

SISTEMA DE BUSCA ESTRUTURADA POR PLÁGIO ACADÊMICO COM USO DA WEB

Ronney Luan Albuquerque Bezerra – rlab@ecomp.poli.br
Universidade de Pernambuco, Escola Politécnica de Pernambuco
Rua Benfca, 455
50.720-001 – Recife – Pernambuco

Sérgio Campello Oliveira – scampello@upe.br
Universidade de Pernambuco, Escola Politécnica de Pernambuco
Rua Benfca, 455
50.720-001 – Recife – Pernambuco

Resumo: *O plágio acadêmico é um dos problemas que prejudica a pesquisa científica. Com o surgimento e posterior popularização da internet, esse problema tornou-se ainda mais evidente, e possivelmente mais intenso, dada a facilidade de acesso às publicações existentes e a falsa noção de anonimato que o mundo virtual propicia. Da mesma forma que a internet facilita o plágio ela pode auxiliar no seu combate, ajudando a identificar a ocorrência e a punir os responsáveis. Softwares de identificação de plágio têm um papel fundamental nesse combate, auxiliando professores no processo de identificação, contudo há custos ou limitações de uso. Este artigo apresenta uma alternativa aos softwares existentes, descrevendo uma metodologia e proposta de aplicativo para identificação de plágio com baixo uso de dados.*

Palavras-chave: *Plágio acadêmico. Software de identificação de plágio.*

1. INTRODUÇÃO

O plágio acadêmico é uma prática que remonta à antiguidade romana (KROKOSCZ, 2011) e com o passar do tempo não só se manteve ativa, como, aparentemente, se tornou generalizada. “Quem o pratica, foge totalmente de um dos pilares da pesquisa científica que é a ética” (KROKOSCZ, 2011). Com a facilidade de acesso à informação propiciada pela internet, a prática tem crescido e vem sendo enfrentado por universidades em todo o mundo (KROKOSCZ, 2011).

Por se tratar de um problema de ordem ética, campanhas educativas são de grande importância na prevenção do plágio em novas publicações. Mas como o problema já ocorre em grande parte das universidades pelo mundo, além de campanhas preventivas é importante combatê-lo de maneira eficiente. O primeiro passo no combate é a identificação do problema por parte dos primeiros responsáveis pela revisão do texto, em geral o professor ou o orientador. Pelo conhecimento que têm do aluno, os professores podem desconfiar e, diante disso, verificar de forma mais atenta o trabalho (ARAÚJO, 2017).

É possível identificar trechos de plágio com o auxílio de *softwares* que fazem varreduras nos textos e buscam por conteúdos iguais ou semelhantes publicados na internet. Contudo, as ferramentas disponíveis para auxílio na identificação e plágio são, muitas vezes, pagas e de alto custo e, dentre as gratuitas, há muitas restrições de uso. Porém, as buscas efetuadas pelos alunos para localizar textos usam, na sua grande maioria, ferramentas de buscas gratuitas. Essas mesmas ferramentas podem, por sua vez, serem utilizadas para buscar

trechos do texto e identificar, de forma efetiva, os locais com texto igual ou muito semelhante àquele copiado por parte dos plagiadores.

Este artigo apresenta uma metodologia eficiente na busca de plágio a partir de trechos retirados do texto sob suspeita e com o auxílio de ferramentas de busca da internet. Uma vez que a metodologia mostrou-se eficiente, este artigo descreve um aplicativo em desenvolvimento para automatizar a metodologia, reduzindo o tempo gasto na busca além de aumentar a quantidade de buscas, tornando a metodologia ainda mais eficiente.

O artigo segue organizado em quatro seções. A Seção 2 apresenta uma revisão da literatura sobre a temática do plágio acadêmico e ferramentas de combate. A Seção 3 descreve o algoritmo utilizado na metodologia para identificação de plágio em textos acadêmicos. Na Seção 4 está a proposta de um sistema web que implementa o algoritmo proposto. Por fim a Seção 5 analisa resultados obtidos com a execução do algoritmo.

2. PLÁGIO

ARAÚJO (2017) explica que são inegáveis os benefícios trazidos pela pesquisa científica. Através dos resultados obtidos pelas pesquisas, grandes melhorias são alcançadas para a sociedade oferecendo diversas soluções para um determinado problema. Para alunos, a pesquisa é importante tanto para os que decidem continuar com a carreira acadêmica, pois será objeto do seu trabalho diário, quanto para os que decidem seguir uma carreira profissional fora da universidade, pois contribui para um melhor embasamento teórico e pensamento crítico.

Entre os muitos benefícios trazidos pela internet, pode-se citar com destaque a facilidade com a qual pesquisas científicas são realizadas em relação ao período pré-internet, pois para a pesquisa científica a comunicação de seus resultados é tão importante quanto a pesquisa em si (ARAÚJO, 2017). Neste contexto a internet cumpre um papel de grande importância na disseminação do conhecimento, possibilitando que a comunidade científica tenha acesso a esses conteúdos de maneira fácil e rápida. Hoje pesquisas que seriam de difícil acesso, ou publicadas em uma região geograficamente distante para alguns pesquisadores e estudantes estão disponíveis de maneira ubíqua.

