Séminaire thématique "*l'Internet des Objets*"
Document de synthèse

Conseil Scientifique de l'INS2I

Groupe de travail : Véronique Cortier, Inbar Fijalkov, Philippe Lamarre, Fabrice Théoleyre (animateurs).

Avec la participation de : Mérouane Debbah (Huawei/Supelec), Andrzej Duda (LIG), Aurélien Francillon (EURECOM), Jean-Marie Gorce (CITI / INSA Lyon), l'axe rescom du GDR RSD (Eric Fleury, Thomas Noel), le pré-GDR sécurité (Marc-Olivier Killijian, Gildas Avoine), Sihem Amer-Yahia (LIG), le GDR MADICS (Christine Collet)

Executive summary : L'Internet connait actuellement une révolution, en intégrant de petits objets intelligents, capables de réagir de façon autonome, donnant naissance à un monde cyber-physique (CHPS). De par le domaine d'application, l'Internet des Objets (IoT) requiert une approche pluridisciplinaire, allant de la conception des objets matériels à la conception d'applications intelligentes autonomes tout en intégrant les propriétés de sécurité, d'efficacité énergétique et de passage à l'échelle. Pour tenter d'identifier les verrous clé de ce domaine d'application, le CSI a organisé le 6 Décembre 2016 un séminaire de réflexion autour de cette thématique, en invitant plusieurs experts autour de quelques grandes thématiques identifiées comme clés pour l'INS2I dans l'IoT.

Table des matières

1 Remarques préliminaires

Nous insistons sur le fait que cet texte sur la recherche dans le domaine de l'Internet des Objets **ne se veut pas comme exhaustif**. Il a pour but de donner des **exemples** d'équipes, thèmes, et verrous scientifiques associés au domaine de l'Internet des Objets. En particulier, certains thèmes et équipes, ne sont peut-être pas mentionnés dans ce document bien qu'ils travaillent sur ou touchent à la problématique de l'Internet des Objets.

Nous encourageons ceux que nous avons oubliés ou même dont nous avons mal décrit les travaux (erreur, incomplétude, etc.) d'utiliser le formulaire à l'adresse https://csins2i.irisa.fr/seminaire-thematique-internet-des-objets/ **pour ajouter des précisions.**

Tout document de prospective scientifique du domaine est par exemple bienvenu. De même, nous encourageons tous les lecteurs à consulter les commentaires issus de ce formulaire à la même adresse.

2 Introduction

Goldman Sachs a pronostiqué une croissance du nombre d'objets connectés à Internet sur (2014- 2017) de 60%, avec au total 28 milliards d'objets connectés [1]. Cisco prédit que le marché de l'Internet des Objets (IoT) générera un chiffre d'affaire de 14,4 milliards de dollars en 2022 [2].

Cependant, la communauté est assez divisée sur la définition exacte à donner à l'Internet des Objets. Un petit équipement, plus ou moins intelligent, se connecte de façon continue ou intermittente à Internet. Un capteur va mesurer une grandeur physique, tandis qu'un actionneur va au contraire déclencher une action physique. De façon générale, il s'agit donc de l'extension de l'Internet, permettant de rendre les mondes virtuels et physiques plus perméables (Cyber Physical Systems).

Les objets peuvent présenter une diversité de caractéristiques très importante :

1. Une TV connectée représente un objet qui va récupérer des informations d'Internet afin de les afficher pour l'utilisateur.
2. Un objet fonctionnant avec une pile va mesurer une grandeur physique (température, pression), qu'il envoie à un entrepôt de données dans Internet, où l'information sera traitée. De même, de tels objets sont particulièrement pertinents pour la maintenance préventive, permettant de prévenir l'opérateur d'une panne (chaudière, vanne).
3. Un tag RFID sert à identifier un objet en transit, permettant de le localiser dans la chaine de logistique. Cependant, il ne s'agit dans ce cas là que d'un objet passif, ne pouvant traiter de l'information.
4. Une voiture connectée permet de récupérer en temps-réel les conditions de circulation, interagit avec son environnement proche pour prévenir les accidents, utilise des données de capteurs embarqués pour faciliter la conduite.

