Offre de Stage de recherche.

Intitulé du stage : RDAP : Frontière entre spécification du protocole et son implémentation.

Domaine du stage : Réseau et Cybersécurité.

Organisme d'accueil : Laboratoire Informatique de Grenoble – Équipe Drakkar

Localisation : Campus Universitaire de Grenoble.

Durée : 6 mois.

Contexte : Le Laboratoire d'Informatique de Grenoble (LIG) héberge plus de 450 chercheuses et chercheurs réparti.e.s dans plus de 22 équipes de recherche. L'équipe DRAKKAR est spécialisée dans l'étude des réseaux. Une partie de l'équipe est plus particulièrement attachée à l'étude du DNS et du « DNS abuse ».

DNS étant l'acronyme de « Domain Name System », notre axe de recherche s'étend sur l'étude des noms de domaine. Les noms de domaine font partie intégrante de notre quotidien. Vous avez déjà rencontré.e des noms de domaine tels que « google.com », « univ-grenoble-alpes.fr », ou encore « amazon.com ». Néanmoins, que pouvez-vous dire de « support-tech-apple.com », « bank-of-america-help.com » ou même « paypall.com » ? Tous ceux-ci sont des noms de domaines malicieux fictifs dont le but est de tromper l'utilisateur afin de faire du phishing.

L'étude des noms de domaine permet alors d'améliorer la sécurité d'Internet. Un des projets précédants de notre équipe est « COMAR », permet de différencier les noms de domaine compromis ou malicieux. Par exemple, si le nom de domaine « univ-grenoble-alpes.fr » héberge du contenu malveillant, faut-il le supprimer ? Et qu'en est-il de paypall.com ? Comar propose alors de faire la distinction entre un domaine compromis (dans notre cas univ-grenoble-alpes.fr) pour lequel il faut prévenir l'administrateur, et paypall.com, pour lequel la solution serait de prendre des mesures contre le nom de domaine

Certains outils d'analyse des noms de domaines ont alors besoin des données d'enregistrement (age du domaine, companie d'enregistrement, ...) collecté via WHOIS. Néanmoins, le protocole WHOIS étant désuet, l'IETF (Internet Engineering Task Force) a donc proposé le protocole RDAP (RFC 7482 et RFC 7483). Les RFC (Requests For Comments) sont des documents définessant les « normes » et « standards » d'internet.

Cependant, est-ce que tout les services proposant RDAP suivent les normes et spécifications de l'IETF ? Ayant déjà utilisé le protocole pour différentes recherches, nous avons constaté qu'il arrive que non. De plus le protocol RDAP, scencé être simple d'utilisation, se retrouve des fois plus compliqué que WHOIS.

Description du stage :

Ce stage est une première introduction dans le monde de la recherche. Le stage vise à explorer en profondeur le protocole RDAP ainsi que sa spécification associée. Le protocole RDAP, conçu pour

remplacer le protocole WHOIS, est devenu une norme majeure pour accéder aux informations d'enregistrement de noms de domaine. L'objectif de ce stage est de développer une compréhension approfondie du fonctionnement du protocole RDAP et de son utilité dans le contexte actuel d'Internet afin de proposer un outil plus « human friendly ».

Missions :

- Création d'un outil open-source en Python pour la récupération des données RDAP.
- Création d'un outil de vérification de la syntaxe RDAP/RFC 7483.
- Vérification que des services suivent les spécifications et, dans le cas contraire, les contacter.

Profil recherché :

Étudiant.e en dernière année de Master 2 en Informatique (MOSIG, CYBERSEC, ...) , avec un fort intêret pour la recherche et une envie de continué dans le monde académique. Le ou la candidat.e doit faire preuve d'une forte curiositée : l'envie de comprendre, de découvrir et de critiquer. Le stage est aussi ouvert au étudiant.es de M1 très fortement motivés.

Modalités de candidature :

Les candidats intéressés sont invités à envoyer leur candidature comprenant un CV, une lettre de motivation et toutes autres informations jugées intéressantes à Maciej KORCZYNSKI"= maciej.korczynski at univ-grenoble-alpes.fr et Olivier HUREAU olivier.hureau at univ-grenoble-alpes.fr

Références :

RFC 7482 : Registration Data Access Protocol (RDAP) Query Format (https://datatracker.ietf.org/doc/html/rfc7482)

RFC 7483 :JSON Responses for the Registration Data Access Protocol (RDAP) (https://datatracker.ietf.org/doc/html/rfc7483)

[1] Aruna Prem Bianzino, Davide Pezzuolo, and Gianluca Mazzini. 2014. Who Is Whois? An Analysis of Results Consistence.

