Identification du type des objets connectés par les informations des protocoles réseaux

Nesrine Ammar, Ludovic Noirie, Sébastien Tixeuil

Identification du type des objets connectés par les informations des protocoles réseaux

Nesrine Ammar[1,2], Ludovic Noirie[1] et Sébastien Tixeuil [2 †]

[1]*Nokia Bell Labs, Nokia Paris Saclay, route de Villejust, 91620 Nozay, France*
[2]*Sorbonne Université, Laboratoire d'Informatique de Paris 6, LIP6, F-75005, Paris, France*

Dans cet article, nous traitons le problème de l'identification du type des objets connectés dans un réseau domestique. Nous proposons une approche basée sur les protocoles réseaux tels que les protocoles de découverte. Les résultats obtenus montrent que, grâce aux informations collectées avec les approches utilisées, nous réussissons à identifier la nature et les fonctionnalités de 30 objets parmi les 33 objets connectés disponibles dans les jeux de données utilisés.

Mots-clefs : Internet des objets, systèmes de recommandation, identification des objets connectés, protocoles réseaux

1 Introduction

Avec le développement de l'internet des objets (*Internet of Things*, IoT), de plus en plus de personnes achètent des objets connectés qui sont de plus en plus variés. En conséquence, le nombre de services utilisant ces objets ne cesse d'augmenter [AIM10]. Dès lors, l'utilisateur n'a pas nécessairement pleine conscience de l'ensemble des objets connectés autour de lui, ni de tous les services IoT dont il pourrait profiter. Il est donc primordial de définir un système de recommandation proposant à l'utilisateur des services IoT personnalisés et l'aidant à choisir les objets connectés adéquats pour l'établissement de ces services, en fonction de ceux qui sont à sa disposition.

Les systèmes de recommandation de services ont été largement utilisés dans de nombreux domaines, mais très peu d'articles traitent de leur usage dans l'IoT [MAC16]. Nous avons proposé un système de recommandation [NPA17, PNP+18] reposant sur la plateforme *Majord'Home* [BBC+15], qui utilise la technologie *software-defined networking* (SDN). Ce système propose à l'utilisateur les objets connectés adéquats pour obtenir un service particulier. Une identification précise et autonome de la nature des objets connectés dont l'utilisateur dispose est nécessaire pour obtenir une recommandation pertinente.

Dans cet article, nous considérons ce problème d'identification du type des objets connectés dans un réseau domestique, un défi majeur dû à l'hétérogénéité de l'IoT. La section 2 décrit la plateforme *Majord'Home* qui intègre notre assistant d'identification des objets connectés, ainsi que les différentes approches existantes pour identifier de tels objets. Ensuite, nous décrivons l'implémentation de notre solution et les résultats obtenus dans la section 3. Finalement, nous concluons par les travaux futurs dans la section 4.

2 Contexte et état de l'art

2.1 La plateforme Majord'Home

Basée sur les principes SDN, la plateforme *Majord'Home* permet aux utilisateurs de gérer plus facilement les objets connectés à leur disposition, en créant des tranches de réseau (*Software-Defined LAN*, SD-LAN, voir [BBC+15] pour plus de détails) permettant de regrouper des objets connectés dans un LAN virtuel pour un usage donné, les isolant des autres objets. Cela nous sera très utile pour identifier tout objet nouvellement connecté, en créant un SD-LAN entre notre assistant d'identification et l'objet nouvellement connecté.

L'architecture de la plateforme *Majord'Home* [BBC+15, PNP+18] est décrite dans la figure 1 :

†Ce travail a été partiellement réalisé dans le cadre du LINCS (*Laboratory of Information, Networking and Communication Science*, http://www.lincs.fr/). Les auteurs remercient aussi leurs collègues de Nokia Bell Labs Paris-Saclay et Anvers pour leur support sur la plateforme *Majord'Home*, en particulier Dinh Thai Bui, Werner Liekens, Michel le Pallec, Pierre Peloso et Frederik Vandeputte.

