生成AIによって生成されたソースコードの同定

Identification of source code generated by generative AI

西村太孝・マネジメント分科会・情報セキュリティ大学院大学

背景と目的

近年では、業務での生成AIの使用の増加、サイバー犯罪者の生成AIの使用がある。これらを適切に管理するためには生成AIによって生成されたコードがどの生成AIによって生成されたか正しく分類される必要がある。
よって本研究では生成AIによって生成されたソースコードがどのAIによって生成されたかの同定を目的とする。

すでに見分けられるモノ(一例)

生成AIと人の書いた文章を見分ける
生成AIと人の書いたコードを見分ける
画像が生成AIによるものか見分ける

今後の課題等

・どのような方法で判断をするのか
機械学習？目視判定できる部分？
・使用する生成AIの設定
クローズドとオープンの両方の予定
・使用するプロンプトの設定
プロンプトと出力の揺らぎがどれぐらいある？
・使用する言語の設定
生成AIと人の書いたコードを見分ける研究ではPythonが使用されていた
・生成AIの進化によってどうなるか
実は本人の一番気になっているポイントであります。