Les actions récentes tendant à vouloir offrir une connectivité spécifique pour les objets. Ainsi, la 5G développe un ensemble de technologies offrant une connectivité sans-fil à faible débit pour des objets potentiellement mobiles, à faible consommation énergétique.

1. http://www.goldmansachs.com/our-thinking/pages/iot-infographic.html.
2. http://www.forbes.com/sites/louiscolumbus/2015/12/27/roundup-of-internet-of-things-forecasts-and-market-estimates-2015/

Les industriels montrent un intérêt croissant dans l'Internet des Objets. Ainsi, l'IoT Valley [3] a vu le jour en 2011 à Labège pour fédérer les industriels et académiques travaillant autour de cette thématique. Un deuxième site est en cours d'aménagement pour faire face à l'engouement industriel, de nombreux acteurs venant rejoindre le projet.

Proposition 1: *Le CNRS devrait étudier l'opportunité de participer activement dans l'IoT Valley. Il existe une opportunité intéressante pour favoriser à travers cette structure les collaborations entre académiques et industriels.*

La communauté se focalise depuis quelques années sur les défis suivants :

Énergie : beaucoup des objets connectés ne sont pas reliés à une source d'alimentation continue. Eventuellement, des dispositifs de récupération d'énergie sont disponibles (solaire, cinétique, etc.) mais ne couvrent qu'une partie des besoins. Une grande partie de l'énergie étant consommée par l'interface radio, il est nécessaire de concevoir des solutions économes.

Mobilité : de nombreuses objets mobiles commencent à se connecter à Internet (ex : géolocalisation de véhicules, e-sport). Ils doivent pouvoir envoyer et recevoir des informations d'Internet tout en restant éteints le plus souvent.

Densité : les usages connaissant une croissance très forte, nous attendons des densités extrêmement importantes. Il est donc nécessaire que les nouveaux réseaux prennent en charge plusieurs milliers d'équipements par centaine de mètres carrés. Bien que chaque équipement envoie peu d'information individuellement, le volume agrégé crée des défis importants.

Nous avons choisi ici de nous focaliser dans ce document sur les activités scientifiques se développant sur l'Internet des Objets plutôt que sur les usages, regroupant un très vaste ensemble de domaines (smart grids électriques, climatologie, etc.). En effet, l'IoT est dans ce cas là considéré comme un *outil* et non comme un objet d'étude à part entière. Ce document est notamment inspiré des présentations réalisées au CSI du 5 décembre 2016 par Mérouane Debbah, Andrzej Duda et Aurélien Francillon [4]. Par ailleurs, le groupe de travail a choisi de se focaliser sur les quatre aspects suivants :

Traitement du signal : l'IoT repose sur des transmission radio, bas débit, à basse consommation pour des réseaux très denses.

Réseau : la contrainte énergétique est un paradigme nouveau dans Internet, requérant à la fois des protocoles frugaux et le support d'une connexion bidirectionnelle intermittente ;

Sécurité : les équipements embarqués présentent des capacités limitées, sans IHM pour leur configuration. Il est donc important de sécuriser l'IoT afin d'en pouvoir étendre les usages sans risque sur leur utilisation quotidienne dans des tâches de plus en plus essentielles ;

Données : l'IoT génère un très grand volume de données, qui doivent être traitées et filtrées au plus près des *producteurs*.

Enfin, pour chacun de ces domaines, nous avons essayé d'identifier des thèmes en rupture dans le domaine, spécifiques à l'Internet des objets. Nous avons également tenté de dresser un panorama des outils et de la structuration nationale de la communauté française sur ces thématiques. Nous sommes restés focalisés dans ce document sur la recherche académique.

3. http://www.iot-valley.fr/fra

4. L'ensemble des présentations est disponible sur le site du CSI — https://csins2i.irisa.fr/seminaire-thematique-internet-des-objets/

3 Transmission radio

Dans cette section, nous nous focaliserons sur les aspects bas-niveau de l'Internet des Objets, ou comment envoyer (et recevoir) avec des transmissions radio des informations générées par de petits objets autonomes en énergie.