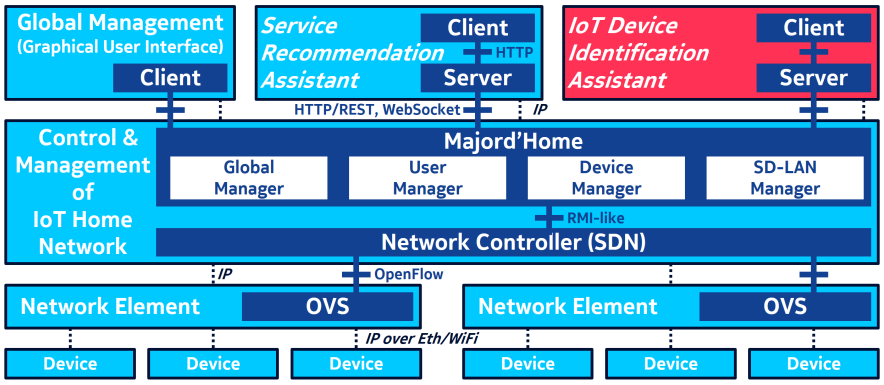

FIGURE 1: Architecture du Majord'Home incorporant l'assistant d'identification des objets connectés

– Les éléments de réseau (box) connectent les objets par Wi-Fi ou Ethernet. Chaque box intègre un *Open vSwitch* (OVS) configuré par le contrôleur de réseau via OpenFlow.
– La plateforme de contrôle et gestion de réseau IoT comprend le contrôleur de réseau et le *Majord'Home* qui gère l'environnement global du réseau domestique, l'autorisation d'accès aux objets pour chaque utilisateur, les données sur les objets connectés et la création/configuration/suppression des SD-LAN.
– L'administrateur du réseau domestique gère la totalité de la plateforme via une interface de gestion globale communiquant avec le *Majord'Home* via HTTP/REST et WebSocket.
– Tout assistant logiciel a deux composants communiquant via HTTP : un client interagissant avec l'utilisateur final et un serveur s'interfaçant avec le *Majord'Home* via HTTP/REST et/ou WebSocket.

2.2 Différentes méthodes d'identification d'objets connectés

Identification par protocoles de découverte de services (Bonjour et UPnP)

Une première approche est basée sur les protocoles de découverte de services (SDP) dans le réseau. Pour annoncer les services fournis par un objet connecté dans un réseau local, l'objet publie des informations sur ses capacités, son emplacement, son nom et la description de ses services dans le réseau. Par exemple, les objets UPnP utilisant le protocole de découverte SSDP (Simple Service Discovery Protocol) [BPW13] transmettent un message de type NOTIFY afin de publier leurs capacités dans le réseau. Ce message contient une URL pointant vers la description XML de l'objet. Concernant les objets utilisant le protocole Bonjour qui utilise le multicast DNS (mDNS) pour publier ses capacités dans le réseau, l'identification est basée sur les informations sur ces capacités, partagées dans les *records* mDNS [CK13]. Collectant ces informations, nous pourrons construire un profil identifiant chaque type d'objet. Cette technique sera combinée avec les autres approches afin de compléter ses lacunes et consolider les résultats.

Identification par empreintes avec DHCP (et adresses MAC)

Chaque objet nouvellement connecté interagit avec le serveur DHCP pour obtenir une adresse IP, utilisant des champs optionnels du protocole pour donner des informations supplémentaires : nom de l'hôte, constructeur (information aussi récupérable grâce à l'adresse MAC [IEE18]), modèle et système d'exploitation [AD97]. La technique par empreintes avec DHCP exploite ces données lorsque elles sont disponibles. Elle est utilisée par exemple pour les objets mobiles, les tablettes, les imprimantes ainsi que les ordinateurs.

Identification par *user-agent* dans l'entête de HTTP

Cette approche d'identification est basée sur le *user-agent* dans l'entête des requêtes du protocole HTTP, elle nécessite donc que l'objet connecté envoie des requêtes HTTP sur le réseau. Le *user-agent* est principalement utilisé dans les applications de livraison de contenu afin d'optimiser l'expérience utilisateur. Une première variante utilise le *Composite Capability/Preferences Profile* (CC/PP) défini par le Consortium World Wide Web (W3C), avec un format XML décrivant les fonctionnalités de l'objet connecté concernant le logiciel, le matériel, le réseau et le navigateur utilisé. Une deuxième variante utilise le *User-Agent-Profile* (UAProf) basé sur CC/PP, dont l'URL est incluse dans l'entête de la requête HTTP.

