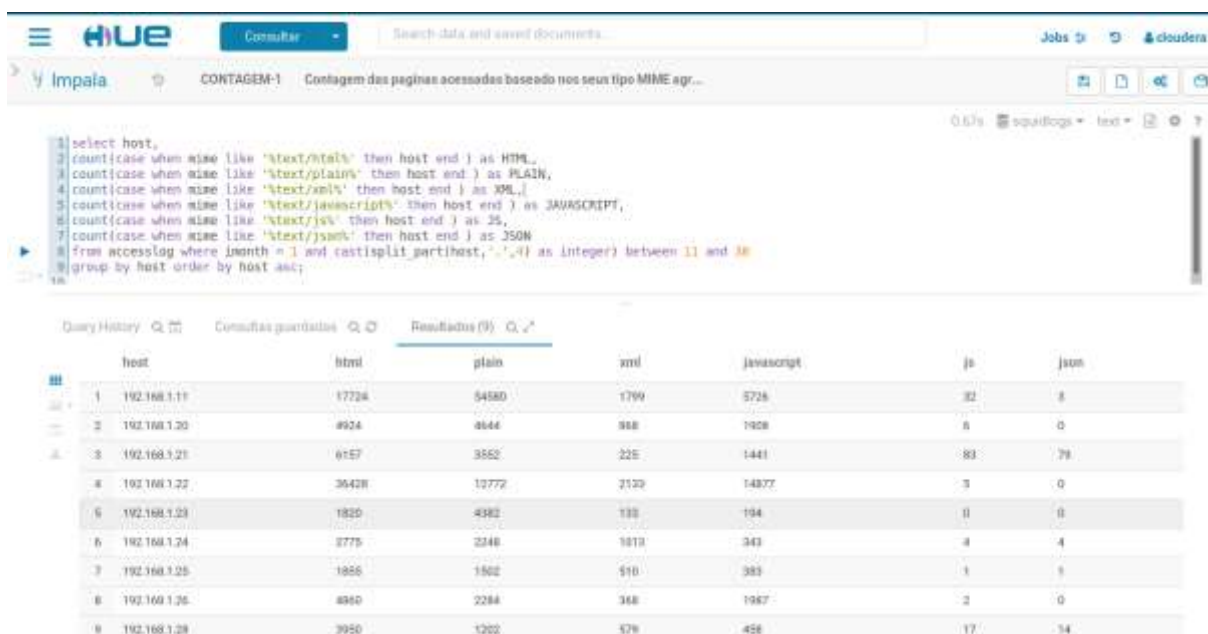


Figura 6: Contagem das páginas acessadas pelo tipo MIME.



CONCLUSÃO E TRABALHOS FUTUROS

Os problemas relacionados ao gerenciamento de logs por nós estudados, podem ter sua complexidade reduzida ao aplicarmos técnicas de Big Data, para o tratamento dos dados. Grande parte da complexidade é atribuída a coleta e centralização dos dados e ao desenvolvimento de algoritmos capazes de efetuar as análises de forma automática e padronizada, já que cada sistema possui particularidades quanto ao log gerado, tanto em formato como nos dados arquivados propriamente ditos.

Ao usarmos as tecnologias sugeridas demonstramos que é possível implementar uma estrutura com baixa complexidade e facilmente reproduzível em ambientes de pequena escala quando necessário.

A utilização do MapReduce mostra-se muito útil quando do tratamento dos dados puros, permitindo sua sanitização a fim de que possamos garantir que apenas dados úteis sejam gravados nas bases de dados para posterior geração de resultados. Apesar da geração das análises não terem sido feitas em tempo real, as mesmas técnicas podem ser adaptadas para ambientes nos quais a detecção e prevenção de intrusões em tempo real sejam prioridade.

Em trabalhos futuros serão abordados os seguintes temas: A implementação de agentes Flume em múltiplos servidores para estudos dos impactos do tráfego massivo de dados na rede e como solucionar os problemas dos gargalos, e o desenvolvimento e otimização de configurações para coleta e transmissão de dados diretamente para o Hadoop HDFS.

REFERÊNCIAS BIBLIOGRÁFICAS

ABNT NBR ISO/EIC 27002. Tecnologia da Informação - Técnicas de segurança - Código de prática para a gestão de segurança da informação. 2005. Primeira Edição: 31.08.2005 Válida a partir de 30.09.2005. Disponível em: <http://www.fieb.org.br/download/senai/nbr_iso_27002.pdf>. Acesso em: 23 Jan 2019.

CLOUDERA. Cloudera. 2018. Disponível em: <<https://www.cloudera.com/>>. Acesso em: 10 Dez 2018.

GARTNER. IT Glossary: What is big data? 2018. Disponível em: <<https://www.gartner.com/it-glossary/big-data>>. Acesso em: 23 Nov 2018.

HUE. Hue Overview. 2018. Disponível em: <<http://http://gethue.com/overview/>>. Acesso em: 13 Dez 2018.

IMPALA. Apache Impala Overview. 2019. Disponível em: <<https://impala.apache.org/>>. Acesso em: 10 Jan 2019.

KENT, K.; SOUPPAYA, M. Guide to computer security log management. NIST special publication, v. 92, 2006. Disponível em: <<https://csrc.nist.gov/publications/detail/sp/800-92/final>>. Acesso em: 01 Jan 2019.

SÊMOLA, M. Gestão da segurança de informação: uma visão executiva. 110 reimpressão. ed. [S.l.]: Rio de Janeiro. Elsevier, 2003.

SIMON, F.; SANTOS, A. L. dos; HARA, C. S. Um sistema de auditoria baseado na análise de registros de log. Escola Regional de Banco de Dados (ERBD'2008), Departamento de Informática - Universidade Federal do Paraná (UFPR), 2008. Disponível em: <https://www.researchgate.net/profile/Aldri_Santos/publication/239928992_Um_Sistema_de_Auditoria_baseado_na_Analise_de_Registros_de_Log/links/53e51d2a0cf21cc29fcb4232/Um-Sistema-de-Auditoria-baseado-na-Analise-de-Registros-de-Log.pdf>. Acesso em: 23 Jan 2019.

TAURION, Cezar. **Big Data: Velocidade, Volume, Variedade, Veracidade Valor**. Rio de Janeiro: Brasport Livros e Multimídia, 2013. (978-85-7452-608-9).

TEOH, S. T. et al. Visual data analysis for detecting flaws and intruders in computernetwork systems. IEEE Computer Graphics and Applications, Citeseer, v. 24, n. 5, 2004. Disponível em: <https://www.usenix.org/legacy/event/lisa02/tech/full_papers/takada/takada_html/>. Acesso em: 23 Abril 2019.

ANÁLISE DO PENSAMENTO COMPUTACIONAL NA EDUCAÇÃO VOLTADO À SERIOUS GAMES

*ANALYSIS OF COMPUTATIONAL THINKING IN EDUCATION AIMED TO SERIOUS
GAMES*

Fernanda Moreira POUZA

fernandampouza@hotmail.com

Curso de Sistemas de Informação, Centro Universitário Padre Anchieta

Carlos Eduardo CÂMARA

ccamara@anchieta.br

Ciência da Computação, Sistemas de Informação, Centro Universitário Padre Anchieta

Resumo

Este trabalho tem por objetivo analisar a importância da construção do Pensamento Computacional aplicado na educação, propor uma nova metodologia baseada em *serious games* e apresentar, por meio de um protótipo, uma melhor alternativa para um desenvolvimento cognitivo mais favorável em alguns anos. Para verificar a eficiência do método deste projeto, foram analisados diversos artigos acerca do desenvolvimento do Pensamento Computacional, a utilização de *Serious Games* na área da educação, os conceitos de algoritmos e o ambiente de programação Karel. Para validação do protótipo desenvolvido, foram realizados testes de própria autoria e validações de resultados esperados. Com os testes realizados, chegou-se à conclusão que o protótipo é útil para utilização como ferramenta para o aprendizado e desenvolvimento do Pensamento Computacional.

Palavras-Chave

Pensamento computacional; *serious games*; algoritmo; Karel; educação.

Abstract

This work aims to analyze the importance of construction of Computational Thinking applied in education, propose a new methodology based on serious games and present, through a prototype, a better alternative for a more favorable cognitive development in some years. In order to verify the efficiency of the method of this project, several articles about the development of Computational Thinking, the use of Serious Games in education area, concepts of algorithms and the Karel programming environment were analyzed. For validation of the developed prototype, tests of own authorship and validations of expected results were performed. With tests carried out, it was concluded that the prototype is useful for use as a tool for learning and development of Computational Thinking.

Keywords

Computational thinking; serious games; algorithms; Karel; education.

INTRODUÇÃO

Nos últimos anos, o meio tecnológico está modificando e se adaptando à vida das pessoas e, as pessoas tendo de se adaptar a estas novas tecnologias, de uma maneira muito acelerada e acessível.

Segundo uma pesquisa realizada pela *Microsoft* e o *McKinsey & Company's Education Practice*, “os avanços na tecnologia devem levar a grandes rupturas no mercado de trabalho na medida em que a automação já pode substituir até 50% dos empregos existentes, somente nos EUA”. Isso fará com que até 11.5 milhões de empregos ligados a níveis de baixa escolaridade deixem de ser ocupados por pessoas até 2030, segundo a mesma pesquisa citada anteriormente.

Muito se percebe que a tecnologia está se tornando uma ferramenta intrínseca ao cotidiano das pessoas, fenômeno denominado como computação ubíqua por Weiser (1991).

Visto isso, a maneira com que estamos acostumadas a trabalhar com a tecnologia irá mudar muito rapidamente e será necessário que novas ideias e novas práticas garantam, principalmente, que as crianças de hoje possam desenvolver competências cognitivas que as auxiliem na forma de pensar, aprender, trabalhar, se relacionar, etc.

“Considerando essa presumível relevância do compreender e do saber criar tecnologias digitais em detrimento de apenas consumi-las, é essencial que a educação das novas gerações seja planejada de modo a desenvolver tais habilidades” (SEVERGNINI, 2018).

Atualmente, as técnicas de ensino estão começando a se adequar às tecnologias, mas em alguns anos como deverá ser o ensino para que as pessoas possam se promover e até mesmo se desenvolver?

Para Resnick (2012) a maioria dos jovens possui pouca experiência em projetar e criar com mídias digitais, sentindo-se mais confortáveis em usar a tecnologia ao contrário de criar suas próprias. “Eles não são verdadeiramente fluentes em tecnologias digitais: é como se eles pudessem ler, mas não escrever” (RESNICK, 2012).

Deste modo, é possível perceber que em pouco tempo será necessário elevar o nível cognitivo em determinadas áreas para suprir as necessidades que irão exigir maiores competências.

Visto que a automação irá ocupar certos empregos das pessoas, haverá a necessidade de tornar as pessoas mais preparadas e especializadas para lidar com as máquinas.

A partir desse ponto que o presente trabalho está desenvolvido. Apresentar um protótipo baseado em *Serious Games* para aprimorar o desempenho educacional. Com isso, serão definidos alguns conceitos sobre *Serious games*, Pensamento Computacional e técnicas para serem aplicadas no desenvolvimento do protótipo a ser apresentado.

CONCEITOS PRELIMINARES

SERIOUS GAMES

Muito se tem se tem pesquisado e concluído que os jogos digitais estão auxiliando como ferramenta no ensino. Segundo Maia (2017) o aumento do estudo relacionado à *Serious Games*, ou jogos Sérios, pode ser atribuído, em parte, ao crescente interesse do público-alvo em um método alternativo para o aprendizado.

É fato, de acordo com Kafai e Burke (2015), que ainda há um interesse considerável em examinar o potencial educacional que existe em jogar *videogames*.

Pode-se perceber que os jogos têm a capacidade de envolver e cativar a atenção de uma pessoa por longos períodos de tempo, enquanto proporciona alguns outros benefícios, como o desenvolvimento de habilidades de pensamento de alto nível (KENWRIGHT, 2017).

Visto que os jogos podem ser utilizados como uma ferramenta de auxílio para a educação e não um obstáculo, é preciso garantir que essa ferramenta seja utilizada da maneira mais proveitosa, possuindo características educacionais, sem deixar de ser atrativo para o jogador.

“Estes jogos, com propósito e conteúdo específicos, são conhecidos como *Serious Games* e permitem apresentar novas situações, discutir soluções, construir conhecimentos e treinar atividades particulares” (MACHADO et al., 2011).

Sendo assim, *Serious Games* são os jogos desenvolvidos com o intuito de verificar problemas, analisar a situação e determinar um meio de alcançar uma solução. Ao utilizar os conceitos de *Serious game* como ferramenta para desenvolvimento de um jogo educacional, é possível motivar o aluno a aprender de uma maneira mais descontraída, sem deixar de lado o enfoque principal do aprimoramento do ensino.

O PENSAMENTO COMPUTACIONAL

As disciplinas relacionadas à computação estão presentes no currículo escolar de diversos países, de acordo com Brackmann (2017), onde a introdução dessa ferramenta de ensino ocorre de forma rigorosa.

A implementação dessa disciplina como obrigatória traz benefícios educacionais, onde as habilidades para raciocínio e solução de problemas se mostram muito mais aprimoradas.

Uma pesquisa feita por Balanskat et. al. (2015), apontou que na Europa o ensino de Ciência da Computação já está integrado no currículo da Educação Básica de 15 países, sendo eles: Áustria, Bulgária, República Tcheca, Dinamarca, Estônia, França, Hungria, Irlanda, Lituânia, Malta, Espanha, Polônia, Portugal, Eslováquia e Inglaterra.