[2] Suqi Liu, Ian Foster, Stefan Savage, Geoffrey M. Voelker, and Lawrence K. Saul. 2015. Who Is .Com?: Learning to Parse WHOIS Records.

[3] Chaoyi Lu, Baojun Liu, Yiming Zhang, Zhou Li, Fenglu Zhang, Haixin Duan, Ying Liu, Joann Qiongna Chen, Jinjin Liang, Zaifeng Zhang, Shuang Hao, and Min Yang. 2021. From WHOIS to WHOWAS: A Large-Scale Measurement Study of Domain Registration Privacy under the GDPR.

Internship position

Internship title: RDAP: Frontier between protocol specification and implementation.

Field: Network and Cybersecurity.

Hosting organization: Grenoble Computer Laboratory - Drakkar Team

Location: Grenoble University Campus.

Duration: 6 months.

Context: The Laboratoire d'Informatique de Grenoble (LIG) is home to over 450 researchers in more than 22 research teams. The DRAKKAR team specializes in the study of networks. Part of the team is dedicated to the study of DNS and DNS abuse.

As DNS stands for "Domain Name System", our research focuses on the study of domain names. Domain names are an integral part of our daily lives. You've already come across domain names such as "google.com", "univ-grenoble-alpes.fr", or "amazon.com". But what about "support-tech-apple.com", "bank-of-america-help.com" or even "paypall.com"? All of these are fictitious malicious domain names whose purpose is to mislead the user for phishing purposes.

By studying domain names, we can improve Internet security. One of our team's previous projects, COMAR, enables us to differentiate between compromised and malicious domain names. For example, if the domain name "univ-grenoble-alpes.fr" hosts malicious content, should it be deleted? And what about paypall.com? Comar suggests distinguishing between a compromised domain (in our case univ-grenoble-alpes.fr), for which the administrator should be notified, and paypall.com, for which the solution would be to take action against the domain name.

Internship description:
This internship is a first introduction to the world of research. The internship aims to explore in depth the RDAP protocol and its associated specification. The RDAP protocol, designed to replace the WHOIS protocol, has become a major standard for accessing domain name registration information. The aim of this internship is to develop an in-depth understanding of how the RDAP protocol works and its usefulness in today's Internet context, in order to propose a more "human-friendly" tool.

Missions :
- Creation of an open-source tool in Python for retrieving RDAP data.
- Creation of a tool to check RDAP/RFC 7483 conformence.
- Check that RDAP services follow the specifications

Profile required:

Student in final year of Master 2 in Computer Science (MOSIG, CYBERSEC, ...), with a strong interest in research and a desire to continue in the academic world. The candidate must demonstrate a strong

curiosity: the desire to understand, discover and criticize. The position is also open to highly motivated M1 students.
How to apply :
Interested candidates are invited to send their application, including CV, covering letter and any other information deemed of interest, to Maciej KORCZYNSKI maciej.korczynski at univ-grenoble-alpes.fr and Olivier HUREAU olivier.hureau at univ-grenoble-alpes.fr

References :

RFC 7482 : Registration Data Access Protocol (RDAP) Query Format (https://datatracker.ietf.org/doc/html/rfc7482)

RFC 7483 :JSON Responses for the Registration Data Access Protocol (RDAP) (https://datatracker.ietf.org/doc/html/rfc7483)

[1] Aruna Prem Bianzino, Davide Pezzuolo, and Gianluca Mazzini. 2014. Who Is Whois? An Analysis of Results Consistence.

[2] Suqi Liu, Ian Foster, Stefan Savage, Geoffrey M. Voelker, and Lawrence K. Saul. 2015. Who Is .Com?: Learning to Parse WHOIS Records.

[3] Chaoyi Lu, Baojun Liu, Yiming Zhang, Zhou Li, Fenglu Zhang, Haixin Duan, Ying Liu, Joann Qiongna Chen, Jinjin Liang, Zaifeng Zhang, Shuang Hao, and Min Yang. 2021. From WHOIS to WHOWAS: A Large-Scale Measurement Study of Domain Registration Privacy under the GDPR.