Dada a importância da pesquisa científica, e pela sua própria natureza metodológica, se faz necessário observar o que colabora para que a pesquisa evolua e seja difundida, visando seu fortalecimento, e, da mesma forma, analisar quais são os problemas que dificultam sua evolução. Dentre os vários problemas enfrentados pela comunidade científica, o plágio acadêmico é um dos mais antigos e que continua sendo praticado (KROKOSZ, 2011). O plágio pode ser definido como “a utilização de obra, ou parte dela, de outrem assumindo-se como sua a autoria” (KROKOSZ, 2011).

São muitos os fatores que colaboram com a manutenção desta prática dentro das universidades, dentre eles a internet colabora com a expansão da problemática, “pois a facilidade de acesso e manipulação da informação e noção de que no campo virtual as regras e os princípios são diferentes, vêm intensificando a reprodução inescrupulosa das palavras e ideias de outros autores como se fossem próprias” (ARAÚJO, 2017). Porém não se pode atribuir culpa a uma tecnologia, já que se trata de um problema ético e da ação humana, sendo a internet apenas um facilitador.

O enfrentamento deste problema se inicia com a abordagem do tema dentro da própria universidade, por exemplo, as principais universidades da Oceania, Europa e América o plágio é discutido “de forma diversificada e abrangente, envolvendo medidas institucionais, preventivas, diagnósticas e corretivas” (ARAÚJO, 2017). No Brasil, contudo, o tema não é abordado de maneira tão efetiva, o que torna ainda mais importante a discussão sobre ele no

cenário nacional. A primeira etapa desse processo é identificar a presença do plágio nas publicações, o que pode ser difícil, pois são várias as formas em que ele pode ser praticado. O conhecimento dos alunos por parte dos professores auxilia na suspeita e posterior identificação da prática, mas a utilização de *softwares* tem sido uma grande aliada nesse combate, agilizando o processo de identificação (ARAÚJO, 2017).

Os *softwares* de identificação de plágio fazem buscas na *web*, fazendo varreduras sobre todo o texto ou em partes específicas tentando encontrar conteúdos idênticos ou muito semelhantes aos encontrados no texto em questão. O professor fica responsável por avaliar os resultados retornados pela ferramenta e verificar se de fato houve plágio ou não. Dentre os programas mais conhecidos na detecção de plágio a literatura destaca oito deles (EDITORIAL DOM BOSCO, 2015):

- Turnitin - <http://www.turnitin.com/>
- iThenticate - <http://www.ithenticate.com/>
- Plagiarism detect - <http://plagiarism-detect.com/>
- Plagius - <http://www.plagius.com/s/br/default.aspx>
- Ephorus - <https://www.ephorus.com/pt/home>
- Farejador de Plágio - <http://www.plagiarismcombat.com/>
- DOC Cop - <http://www.doccop.com/index.html?nc=34411419>
- Plagium - <http://www.plagium.com/pt/detectordeplagio>

Dos *softwares* listados é possível observar que alguns deles não são de uso gratuito, como Turnitin e iThenticate, outros como Plagius e Farejador de Plágio precisam ser baixados e instalados. Plagiarism detect estava fora do ar, durante as pesquisas realizadas para a escrita deste artigo, e o Ephorus foi fundido ao Turnitin. Essas características de alguma forma limitam a utilização dos *softwares* listados, seja por questão financeira, para os que são pagos, seja por dificuldade de instalação e limitação de sistemas operacionais, para os que precisam ser baixados e instalados, sendo que, alguns desses, disponibilizam apenas parte de seus recursos de maneira gratuita. Essas limitações aliadas a usabilidade confusa em alguns casos e pouca eficiência do algoritmo em outros, evidenciam a necessidade do desenvolvimento de alternativas gratuitas e eficientes.

3. ALGORITMO ALTERNATIVO PARA DETECÇÃO DE PLÁGIO

Conforme já mencionado, o advento da internet serviu como um impulsionador para a prática do plágio, uma vez que ela fornece acesso à informação de maneira rápida e ubíqua. A forma mais comum de buscar por essas informações é com o uso de palavras chave em motores de busca disponíveis por provedores de conteúdo. As buscas efetuadas pelos alunos para localizar textos usam, na sua grande maioria, esses motores de buscas gratuitas.

Da forma explicitada por ARAÚJO (2017), professores e orientadores conseguem, pela convivência próxima com os alunos, identificar trechos do texto com características de escrita diferentes daquelas usadas corriqueiramente pelos estudantes. Na prática diária a 11 anos de ensino na disciplina de Projeto de Final de Curso, um dos autores deste trabalho corrobora a afirmação de ARAÚJO (2017).

Uma vez identificado o trecho sob suspeita, os mesmos motores de busca gratuita utilizados pelos alunos podem, por sua vez, serem utilizados, por parte do professor, para buscar trechos do texto e identificar, de forma efetiva, os locais com texto igual ou muito semelhante àquele copiado por parte dos plagiadores.