No Brasil, o Pensamento Computacional não está definido como parte do currículo das escolas do Ensino Básico (Brackmann, 2017). Os documentos existentes relacionados à tecnologia estão restritos à abordagem de letramento e inclusão digital.

O termo “Pensamento Computacional” começou a ser popularizado no ano de 2006 por meio do artigo “*Computational Thinking*” de Jeannette Wing (2006) e definido de várias maneiras, entre elas, a de que “Pensamento computacional envolve a resolução de problemas, projeção de sistemas, e compreensão do comportamento humano, através da extração de conceitos fundamentais da ciência da computação” (WING, 2006). Ainda segundo Wing (2006), o pensamento computacional é a reformulação de um problema que consideramos difícil em um problema que sabemos como resolver.

Desde então, as definições estão se aperfeiçoando, como a de que o Pensamento Computacional “*É uma abordagem usada para solução de problemas utilizando o que se sabe sobre Computação*” (GOOGLE FOR EDUCATION, 2015) e mais atualmente sendo caracterizado por Brackmann (2017) como sendo o Pensamento Computacional:

[...] Uma distinta capacidade criativa, crítica e estratégica humana de saber utilizar os fundamentos da Computação, nas mais diversas áreas do conhecimento, com a finalidade de identificar e resolver problemas, de maneira individual ou colaborativa, através de passos claros, de tal forma que uma pessoa ou uma máquina possam executá-los eficazmente.

O Pensamento Computacional não é uma habilidade que pode ser ligada somente às pessoas que estão relacionadas ao desenvolvimento da tecnologia (BOUCINHA, 2017). O processo de resolução de problemas utilizados em áreas como Ciência da Computação pode ser generalizado e aplicado para diferentes áreas de pesquisa e até mesmo para a vida cotidiana (WING, 2006) e, como todos sabemos já está atuando em todas as áreas do conhecimento, como uma disciplina básica.

A grande ideia para a resolução de qualquer problema dentro do Pensamento Computacional se baseia, principalmente, na abstração e decomposição de uma tarefa grande ou complexa (WING, 2006).

Um estudo desenvolvido pela Code.Org (2016), Liukas (2015) e *BBC Learning* (2015), com elementos citados por Grover e Pea (2013) e o guia da *Computer at School* (CSIZMADIA et al., 2015), chegaram aos “Quatro Pilares do Pensamento Computacional”: Decomposição, Reconhecimento de Padrões, Abstração e Algoritmos. (Brackman, 2017 Grover, 2013))

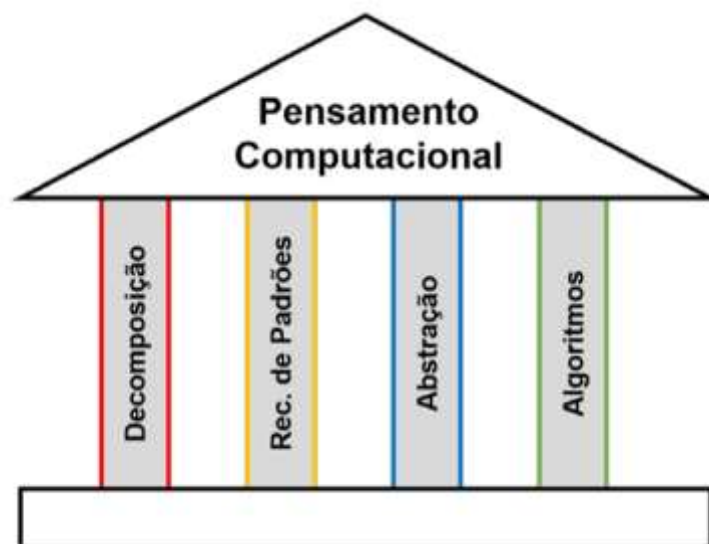
No seguinte tópico são abordados os conceitos básicos dos Pilares citados anteriormente.

QUATRO PILARES DO PENSAMENTO COMPUTACIONAL

Brackmann (2017) cita o Pensamento Computacional como uma forma de identificar um problema complexo e dividi-lo em pedaços, ou seja, fazer sua decomposição. Cada pedaço pode ser analisado individualmente com maior profundidade, verificando os problemas que já foram vistos anteriormente e, assim, poder fazer o reconhecimento de padrões. Os detalhes mais importantes devem ter um enfoque maior, enquanto as informações irrelevantes devem ser descartadas, proporcionando a abstração das informações. Por último, uma sequência de passos pode ser montada para levar a uma resolução à cada um dos subproblemas verificados, tendo assim, um algoritmo.

Os pilares que formam a base do Pensamento Computacional podem ser observados na figura a seguir.

Figura 1. Os Quatro Pilares do Pensamento Computacional



Decomposição

O Dicionário Aurélio nos dá as seguintes definições para a palavra decomposição: “Examinar por partes. Separar ou separar-se (um corpo ou um conjunto) em seus elementos ou partes; [...]” (Aurélio).

O Dicionário Michaelis também nos dá os seguintes significados para a mesma palavra: “Ação, processo ou efeito de decompor. Separação de um todo em suas partes constitutivas.” (Michaelis).

No caso do Pensamento Computacional, a decomposição é o processo onde os problemas são separados em partes menores, para uma melhor compreensão ao serem examinados separadamente.

É possível citar, como fez Brackmann (2017), a ligação que a decomposição tem com os programadores, que utilizam essa técnica para dividir um problema em instruções, baseadas em funções, métodos, objetos que compõem um algoritmo e seu consequente, código-fonte. Quando essa técnica é aplicada no desenvolvimento de programas, o entendimento e a correção de possíveis erros no código acabam se tornando um processo mais fácil e rápido, já que o erro pode se encontrar em apenas um módulo do projeto. Isso não seria possível caso a técnica de decomposição não fosse aplicada, pois seria necessário alterar várias linhas diferentes de códigos pelo projeto.

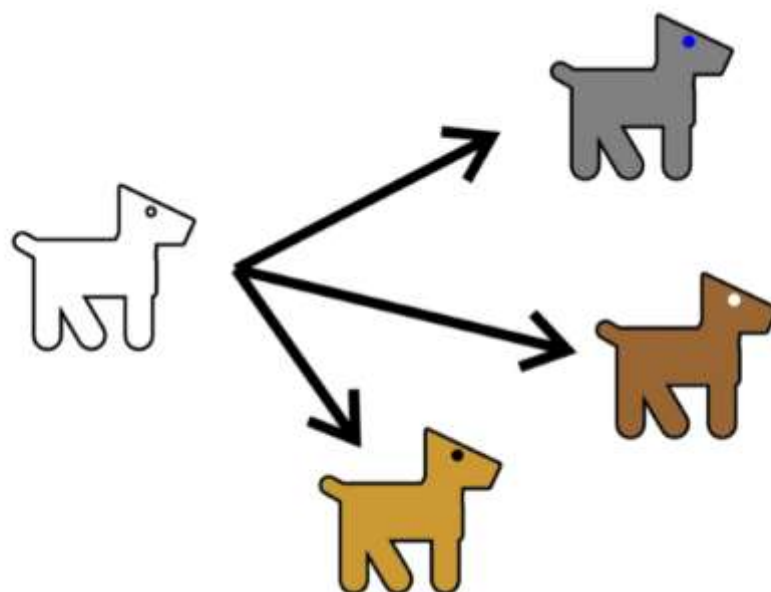
Reconhecimento de Padrões

O Reconhecimento de Padrões se dá a partir da finalização da etapa de decomposição. Ao quebrar o problema em partes menores é possível verificar a existência de padrões. Segundo Brackmann (2017), é uma forma de resolver problemas rapidamente, fazendo uso de soluções previamente definidas em outros problemas e com base em experiências anteriores.

Visto isso, é possível perceber que pelo reconhecimento de padrões, é possível tornar uma solução de um problema mais fácil e aplicá-la a problemas semelhantes.

Um exemplo que pode ser utilizado para ilustrar como se dá o reconhecimento de padrões é pelas características de um cachorro. Pode-se identificar que os cachorros possuem características semelhantes, algumas delas sendo os olhos, pelagem, rabo e orelhas.

Identificado um padrão inicial, como dado no exemplo anterior, é visto que caso haja necessidade de identificar um outro cachorro, seria preciso replicar este padrão como um modelo genérico para os padrões que forem encontrados, como é possível verificar na Figura 2.



Abstração

A abstração é uma das bases citadas inicialmente por Wing (2006) como sendo um dos conceitos principais para o Pensamento Computacional.

De acordo com Brackmann (2017), este pilar abrange a filtragem dos dados, onde deve-se ignorar elementos que não sejam necessários e que se possa concentrar a atenção, em dados que sejam mais importantes para desenvolver a ideia do problema a ser resolvido.

A abstração está muito presente na área da computação, onde os profissionais precisam criar abstrações de problemas para desenvolver sistemas. Podemos citar como um exemplo de abstração um algoritmo, onde o processo recebe uma entrada, executa uma sequência de passos e produz uma saída (BRACKMANN, 2017).

Algoritmos

Guimarães e Lages (1994) apresentam a definição de algoritmo como sendo uma *“descrição de um padrão de comportamento, expressado em termo de um repertório bem definido e finito de ações “primitivas”, das quais damos por certo que elas podem ser executadas”*.

Outra definição de algoritmos apontada por Farrer et al. (2008) é dada como uma *“descrição de um conjunto de comandos que, obedecidos, resultam numa sucessão finita de ações”*.

Em outras palavras, um algoritmo é um texto contendo um conjunto de instruções, regras ou passos que devem ser executados em determinada ordem para atingir um resultado ou solução de certo problema.

Os algoritmos estão muito presentes no dia a dia das pessoas, porém sem que elas mesmas se deem conta disso. A comparação mais comum é com uma receita de bolo. Uma receita possui, um problema inicial a ser resolvido, a lista dos ingredientes, alguns fixos e outros variáveis, uma sequência de passos a serem seguidos e, ao final, se todos os passos forem seguidos corretamente, haverá um bolo pronto, ou seja, o problema estará resolvido.

A compra de um produto pela Internet, a substituição de uma lâmpada queimada e até mesmo as ações ao acordar podem seguir um algoritmo, conforme o dia da semana é possível tomar a decisão de sair da cama em um determinado horário ou dormir por mais algum tempo.

Entretanto, existem algoritmos muito mais complexos do que uma receita de bolo e que também estão muito presentes no cotidiano das pessoas. É o caso de como as publicações das redes sociais aparecem na linha do tempo. Entretanto, esses modelos de algoritmos não serão tratados no presente trabalho.

KAREL

“Na década de 70, o então estudante Richard E. Pattis decidiu projetar um ambiente para ensinar de uma maneira mais simples os fundamentos das linguagens de

Revista Ubiquidade, ISSN 2236-9031 – v.2, n.1 – jan. a jul. de 2019, p. 69

programação. Surgiu então um projeto onde deve-se ensinar um robô a resolver problemas simples. Para o robô, deu-se o nome de Karel, em homenagem ao tcheco Karel Capek, cujo em um dos trabalhos mencionou pela primeira vez a palavra “robot”, que posteriormente popularizou-se pelo mundo” (ROBERTS, 2005).

O Karel é um robô que vive em um mundo bem simples, quadriculado, plano e infinito (GRATÉROL; MARÍN, 2006) onde irá executar uma série de instruções, sendo possível direcioná-lo para determinadas tarefas dentro deste mundo.

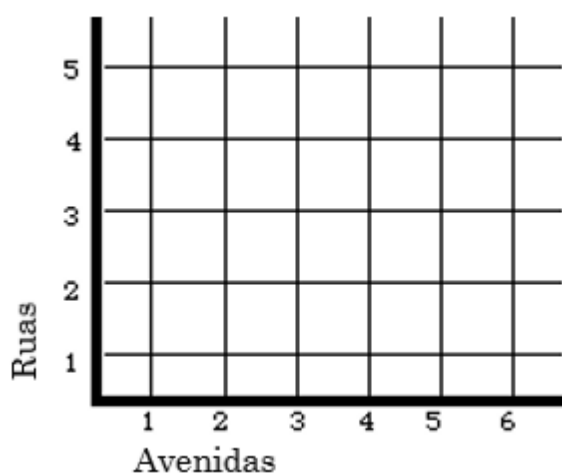
METODOLOGIA E DESENVOLVIMENTO DO PROTÓTIPO

O protótipo a ser desenvolvido se caracteriza por um *serious game* baseado no ambiente de programação chamado Karel, o robô, com o objetivo de aprimorar o processo de aprendizado dos alunos, proporcionando um estudo mais dinâmico a partir dos jogos e motivando-os com técnicas diferenciadas de ensino. Para a construção da proposta do jogo, foram utilizadas como base algumas mecânicas do Mundo de Karel, descritas a seguir.

Quando é definido o processo de especificação desses comandos, inicialmente o robô Karel entende muito pouco desses comandos, portanto, não é indicado, a princípio, que sejam definidos objetivos muito complexos. Porém, conforme se dá o processo de programação, o robô pode ir aprimorando sua capacidade e estender o nível de dificuldade das tarefas a serem executadas a cada novo objetivo que possa surgir.

Os comandos devem ser bem especificados, como será mostrado mais adiante, para que o robô consiga interpretar qual a tarefa a ser cumprida e também para desenvolver os quatro pilares do Pensamento Computacional no aluno.