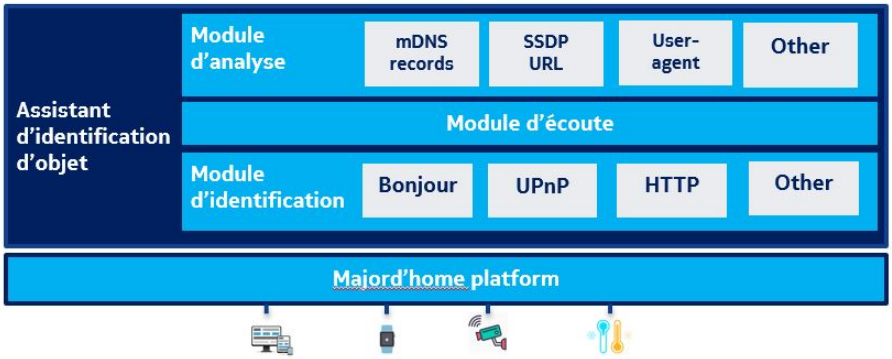

FIGURE 2: Architecture de l'assistant d'identification des objets connectés dans le réseau domestique

3 Solution implémentée et résultats

3.1 Architecture de l'assistant d'identification des objets connectés

Nous proposons un assistant logiciel identifiant le type des objets connectés au réseau domestique, basé sur les informations récupérées des protocoles de la section 2.2 et s'intégrant à la plateforme *Majord'Home* de la sous-section 2.1. Son architecture est constituée de trois éléments principaux (figure 2) :

1. Le module d'identification de protocole est chargé de détecter les éventuels protocoles utilisés par l'objet connecté. Notamment, pour les protocoles de découverte, il envoie des requêtes sur les adresses *multicast* adéquates. Si l'objet est capable d'écouter une requête spécifique, il enverra une réponse annonçant ses capacités. Ce module tire parti du *Majord'Home* : l'assistant configure un SD-LAN spécifique entre lui et l'objet à identifier, ce qui permet l'utilisation des protocoles de découverte même si l'objet n'est pas dans le même LAN physique que l'assistant.
2. Le module d'écoute consiste à récupérer les réponses envoyées par les objets.
3. Le dernier module consiste à analyser les réponses reçues sous format ".pcap". L'analyse des réponses est différente d'un protocoles à un autre vu qu'ils utilisent différents formats de données.

3.2 Résultats

Afin d'évaluer l'efficacité de notre solution, nous avons utilisé les captures de 21 objets connectés venant de [MMH+17] (table 1) ainsi que celles de 12 objets de notre laboratoire [PNP+18] (table 2).

Type	Nb	Vendeurs	mDNS	SSDP	DHCP	User-agent	Aucune
Caméra	6	D-Link, Edimax, Ednet	2	4	2	0	2/6
Lampe	2	Philips, Lightify	1	1	1	0	0/2
Électroménager	2	iKettle2, Espressif	0	0	2	0	0/2
Capteur	3	D-Link	3	0	2	0	0/3
Mobile	3	Apple, Samsung	0	0	0	2	1/3
Interrupteur	3	D-Link, Philips, WeMo	1	3	1	0	0/3
Sirène	1	D-Link	1	0	1	0	0/1
Hub	1	D-Link	1	1	1	0	0/1
Total	**21**		**9**	**9**	**10**	**2**	**3/21**

Table 1 : 21 objets connectés de [MMH+17] analysés par les différentes techniques

Type	Nb	Vendeurs	mDNS	SSDP	DHCP	User-agent	Aucune
Caméra	4	D-Link, Panasonic, TRENDnet	1	4	2	0	0/4
Lampe	1	Philips	0	0	1	0	0/1
Tablette	2	Asus, Pixel C	0	0	2	2	0/2
PC	3	HP	0	1	3	3	0/3
Haut-parleur	1	Chromecast	0	0	1	0	0/1
TV	1	Chromecast	0	1	1	0	0/1
Total	**12**		**1**	**6**	**10**	**2**	**0/12**

Table 2 : 12 objets connectés de notre laboratoire [PNP+18] analysés par les différentes techniques

Nous analysons d'abord la couverture de nos approches sur les objets de [MMH+17] (table 1). Parmi les 21 objets, 9 sont identifiables par mDNS, 9 par SSDP, 10 par DHCP et 2 par *user-agent*, tandis que 3 objets ne sont identifiables par aucune de ces quatre techniques. Pour les résultats concernant les 12 objets dans notre laboratoire [PNP+18] (table 2), 1 seul objet identifié par mDNS, 6 objets par SSDP, 10 objets par DHCP et enfin 5 objets par user-agent, tous sont identifiables par au moins une de ces techniques.