3.1 Éléments de prospective

Codage de canal : les objets envoient des paquets contenant très peu d'information (quelque fois, une seule mesure, par exemple pour de la télé-relève). Il est donc très complexe dans ce cadre de faire un codage de canal efficace avec des codes courts ;

Bande étroite avec signaux recouvrants : la densité d'accès est très importante, avec des milliers de sources potentiellement en contention. Les nouvelles solutions essaient de gérer un accès en bande étroite, avec des signaux partiellement recouvrant : il s'agit de pouvoir décoder les différents signaux correctement : une retransmission coute à la fois de la bande passante et de l'énergie. De nouvelles techniques d'accès non orthogonales restent donc à proposer ;

Accès asynchrone : ces très nombreux capteurs ne peuvent être synchronisés en temps pour accéder à la transmission, il faut chercher de nouveaux mécanismes d'accès multiple ;

Géolocalisation : la plupart des services innovants reposent sur une connaissance précise de l'objet émettant les données (tracking de flotte, etc.). Il s'agit donc d'utiliser la technologie radio de transmission pour effectuer également de la géolocalisation précise en 3D (hauteur et position géographique), en indoor ou outdoor (d'une cave à un parc) ;

Bidirectionnalité et Full-duplex : de plus en plus de noeuds combinant capteurs et actionneurs sont intégrés. Ainsi, une communication bidirectionnelle doit être établie, poussant les mesures, et récupérant les commandes à exécuter. Puisque l'équipement embarqué doit *s'endormir* (i.e. couper son interface radio), une transmission full-duplex serait particulièrement recommandée, afin de pouvoir simultanément émettre et recevoir et donc réduire les instants de réveil.

Efficacité énergétique : le problème énergétique est encore actuellement une barrière : à la fois concernant la consommation en pic, et en terme de longévité. Les techniques de machine learning sont prometteuses pour apprendre quand émettre, et comment.

Latence faible : certaines applications telles les voitures connectées demandent des délai d'accès au canal très faible, requérant des techniques non OFDM pour la transmission.

3.2 Structuration nationale

3.2.1 Plateformes

Nous avons identifié les plateformes suivantes pouvant servir d'outil de recherche à la communauté radio :

FIT Cortex-Lab (Lyon) [5] : la plateforme est la première à intégrer le software defined radio (SDR), la radio cognitive (CR) et les réseaux auto-optimisant (SON). La plateforme est ouverte en libre accès à la communauté ;

5. http://www.cortexlab.fr/

OpenAirInterface (Sophia-Antipolis)[6] : Eurecom a initié la création de cette plateforme, ayant pour but de virtualiser les plateformes de télécommunications 5G, incluant les aspects IoT.

3.2.2 GDR

Par ailleurs, les aspects *transmission* de l'IoT sont notamment traités par les GDR suivant :

ISIS : l'Internet des Objets est couvert principalement par l'axe D (Télécommunications : compression, protection, transmission), mais également les axes A (Méthodes et modèles en traitement du signal et de l'image) et C (Algorithme-architecture en traitement du signal et des images) ;

Ondes : devrait couvrir les nouvelles fréquences mmWaves.

3.2.3 Laboratoires & Équipes

Sans pouvoir être exhaustifs, nous avons identifié les laboratoires suivants travaillant dans le domaine :

CITI : l'équipe Socrate possède une expertise forte dans les communications radio et la radio logicielle pour l'IoT. Elle héberge Cortex-Lab, équipex pour les experimentations radio dans l'IoT ;

ETIS : chaire IoT avec OrangeLabs sur la dynamique temporelle (théorie des jeux), codage pour paquets courts, one-bit massive MIMO ;

LEAT : les équipes CMA et MCSOC au LEAT s'intéressent au développement d'antennes, aux protocoles basse consommation et à l'interface matérielle/logicielle pour les objets communicants.

IRIT : certains chercheurs de l'équipe réseaux ont étudiés les problématiques de l'IoT, notamment les aspects de codage de canal ;

LIG : l'équipe Polaris s'intéresse notamment aux problèmes d'optimisation des communications radio.