Figura 3. O mundo de Karel



A interface do mundo de Karel é bem simples, como visto na Figura 3. Relembrando um plano cartesiano, o mundo possui as linhas horizontais como sendo as “Ruas” e as linhas verticais sendo as “Avenidas”. Os pontos onde uma Rua e uma Avenida se cruzam, são chamados de esquina ou nós. Dentro desse mapa o robô irá percorrer os nós para chegar ao objetivo final.

MODO DE JOGO

Tendo como base o mundo de Karel, foi desenvolvido um protótipo de *Serious Game* com o intuito de aprimorar o pensamento computacional, propondo que o jogador entenda os conceitos e execute, por meio dos pilares, os passos para alcançar o objetivo de cada fase. Também foram utilizados como base para este *Serious game*, conceitos de Algoritmos para Busca de Caminho (A*). Esses algoritmos são baseados na ideia de sair de um ponto inicial, fazendo uma estimativa de movimentação ao próximo nó que será mais viável, acumulando o custo real do caminho já percorrido, para alcançar o estado final da maneira mais rápida e, com menor custo.

No jogo em questão, o robô iniciará cada fase em um ponto específico do mapa (0,0) e movimenta-se pelas linhas das ruas e avenidas do mapa, percorrendo o caminho através das intersecções.

O jogador deverá escolher executar dentre quatro comandos disponíveis, tais como: mover para frente, rotacionar para esquerda, rotacionar para a direita ou retornar, para encontrar a solução do problema proposto e chegar ao ponto final desejado. Deve-se cumprir os objetivos para avançar de nível, e assim, avaliar o conteúdo colocado em prática por meio dos resultados obtidos.

Durante a partida, conforme o jogador realizar as movimentações, um painel com cartas irá apresentar informações de jogadas para combinar as movimentações anteriores com as possíveis jogadas que irão suceder, ou seja, a cada movimentação, o painel irá apresentar algumas opções de movimentações que podem ser feitas e assim estabelecer um critério de pontuação onde, caso o jogador faça uma movimentação que deixe o robô mais perto do objetivo, sua pontuação aumente e, caso faça um movimento em que o robô tenha que percorrer mais que o desejado, sua pontuação diminua.

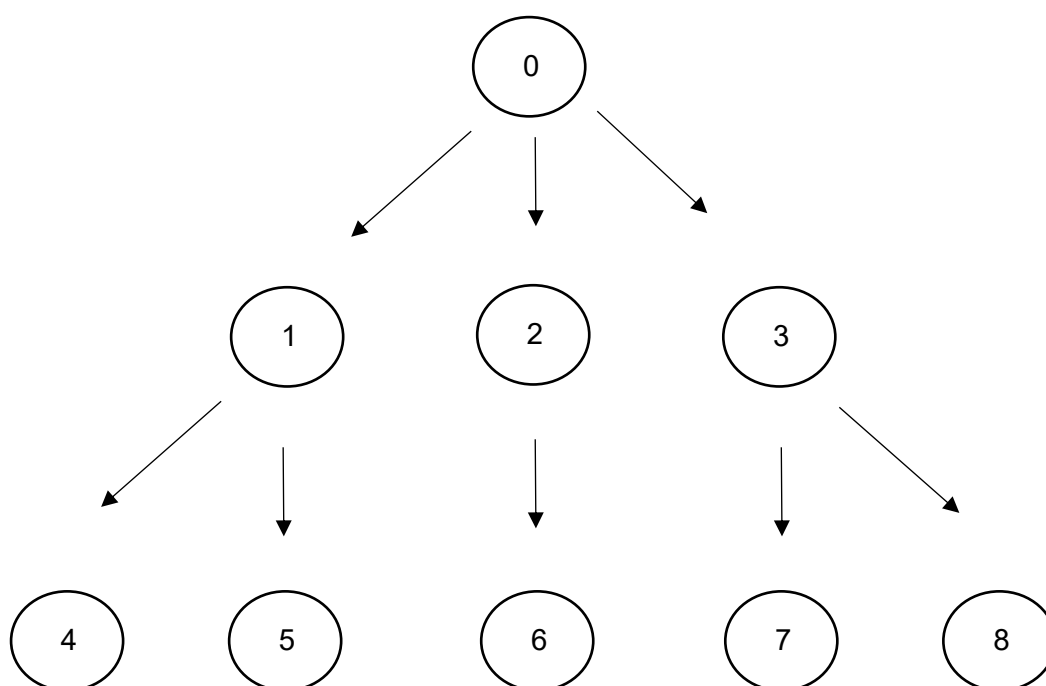
Ao completar cada objetivo, o nível de dificuldade das próximas fases é aprimorado, passando a conter mais números de movimentações do robô, mais elementos no ambiente de jogo e cenários diferentes onde o robô poderá diminuir sua pontuação, ou perder vidas, caso passe por algum elemento que cause danos a sua estrutura. Nessas próximas fases o jogador deverá aplicar novas instruções de algoritmos, como os laços de repetição, comparações, encadeamento de instruções, para construir repetições sequentes de instruções, na busca por atingir um objetivo. Um exemplo apresentado é, enquanto o robô não chegar ao ponto final, o jogador deverá executar mais um movimento.

Também há a possibilidade de determinação de tempo para atingir determinado ponto do objetivo, deixando o jogo com certa dificuldade após algumas fases de jogo.

Uma forma de solucionar o problema é representando o ambiente em forma de um grafo, onde o jogador deve verificar qual será o menor caminho que o robô poderá percorrer a partir do ponto inicial até o ponto final.

Para isso, serão considerados alguns elementos básicos, como o ponto inicial, ponto final, os comandos de direção (avançar, rotacionar para a esquerda, rotacionar para a direita e voltar) a serem utilizados para movimentar o robô pelo percurso.

Figura 4. Modelo de grafo



Observando a figura 4 como exemplo, o jogo se encontra na posição inicial 0 e a partir de então são possíveis 3 tipos de movimentação: avançar, seguir para direita e seguir para a esquerda. Seja qualquer posição que se seguir, serão possíveis outras 4 opções de movimentação, as três anteriores e a opção de retornar ao ponto anterior. Assim, cada movimentação abre uma série de sequências que podem ser definidas para avançar e chegar à posição final.

Ao evoluir de fase, é possível verificar que esse grafo também irá evoluir para outro tipo de grafo, conforme a seguinte imagem.

Figura 5. Modelo de grafo 2

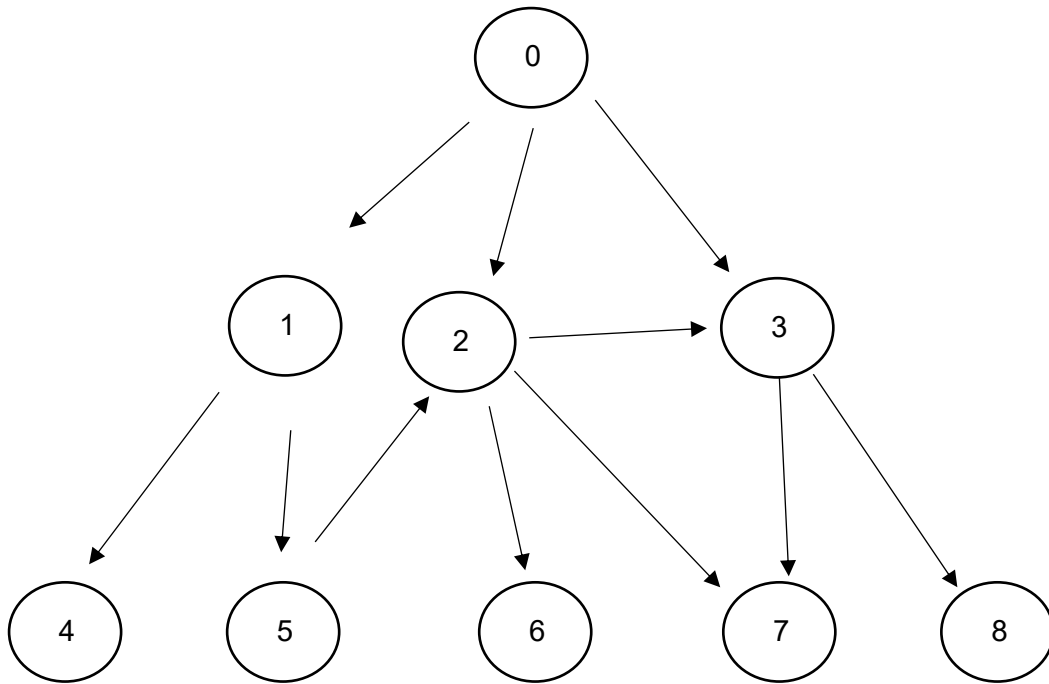
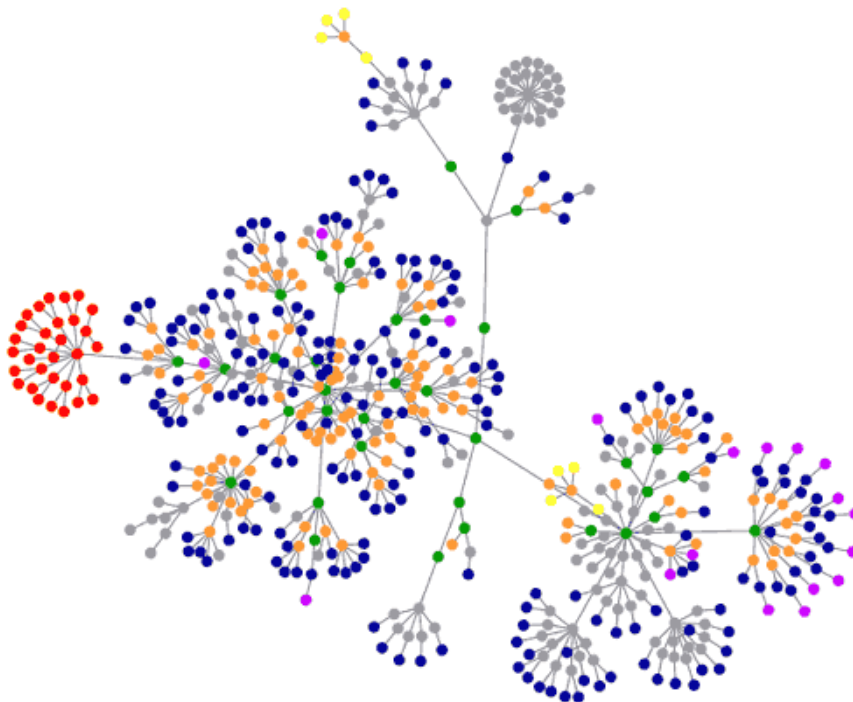


Figura 6. Conjunto de árvores formando uma floresta



TESTES E ANÁLISE DE VIABILIDADE

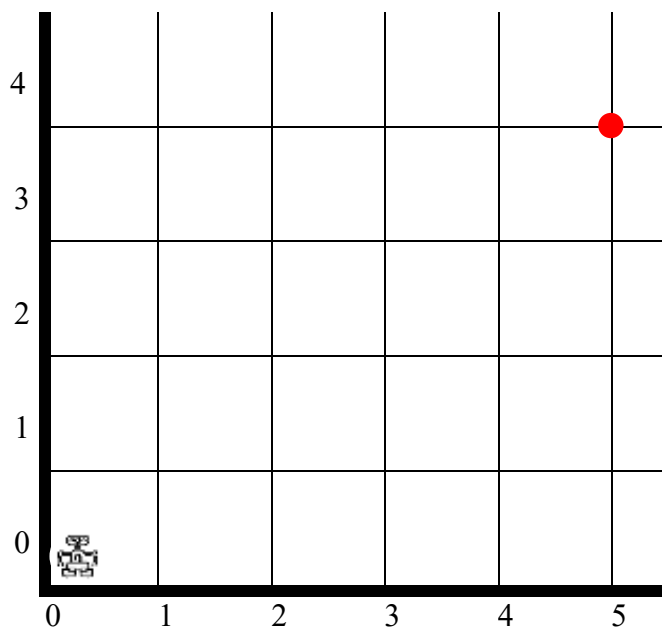
A primeira fase se dá de maneira com que o jogador se familiarize com o ambiente e o modo de jogo, onde ele possa entender como funcionam as jogadas, como movimentar o robô e como atingir o objetivo. Mesmo que simples, essa fase é de suma importância, pois é onde o jogador irá ter a visualização inicial de como poderá aplicar os conceitos básicos nas próximas fases.

O ponto inicial do robô é definido na Rua Y, Avenida X. O mundo é simples e o primeiro objetivo possui o nível de dificuldade mais baixo, pois as jogadas são menos complexas e não há obstáculos no campo de jogo.

É possível observar que há mais de uma maneira de chegar aos objetivos, como é visto nos exemplos a seguir. Porém no exemplo 1, o algoritmo possui 10 jogadas, enquanto o exemplo 2 possui 17 jogadas. Evidentemente, o exemplo de algoritmo 1 receberia uma maior pontuação em um caso de jogo, visto que apresentou menos movimentações para atingir o ponto final.

FASE 1

Figura 7. KAREL, FASE 1



A fase 1 inicia-se com o robô no estado (0,0) e o cenário apresentado é de caráter mais básico, sem obstáculos entre o robô e o estado final (4,5).