A forma como funcionam os motores de busca permite a busca por trechos exatos (colocados entre aspas) do texto. Ao longo dos 11 anos de docência citados, a metodologia mostrou-se efetiva na identificação do plágio. Contudo, o trabalho manual necessário para a efetiva identificação pode ser longo, pois por vezes, os alunos alteram palavras isoladas no trecho copiado, o que pode exigir a repetição das buscas em diversos trechos do texto até que uma das buscas retorne *links* para os locais onde o texto está disponível.

O algoritmo que se propõe neste trabalho consiste em coletar do usuário informações sobre a forma de seleção de trechos do texto sob investigação da prática de plágio, como tamanho das frases, número de frases a serem selecionadas e em qual parte do texto a seleção deve ser feita. Com essas informações alguns trechos do texto são escolhidos, de forma aleatória, e utilizados em buscas na web por sua forma exata. Por fim, os *links* retornados nas buscas são associados aos trechos e mostrados ao usuário. A Figura 1 ilustra o funcionamento do algoritmo. Em outras palavras, este algoritmo pretende automatizar o processo de busca descrito anteriormente facilitando e tornando mais eficiente o processo já utilizado de modo manual.

Inicialmente o usuário fará um *upload* do texto que se deseja investigar. Em seguida deve ser informado qual o trecho do texto que será analisado. Isso é importante pois, em grande parte dos trabalhos, as cópias se concentram nas seções de revisão da literatura e fundamentação teórica. A seleção de trechos será realizada, então, apenas na parte do texto marcada. A quantidade e o tamanho (em número de palavras) dos trechos, parâmetros passados pelo usuário, serão selecionados e colocados, de forma sequencial e entre aspas, em motores de busca gratuitos. Ao se buscar por trechos entre aspas os motores de busca retornam *links* onde há coincidências exatas do trecho buscado ou retornam a mensagem de que não foram encontrados locais com o trecho exato buscado.

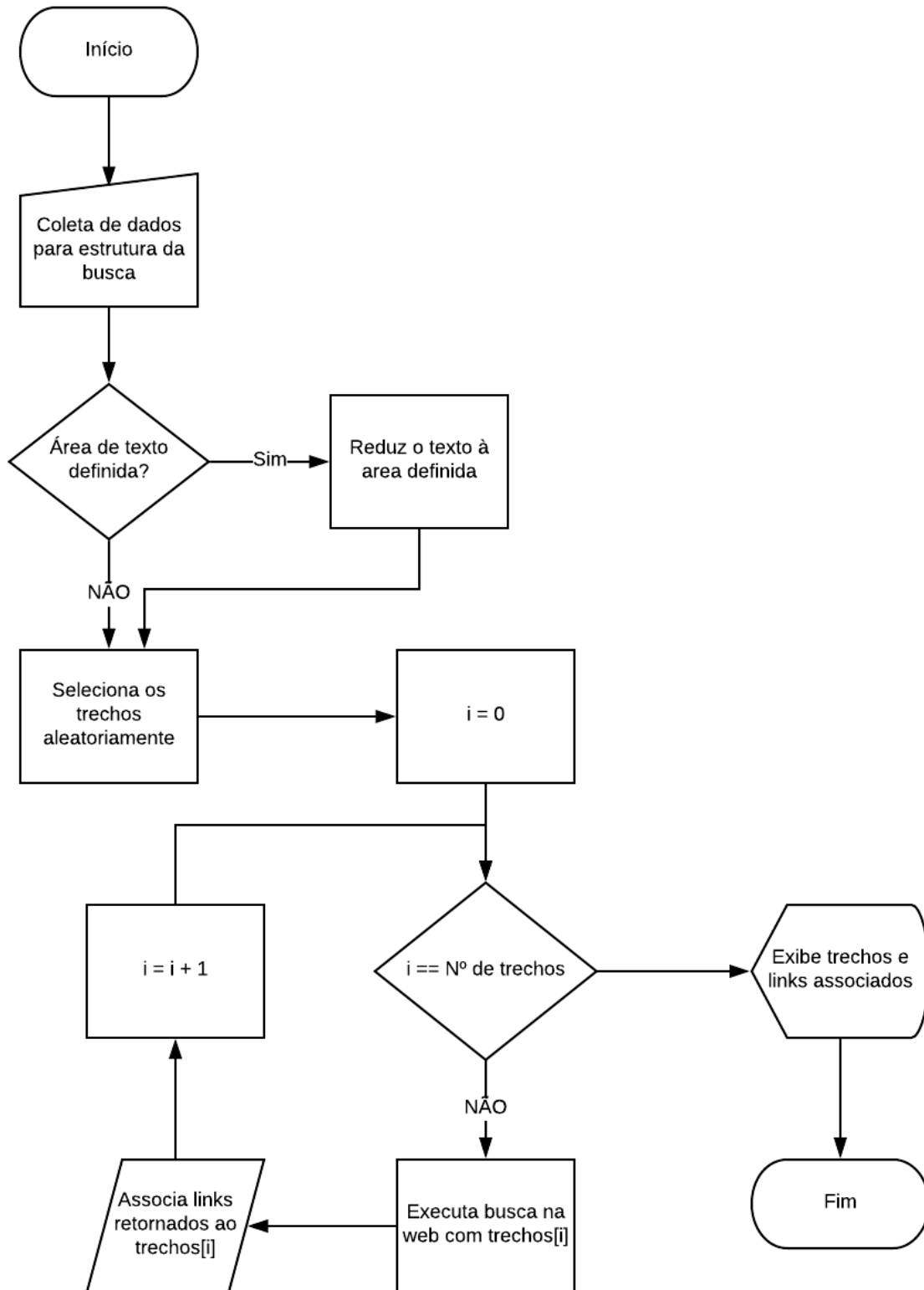
O aplicativo irá associar os trechos buscados com os *links* ou com a ausência deles para que o usuário, professor responsável, possa analisar individualmente os resultados. Caso nenhum *link* seja retornado é bastante provável que não tenha ocorrido plágio. Caso diversos *links* para diferentes trechos sejam encontrados é bastante provável que o plágio tenha ocorrido. Caso poucos *links* sejam retornados, uma investigação mais aprofundada, possível pois os *links* serão disponibilizados, merece ser realizada para afirmar sobre a ocorrência ou não do plágio.