LSS : le groupe LANEAS possède une compétence forte en radio, MIMO, et techniques d'accès pour l'IoT ;

Télécom Lille l'équipe *signal processing* étudie les problématiques liées à la modélisation du canal et de la couche physique ;

Télécom ParisTech : l'ERC starting grant de Michèle Wigger s'intéresse notamment à la 5G et à l'optimisation des communications d'objets radio coopérant entre eux ;

4 La mise en réseau

Dans ce sous-thème, nous nous focalisons aux aspects réseaux de l'Internet des Objets : quels protocoles et algorithmes sont nécessaires pour qu'une collection d'objets plus ou moins intelligents puissent échanger des informations via une interface radio tout en optimisant la consommation d'énergie ?

6. http://www.openairinterface.org/

4.1 Éléments de prospective

Nous avons identifié les points suivants, représentant des défis de rupture par rapport aux réseaux classiques, et l'Internet en général :

Accès massif : les scénarios actuels envisagent des densités de plus en plus élevées. Les technologies longues portée telles que LoRa ou Sigfox envisagent même de couvrir des zones de plusieurs kilomètres carré. Les techniques algorithmiques traditionnelles permettaient plutôt de gérer l'accès pour un nombre faible de flots envoyant beaucoup de données.

Le paradigme dans l'IoT change : il s'agit de gérer l'accès de très petits flots (quelques paquets par heure ou par jour) en très grand nombre (quelques milliers).

Auto-configuration : un équipement connecté à l'IoT peut avoir des capacités très limitées, et ne possède souvent pas d'IHM. Une configuration manuelle n'est donc pas envisageable, du fait même de la scalabilité des solutions ;

Auto-adaptation : les scénarios d'utilisation ont tendance actuellement à se multiplier, et de nouveaux usages sont chaque jour trouvé. Il est donc inconcevable de fixer a priori les conditions d'utilisation.

La propriété de versatilité est donc primordial : il s'agit de mesurer l'environnement, en estimer les caractéristiques, et basculer sur un algorithme / protocole permettant d'avoir des performances optimales dans la situation considérée.

Garantie de performances : de plus en plus d'applications demandent à ce que des contraintes applicatives soient respectées. Le service définit par exemple le délai et le taux de perte de bout en bout acceptables, et le réseau doit respecter ces contraintes. En particulier, le réseau doit pouvoir supporter une qualité de service différenciée, ne consommant de l'énergie que pour que les contraintes soient respectées ;

Garantie de fonctionnement : certaines applications critiques doivent pouvoir garantir qu'elles fonctionnent correctement quelle que soit la situation (véhiculaire, réseaux industriels de contrôle-commande). Les techniques de type vérification formelle sont dans ce cadre nécessaires pour vérifier le bon fonctionnement des protocoles mis en place.

4.2 Structuration nationale

4.2.1 Plateformes

Nous avons identifié les plateformes suivantes pouvant servir d'outil à la communauté réseau :

FIT IoT-Lab[7] (Grenoble, Lille, Paris, Rennes, Strasbourg) : 2 700 capteurs sont répartis sur 6 sites géographiques. Un libre accès est donné aux chercheurs, afin d'évaluer *in vivo* les performances des algorithmes et protocoles qu'ils proposent. La plateforme intègre à la fois des noeuds statiques dans une large variété de topologies, et des noeuds mobiles, dont la trajectoire est contrôlable à distance. IoT-Lab ets labellisé IR ;

DOMUS[8] (LIG) : un appartement équipé permet de tester des hypothèses sur *"l'écosystème numérique et humain"*. Bien que le focus soit donné aux usages, sur le comportement humain-numérique, des aspects réseau peuvent être introduits ;

IS[9] (LORIA) : la plateforme considère de nouvelles relations entre un système informatique, son utilisateur humain et son environnement. Un des axes de recherche de la plateforme concerne les "réseaux de capteurs".