Cada movimento que o faça chegar mais perto de seu objetivo são somados para contabilizar a progressão à próxima fase. Quanto menor a quantidade de movimentos feitos, maior será a sua pontuação. Alguns alertas podem surgir no painel para advertir ao jogador de que fazer movimentos em que o robô volte estados não são aconselhados. Caso o jogador não atinja uma pontuação suficiente, ele deverá repetir a fase para percorrer um novo caminho.

Exemplo de algoritmo 1: Ponto inicial (r0, a0) e ponto final (r4, a5).

INICIO

Avançar (0,1)

Avançar (0,2)

Avançar (0,3)

Avançar (0,4)

Avançar (0,5)

Virar para a esquerda (0,5)

Avançar (1,5)

Avançar (2,5)

Avançar (3,5)

Avançar (4,5)

FIM ALGORITMO

Exemplo de algoritmo 1 com laço de repetição:

INICIO

Enquanto $r < 4$ faça

Avançar;

Fim enquanto

Virar para a direita; (4,0)

Enquanto $a < 5$

Então Avançar;

Fim enquanto

FIM ALGORITMO

Exemplo de algoritmo 2: Ponto inicial (r0, a0) e ponto final (r4, a5).

INICIO

Avançar (0,1)

Virar para a esquerda (0,1)

Avançar (1,1)

Virar para a direita (1,1)

Avançar (1,2)

Virar para a esquerda; (1,2)

Avançar (2,2)

Virar para a direita; (2,2)

Avançar (2,3)

Virar para a esquerda; (2,3)

Avançar (3,3)

Virar para a direita; (3,3)

Avançar (3,4)

Virar para a esquerda; (3,4)

Avançar (4,4)

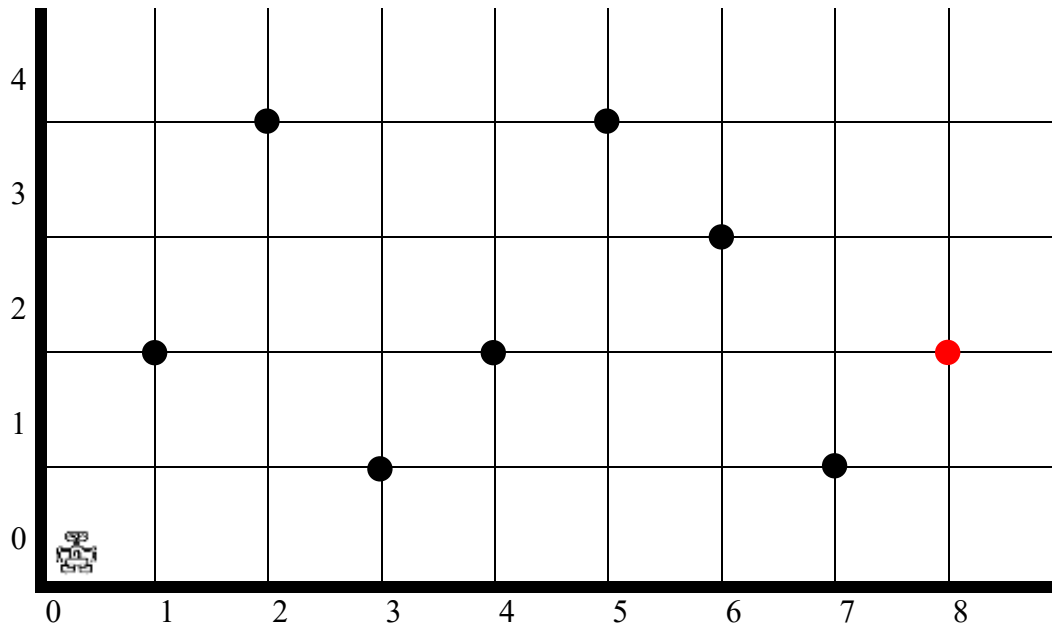
Virar para a direita; (4,4)

Avançar (4,5)

FIM ALGORITMO

FASE 2

Figura 8. KAREL, FASE 2



A fase 2 origina-se com o robô no estado (0,0) e o cenário começa a possuir elementos pelo percurso onde o robô deverá cruzar para agregar pontos a contagem e atingir o estado final a ser alcançado.

As mesmas regras da primeira fase são aplicadas, onde cada movimento é somado para contabilizar a progressão à próxima fase e quanto menor a quantidade de movimentos realizados, maior será a pontuação. Ao mesmo passo, ao movimentar o robô para um estado anterior ao atual, a pontuação do jogador será reduzida, fazendo com que ele evite voltar as jogadas. Para isso, os alertas ainda se farão presentes no painel, advertindo ao jogador que fazer movimentos de retorno não são aconselhados.

Caso o jogador não atinja uma pontuação considerada suficiente, ele deverá repetir a fase para percorrer um novo caminho.

Diante desse cenário de jogo, pode-se dar o seguinte algoritmo como exemplo para chegada ao objetivo.

Exemplo de algoritmo 1:

INICIO

Avançar (0,1)

Virar para a esquerda (0,1)

Avançar (1,1)

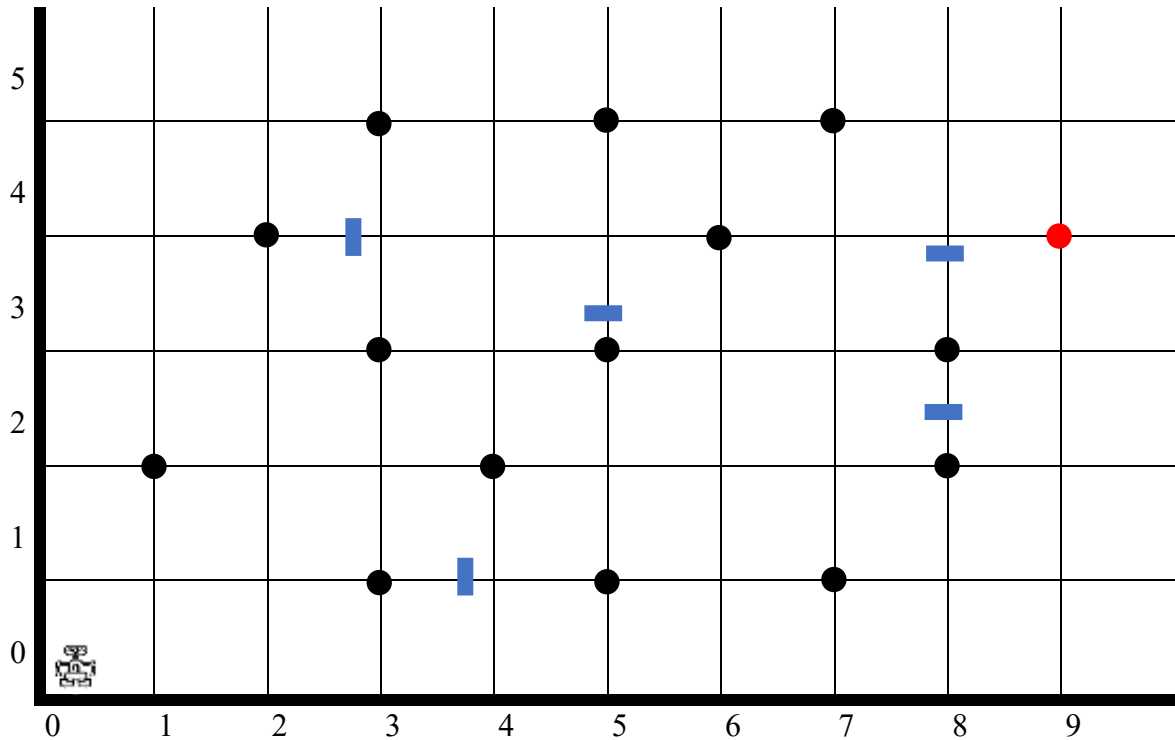
Avançar (2,1)
Virar para a direita (2,1)
Avançar (2,2)
Virar para a esquerda (2,2)
Avançar (3,2)
Avançar (4,2)
Virar para a direita (4,2)
Avançar (4,3)
Avançar (4,4)
Avançar (4,5)
Avançar (4,6)
Virar para a direita (4,6)
Avançar (3,6)
Avançar (2,6)
Avançar (1,6)
Virar para a esquerda (1,6)
Avançar (1,7)
Avançar (1,8)
Retornar (1,7)
Virar para a esquerda (1,7)
Avançar (2,7)
Virar para a direita (2,7)
Avançar (2,8)

FIM ALGORTIMO

A partir desse ponto, o jogo começa a expandir e as fases necessitarão de mais raciocínio. A estrutura para os algoritmos a serem desenvolvidos pelo jogador será avaliada e a árvore de resolução, como visto na Figura 6, irá elaborar um conjunto de padrões para avaliar a possibilidade de movimentação do jogador, indicando no painel as combinações que podem ser feitas a partir das movimentações já realizadas com as possíveis movimentações no estado atual.

FASE 3

Figura 9. KAREL, FASE 3



A fase 3 possui alguns elementos no percurso que são barreiras onde o robô não conseguirá atravessar. Esses elementos são os retângulos azuis dispostos nas ruas e avenidas do mundo de Karel. Essas barreiras estão presentes nessa fase de jogo para aprimorar o raciocínio e o pensamento através da inserção de um novo objeto no campo de jogo.

As regras principais continuam, porém como agora há um objeto que impossibilita o robô de cruzar determinados caminhos, o jogador deverá voltar um estado do robô, caso tenha chegado até esse ponto, e percorrer outro caminho para alcançar o estado final. Caso o jogador se atente de antemão a respeito do obstáculo, deverá evitar aquele caminho para não sofrer perda de pontuação.

Nessa fase também o tempo irá interferir na pontuação, fazendo com que se o jogador demorar a chegar ao objetivo, sua pontuação seja menor.

Caso o jogador não atinja uma pontuação considerada suficiente, ele deverá repetir a fase para percorrer um novo caminho.

Diante desse cenário de jogo, pode-se dar o seguinte algoritmo como exemplo para chegada ao objetivo.

Exemplo de algoritmo fase 3:

INICIO

Avance (0,1)
Virar para a esquerda (0,1)
Enquanto $r < 2$ faça
 Avançar;
Fim enquanto
Virar para a direita (2,1)
Enquanto $a < 4$ faça
 Avançar;
Fim enquanto
Virar para a esquerda (2,4)
Avance (3,4)
Virar para a direita (3,4)
Avance (3,5)
Virar para a direita (3,5)
Enquanto $r \neq 1$ faça
 Avançar;
Fim enquanto
Virar para a esquerda (1,5)
Enquanto $a < 8$ faça
 Avançar;
Fim enquanto
Virar para a esquerda (1,8)
Avance (2,8)
Virar para a direita (2,8)
Avance (2,9)
Virar para a esquerda (2,9)
Enquanto $r < 4$ faça
 Avançar;

Fim enquanto
FIM ALGORTIMO

CONCLUSÕES

Este presente trabalho, juntamente com o protótipo desenvolvido para a construção do pensamento computacional, baseado em um *Serious game*, pôde apresentar através de modelos de jogo e a partir dos resultados desse estudo, a validação da ideia para o desenvolvimento de novas metodologias de aprendizado.

Verificando a futura ampliação do ensino de matérias relacionadas à Computação, visto a necessidade que será exigida, é incontestável a aplicação de formas mais atrativas e envolventes de aprendizado, como a apresentada.

Para elaboração deste trabalho, foram consultados diversos artigos, livros e pesquisas referentes ao pensamento computacional e suas respectivas áreas de abrangência, principalmente da área educacional, para avaliar alternativas de uma melhor metodologia para aplicação dessa ideia deste conceito.

Também foram pesquisados conceitos e aplicações a cerca de *Serious Games*, levantadas informações acerca dos Serious games, os modos de aplicações nos diversos projetos encontrados e os resultados após as respectivas aplicações.

A ideia, a princípio, se baseia em implementar o Pensamento Computacional como forma de resolução de problemas na área educacional, entretanto, não é descartada a possibilidade de posteriormente ser expandido para outras áreas de aplicação.

É esperado que, seja possível fornecer uma nova metodologia para introdução do Pensamento Computacional entre as metodologias de ensino acadêmicas futuras, onde a tecnologia possa ser inserida para facilitar e auxiliar na resolução de problemas.

Vale ressaltar que o protótipo desenvolvido é apenas um modelo, podendo ser modificado e aperfeiçoado, para assim, poder aplicá-lo em uma amostra maior. As fases apresentadas se dão como base para um teste de aplicação onde os pilares do Pensamento Computacional que foram apresentados podem ser adquiridos e aprimorados conforme a manipulação do jogo para resolução das situações, modificando o modo e tempo de raciocínio do jogador.

Para os trabalhos futuros, já é considerada a ampliação do protótipo, onde será desenvolvida com mais cuidado a heurística do jogo e a inclusão de Inteligência Artificial, para verificar padrões de jogada dos usuários, identificar possíveis jogadas “viciadas”, avaliar a possibilidade de saltos entre as fases do jogo. Caso sejam identificadas jogadas mais elaboradas do jogador, habilitar a visão total do mapa somente na introdução do jogo

(ao mapa) e, após o início, a tela poderá se resumir a poucos pontos ao redor do robô e outras funcionalidades interessantes que possam surgir mais adiante.

REFERÊNCIAS BIBLIOGRÁFICAS

BERGIN, Joseph et al. **Karel++: A Gentle Introduction to the Art of ObjectOriented Programming**. JW, 1996.