D'après nos résultats, l'identification basée sur DHCP montre un taux plus élevée que les autres techniques. Les résultats montrent aussi que le *user-agent* réussit à identifier des types spécifiques d'objets connectées tels que les tablettes, les ordinateurs et les smart phones. Pour mDNS et SSDP, nous identifions chaque objet par l'ensemble des informations extraites des paquets de ces protocoles. Si un objet peut être identifié par plusieurs méthodes, nous combinons les informations afin de renforcer leur fiabilité.

Pour SSDP, les objets publient dans leur message une URL pointant vers la description XML de chaque objet, ce qui nous permet d'identifier ce dernier. Ces informations permettent d'identifier le vendeur, le modèle et les instances de services fournies par l'objet. Le nom de la même instance de service peut être différent d'un vendeur à un autre, mais cela n'affecte pas les résultats obtenus. Ces informations nous permettent ainsi de construire nos données d'identification pour les futures objets.

Les noms utilisés pour un même service fourni par le même type d'objet varient fortement. Pour améliorer l'identification des objets, nous représentons chaque type d'objet par un sac de mots. Cette représentation est une description très utilisée en recherche d'information. Donc, nous extrayons les mots utiles à partir des instances de services et du nom local de l'objet afin de construire une description textuelle identifiant chaque type d'objet. Cette technique sera très utile dans les travaux futurs utilisant l'apprentissage automatique pour identifier les nouveaux objets.

4 Conclusion

Dans cet article, nous avons présenté les différentes approches permettant d'identifier les objets connectés dans un environnement hétérogène, leur combinaison dans une architecture permettant leur utilisation en parallèle, ainsi que nos premiers résultats. Ces derniers montrent que grâce aux informations contenues dans les messages échangés par les protocoles réseaux, nous pouvons identifier 30 objets parmi les 33 objets connectés disponibles dans nos jeux de données.

Dans les travaux futurs, nous prévoyons d'explorer d'autres protocoles de découverte (par exemple CoAP) et d'autres technologies sans fil pour des objets plus contraints en énergie (par exemple, Bluetooth Low Energy). Enfin, nous envisageons d'utiliser aussi des techniques d'apprentissage automatique [MMH+17] pour améliorer les résultats déjà obtenus.

Références

[AD97] S. Alexander and R. Droms. DHCP options and BOOTP vendor extensions. RFC 2132, RFC Editor, March 1997.

[AIM10] L. Atzori, A. Iera, and G. Morabito. The Internet of Things : a survey. *Computer Networks*, 54(15) :2787 – 2805, 2010.

[BBC+15] M. Boussard, D. T. Bui, L. Ciavaglia, R. Douville, M. L. Pallec, N. L. Sauze, L. Noirie, S. Papillon, P. Peloso, and F. Santoro. Software-Defined LANs for interconnected smart environment. In *2015 27th International Teletraffic Congress*, pages 219–227, September 2015.

[BPW13] M. Boucadair, R. Penno, and D. Wing. Universal Plug and Play (UPnP) Internet Gateway Device - Port Control Protocol InterWorking Function (IGD-PCP IWF). RFC 6970, RFC Editor, July 2013.

[CK13] S. Cheshire and M. Krochmal. Multicast DNS. RFC 6762, RFC Editor, February 2013.

[IEE18] IEEE.org. Public listing for IEEE standards registration authority. `https://regauth.standards.ieee.org/standards-ra-web/pub/view.html#registries`, IEEE standards association, 2018.

[MAC16] I. Mashal, O. Alsaryrah, and T. Y. Chung. Analysis of recommendation algorithms for Internet of Things. In *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 181–186, April 2016.

[MMH+17] M. Miettinen, S. Marchal, I. Hafeez, T. Frassetto, N. Asokan, A. R. Sadeghi, and S. Tarkoma. IoT Sentinel demo : automated device-type identification for security enforcement in IoT. In *IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pages 2511–2514, June 2017.

[NPA17] L. Noirie, M. Le Pallec, and N. Ammar. Towards automated IoT service recommendation. In *20th Conference on Innovations in Clouds, Internet and Networks (ICIN)*, pages 103–106, March 2017.

[PNP+18] M. Le Pallec, L. Noirie, P. Peloso, D. T. Bui, and N. Le Sauze. Digital assistance for the automated discovery and deployment of IoT services. In *21st Conference on Innovations in Clouds, Internet and Networks (ICIN)*, February 2018.