4. SISTEMA DE BUSCA ESTRUTURADA

“O objetivo principal do desenvolvimento de um produto de *software* para alguns autores é obter como resultado um sistema de alta qualidade” (SILVA, 2016). Baseado nesta afirmação é importante que ao se iniciar o processo de desenvolvimento de um *software*, seja levado em consideração aspectos de qualidade, tais como usabilidade, cobertura de todos os requisitos e performance dentro da esperada pelos usuários. SOUZA (2011) cita que uma das definições para qualidade de *software* é “o grau pelo qual um sistema, um componente ou um processo satisfazem seus requisitos especificados, e as necessidades ou expectativas de clientes ou usuários”.

O sistema proposto será desenvolvido para funcionamento em plataforma *web*, possibilitando o acesso por qualquer computador com acesso à internet. Na fase de análise foram desenvolvidos diagramas de caso de uso e de classe utilizando Linguagem de Modelagem Unificada (UML) (CHONOLLES, 2003) com auxílio da ferramenta *Lucidchart* (2018) e protótipos de telas com auxílio da ferramenta *Axure* (2018). Esses documentos são base para o desenvolvimento da aplicação em si.

Figura 1 - Fluxograma do algoritmo de busca de trechos de texto na web.

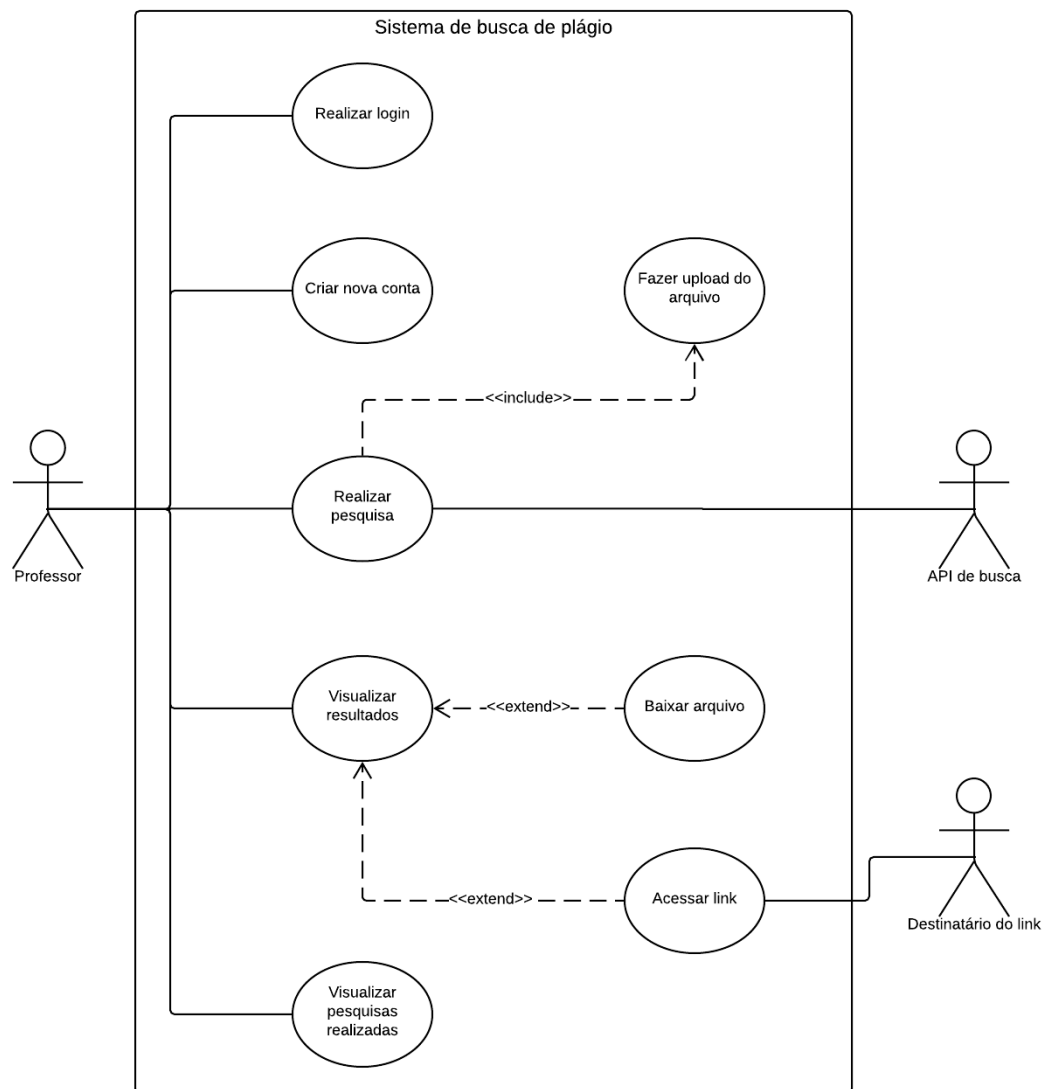


Fonte: Elaboração Própria.

4.1. Diagrama de casos de uso

Como definido em CHONOLES (2003), casos de uso são propósitos que podem ser alcançados pelos usuários utilizando um sistema. Esses propósitos são convertidos em funcionalidades e o diagrama de caso de uso é a ferramenta UML utilizada para representar essas funcionalidades de maneira clara e direta. A Figura 2 traz o diagrama de caso de uso do sistema proposto neste artigo.

Figura 2 - Diagrama de caso de uso do sistema de busca por plágio.



Fonte: Elaboração Própria.

Além da principal funcionalidade do sistema que é realizar a pesquisa, o diagrama ilustra outras funcionalidades como visualizar o conteúdo dos *links* retornados na busca, criar uma conta de acesso ao sistema, fazer *login* no sistema e visualizar a lista de pesquisas realizadas anteriormente.

4.2. Protótipos de telas

Baseado nos casos de uso identificados, foram construídos protótipos para as principais telas do sistema. A Figura 3 ilustra a tela de lista de pesquisas realizadas pelo

professor, mostrando as configurações da busca e a data em que foi feita. O histórico facilitará o compartilhamento de informações entre os demais atores envolvidos, como os demais professores de uma eventual banca e o próprio aluno. Uma vez que se pretende instalar o aplicativo nos servidores da própria universidade seus resultados poderão servir como ferramenta acadêmica para intervenções de combate ao plágio e definição dos mecanismos oficiais de procedimentos após sua identificação.

