Avantages

Good programming skills: python, scripting, java/C++, etc.

Compétences

models will be validated over real datasets, either proprietary ones as the ACQUA dataset or the IPL BetterNet dataset, or public datasets as the ones
measurements will be shortened to only keep the most relevant ones for the targeted classification problem. In addition to lab experimentation, the
developed to perform classification of network anomalies from passive measurements collected in the wild. Along this study, the long list of possible
the lab (OneLab, R2Lab) to understand their signatures and their most important features. Based on that, Supervised Machine Learning models will be
developed to perform classification of network anomalies from passive measurements collected in the wild. Along this study, the long list of possible
result in an increase in the packet delay, the loss rate of packets and a drop in the network throughput. Wireless problems result in almost same behaviour,
and the user Quality of Experience.

Mission confiée

In this thesis we want to study a new approach consisting in leveraging the wealth of data available in the end user device as a result of her/his normal
activity (passive traffic captures) and from passive measurements collected freely inside his device. These data, collected at almost no cost, is shaped
so as to be able to take the appropriate actions to counter their origins in the limit of possible.

Principales activités

The thesis will address the above problem by following a data driven approach, where the anomalies will be first categorized and profiled, then tested in
the lab (OneLab, R2Lab) to understand their signatures and their most important features. Based on that, Supervised Machine Learning models will be
developed to perform classification of network anomalies from passive measurements collected in the wild. Along this study, the long list of possible
measurements will be shortened to only keep the most relevant ones for the targeted classification problem. In addition to lab experimentation, the
models will be validated over real datasets, either proprietary ones as the ACQUA dataset or the IPL BetterNet dataset, or public datasets as the ones
released by MLab and ZDnet. By the end, this PhD work is supposed to lead to concrete solutions of help to the end user transforming her/his local
measurements to indicators of her/his access performance and pinpointing the origins of service degradation with the help of the ML models that will be
developed for this purpose.

Compétences

Good knowledge in network protocols and network measurement, in data analytics, and in machine learning.

Good programming skills: python, scripting, java/C++, etc.

A propos du centre ou de la direction fonctionnelle

The Inria Sophia Antipolis - Méditerranée center counts 24 research teams as well as 8 support departments. The center's staff (about 500 people
including 350 Inria employees) is made up of scientists of different nationalities (250 foreigners of 50 nationalities), engineers, technicians and
administrative staff. 1/3 of the staff are civil servants, the others are contractual agents. The majority of the center's research teams are located in Sophia
Antipolis and Nice in the Alpes-Maritimes. Four teams are based in Montpellier and two teams are hosted in Bologna in Italy and Athens. The Center is a
foundering member of Université Côte d'Azur and partner of the I-site MUSE supported by the University of Montpellier.

Contexte et atouts du poste

Despite the considerable improvement we saw last years in terms of internet access performance and the quality of the physical and virtualised
infrastructures hosting the internet services, we are still facing a lot of situations where the internet service degrades and the end user Quality of
Experience (QoE) is less than expected. The reasons are many, from the slowness of the device of the user, to the bad configuration of the WiFi at home, to
the interference caused by the neighbouring WiFi networks, to the saturation of the access link by the many devices and applications running at home, to
the congestion in the ISP network especially on its peering links, till the overload of the servers of the content providers following a sudden increase in the
users' activities. There are also situations where the QoE degrades for other reasons than congestion or lack of resources, as when the ISP or the content
providers decide to reduce the quality of their service to prioritise some part of the traffic over the rest a.k.a. network traffic differentiation, or to face
scenarios of heavy service usage (video resolution reduction by major video stream platforms in current confinement period). Those situations, and many
others, exist well today and will not be solved in the immediate future despite the considerable advances seen and foreseen both at the network and the
cloud levels. The problem is not only in the frustration they cause for the end user, but also in the difficulty for the end user to distinguish between them
so as to be able to take the appropriate actions to counter their origins in the limit of possible.

A long list of solutions and tools has been proposed in the last years to shed light on some of these problems. Network access performance can be analysed
with tools such as SpeedTest, MobiPerf or ACQUA. WiFi access performance can be analysed with tools such as Wireshark, WiFi Analyzer of MS Windows and
WiFi Scanner of Apple. Device performance and its relationship with the Quality of Experience can be monitored with tools such as QoE Doctor. Tools for
detecting the congestion in the core of the network and at peering links are mostly of the research domain, as for example the recent paper in ACM
SIGCOMM 2018 by Dhamdhere et al. Inferring traffic differentiation is another challenging topic that has attracted a lot of research, with most of proposed
solutions consisting in mimicking the behaviour of the suspected differentiated application, e.g., video flow, and in comparing the performance of the two
flows, the original one and the emulated one. Any noticeable difference is then interpreted as sign of traffic differentiation in the network. All these tools
and many others contribute to answer the questions of the end user and the network and service providers, but they are on one side limited to the specific
problem to which they are designed thus requiring the user to install and master all of them, and second, they are in an important part of them intrusive
thus requiring the installation of a long list of tools, and the injection of traffic into the network thus causing overload on a network which is already loaded
at the moment of the problem.

Mission confiée

In this thesis we want to study a new approach consisting in leveraging the wealth of data available in the end user device as a result of her/his normal
activity (passive traffic captures) and from passive measurements collected freely inside his device. These data, collected at almost no cost, is shaped
according to what is going on inside the network, in the device of the user, and on the other side of the service provider. For example, a congestion will
result in an increase in the packet delay, the loss rate of packets and a drop in the network throughput. Wireless problems result in almost same behaviour,
but in addition, the wireless signal will show signs of weakness or high values of the interfering noise. Traffic differentiation will show signs of slowness of
some parts of the traffic whereas the other part behave normally, even better. All these signatures exist together, and the challenge is in defining them,
then in detecting and isolating them from each other in an effort to understand the origins behind any problem causing the drop in the internet service
and the user Quality of Experience.

2020-02570 - PhD Position F/M [CORDIC2020-DIANA] A lightweight data-driven approach
for network monitoring and troubleshooting

Type de contrat : CDD
Niveau de diplôme exigé : Bac + 5 ou équivalent
Fonctio : Doctorant
• Subsidized meals
• Partial reimbursement of public transport costs
• Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
• Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
• Professional equipment available (videoconferencing, loan of computer equipment, etc.)
• Social, cultural and sports events and activities
• Access to vocational training
• Social security coverage

Rémunération
Duration: 36 months
Location: Sophia Antipolis, France
Gross Salary per month: 1982€ brut per month (year 1 & 2) and 2085€ brut/month (year 3)

Informations générales
• Thème/Domaine : Réseaux et télécommunications
  Système & réseaux (BAP E)
• Ville : Sophia Antipolis
• Centre Inria : CRI Sophia Antipolis - Méditerranée
• Date de prise de fonction souhaitée : 2020-11-01
• Durée de contrat : 3 ans
• Date limite pour postuler : 2020-05-24

Contacts
• Equipe Inria : DIANA
• Directeur de thèse :
  Barakat Chadi / Chadi.Barakat@inria.fr

A propos d'Inria
Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3500 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 180 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Consignes pour postuler
Sécurité défense :
Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :
Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.