

**METRICMINER: UMA FERRAMENTA  
WEB DE APOIO À MINERAÇÃO DE  
REPOSITÓRIOS DE SOFTWARE**

FRANCISCO SOKOL

ORIENTADOR: MARCO AURÉLIO GEROSA

CO-ORIENTADOR: MAURÍCIO FINAVARO ANICHE

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contextualização . . . . .	1
1.2	Motivação . . . . .	1
1.3	Estrutura da monografia . . . . .	2
<b>2</b>	<b>Mineração de Repositórios de Software</b>	<b>3</b>
2.1	Evolução de software . . . . .	3
2.2	Métricas de código . . . . .	3
2.3	Mineração de repositórios de software . . . . .	3
2.4	Trabalhos relacionados . . . . .	3
<b>3</b>	<b>Arquitetura do MetricMiner</b>	<b>5</b>
3.1	Tecnologias envolvidas . . . . .	5
3.2	Decisões arquiteturais . . . . .	5
3.3	Estendendo e contribuindo com o projeto . . . . .	5
<b>4</b>	<b>Avaliação da ferramenta</b>	<b>7</b>
4.1	Resultados obtidos com a mineração de repositórios públicos . . . . .	7
<b>5</b>	<b>Conclusão e trabalhos futuros</b>	<b>9</b>
<b>6</b>	<b>Parte subjetiva</b>	<b>11</b>

# Introdução

## 1.1 Contextualização

Evolução de Software é uma área da Engenharia de Software que estuda as atividades de desenvolvimento em um sistema de software após a sua concepção inicial e implantação em produção. Esse termo foi usado pela primeira vez em um artigo publicado por Manny Lehman [Mens 2008]. Neste trabalho, o autor enuncia as “leis da evolução de software”, defendendo que programas que representam alguma atividade do mundo real evoluem continuamente, caso contrário, se tornam menos úteis e perdem seu valor [Lehman 1980]. Além disso, Lehman afirma que este processo de mudança contínua faz com que a complexidade do software cresça inevitavelmente, tornando sua estrutura cada vez mais pobre e seu custo de manutenção maior.

Considerando ainda diversos trabalhos indicando que o custo de manutenção de um software ultrapassa 50% do custo total de um projeto, observa-se que encontrar meios de se manter a qualidade interna de um software é muito importante. Por meio do desenvolvimento ferramentas e métodos, a evolução de software busca entender e controlar esse processo para o tornar mais eficiente.

Nesse contexto, a Mineração de Repositórios de Software estuda esse processo de evolução de forma empírica, por meio da análise dos artefatos envolvidos no seu desenvolvimento como código fonte, dados do sistema de controle de versão e sistemas de rastreamento de bugs [Kagdi et al. 2007].

## 1.2 Motivação

Para desenvolver um trabalho em mineração de repositórios, o pesquisador é obrigado a carregar diversos projetos em sua estação de trabalho e realizar uma série de cálculos sobre o código dos projetos e sobre os metadados de seu repositório. Esse processo requer a instalação de diversas ferramentas e bibliotecas localmente para reaproveitar as ferramentas desenvolvidas nessa área, tornando o processo trabalhoso e demorado.

Além de ser um processo complexo, esse tipo de pesquisa consome muitos recursos computacionais. Baixar os repositórios a serem minerados consome um volume considerável de banda. Depois, os dados devem ser processados e persistidos em um banco de dados, ocupando um grande volume de disco. Só então o pesquisador pode calcular métricas sobre esses dados, além de extrair relações entre os metadados do histórico do sistema de controle de versão, gastando uma quantidade grande de processamento de CPU. Só depois de passar por todas essas etapas, é possível extrair dados e avaliar hipóteses por meio de análises estatísticas.

Dessas dificuldades surgiu a motivação para o desenvolvimento do MetricMiner, uma aplicação web que realiza todas as etapas da mineração de um repositório de software. Essa ferramenta disponibiliza um grande volume de dados já processados prontos para serem extraídos e analisados pelo pesquisador, poupando tempo e recursos computacionais.

## 1.3 Estrutura da monografia

O restante desta monografia está estruturada da seguinte forma:

- Seção 2: nesta seção são abordados temas envolvidos no desenvolvimento desse trabalho, em um breve levantamento bibliográfico.
- Seção 3: apresenta a arquitetura e as tecnologias envolvidas no desenvolvimento do MetricMiner.
- Seção 4: expõe os resultados obtidos com a ferramenta na mineração de repositórios de projetos de código aberto.
- Seção 5: análise dos resultados e levantamento de possíveis extensões futuras do projeto.
- Seção 6: apresenta as impressões do aluno na realização desse trabalho e a sua relação com o curso de Bacharelado em Ciência da Computação.

SEÇÃO 2

# Mineração de Repositórios de Software

2.1 Evolução de software

2.2 Métricas de código

2.3 Mineração de repositórios de software

2.4 Trabalhos relacionados



SEÇÃO 3

# Arquitetura do MetricMiner

3.1 Tecnologias envolvidas

3.2 Decisões arquiteturais

3.3 Estendendo e contribuindo com o projeto



SEÇÃO 4

## Avaliação da ferramenta

### 4.1 Resultados obtidos com a mineração de repositórios públicos



SEÇÃO 5

## Conclusão e trabalhos futuros



SEÇÃO 6

# Parte subjetiva



# Referências Bibliográficas

- [Kagdi et al. 2007] Kagdi, H., Collard, M. L., and Maletic, J. I. (2007). A survey and taxonomy of approaches for mining software repositories in the context of software evolution. *J. Softw. Maint. Evol.*, 19(2):77–131.
- [Lehman 1980] Lehman, M. M. (1980). Programs, life cycles, and laws of software evolution. *Proc. IEEE*, 68(9):1060–1076.
- [Mens 2008] Mens, T. (2008). Introduction and roadmap: History and challenges of software evolution. In Mens, T. and Demeyer, S., editors, *Software Evolution*, pages 1–11. Springer.