4.2.2 GDR

Les GDR suivants comprennent dans leur spectre de thèmes explorés la problématique de la mise en réseau pour l'Internet des Objets :

MACS : l'axe *Systèmes de commande et interactions* est particulièrement concerné, regardant les aspects contrôle/commande dans l'Internet des Objets ;

RSD : le pôle *Rescom* étudie les aspects réseaux, tandis que le pôle *Systèmes Distribués* se focalise plus sur les aspects intergiciels, dont l'Internet des Objets représente un des domaines applicatifs.

Une réunion inter-GDR a été organisée en 2016 entre les GDR RSD et SoC-SiP sur la thématique de l'IoT.

Proposition 2: *Le CSI propose aux GDR d'organiser régulièrement des journées thématiques inter-GDR sur le problème de l'IoT, comme l'ont fait RSD & SoC-SiP en 2016.*

4.2.3 Laboratoires & Équipes

Nous avons notamment identifié les laboratoires suivants traitant cet aspect des problèmes :

CEA-LETI s'intéresse aux petits objets communicants, avec une partie de ses chercheurs s'intéressant aux protocoles et à la conception d'objets efficients, auto-alimentés (energy harvesting) ;

CITI : l'équipe Agora étudie les réseaux urbains, incluant les aspects IoT ;

Eurecom : David Gesbert porte une ERC advanced grant qui vise à mieux utiliser les ressources "calcul et mémoire" de tout objet connecté,

ICube : l'équipe réseau pour la partie sans-fil étudie le problème de l'IoT dans son cadre protocolaire et algorithmique. Elle héberge notamment une partie de la plateforme équipex IoT-Lab.

INRIA Lille : l'équipe FUN étudie les aspects IoT, et héberge une partie de la plateforme équipex IoT-Lab. Elle possède une compétence forte notamment en RFID, et en réseaux mobiles pour l'IoT ;

INRIA Paris : l'équipe EVA s'intéresse aux réseaux sans-fil non organisés, tels que les réseaux de capteurs ;

INRIA Saclay : l'équipe INFINE regarde les nouveaux paradigmes de communication, s'intéressant notamment aux problématiques liées au content-centric networking (ou comment utiliser les données comme paradigme de communication) ;

IRISA : l'équipe OCIF (Réseaux et Architectures pour les Technologies Innovantes de Communication) s'intéresse aux problématiques réseau dans le cadre des smart grids et du transport intelligent, touchant notamment à l'IoT.

L'équipe Dionysos s'intéresse également à l'IoT, avec un focus sur les aspects évaluation de performances ;

LIG : l'équipe Drakkar explore les problématiques des réseaux sans-fil, touchant notamment à l'Internet des Objets ;

LIMOS : le thème Réseaux de Capteurs de l'axe Systèmes d'Information et de Communication se focalise sur l'IoT ;

LORIA : l'équipe Madynes s'intéresse à la configuration, la surveillance et la sécurité du nouvel Internet, s'intéressant en partie à l'IoT ;

SAMOVAR : l'équipe R3S a travaillé sur la conception de protocoles et algorithmes efficaces en énergie.

5 La sécurité pour des objets

La sécurité des objets connectés peut se tourner en partie vers des technologies existantes. Certains problèmes sont relativement simples et peuvent bénéficier de solutions existantes, notamment lorsque les objets connectés sont en fait comparables à des ordinateurs (smartphones, tablettes, ...). D'autres problèmes sont plus spécifiques et demandent des solutions adaptées à des objets qui disposent de peu de moyens de calculs et de peu de ressources en énergie. Aux défis purement technologiques s'ajoutent plusieurs difficultés plus pratiques :

— Les constructeurs d'objets connectés ne sont pas spécialistes de la sécurité. Les personnes qui déploient des objets connectés sont encore moins spécialistes et n'activent pas toujours les protections (a minima, changement des mots de passe, génération d'une clé publique, etc.)

— Les techniques utilisées sont en général couvertes par le secret industriel. Il est difficile voire impossible pour le monde académique d'auditer les solutions et de les faire progresser.

