

Ficha da Pesquisa – Carlos Eduardo de Albuquerque Moreno

Título do Trabalho	Estamos repetindo os mesmos erros? Análise de funções potencialmente vulneráveis em dataset vs observação atual de códigos identificados e classificados em Common Weakness Enumeration (CWE)
Autor(res) e instituição	Carlos Eduardo de Albuquerque Moreno, (UEM) Dr. Edson Alves de Oliveira Junior, (UEM) Dr. Igor Wiese, (UTFPR)
Email do autor(res)	pg405622@uem.br
Orientador(es)	Dr. Edson Alves de Oliveira Junior
Tipo	() TCC (X) Só para cumprir a disciplina
Contexto	O trabalho tem como foco analisar o vasto dataset de funções mineradas de softwares open source e relacionar a semelhança destas funções observadas com as identificadas e classificadas na Common Weakness Enumeration (CWE).
Objetivo da pesquisa	Analisar o dataset e suas 1.27 milhões de funções correlacionando similaridades entre elas e as catalogadas na lista desenvolvida pela comunidade em Common Weakness Enumeration (CWE), explorando as potenciais vulnerabilidades e meios de prevenção.
Referencial teórico	Draper VDISC Dataset – Vulnerability Detection in Source Code
Trabalhos relacionados	Ainda procurando.
Problema de pesquisa	Analisar a enorme quantidade de funções e encaixar as mesmas nas classificações criadas na CWE.
Justificativa e relevância	Através do tempo a Engenharia de Software evolui cada vez mais em todas as suas áreas e por mais que mais pesquisas e análises nos ajudem a identificar falhas atuais ou possíveis falhas futuras, então esse trabalho visa observar se esse conhecimento gerado realmente ajudou ou estamos em um ciclo repetitivo de identificar falhas e repeti-las.
Evidências do problema	O rápido crescimento do mercado e a necessidade de que um software/código deve estar pronto e entregue. O simples descuido ao desenvolver um código ao qual leva a problemas futuramente. Falta de conhecimento sobre o material e o desenvolvimento de código vulnerável. <u>Vou procurar trabalho ainda para embasar evidência do problema.</u>
Hipótese/Questão	Estamos fadados aos mesmos erros identificados a muito tempo devido a desatenção ou despreparo?
Proposta de Solução	Para responder a questão, as funções contidas no dataset serão analisadas dentro da possibilidade de comparação com as trabalhadas em Common Weakness Enumeration (CWE) e uma solução será proposta a partir do resultado da observação.
Projeto de Avaliação	Resultados serão comparados com outros artigos da área para observar se outros pesquisadores também encontraram erros diversos sendo repetidos através do tempo.
Falseamento	
Abordagem	() Teórica (X) Empírica Se empírica: () Quantitativa () Qualitativa (X) Mista
Finalidade	() Explanatória () Exploratória (X) Descritiva
Método a ser seguido na pesquisa	() Experimento (X) Levantamento/survey () Estudo de caso () Grounded Theory () Etnografia () Pesquisa-ação () Outro:
Técnicas de coleta de dados	() Medição () Questionário () Entrevista () Grupo focal () Observação direta (X) Coleta de Documentos (X) Benchmark () Outras:
Dados	Os dados do dataset consiste no código fonte de 1.27 milhões de funções mineradas de softwares open source catalogadas por análise estatística de potenciais vulnerabilidades. Disponível em três arquivos HDF5.
Técnicas de análise dos dados coletados	() Estatística descritiva () Estatística inferencial () Análise do discurso (X) Análise de conteúdo () Codificação de dados () Outra:
Validade Como você auto-avalia a validade da pesquisa?	Validade interna (o quão robusto é o estudo em relação à possível introdução de erros ou vieses)? Péssimo () 0 () 1 () 2 () 3 (X) 4 () 5 Excelente Validade externa (o quão generalizável ou aplicável em contextos diferentes são os resultados)? Péssimo () 0 () 1 () 2 (X) 3 () 4 () 5 Excelente
Contribuições científicas	Analisar se o cenário mudou desde a criação do dataset utilizado e caso tenha-se mantido no mesmo ponto, quais ações poderiam nos colocar na direção correta.
Contribuições tecnológicas	
Estágio do trabalho	(X) Proposta não defendida () Proposta já defendida
Outros comentários	

Referências citadas neste documento.