BERGIN, Joseph. **Karel++ World for Macintosh: (Preliminary Documentation)**. 1997-1998. Disponível em: <<https://csis.pace.edu/~bergin/KWorld/KWorld.html>>. Acesso em: 12 nov. 2018.

BOUCINHA, Rafael Marimon et al. **CONSTRUÇÃO DO PENSAMENTO COMPUTACIONAL ATRAVÉS DO DESENVOLVIMENTO DE GAMES**. **Renote**, v. 15, n. 1, 28 jul. 2017. Universidade Federal do Rio Grande do Sul. <http://dx.doi.org/10.22456/1679-1916.75146>.

BRACKMANN, Christian Puhlmann. **DESENVOLVIMENTO DO PENSAMENTO COMPUTACIONAL ATRAVÉS DE ATIVIDADES DESPLUGADAS NA EDUCAÇÃO BÁSICA**. 2017. 226 f. Tese (Doutorado) - Curso de Informática na Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.

FEOFILOFF, Paulo. **Algoritmos para Grafos em C**. 2018. Disponível em: <www.ime.usp.br/~pf/algoritmos_para_grafos/>. Acesso em: 30 nov. 2018.

FARRER, Harry et al. **Algoritmos Estruturados**. 3. ed. Rio de Janeiro: LTC, 2008. 284 p.

GOOGLE FOR EDUCATION. What is Computational Thinking? Computational Thinking for Educators, 2015. Disponível em: <<https://computationalthinkingcourse.withgoogle.com/unit?lesson=8&unit=1>>. Acesso em: 30/10/2015.

GUIMARÃES, Ângelo de Moura; LAGES, Newton Alberto de Castilho. **Algoritmos e estruturas de dados**. Rio de Janeiro: LTC, 1985-2008. 216 p.

KAFAI, Yasmin B.; BURKE, Quinn. **Constructionist Gaming: Understanding the Benefits of Making Games for Learning**. **Educational Psychologist**, v. 50, n. 4, p.313-334, 2 out. 2015. Informa UK Limited. <http://dx.doi.org/10.1080/00461520.2015.1124022>.

KENWRIGHT, Ben. Brief review of video games in learning & education how far we have come. **Siggraph Asia 2017 Symposium On Education On - Sa '17**, Bangkok, nov. 2017. ACM Press. <http://dx.doi.org/10.1145/3134368.3139220>.

MACHADO, L. S. et al. Serious games baseados em realidade virtual para educação médica. *Revista Brasileira de Educação Médica*. Rio de Janeiro, v. 35, n. 2, p. 254-262, jun. 2011.

MAIA, Estevan A. P. **SERIOUS GAMES: UMA APLICAÇÃO NO CENÁRIO DA PATOLOGIA CLÍNICA VETERINÁRIA**. 2017. 36 f. TCC (Graduação) - Curso de Sistemas de Informação, Centro Universitário Padre Anchieta, Jundiaí, 2017.

MARIN, Hernando Castaneda; GRATÉROL, Wladimir Rodríguez. **Programación Orientada a Objetos en el Micro mundo del Robot Karel**. 2006. Disponível em: <https://www.researchgate.net/publication/235004760_Programacion_Orientada_a_Obj etos_en_el_Micro_mundo_del_Robot_Karel_libro_1>. Acesso em: 12 nov. 2018.

MICROSOFT. **O Ensino em 2030 e o aprendizado pronto para a vida: O Imperativo tecnológico**. 2018. Disponível em: <https://info.microsoft.com/LA-DIGTRNS-CNTNT-FY19-08Aug-01-ConhecaaSaladeAulade2030-MGC0002878_01Registration-ForminBody.html>. Acesso em: 13 ago. 2018.

RESNICK, Mitchel. **Mother's Day, Warrior Cats, and Digital Fluency: Stories from the Scratch Online Community**. 2012. Disponível em: <<https://web.media.mit.edu/~mres/papers/mothers-day-warrior-cats.pdf>>. Acesso em: 17 nov. 2018.

ROBERTS, Eric. **KAREL THE ROBOT LEARNS JAVA**. 2005. Disponível em: <<https://web.stanford.edu/class/cs106a/book/karel-the-robot-learns-java.pdf>>. Acesso em: 12 nov. 2018.

SEVERGNINI, Luís Filipe. **Alice e o Mistério dos Algoritmos: um serious game como ferramenta de aprendizagem de lógica de programação para crianças**. **Renote**: Revista Novas Tecnologias na Educação, v. 16, n. 1, 2018. Disponível em: <<https://seer.ufrgs.br/renote/article/view/86049>>. Acesso em: 11 nov. 2018.

WEISER, Mark. **The Computer for the 21st Century**. 1991. Disponível em: <<https://www.ics.uci.edu/~corps/phaseii/Weiser-Computer21stCentury-SciAm.pdf>>. Acesso em: 19 out. 2018.

WING, Jeannette. **PENSAMENTO COMPUTACIONAL – Um conjunto de atitudes e habilidades que todos, não só cientistas da computação, ficaram ansiosos para aprender e usar**. Revista Brasileira de Ensino de Ciência e Tecnologia, v. 9, n. 2, 2016. Disponível em: <<https://periodicos.utfpr.edu.br/rbect/article/view/4711>>. Acesso em: 16 ago. 2018.

ANÁLISE DE DESEMPENHO DE UM ALGORITMO DESENVOLVIDO PARA SOLUÇÃO DE DEEP LEARNING UTILIZANDO REDES NEURAIS CONVOLUCIONAIS PARA ANÁLISE DE CONTRASTE DE IMAGENS

*PERFORMANCE ANALYSIS OF AN ALGORITHM DEVELOPED FOR DEEP LEARNING
SOLUTION USING CONVOLUCIONAIS NEURAL NETWORKS FOR IMAGE CONTRAST
ANALYSIS*

Patrick Francisco OLIVEIRA

patrickoliveira58@gmail.com

Sistemas de informação, Centro Universitário Padre Anchieta

Carlos Eduardo CÂMARA

ccamara@anchieta.br

dinhocamara@gmail.com

Sistemas de Informação, Ciência da Computação, Centro Universitário Padre Anchieta

Resumo

O aprendizado profundo (*deep learning*) pertence a um tipo de aprendizado de máquina que utiliza as redes neurais artificiais para a resolução de problemas complexos como o reconhecimento de imagens. A rede neural convolucional (*RNC*) é uma arquitetura baseada em redes neurais artificiais, e vem sendo muito aplicada para o aprendizado profundo, trazendo bons resultados no processamento e análise de imagens. Nesse artigo será apresentado conceitos preliminares para o entendimento das *RNCs*, analisando as técnicas para a melhora do contraste das imagens, mostrando o processo de aprendizagem e seus resultados.

Palavras-Chave

Aprendizado profundo. Rede neural convolucional. Contraste.

Abstract

Deep learning belongs to a type of machine learning that uses artificial neural networks for solving complex problems such as the recognition of images. Convolutional neural network (*RNC*) is an architecture based on artificial neural networks, and has been much applied for deep learning, bringing good results in the processing and analysis of images. This article will be presented preliminary concepts for the understanding of the *RNCs*, analyzing techniques for the improvement of the contrast of images, showing the learning process and its results.

Keywords

Deep learning. Convolutional neural network. Contrast.

INTRODUÇÃO

A capacidade dos computadores em realizarem tarefas vem crescendo cada vez mais com o passar dos anos e através disso estão surgindo novas possibilidades. Os estudos voltados ao aprendizado de máquinas é algo recente comparado a outros campos, tendo início após a segunda guerra mundial, com o objetivo de entender como nós seres humanos pensamos e, indo mais além, construindo algoritmos utilizando mecanismos que permitem emular a forma de aprender biologicamente, a Inteligência Artificial (IA) que pode executar as nossas atividades de forma ágil e eficiente.

Diversas áreas estão envolvidas para se desenvolver a IA, disciplinas de filosofia que auxiliam a entender de onde vem o conhecimento e como ele conduz a ação, a psicologia que mostra como nós seres humanos pensamos e agimos, a matemática que é uma das bases do desenvolvimento e que determina o que pode ser computado. Quais regras e fórmulas determinam os modelos que serão utilizados para fazer os cálculos e as probabilidades. A biologia que mostra como o cérebro funciona, como é feito o processamento das informações, entre outras que contribuem para a construção da inteligência artificial. (VON ZUBEN, 2015)

A Inteligência Artificial, hoje, abrange diversos segmentos onde podem ser aplicadas, sendo eles, geralmente, de grande escala, como os que são usados na área da saúde, automação, ou para uma simples tarefa como um jogo de xadrez.

Com a grande massa de fotos capturadas por câmeras de smartphones, de variadas qualidades, é possível notar que algumas imagens não possuem iluminação necessária, o contraste, para que a visualização do conteúdo presente na foto possa ser identificada da melhor forma.

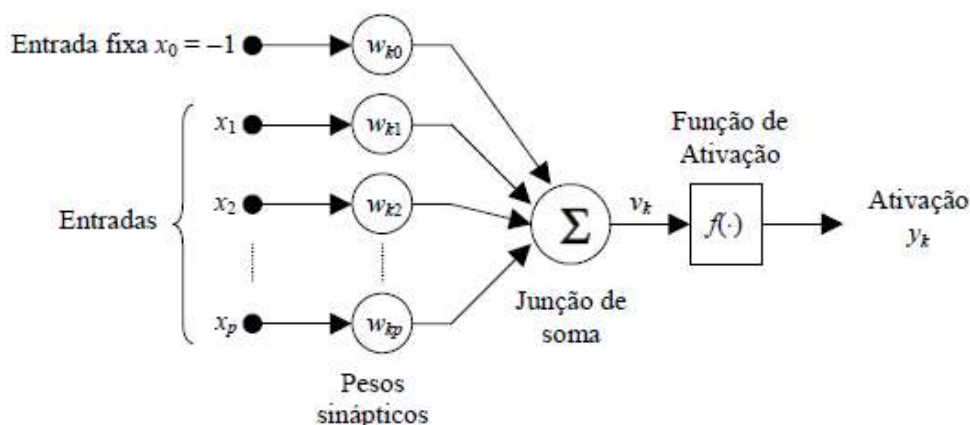
Para se ter uma ideia, o *Google Fotos*, aplicativo de galeria de fotos da google, que possui mais de 500 mil usuários, tem a visualização de 5 bilhões de fotos por dia. Com essa grande quantidade de imagens se torna possível a utilização de IA para poder avaliar essas fotos e realizar algumas ações em cima delas.

Em volta dessa questão, nesse projeto será feita a utilização de deep learning para a implementação de redes neurais convolucionais com o objetivo de melhorar o contraste das fotos, corrigindo a iluminação daquela imagem que foi capturada em condições de pouca iluminação.

REDES NEURAIAS

O estudo de redes neurais artificiais começou a partir do interesse de como o cérebro humano realiza diversas atividades complexas, com alto grau de agilidade, tendo como base que, as funções neurais biológicas são armazenadas nos neurônios e nas conexões entre eles. O processo de aprendizagem se dá com a geração de novas conexões entre esses neurônios e a alteração nas conexões existentes. A partir desse conceito foi construído o conceito, um modelo matemático denominado *redes neurais artificiais* que são um conjunto de neurônios simples agrupados. (HAYKIN, 2009)

Figura 1. Modelo de neurônio artificial.



A figura 1, mostra o neurônio artificial que é a unidade básica de uma rede neural, definido por três elementos básicos:

- **Conjunto de sinapses:** ou links de conexões, caracterizadas por um peso (w_{ki}) ou a sua intensidade, que são definidas como um peso da entrada naquele ramo. Um sinal de entrada x_j conectado ao neurônio k , é multiplicado pelo peso da sinapse w_{kj} . A entrada fixa, x_0 , é um parâmetro externo do neurônio em k , que determina uma tendência.
- **Somatório:** ou junção. Realiza a adição do resultado da multiplicação dos sinais de entrada x_i pelas sinapses, w_{ki} , do neurônio. Essa operação constitui um combinador linear.
- **Função de ativação:** $\varphi(\cdot)$ ($f(\cdot)$) define, ou limita, a amplitude do sinal de saída a um valor finito. Também é definida como uma função limitadora da saída v_k do neurônio. (VARGAS et. al., 2016)

A rede neural é construída através de camadas, que é o agrupamento dos neurônios. Essas camadas são agrupadas, definindo assim a profundidade da rede, de modo que, dependendo de como é feita a organização das camadas, a rede se comportará de forma diferente, podendo apresentar diversos tipos de arquiteturas.

Redes neurais convolucionais

As redes neurais convolucionais são redes artificiais que possuem em uma de suas camadas a operação de convolução. A dificuldade para se treinar uma rede neural de forma eficiente era visível e por isso, outras arquiteturas de redes neurais foram modeladas. Uma dessas redes que apresentou uma maior facilidade para ser treinada comparado a redes completamente conectadas foi a **Rede Neural Convolucional (RNC)**. (FERREIRA. 2017)

As redes neurais convolucionais são muito utilizadas quando existe uma grande quantidade de amostras rotuladas. Algumas vantagens da *RNC*:

- Possuem a capacidade de extrair características relevantes através de aprendizado de transformações
- Depende de um número menor de parâmetros de ajustes do que redes totalmente conectadas com o mesmo número de camadas ocultas.