Figura 3 - Protótipo da tela de lista de pesquisas realizadas.

Data	Número de Frases	Tamanho das frases	Número total de links	Tipo da pesquisa
01/01/2018	3	4 Palavras	0	Todo texto
01/01/2018	3	4 Palavras	10	Início do texto
01/01/2018	4	4 Palavras	2	Todo texto
01/01/2018	5	4 Palavras	7	Todo texto
01/01/2018	2	4 Palavras	5	Início do texto
01/01/2018	4	4 Palavras	6	Todo texto
01/01/2018	5	4 Palavras	3	Todo texto
01/01/2018	6	4 Palavras	8	Todo texto
01/01/2018	7	4 Palavras	9	Início do texto
01/01/2018	6	4 Palavras	23	Início do texto

Fonte: Elaboração Própria.

A Figura 4 ilustra a tela de configuração para realizar novas pesquisas. Haverá um local para fazer o upload, numa primeira versão, de um arquivo para averiguação. O número e o tamanho da frase, medido em palavras, poderão ser passados como parâmetro, para a realização de buscas. O percentual e a região onde as buscas devem se concentrar também serão passados como parâmetros. Para melhorar a usabilidade, a última configuração usada será apresentada como opção de preenchimento automático, evitando ações repetitivas no uso do aplicativo.

Figura 4 - Protótipo da tela de realizar nova pesquisa.

eCOMP UPE - POLI
Engenharia de Computação

Nova Pesquisa

UPE
UNIVERSIDADE
DE PERNAMBUCO

Arquivo
Clique para fazer upload

Tamanho da frase (Nº de palavras)

Nº de frases

Percentual do texto
20%

Área do texto
Início

Cancelar Confirmar

Fonte: Elaboração Própria.

A Figura 5 ilustra a tela de resultado de uma pesquisa realizada. Nela será possível observar os diversos *links* encontrados para cada trecho pesquisado. Clicando em cada link será possível visualizar o conteúdo e comparar com o trecho selecionado, bem como com o restante do parágrafo, de modo a deixar evidente o plágio realizado.

Figura 5 - Protótipo da tela de resultado de uma pesquisa.

eCOMP UPE - POLI
Engenharia de Computação

Resultado da pesquisa

UPE
UNIVERSIDADE
DE PERNAMBUCO

Data: 01/01/2018

- ▼ "Frase selecionada no texto 1" (5 links encontrados)
 - <https://link-resultado-1>
 - <https://link-resultado-2>
 - <https://link-resultado-3>
 - <https://link-resultado-4>
 - <https://link-resultado-5>
- ▶ "Frase selecionada no texto 2" (10 links encontrados)
- ▶ "Frase selecionada no texto 3" (Nenhum link encontrado)
- ▼ "Frase selecionada no texto 4" (3 links encontrados)
 - <https://link-resultado-1>
 - <https://link-resultado-2>
 - <https://link-resultado-3>

Voltar

Fonte: Elaboração Própria.

