Université de Rennes

Direction de la Recherche et de l'Innovation

Appel à projets - Année 2025

Contrats doctoraux «Ordinaires»

ECOLE DOCTORALE : Matisse **POLE DE RECHERCHE : Mathématiques-numérique**

Unité de recherche d'accueil

Unité de Recherche : UMR 6074

Equipe : Expression (Lannion)

Encadrement

Directeur de thèse : Jonathan Chevelu **Unité de recherche** : UMR 6074

Employeur/Établissement d'origine : Univ Rennes

HDR : • ~~oui~~ • non **Nb de thèses en cours** : 2

Mail : jonathan.chevelu@irisa.fr **Tél**. : 02 96 46 91 84

Co-directeur de thèse : **Unité de recherche** :

Employeur/Établissement d'origine :

HDR : • oui • non **Nb de thèses en cours** :

Mail : **Tél.** :

Sujet

Titre : Stéganographie Textuelle par Paraphrase (STP)

Mots clés : stéganographie, paraphrase, production automatique de texte, LLM

Cotutelle internationale (~~oui~~/non) : **Si oui, préciser l'établissement et le pays de rattachement** :

Descriptif (15-30 lignes) :
Contexte socioéconomique et scientifique ; Les hypothèses et questions posées ; Les grandes étapes de la thèse et démarche ; Approches méthodologiques et techniques envisagées

La stéganographie est une technique qui consiste à dissimuler une information sensible dans un média de couverture à l'apparence anodine, ici un texte. Contrairement au tatouage de documents, qui vise à rendre l'information inamovible des données de couverture, la stéganographie cherche à rendre cette information indétectable tout en maximisant la quantité d'information dissimulée dans un document de taille donnée.

Les approches récentes en stéganographie textuelle s'appuient souvent sur la génération de textes de couverture respectant une thématique prédéfinie. Cependant, ces textes générés peuvent manquer de crédibilité dans le contexte d'utilisation.
De plus, la quantité d'informations qui doit être pré-partagée entre l'émetteur et le récipiendaire est parfois disproportionnée (modèle informatique, thématique du message de couverture, etc.), ce qui peut rendre les méthodes proposées irréalistes.

Par ailleurs, la production de paraphrases a considérablement progressé ces dernières années, notamment grâce aux avancées des grands modèles de langage (LLM). Si la réécriture de textes en conservant leur sens est de plus en plus performante, le contrôle précis de cette réécriture pour des tâches spécifiques reste un verrou scientifique difficile.
La stéganographie textuelle représente une application ambitieuse dans ce domaine, car elle impose un critère strict au texte généré : le message secret doit être retrouvé par le récipiendaire, et uniquement par lui.

L'objectif de cette thèse est donc d'explorer la possibilité de modifier un texte existant pour y dissimuler un message tout en préservant son sens original. Des résultats collatéraux sur la détection de stéganographie textuelle sont également attendus.

➢ **Ce document est à renvoyer par la direction de l'unité de Recherche aux animateurs du Pole Recherche et à l'école Doctorale pour le 13 février 2025 au plus tard**