As *RNCs* podem ser utilizadas na classificação de imagens através do conhecimento sintetizado via *RNC* ou para reconhecer objetos, pessoas entre outros. (FERREIRA. 2017)

A seguir será destacado alguns fundamentos básicos e os componentes de uma rede neural convolucional.

Lenet

Proposta por *Yann LeCun*, foi um dos primeiros projetos de redes neurais convolucionais, utilizado inicialmente para reconhecimento de caracteres manuscritos, tendo surgido propostas para melhoramento dessa rede.

As *RNCs* são formadas por sequências de camadas e cada uma destas possui uma função específica na programação do sinal de entrada. A arquitetura *LeNet* possui três camadas principais: a de *convolução*, de *pooling* e *totalmente conectadas*, sendo que a camada (MAZZA, 2016):

- **Convolutacional:** é responsável por extrair atributos dos dados de entrada.
- **Pooling:** são responsáveis por reduzir a dimensionalidade dos dados resultantes da camada convolucional.
- **Totalmente conectadas:** são responsáveis pela programação do sinal por meio da multiplicação ponto a ponto e o uso de uma função de ativação (x_0 na figura 1).

O resultado da análise da rede neural convolucional será a probabilidade da imagem pertencer a uma determinada classe para qual ela foi treinada. Abaixo será explicado cada uma das camadas.

Camadas Convolucionais

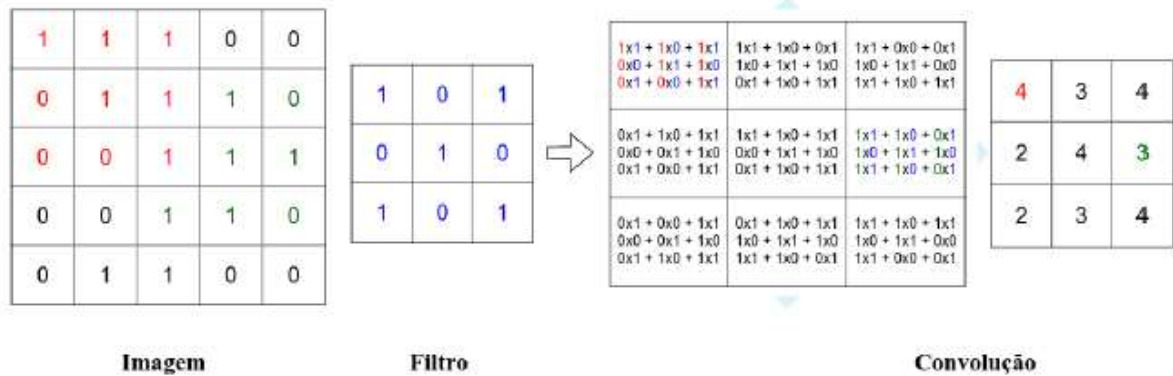
A convolução é uma operação matemática entre duas funções f e g , gerando uma terceira função, que seria uma modificação de f . A expressão que define a convolução, tendo as funções f e g , com a variável contínua x e tendo o símbolo $*$ como operador de convolução, é dada por (FERREIRA. 2017):

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\tau) \cdot g(x - \tau) d\tau$$

No processamento de imagem, apresentado na figura 3, onde a imagem é definida como uma função bidimensional, a convolução funciona muito bem para detectar as bordas, suavização de imagem, entre outras aplicações. Nas imagens a convolução funciona da seguinte forma: o somatório da multiplicação de cada elemento da imagem,

com seus vizinhos locais, pelos elementos da matriz que representam o filtro de convolução.

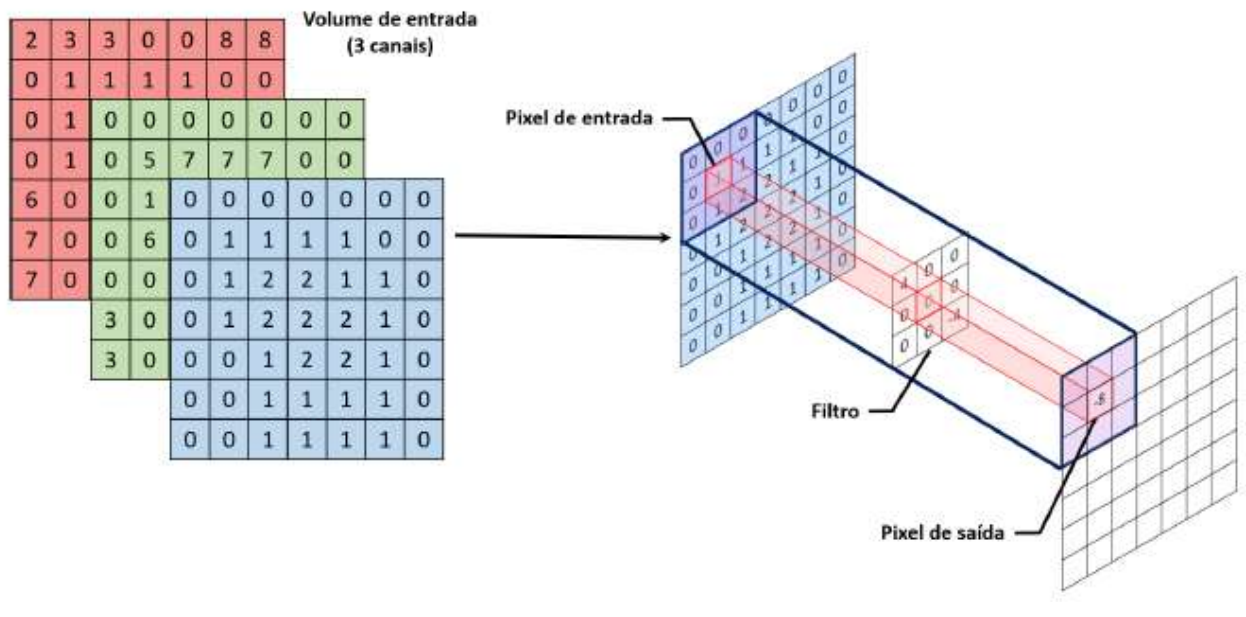
Figura 2: Exemplo de convolução em uma imagem. (FERREIRA. 2017)



Essa camada possui um conjunto de filtros que recebem dados de entrada, chamados de volume. Cada filtro tem dimensão reduzida, mas pode se estender. Se a imagem for colorida ela possui 3 canais e o filtro da primeira camada convolucional terá o tamanho de 5 pixels de altura, 5 de largura, 3 de profundidade e durante o processo de treinamento a dimensão do filtro vai se ajustar de acordo com as características mais importantes dos dados de entrada.

Cada filtro dá origem a uma estrutura conectada que percorre toda a extensão dos dados de entrada. A convolução é feita pela somatória do produto ponto a ponto entre os valores de um filtro e cada posição do volume de entrada. Os resultados passam por uma função de ativação chamada de *Rectified Linear Units*.

Figura 3: Demonstração da convolução entre um filtro 3x3 e o volume de entrada.
(ARAÚJO et. al., 2017)

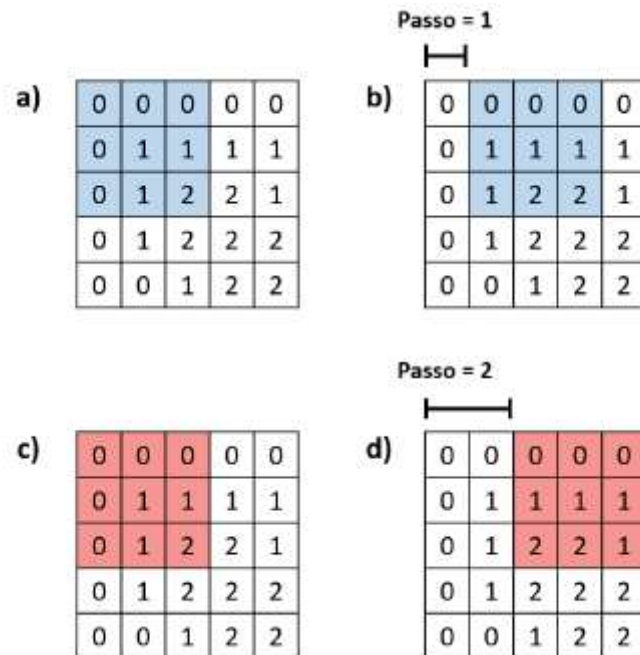


Tendo três parâmetros que controlam o tamanho dos resultados da camada convolucional (ARAÚJO et. al., 2017):

- **Profundidade:** depende do número de filtros utilizados.
- **Passo:** tamanho do salto na operação de convolução.
- **Zero-Padding:** preenche a borda do volume de entrada.

Cada um desses filtros será responsável por extrair as características diferentes dos dados de entrada. Quanto o maior número de filtros, maior será o número de características extraídas dos dados de entrada, porém a complexidade computacional, relativa ao tempo e ao uso de memória será maior.

Figura 4: Ilustração de como o passo influencia o deslocamento de um filtro 3x3 em duas etapas sucessivas da convolução.

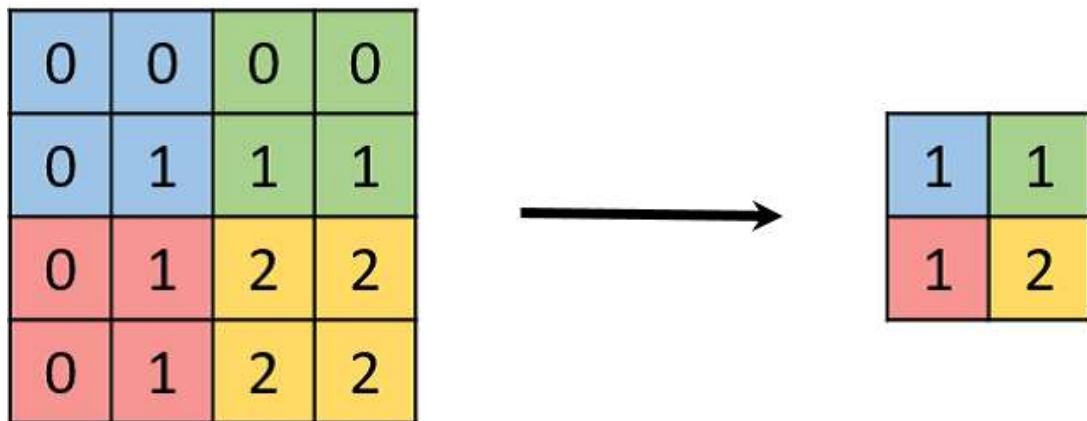


Camada Pooling

Após a camada convolucional existe uma camada de *Pooling*. Esta camada tem como objetivo, reduzir a dimensão espacial do volume de entrada, reduzindo o custo computacional da rede.

A camada *Pooling* (figura 5), recebe os valores de uma parte da região do mapa de atributos, que foram gerados pela camada convolucional, e os substitui por alguma métrica dessa região. A operação *max pooling*, onde é feita a substituição dos valores pelos valores máximos é muito utilizada. Essa operação realiza a eliminação de valores baixos, reduzindo assim a dimensão da representação dos dados, acelerando a computação necessária para as próximas camadas. (ARAÚJO et. al., 2017).

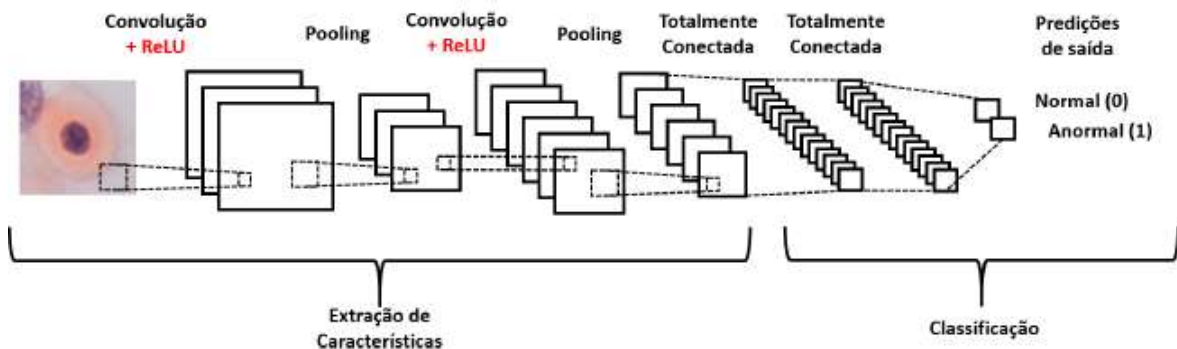
Figura 5: Aplicação de *max pooling* em uma imagem 4x4 utilizando um filtro 2x2.



Camada Totalmente Conectada

Com os resultados das camadas convolucionais e de *Pooling* que irão trazer as características da imagem de entrada, a camada totalmente conectada realiza a classificação da imagem utilizando uma classe pré-determinada para avaliar essas características, como demonstrado na figura 6. (ARAÚJO et. al., 2017)

Figura 6: Demonstração da extração das características de uma imagem pela RNC e a classificação baseadas nessas informações.



A camada é formada por unidades de processamento e essas unidades estão conectadas com a camada anterior e a posterior, por isso o nome “totalmente conectadas”.

É utilizada uma função de ativação chamada de *softmax*. Esta função, recebe um vetor como input e realiza o processamento desse valor analisando a probabilidade de ele pertencer a uma das classes na qual a rede foi treinada.

