

# Sujet : Simulation et préparation de traces d'un réseau domestique

## Contexte

On s'intéresse au scénario d'une maison connectée dotées de N équipements accédant à Internet via un point d'accès. Le point d'accès est donc l'équipement central du réseau domestique et tous les équipements y sont connectés en WiFi (standard 802.11ac). On suppose que les N équipements sont de  $K \leq N$  types différents (par exemple caméra de surveillance, capteur de température, capteur d'humidité, ordinateurs, etc.), et que les équipements d'un type donné exécutent tous la même application.

Chaque application génère un type de trafic qui lui est propre, en termes de la nature du trafic (HTTP, IPTV, VoIP, COAP, etc.), du protocole de transport utilisé (TCP, UDP), des périodes d'activité (par exemple, envoi ou réception des données une fois chaque 10 millisecondes, 5 secondes, 2 minutes, etc.), du nombre et de la taille des paquets envoyés ou reçus pendant chaque période d'activité. Par exemple, on peut imaginer qu'une caméra de surveillance va principalement générer du trafic UDP et régulier (presque continu) composé de paquets assez gros, ce qui ne sera pas le cas d'un simple capteur de température qui générera typiquement du trafic sporadique, probablement en TCP et pas très gourmand en ressources.

Pour chaque application, il existe un serveur distant (ne faisant pas partie du réseau domestique) et raccordé au point d'accès de façon filaire (Ethernet) et qui reçoit les données de tous les équipements du type considéré.

La figure 1 ci-dessous récapitule ce scénario.

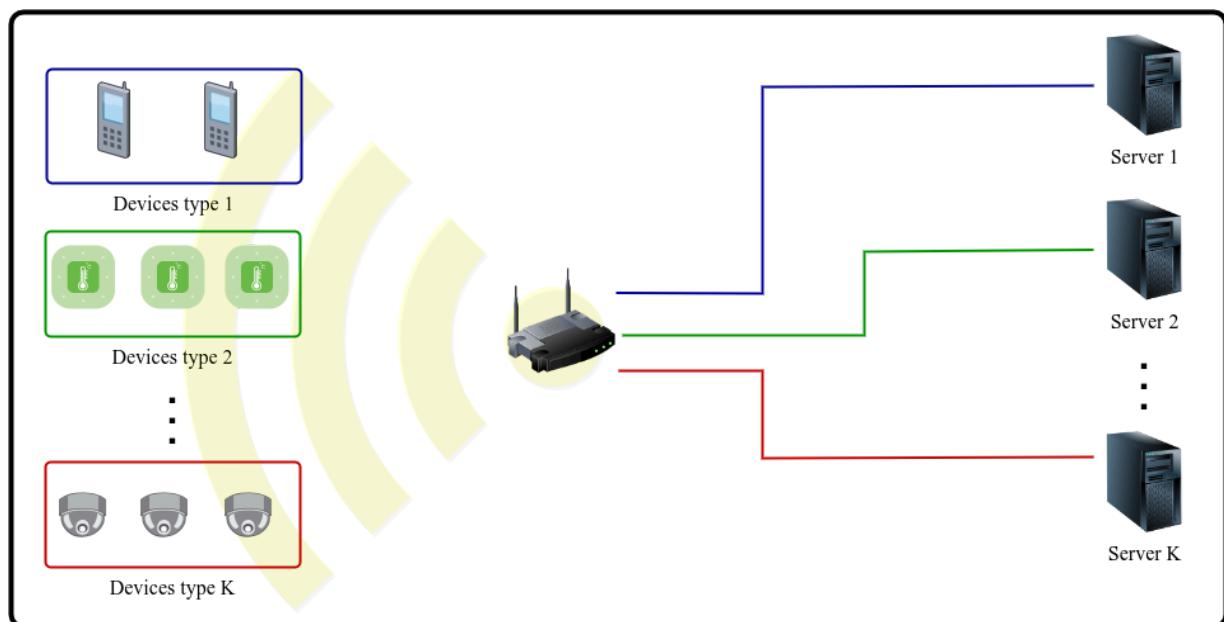


Figure 1: Scénario à simuler

## Travail demandé

On vous demande de simuler ce scénario et d'en sauvegarder les traces d'exécution. Vous devez configurer votre simulation pour collecter le maximum d'informations au niveau du point d'accès.

Ensuite, vous devez utiliser ces traces pour préparer un jeu de données (dataset) qui peut être utilisé pour entraîner un modèle de Machine Learning (ML). On suppose que ces traces vont être utilisées pour construire un modèle ML permettant, au niveau du point d'accès, de classifier chaque trafic comme appartenant à une application donnée parmi les K que vous aurez mises en place, la classification se faisant sur des morceaux de trafic d'une longueur que vous pouvez définir par vous-mêmes. *Vous ne devez pas entraîner le modèle, mais juste préparer le dataset.*

## Quelques consignes

- Pour la simulation, vous devez utiliser le simulateur ns-3
- Les traces simulées doivent contenir beaucoup de variabilité, à la fois sur les paramètres mentionnés dans le contexte et éventuellement sur d'autres paramètres de votre choix. Il vous est laissé le choix pour les valeurs de ces paramètres.
- Vous êtes libres de fixer les valeurs des nombres N d'équipements et K de types d'équipements, ainsi que les paramètres de communication de chacune des applications. Plus N et K sont grands, plus il y a de la variabilité, et mieux c'est. Rassurez-vous d'exécuter la simulation suffisamment longtemps pour avoir assez de données
- L'instant de début d'émission de chaque équipement doit être aléatoire et compris dans les 5 premières secondes.

## Rendus attendus

- Un rapport décrivant les choix effectués sur toutes les parties du travail et les résultats obtenus (résultats intermédiaires et résultats finaux). *Le rapport ne doit pas dépasser 5 pages. Il doit cependant permettre de bien comprendre votre travail et les résultats obtenus.*
- Tous les codes sources produits permettant de reproduire votre travail et vos résultats, ainsi que les traces générées, idéalement sur un dépôt github.

**Critères d'évaluation :** votre capacité de prise en main des outils, la rigueur de votre approche scientifique, les outils utilisés ainsi que la variabilité introduite dans les données et dans les analyses. *Le but pour vous doit être d'aller le plus loin possible.*

## Ressources :

- Simulateur ns-3 : <https://www.nsnam.org/>
- Toutes autres ressources de votre choix, à condition de les citer explicitement et d'être capable d'expliquer en détail ce que fait votre code/solution.