4.3. Implementação do sistema

O sistema foi implementado para funcionamento na plataforma web com a linguagem de programação *Ruby* utilizando o framework *Ruby on Rails* e banco de dados *PostgreSQL*. A implementação foi baseada nos documentos listados anteriormente, cobrindo os casos de uso identificados e com uma interface muito próxima à proposta nos protótipos desenvolvidos. O controle de versão do código foi feito utilizando o sistema *GIT*, mantendo o histórico de mudanças realizadas ao longo do tempo.

A arquitetura seguida foi a *Model View Controller (MVC)* que é facilitada com o uso do *Ruby on Rails*. Para desenvolvimento das telas foi utilizada a biblioteca *Bootstrap* que provê uma série de componentes para composição das telas com uma interface agradável e funcional, além disso é uma biblioteca muito popular o que facilita futuras mudanças feitas por outros desenvolvedores.

Para realizar as buscas utilizando os trechos selecionados aleatoriamente pelo algoritmo, foi feita uma integração com uma *Application Programming Interface (API)* externa ao sistema desenvolvido, provida por um motor de busca. Inicialmente pretendia-se utilizar a API do *Google Search*, porém não é mais possível realizar buscas genéricas utilizando essa ferramenta, atualmente é necessário determinar um grupo específico de domínios sobre os quais a busca deve ser realizada, o que impossibilita realizar uma busca mais abrangente. Como alternativa o sistema de buscas do *Yahoo* foi cogitada, porém a sua API foi descontinuada em 31/03/2016. A opção utilizada foi a API de buscas da *Azure*, desenvolvida pela *Microsoft*, que permite realizar buscas sem limitação de domínios. Nenhuma outra ferramenta de busca além das citadas foi verificada.

Para um funcionamento mais preciso do algoritmo é feita uma remoção de referências no texto, de acordo com o padrão de referências da ABNT, antes de selecionar os trechos e realizar as buscas na *WEB*.

A publicação do sistemas para testes foi realizada utilizando o *Heroku*, ferramenta que permite fazer *deploy* de sistemas com muita facilidade, já integrado com o sistema de controle de versão e com pouca necessidade de configuração.

5. RESULTADOS

Para validação do algoritmo foram realizados testes com três grupos de textos, cada um com cinco arquivos:

- Grupo 1: textos com plágio conhecido;
- Grupo 2: textos sem nenhuma informação de plágio, ainda não publicados, ou seja, podem ter plágio ou não;
- Grupo 3: textos sem nenhuma cópia e ainda não publicados.

Para cada grupo foram realizados 3 testes utilizando a aplicação descrita anteriormente e para cada teste foram utilizadas configurações diferentes. No primeiro teste foram selecionadas 10 frases, cada uma delas com 10 palavras, selecionadas em 50% do texto a partir do início das publicações. No segundo teste foram selecionadas 5 frases e no terceiro foi selecionada apenas uma frase. As demais configurações se repetem para os 3 testes.

As tabelas 1, 2 e 3 trazem os resultados obtidos nos 3 testes realizados em cada grupo:

Tabela 1 - Resultados obtidos no Grupo 1

	Teste 1		Teste 2		Teste 3	
	Nº de frases	Links retornados	Nº de frases	Links retornados	Nº de frases	Links retornados
Publicação 1	10	46	5	39	1	10
Publicação 2	10	38	5	22	1	10
Publicação 3	10	27	5	7	1	0
Publicação 4	10	73	5	36	1	10
Publicação 5	10	50	5	26	1	1

Fonte: Elaboração Própria.

Tabela 2 - Resultados obtidos no Grupo 2

	Teste 1		Teste 2		Teste 3	
	Nº de frases	Links retornados	Nº de frases	Links retornados	Nº de frases	Links retornados
Publicação 1	10	5	5	2	1	0
Publicação 2	10	10	5	0	1	0
Publicação 3	10	2	5	0	1	1
Publicação 4	10	3	5	0	1	0
Publicação 5	10	0	5	0	1	0

Fonte: Elaboração Própria.

Tabela 3 - Resultados obtidos no Grupo 3

	Teste 1		Teste 2		Teste 3	
	Nº de frases	Links retornados	Nº de frases	Links retornados	Nº de frases	Links retornados
Publicação 1	10	1	5	0	1	0
Publicação 2	10	0	5	1	1	0
Publicação 3	10	4	5	0	1	0
Publicação 4	10	0	5	0	1	0
Publicação 5	10	5	5	0	1	0

Fonte: Elaboração Própria.

Pode-se observar que o grupo com textos de plágio conhecido retornaram *links* para todos os tamanhos de frases definidos nos testes aqui realizados e para todos os trechos selecionados, com exceção de apenas uma publicação no terceiro teste realizado para o grupo. O segundo grupo teve *links* retornados para 4 das 5 publicações, tendo uma delas retornado mais de 10, sendo todos de uma mesma frase. O terceiro grupo teve um número muito menor de *links* retornados, para 4 das 5 publicações, nenhuma delas retornando mais de 5, sendo todos eles de trechos de referências, o que a princípio descaracterizaria o plágio. Contudo, é preciso que o professor observe se as normas para transcrição foram seguidas, deixando claro que as palavras usadas são transcrições.