Outra função muito utilizada é a dropout que reduz o tempo de treinamento, sendo que, ela remove, aleatoriamente a cada iteração de treinamento uma porcentagem das unidades de processamento de uma determinada camada e as recoloca na próxima

camada, evitando overfitting que seria o ajuste excessivo aos dados de treinamento, reduzindo a taxa de acerto. (ARAÚJO et. al., 2017)

PROCESSAMENTO DE IMAGEM

O processamento de imagem consiste em realizar operações matemáticas nos dados, procurando melhorar os aspectos espectrais e espaciais da melhor forma possível, atendendo uma determinada necessidade de uma aplicação. Podem ser separados em dois tipos de técnicas de processamento, sendo elas: de realce e de classificação. Essas técnicas são utilizadas após o pré-processamento, tendo como objetivo corrigir os erros que foram encontrados nos dados processados. (MENESES & ALMEIDA, 2012)

Existem diversas causas para os erros encontrados nas imagens digitais, podendo ser um problema de sensor, que com o passar do tempo acaba perdendo sua precisão ou até o equipamento que compõem o sensor. Outras causas que não envolve a ferramenta utilizada para captura, mas o ambiente que será capturado é a *atmosfera*. Ela pode gerar diversas modificações na imagem, ocorrendo distorções na escala, imprecisão na posição espacial dos pixels, detrimento no contraste ou armazenar incorretamente os valores dos pixels. Ao extrair os dados de uma determinada imagem, na maioria das vezes, é preciso realizar correções para compensar os erros que ela traz, obtendo dados mais precisos. Algumas dessas correções são realizadas em uma das etapas do processamento chamada pré-processamento. Visando corrigir certas imperfeições da imagem, sendo que esses erros são identificados a partir da causa da distorção. Essa etapa é muito importante para o decorrer do processamento, pois, auxilia as próximas fases do processamento, onde são aplicadas as técnicas de transformação, fazendo assim com que as modificações não realçassem as falhas da imagem. (MENESES & ALMEIDA, 2012) (MARTINS, 2017)

Os algoritmos tratam a imagem de forma matemática, com dados diretamente relacionados ao processo físico, porém a representação desse processamento não é totalmente fiel a do mundo real, tendo que ser analisada e interpretada pelo usuário para um melhor entendimento dos dados. Não possuindo um fluxo fixo de processamento para que o usuário siga, dando assim, diversas possibilidades na hora da extração dos dados, atendendo os requisitos necessários naquela situação e que não necessariamente irá atender outras áreas.

Contraste

O contraste é uma medida qualitativa e que está relacionada com a distribuição dos tons de cinza de uma imagem. Esses tons de cinza, às vezes, podem ser difíceis de serem percebidos, pois em determinadas imagens a variação de pixels cinzas possuem uma pequena diferença, sendo complicado perceber os detalhes da textura e do espectro. (MARTINS, 2017)

Por isso existem algumas maneiras básicas para se realçar o contraste da imagem, analisando o ponto de vista espectral e espacial da imagem. Utilizando o espectro da imagem como base é possível realçar o contraste elevando o valor dos pixels em relação ao seu vizinho, desconsiderando o valor que esse pixel vizinho possui.

Baseando-se no espaço, os realces são feitos levando em consideração o valor do pixel vizinho, dependendo assim da distribuição espacial dos pixels que estão próximos ao pixel (ou a uma região de pixels) no qual será feito o reajuste do valor do contraste. Essas técnicas utilizadas para melhorar o realce do contraste das imagens são consideradas transformações de espaço, pois, realizam alterações nos espaços espectrais e espaciais da imagem. Uma das principais técnicas para transformar o espaço da imagem, é o realce de contraste por expansão histográfica (LAKSHMI et. al., 2016).

DESENVOLVIMENTO

O desenvolvimento da aplicação que está sendo realizada neste trabalho foi baseado no trabalho de Mazza (MAZZA, 2016): “Aplicação de redes neurais convolucionais densamente conectadas no processamento digital de imagens para remoção de ruído gaussiano”, onde emprestamos este programa para fazer o estudo do comportamento de uma simulação usando *Deep Learning* para o tratamento de imagens, alternado o contraste. Descrevemos a seguir como é montado um programa em python usando as bibliotecas Keras e TensorFlow.

Tensor Flow

O *tensorflow* é uma biblioteca de código aberto utilizada para computação numérica e aprendizado de máquina, disponibilizado pela Google em novembro de 2015. Atualmente uma das bibliotecas mais utilizadas para o aprendizado profundo, ela reúne uma variedade de modelos e algoritmos de *machine learning* (aprendizado de máquina) e *deep learning* (aprendizado profundo). (HOPE et. al., 2017)

Com esta biblioteca é possível treinar e executar redes neurais profundas em diversas aplicações como a classificação de dígitos manuscritos (*Le Cun*), reconhecimento de imagens, processamento de linguagem natural. Possuindo portabilidade, permitindo que os cálculos e gráficos gerados possam ser executados em uma grande variedade de ambientes e plataformas de hardware, tendo a possibilidade, por exemplo de, com o código idêntico, utilizando a mesma rede neural, realizar o treinamento na nuvem distribuindo em um cluster de várias máquinas ou em um único notebook ().

O núcleo do *tensorflow* é construído em C++, tendo duas interfaces principais de alto nível, com indicadores e interfaces que expressam e executam os gráficos. Sendo desenvolvido o *front-end* em grande parte na linguagem *python*, utilizado pela maioria dos desenvolvedores e cientista de dados. Outro ponto importante é a flexibilidade, que permite expressar os modelos com facilidade, substituindo os blocos da rede por outros, analisando os resultados e projetando novos blocos.

Keras

O *keras* é uma API de redes neurais de alto nível, desenvolvida utilizando a linguagem *python*, podendo ser executado junto ao *tensorflow*. Desenvolvido para se ter uma rápida realização dos testes, possibilitando a criação de protótipos de forma fácil e rápida, tendo suporte a redes neurais convolucionais e recorrentes, podendo ser executado utilizando a CPU e GPU em conjunto ou somente a GPU do computador.

Oferece APIs consistentes e simples, que facilitam no processo de desenvolvimento, dando um *feedback* claro sobre os erros recorrentes. Possuindo modularidade, tal que as camadas neuronais, otimizadores, funções de ativação, entre outros, são módulos independentes que podem ser combinados para a criação de novos módulos ou de novos modelos. A facilidade para adicionar novos módulos, e a grande quantidade de exemplos dos módulos existentes, auxiliam no desenvolvimento. Utilizando-se da linguagem *python* para descrever os modelos, tornando-os mais fáceis para depurar, além de auxiliar na extensibilidade. ()

Criação da base de dados

Para a criação da base de dados (*dataset*) é necessário realizar algumas importações de determinadas bibliotecas (fig 7). A biblioteca *numpy*, que permite a criação de arranjos, vetores e matrizes com *n* dimensões. A biblioteca *matplotlib* utilizada para a geração de gráficos 2D a partir de arranjos, que irão exibir os resultados do treinamento da rede neural. O módulo *os* do próprio *python*, que traz as informações sobre o sistema operacional e serão utilizados para trabalhar com o sistema de arquivos, localizando os diretórios onde estão localizadas as imagens da base de dados. O módulo *cv2*, da biblioteca OpenCV, que possibilita a manipulação de imagens e vídeos. A biblioteca *tqdm*, que possibilita a criação de uma barra de progressão, utilizada para a melhor visualização e verificação do processo de criação da base de dados.

Figura 7: Importações das bibliotecas que serão utilizadas na criação do *dataset*.

```
import numpy as np
import matplotlib.pyplot as plt
import os
import cv2
from tqdm import tqdm
```

Após as importações é definido o diretório onde se encontra as imagens para a construção da base de dados, além da categoria que os dados estão divididos (fig 8). Em seguida é criada uma função que possui uma estrutura de repetição que irá realizar uma varredura no diretório em que será feita a criação do caminho das categorias e a captura do índice dessas categorias. A outra estrutura de repetição contém uma variável que será

responsável por converter o vetor com os dados da imagem, e outra variável para redimensionar a imagem para padronizar o tamanho. Depois de realizar a tratativa dos dados, eles são inseridos dentro do vetor e a base de dados é criada.

Figura 8: Função de construção da base de dados.

```
DIRETORIODATA = "C:/Users/Patrick/Desktop/TCC/Base de Dados/basededados"
CATEGORIA = ["test", "train"]

dados_treinamento = []

def criacao_base_de_dados():
    for category in CATEGORIA:
        path = os.path.join(DIRETORIODATA, category)
        class_num = CATEGORIA.index(category)

        for img in tqdm(os.listdir(path)):
            try:
                img_array = cv2.imread(os.path.join(path, img)
, cv2.IMREAD_GRAYSCALE)
                new_array = cv2.resize(img_array, (IMG_SIZE, IMG_SIZE))
                dados_treinamento.append([new_array, class_num])
            except Exception as e:
                pass

criacao_base_de_dados()

print(len(dados_treinamento))
```

Depois da criação da base de dados é necessário embaralhar os dados, para evitar que predição dos dados seja previsível. Em seguida é feita a criação do modelo neural, utilizando o módulo *pickle*, que permite realizar a serialização do objeto, transformando-os em uma sequencias de bytes. O método *dump* é responsável por capturar os dados do objeto para a geração de uma sequência de bytes que representem esses dados, possibilitando a transmissão pela rede.

Figura 9: Embaralhamento dos dados e criação do modelo.

```
import random

random.shuffle(dados_treinamento)
for sample in dados_treinamento[:10]:
    print(sample[1])

X=[]
y=[]

for features, label in dados_treinamento:
    X.append(features)
y.append(label)

X = np.array(X).reshape(-1, IMG_SIZE, IMG_SIZE, 1)

import pickle

pickle_out = open("X.pickle","wb")
pickle.dump(X, pickle_out)
pickle_out.close()

pickle_out = open("y.pickle","wb")
pickle.dump(y, pickle_out)
..
```

Criação do modelo de rede neural convolucional

Primeiramente é necessário realizar algumas importações de bibliotecas que serão utilizadas na criação da rede neural convolucional. A biblioteca do Keras possui funções que auxiliam na construção da rede neural convolucional, além de disponibilizar alguns *datasets* para servir como base para o treinamento. O modelo *Sequential* é um modelo pré-construído do keras (fig 10), onde serão adicionadas as camadas da rede neural convolucional que são: a Camada de convolução (*Conv2D*), e a Camada de Pooling (*MaxPolling2D*). Também é feita a importação da camada *Dense* usada para predição dos rótulos, a camada de *Dropout* que realiza a redução do *overfitting* (ajuste excessivo aos

dados de treinamento, reduzindo a taxa de acerto), e a camada *Flatten* que tem a função de expandir o vetor tridimensional em um vetor unidimensional.

Figura 10: Importações das bibliotecas para a construção da rede neural.

```
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout, Activation,
Flatten
from tensorflow.keras.layers import Conv2D, MaxPooling2D
from tensorflow.keras.callbacks import TensorBoard
import pickle
import time
```

A variável *batch size* indica o número total de exemplos de treinamento presentes em um único lote, a *num_classes* detém o número de classes que será utilizado do *dataset* (base de dados) e a variável *epochs* refere-se a quantas vezes os dados de entrada serão usados para atualizar os pesos durante o treinamento do modelo. Quando se tem uma época (*epoch*), o modelo vai utilizar o valor da entrada somente uma única vez. Com o aumento desse valor, os pesos serão atualizados ao término de cada época, gerando melhores resultados. As variáveis *img rows* e *imgcols* correspondem a largura e altura da imagem respectivamente (fig 11).

Figura 11: Declaração das variáveis para utilização no modelo.

```
batch_size = 128
num_classes = 10
epochs = 6
img_rows, img_cols = 28, 28
```

Para o treinamento foi utilizado dois *datasets* de tamanhos diferentes, sendo que o primeiro é composto por 180 imagens no total, sendo 150 imagens com contraste distorcido a partir de 30 imagens originais que serviram como base para as respectivas distorções. Esse *dataset* foi criado utilizando a base de dados de imagens CSIQ, famosa para a realização de testes de algoritmos de avaliação da qualidade de imagem e outros aspectos relacionados a qualidade de imagem. (LARSON & CHANDLER, 2010)

O segundo *dataset* é composto por 923 imagens, utilizando 800 imagens de aviões e 123 de carros, utilizando a base de imagens Caltech 101, que é composta por 101 categorias de objetos, sendo que cada categoria tem por volta de 40 a 800 imagens.

Cada um dos *datasets* foram preparados para terem imagens com tamanhos e cores iguais. O tamanho foi fixado em 150 de largura e 150 de altura e as imagens receberam um efeito de *grayscale*, transformando-as em uma escala de cinza. Após as importações, para facilitar o entendimento e uma melhor visualização dos resultados do treinamento foi utilizado a ferramenta do tensorflow chamada TensorBoard. Para utilizar essa ferramenta é necessário importar uma função da biblioteca keras chamada de *callbacks*, um conjunto de funções que são aplicados durante os estágios do treinamento para obter uma visão dos estados internos e estatísticos do modelo.