5.1 Éléments de prospective

cryptographie à basse consommation / capacité : les équipements intégrés à l'IoT possèdent des capacités très limitées. A l'inverse, les équipements mis en oeuvre pour décrypter ont des ressources importantes (ex : cluster de calcul). Comment assurer un (dé)chiffrement facile sur de petits équipements, tout en rendant le décrtyptage compliqué ?

zéroconf : un grand nombre d'objets doivent pouvoir être configurés sans intervention humaine.

analyse de vulnérabilité : chaque objet est déployé en très grand nombre, rendant une attaque *rentable* (i.e. une vulnérabilité détectée peut compromettre un grand nombre d'objets). Comment détecter ces vulnérabilités et circonscrire leur impact ?

frontière avec les SHS : comment créer de la confiance dans un environnement distribué ? Il s'agit d'intégrer des usagers qui doivent accepter d'être surveillées numériquement, tout en gardant le contrôle sur les données générées et leur usage ;

confiance : des objets doivent pouvoir créer une confiance entre eux. Ainsi, un consommateur de données doit pouvoir avoir confiance dans un producteur. Cette confiance non binaire doit se traduire dans les algorithmes de traitement des données.

5.2 Structuration nationale

5.2.1 Plateformes

Aucune plateforme n'est réellement spécifique aux aspects sécurité. Les plateformes utilisées pour valider les protocoles peuvent par exemple très bien être utilisées également pour

évaluer la sécurité (ex : sécurisation d'un protocole, implémentation d'algorithmes cryptographiques, etc.) Ainsi, la section 4.2.1 reste ici pertinente pour le domaine de la sécurité.

5.2.2 GDR

Au niveau de la communauté, le pré-GDR Sécurité a été créé en 2016 (http://gdr-securite.irisa.fr/). Un sous-groupe animé par Carlos Aguilar Melchor (http://gdr-securite.irisa.fr/gt-sri.html) se focalise notamment sur les aspects Sécurité des réseaux et des infrastructures, dont l'IoT. D'autres sous-groupe tels que *vie privée*, traitent des questions également présentent au sein des objets connectés.

5.2.3 Laboratoires & Équipes

CITI : l'équipe Privatics s'intéresse à la vie privée dans l'Internet en général, incluant les aspects IoT en partie ;

CEA-LIST s'intéresse à la sécurité, la gestion et la mise à jour de réseaux de capteurs sans-fil étendus.

CRISTAL l'équipe 2XS se focalise sur la conception de logiciels fiables ;

ENS Paris : l'équipe ISG possède une expertise forte en attaques (génération automatique d'exploits), botnets, etc. L'équipe s'est lancé depuis peu de temps dans le domaine de l'IoT ;

GREYC : l'équipe Monétique & biométrie s'intéresse à la sécurisation des transactions électroniques, rentrant ainsi dans le domaine de l'IoT (RFID) ;

IRISA : le LHS s'intéresse en particulier à la sécurité des cartes à puces, domaine pouvant être intégré dans le périmètre de l'IoT ;

L'équipe EMSEC cible les questions de recherche liées à la sécurité des systèmes informatiques et électroniques embarqué (RFID, cartes à puce, FPGA).

L'équipe CIDRE étudie les problèmes de détection d'intrusion, de préservation de la vie privée, et de la confiance ;

Heudyasic : l'équipe RO s'intéresse à la fois aux aspects infrastructure réseau et à la sécurité, intégrant notamment les aspects confiance ;

INRIA Paris : l'équipe Secret s'intéresse aux aspects cryptographie, notamment pour créer des fonctions sur du matériel à *bas-coût*.

LAAS : l'équipe TSF s'intéresse à la sécurisation d'objets intelligents (véhicules, TV, etc.)

LABRI : l'équipe MUSE (Mobilité, Ubiquité, Sécurité) s'intéresse à la sécurisation d'une flotte de drones, pouvant être considéré comme un IoT mobile ;

Labsticc : l'équipe SFIIS s'intéresse à la sécurisation, tant les aspects cloud (laaS cloud) que communications radio (e.g. Wireless HART)

LIRIS : l'équipe DRIM s'intéresse au problème de la confiance et de la préservation de la vie privée ;

LMV : l'équipe CRYPTO s'intéresse à la cryptographie sur matériel embarqué.