6. CONCLUSÃO

Como observado neste artigo, o problema do plágio acadêmico merece atenção devido ao prejuízo causado à pesquisa de uma forma geral. Embora existam ferramentas que auxiliam a identificação do plágio, limitações, como o fato de algumas não serem gratuitas, abrem espaço para o surgimento de alternativas.

As universidades necessitam de mecanismos eficazes para a identificação de plágio para subsidiar suas políticas de combate. O algoritmo aqui proposto mostrou-se útil e eficaz. O aplicativo proposto também mostra-se viável e poderá ser facilmente instalado nos servidores de serviços de internet das universidades. Uma vez que as buscas serão realizadas com acesso registrado e realizado dentro das intranets, todo o processo poderá ser auditável e por isso confiável para servir como subsídio necessário para as políticas anti-plágio de universidades.

Com as estatísticas observadas a partir da primeira versão do sistema, pretendem-se novos desenvolvimentos como, por exemplo, testes de diferentes motores de busca e análises de evolução de detecção, redução ou de identificação de novas práticas de plágio.

Para trabalhos futuros pode-se continuar aprimorando o algoritmo de extração de trechos do texto enviado pelo usuário para serem utilizados nas buscas. Foi verificado nos testes que trabalhos com sumário e muitas referências tendem a ter trechos selecionados que não retornam resultados úteis, nesses casos a área de seleção dos trechos se torna uma entrada importante. O algoritmo de remoção de referências pode ser melhorado buscando remover diferentes tipos de referências e citações.

Além disso é importante verificar alternativas para a ferramenta de busca, analisando se outras retornam resultados mais precisos e se a atual está realizando a busca da maneira esperada. Experimentos podem ser feitos para verificar se o sistema aqui implementado tem resultados melhores ou não do que outros existentes.

REFERÊNCIAS

ARAÚJO, E. R. de O., O plágio na pesquisa científica do ensino superior, **Revista Conhecimento em Ação**, Rio de Janeiro, v. 2, n. 1, pp. 39-107, 2017.

AXURE. Axure Software Solutions. Disponível em: <<https://www.axure.com/>>. Acessado em 18/04/2018.

CHONOLES, M. J.; Schardt J. A. UML for dummies. Hungry Minds, 2003.

KROKOSZ, M, Abordagem do plágio nas três melhores universidades de cada um dos cinco continentes e do Brasil, **Revista Brasileira de Educação**, v. 16, n. 48, pp. 745-818, 2011.

LUCIDCHART. Lucid Software inc. Disponível em: <<https://www.lucidchart.com>>. Acessado em 20/04/2018.

EDITORIAL DOM BOSCO, **Conheça 8 programas que detectam plágio em trabalhos acadêmicos**. Disponível em: <<https://www.domboscoead.com.br/pos-graduacao/noticias/conheca-8-programas-que-detectam-plagio-em-trabalhos-academicos/101>>. Acessado em 26/04/2018. 2015.

SILVA, V. M.; BARBOSA, R. de M.; ADAMATTI, D. F. **Princípios de Usabilidade e a Importância do Usuário no Projeto de Interfaces**. 2016. Disponível em: <https://www.researchgate.net/publication/304014230_Principios_de_Usabilidade_e_a_Importancia_do_Usuario_no_Projeto_de_Interfaces>. Acessado em: Acessado em 25/04/2018.

SOUZA, N. B. O. de, **Caracterização de Software Científico: Um Estudo de Caso em Modelagem Computacional**. 2011. 73f. Dissertação de Mestrado - Universidade Federal de Juiz de Fora, Juiz de Fora, 2011.

CÓDIGO FONTE DO SISTEMA, **Bitbucket**. Disponível em: <<https://bitbucket.org/ronneyluan/plag-finder/src/master>>

RUBY, **Linguagem de programação Ruby**. Disponível em: <<https://www.ruby-lang.org/pt/>>

RUBY ON RAILS, **Ruby on Rails Framework**. Disponível em: <<https://rubyonrails.org/>>

POSTGRESQL, **Banco de dados PostgreSQL**. Disponível em: <<https://www.postgresql.org/>>

GIT, **GIT sistema de controle de versão distribuído**. Disponível em: <<https://git-scm.com/>>

MVC, **Model View Controller**. Disponível em: <<https://pt.wikipedia.org/wiki/MVC>>

HEROKU, **Heroku ferramenta de deploy**. Disponível em: <<https://www.heroku.com/>>

AZURE, **Plataforma e serviços de computação em nuvem do Microsoft Azure**. Disponível em: <<https://azure.microsoft.com/pt-br/>>