Figura 12: Carregando a base de dados.

```
NOME = "RNC-Model-ContrastDataset-{}".format(int(time.time()))
tensorboard = TensorBoard(log_dir='logs/{}'.format(NOME))

pickle_in = open("X.pickle","rb")
X = pickle.load(pickle_in)
pickle_in = open("y.pickle","rb")
y = pickle.load(pickle_in)

X = X/255.0
```

É construído um modelo sequencial (*Sequential*), (fig 16) e, adicionado a ele, as camadas de rede neural convolucional, sendo elas a de convolução (*Conv2D*) e de pooling (*MaxPooling*). Também é adicionado no modelo a camada de *dropout* que irá desativar alguns neurônios da rede aleatoriamente, forçando a encontrar novos caminhos, reduzindo assim o sobre ajuste (*overfitting*), e no final é adicionado a camada *dense* usada para prever a classe apresentado na Figura 13.

Figura 13: Construção da rede neural convolucional.

```
model = Sequential()
model.add(Conv2D(32, kernel_size=(3, 3),
                activation='relu',
                input_shape=(28,28,1)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(num_classes, activation='softmax'))
```

Por fim o modelo é compilado, utilizando uma função de perda, um otimizador e a métrica de acerto. Em seguida é feito um ajuste no conjunto de dados usando a função *fit*, treinando o modelo por um número determinado épocas. Ao final do treinamento é feita a avaliação de perda e de acerto do modelo e, então, os valores são exibidos.

Figura 14: Ajuste dos dados e exibição dos resultados.

```
model.compile(loss=keras.losses.categorical_crossentropy,optimizer=k
eras.optimizers.Adadelta(),metrics=['accuracy'])
model.fit(x_train, y_train,
        batch_size=batch_size,
        epochs=epochs,
        verbose=1,
        validation_data=(x_test, y_test))
score = model.evaluate(x_test, y_test, verbose=0)
print('Test loss:', score[0])
print('Test accuracy:', score[1])
```

TESTES E RESULTADOS

Os testes foram realizados com uma variação na estrutura do modelo, alterando a quantidade de camadas de convolução e profundidade e o número de nós.

Para o treinamento da *RNC* foram utilizados os dois *datasets* criados, utilizando o período de 30 épocas e posteriormente de 50 épocas para analisar a curva de aprendizado da rede nessa progressão de treino. Sendo utilizados para o primeiro *dataset* 125 exemplos para treinamento e 55 para validar e o segundo *dataset* 646 exemplos para testes e 277 para validar. Com o primeiro *dataset* foram obtidos melhores resultados utilizando o modelo com 1 camada de convolução, 64 nós e 1 camada de profundidade. Os modelos com duas camadas de convolução tiveram resultados apresentados na Tabela 1, concluindo que neste caso, o aumento de camadas convolucionais não necessariamente trará melhores resultados na aprendizagem. Com o aumento do número de épocas, os resultados melhoraram, tendo uma acurácia de 1.0 em todos os modelos e uma consequente diminuição na perda.

Figura 15: Gráfico do modelo de treinamento com melhor resultado, que demonstra a progressão do treinamento, tendo no eixo x o número de épocas e no eixo y os valores que vão de 0.00 a 1.00.



Tabela 1: Resultados de acurácia e perda no aprendizado utilizando o primeiro *dataset*, com o período de 30 épocas.

<i>Modelo de treinamento</i>	<i>Acurácia</i>	<i>Perda</i>	<i>Tempo</i>
<i>Conv=1 Nós=32 Prof=0</i>	0.96	0.09	6m 46s
<i>Conv=1 Nós=64 Prof=0</i>	1.0	0.03	8m 51s
<i>Conv=1 Nós=32 Prof=1</i>	1.0	0.04	9m 32s
<i>Conv=1 Nós=64 Prof=1</i>	1.0	0.01	12m 29s
<i>Conv=2 Nós=32 Prof=0</i>	0.93	0.16	6m 20s

<i>Conv=2 Nós=64 Prof=0</i>	0.95	0.11	8m 15s
<i>Conv=2 Nós=32 Prof=1</i>	0.88	0.25	9m 58s
<i>Conv=2 Nós=64 Prof=1</i>	0.97	0.06	12m 43s

Figura 16: Um dos gráficos do modelo de treinamento com melhor resultado, que demonstra a progressão do treinamento, tendo no eixo x o número de épocas e no eixo y os valores que vão de 0.00 a 1.00.



Tabela 2: Resultados de acurácia e perda no aprendizado utilizando o primeiro *dataset*, com o período de 50 épocas.

Modelo de treinamento	Acurácia	Perda	Tempo
<i>Conv=1 Nós=32 Prof=0</i>	1.0	0.03	9m 34s
<i>Conv=1 Nós=64 Prof=0</i>	1.0	0.01	13m 7s
<i>Conv=1 Nós=32 Prof=1</i>	1.0	0.00	13m 48s
<i>Conv=1 Nós=64 Prof=1</i>	1.0	0.00	18m 52s
<i>Conv=2 Nós=32 Prof=0</i>	1.0	0.04	11m16s
<i>Conv=2 Nós=64 Prof=0</i>	1.0	0.01	14m26s
<i>Conv=2 Nós=32 Prof=1</i>	1.0	0.01	14m 58s
<i>Conv=2 Nós=64 Prof=1</i>	1.0	0.00	18m 38s

Com o segundo *dataset*, apresentado nas Tabelas 3 e 4, os resultados tiveram uma média tanto da acurácia quanto da perda, se mantendo entre 1.0 e 0.0 respectivamente, tendo uma curva de aprendizado ao passar das épocas crescendo e estabilizando, da mesma forma a perda que diminuiu nas primeiras épocas e estabilizou. Isso se repetiu com o aumento no período de épocas.

Figura 17: Um dos gráficos do modelo de treinamento com melhor resultado, que demonstra a progressão do treinamento, tendo no eixo x o número de épocas e no eixo y os valores que vão de 0.00 a 1.00.



Tabela 3: Resultados de acurácia e perda no aprendizado utilizando o segundo dataset, com o período de 30 épocas.

<i>Modelo de treinamento</i>	<i>Acurácia</i>	<i>Perda</i>	<i>Tempo</i>
<i>Conv=1 Nós=32 Prof=0</i>	1.0	0.0	29m38s
<i>Conv=1 Nós=64 Prof=0</i>	1.0	0.0	40m 54s
<i>Conv=1 Nós=32 Prof=1</i>	1.0	0.0	40m 23s
<i>Conv=1 Nós=64 Prof=1</i>	1.0	0.0	59m 8s
<i>Conv=2 Nós=32 Prof=0</i>	1.0	0.0	30m 33s
<i>Conv=2 Nós=64 Prof=0</i>	1.0	0.0	42m 46s
<i>Conv=2 Nós=32 Prof=1</i>	1.0	0.0	1h 4m 27s
<i>Conv=2 Nós=64 Prof=1</i>	1.0	0.0	1h 3m 39s

Figura 18: Um dos gráficos do modelo de treinamento com melhor resultado, que demonstra a progressão do treinamento, tendo no eixo x o número de épocas e no eixo y os valores que vão de 0.00 a 1.00.

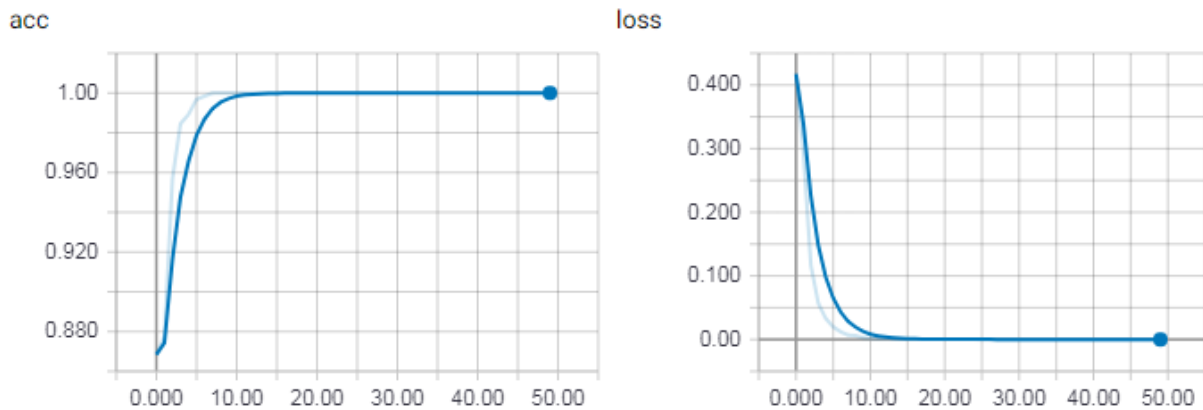


Tabela 4: Resultados de acurácia e perda no aprendizado utilizando o segundo *dataset*, com o período de 50 épocas.

<i>Modelo de treinamento</i>	<i>Acurácia</i>	<i>Perda</i>	<i>Tempo</i>
<i>Conv=1 Nós=32 Prof=0</i>	1.0	0.0	54m 19s
<i>Conv=1 Nós=64 Prof=0</i>	1.0	0.0	1h 13m 24s
<i>Conv=1 Nós=32 Prof=1</i>	1.0	0.0	1h 12m 24s
<i>Conv=1 Nós=64 Prof=1</i>	1.0	0.0	1h 27m 45s
<i>Conv=2 Nós=32 Prof=0</i>	1.0	0.0	53m 24s
<i>Conv=2 Nós=64 Prof=0</i>	1.0	0.0	1h 05m 59s
<i>Conv=2 Nós=32 Prof=1</i>	1.0	0.0	1h 31m 03s
<i>Conv=2 Nós=64 Prof=1</i>	1.0	0.0	1h 43m 10s

CONCLUSÕES

Neste trabalho foi apresentado o aprendizado profundo utilizando redes neurais convolucionais para o processamento de imagens com variação de contraste. Foram realizados diversos testes com modelos de rede neural convolucional baseado no modelo de rede neural do trabalho do Ferreira (FERREIRA. 2017): “Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja”, com uma quantidade de camadas variadas, utilizando dois tipos de *datasets*. O primeiro foi construído com uma base de imagens menor, sendo composto por imagens de contraste distorcido. O segundo *dataset* utilizou uma base de dados maior, com imagens possuindo uma variação no contraste.

Esses modelos foram comparados para identificar o modelo com melhor resultado de aprendizado. Os modelos treinados com o *dataset* 1, treinados em um período com o menor número de épocas, teve resultados melhores com redes que tinham uma quantidade de camadas convolucionais menor, demonstrando que em alguns casos o aumento no número de camadas convolucionais não necessariamente trará uma melhora nos resultados. Com o aumento no período das épocas os resultados foram equivalentes em todos os modelos testados, indicando que quanto maior o período, melhor será o aprendizado.

O treinamento feito utilizando o *dataset* 2 teve resultados que foram semelhantes nos dois períodos de épocas, com modelos com quantidade de camadas diferentes, mostrando que, o aumento da base de dados auxilia numa melhor progressão de aprendizagem da rede neural. Pode se perceber com os resultados de treinamento do *dataset* 2 aconteceu *overfitting*, uma especialização por parte da rede neural, pela baixa quantidade de exemplos de treinamento. Conclui-se que a utilização das redes neurais convolucionais possuem um bom desempenho para análise e processamento de imagens, podendo sugerir como trabalho futuro a extração de característica das imagens, utilizando descritores de imagens, para selecionar os melhores parâmetros gerando assim de forma eficiente e rápida, imagens com uma melhor qualidade.

REFERÊNCIAS BIBLIOGRÁFICAS

ARAÚJO, Flávio H.D; CARNEIRO, Allan C; SILVA, Romuere R.V; MEDEIROS, Fátima N.S; USHIZIMA, Daniela M; Redes Neurais Convolucionais com Tensorflow: Teoria e Prática. III Escola Regional de Informática do Piauí, 2017.

BEZERRA, Eduardo. Introdução à Aprendizagem Profunda, 2016.

FERREIRA, Alessandro dos Santos. Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja. Campo Grande: UFMS, 2017.

HAYKIN, S. Neural Networks and Learning Machines, Prentice Hall, 2009.

HOPE, TOM; RESHEFF, Y; LIEDER, I; Learning TensorFlow: A guide to building deep learning systems", 2017.

LARSON, E. C.; CHANDLER, D. M. Most apparent distortion: full-reference

Image quality assessment and the role of strategy. *Journal of Electronic Imaging*, International Society for Optics and Photonics, v. 19, n. 1, p. 011006-011006, 2010.

LAKSHMI, Hyma T.V; MADHU, T; KAVYA, Sri; DEVI, Geetha. Image Resolution and Contrast Enhancement Using Wavelet Transforms and Contrast Limited Adaptive Histogram Equalization. *International Journal of Computer Science and Information Security*, 2016.

MAZZA, Leonardo Oliveira. Aplicação de redes neurais convolucionais densamente conectadas no processamento digital de imagens para remoção de ruído gaussiano. Rio de Janeiro, 2016.

MARTINS, Samuel Botter. Introdução ao Processamento Digital de Imagens: Definições Básicas, Espaço de Cores e Histogramas. Universidade Estadual de Campinas.

MENESES, Paulo Roberto; ALMEIDA, Tati. Introdução ao processamento de imagens de sensoriamento remoto. Brasília, 2012.

VARGAS, A.; PAES, A; VASCONCELOS, C; Um Estudo sobre Redes Neurais Convolucionais e sua Aplicação em Detecção de Pedestres. Disponível em <http://sibgrapi.sid.inpe.br/col/sid.inpe.br/sibgrapi/2016/09.12.15.44/doc/um-estudosobre.pdf> 2016.

VON ZUBEN, F.J. Síntese automática de redes neurais artificiais com conexões arbitrárias, Universidade Estadual de Campinas, 2015.