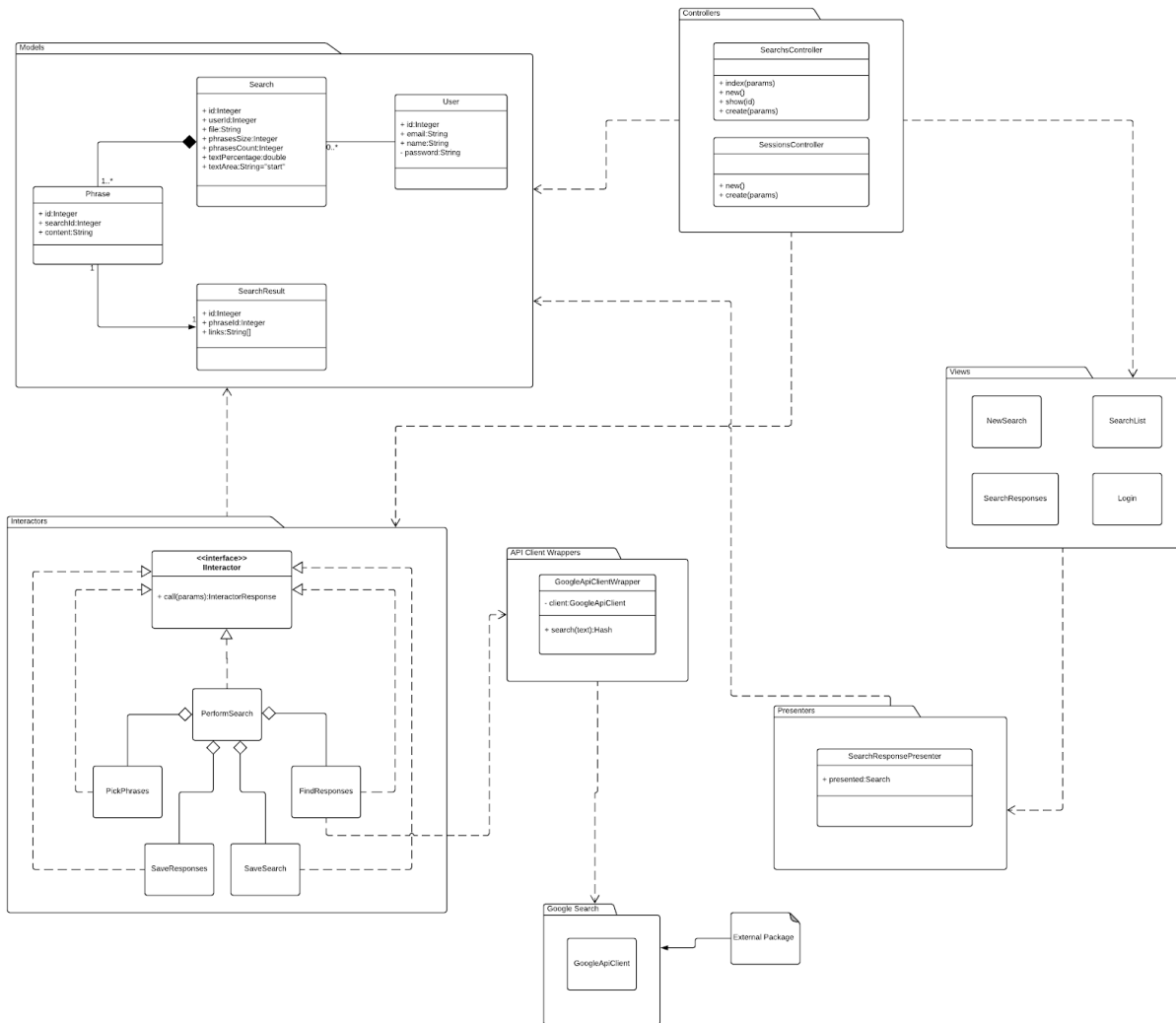
ACADEMIC PLAGIARISM STRUCTURED SEARCH WEB SYSTEM

***Abstract:** The academic plagiarism is one of the great problems for the scientific research. The internet rising and spreading increase the plagiarism problem because it turns free access publications almost fully available. On the other hand internet can also be used to*

plagiarism detect and punishment. Plagiarism identification on-line platforms are available and have a role play in plagiarism combat once they provide help to professors on plagiarism problem detection, but all available platforms imposes limitations to its use or are too expensive. This paper presents an alternative to these platforms, it describes a methodology and a methodology automation proposal for plagiarism detection.

Key-words: *Academic plagiarism. Plagiarism identification software.*

APÊNDICE A – Diagrama de classes.



MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 1 de outubro de 2018, às 9:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **RONNEY LUAN ALBUQUERQUE BEZERRA**, orientado pelo professor **Sérgio Campello Oliveira**, sob título **SISTEMA DE BUSCA ESTRUTURADA POR PLÁGIO ACADÊMICO COM USO DA WEB**, a banca composta pelos professores:

Bruno José Torres Fernandes

Sérgio Campello Oliveira

Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada Aprovada com Restrições* Reprovada

e foi-lhe atribuída nota: 9,5 (*note e meio*)

*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O discente terá 07 dias para entrega da versão final da monografia a contar da data deste documento.



BRUNO JOSÉ TORRES FERNANDES



SÉRGIO CAMPELLO OLIVEIRA