LORIA : l'équipe Madynes s'intéresse aux problèmes de la sécurité des protocoles de l'IoT (déni de service, homme du milieu, etc.) L'équipe possède également une expertise forte en monitoring de réseaux distribués, dont l'IoT ;

L'équipe CARAMBA regarde les aspects cryptographiques, Marine Minier ayant regardé les aspects sécurité protocolaire pour l'IoT ;

SAMOVAR : l'équipe R3S s'intéresse à la sécurité dans l'IoT, étudiant des solutions adaptées à des terminaux et environnements contraints, ainsi que la sécurité des données personnelles ;

XLIM : l'équipe Cryptis regarde la sécurité et la cryptographie, appliquée à l'IoT par certains de ses membres.

Proposition 3: *Le CSI encourage l'émergence de collaborations focalisées sur l'Internet des Objets, au sein du pre-GDR sécurité. Ce domaine recouvre plusieurs des axes du pré-GDR. Ces interactions pourraient être favorisées par exemple par des journées thématiques, la création d'un axe, etc.*

6 La gestion et le traitement des données

L'Internet des Objets génère un large volume de données qui doivent être traitées, éventuellement en temps-réel afin de pouvoir interagir avec l'environnement. La problématique recoupe partiellement les problèmes du big data, avec des contraintes supplémentaires.

6.1 Éléments de prospective

Vélocité : dans l'Internet des Objets, les données sont produites en continu, et potentiellement en très grand nombre. Étant donné un problème, il est nécessaire de déterminer rapidement les données à considérer et de les traiter pour fournir un résultat fiable en un temps acceptable. En particulier, pour certaines applications comme par exemple le contrôle commande dans l'Internet des Objets Industriels (IIoT), les temps de réponse sont bornés.

Véracité : au-delà de la diversité et de l'hétérogénéité, les données produites par l'Internet des Objets sont souvent sujettes aux imprécisions. Les causes sont nombreuses : qualité des matériels, conditions d'usage, impact de l'environnement, etc. Dans ces conditions, obtenir des résultats de qualité (si possible qualifiable ou quantifiable) à un coût (ressources, temps) raisonnable, tout en conservant la possibilité d'adresser à la fois le volume et la vélocité, nécessite des méthodes proposant un équilibre entre efficacité et exactitude.

Composante temporelle : dans l'Internet des Objets les données sont produites par des objets différents avec des fréquences indépendantes. Elles sont acquises de manière asynchrones et ne sont pas temporellement alignées. Cela complique l'obtention de résultats de qualité à partir de la composition de données issues de sources différentes.

Composante spatiale : comme pour la dimension temporelle, les données (mesures) produites par les capteurs sont également très hétérogènes au niveau spatial (granularités différentes, relations spatiales complexes) et souvent difficile à intégrer avec une interprétation précise.

Adaptabilité : Le volume, la vélocité et la volatilité des données et des sources de l'IoT rendent indispensable l'adaptabilité des solutions génériques. Les questions liés à adaptabilité sont actuellement très majoritairement abordées au niveau des traitements. Cependant,

d'autres éléments ont un impact non négligeable sur cette dimension. Par exemple, la collecte des données a un effet évident sur la complétude et le volume des données à traiter. Les caractéristiques de l'IoT peuvent permettre/nécessiter de développer des approches plus globales de l'adaptabilité.

6.2 Structuration nationale

Ce thème semble peu couvert en France, le big data pour l'IoT étant visiblement peu étudié dans ses spécificités dans nos laboratoires. Néanmoins la France acceuille des conférences du domaine comme ICLR 2017 (1000 participants dont les GAFA http://www.iclr.cc).

Une action interdisciplinaire ouverte par la MI CNRS MASTODONS en 2012, et poursuivie dans l'action EADM du GDR MADICS http://sabiod.org/EADM, fédère une trentaine de chercheurs en traitement des masses de données environnementales ((bio)acoustique, qualité air, luminosité,...).

6.2.1 Plateformes

SMIoT [10] (PROTEE & LIS Toulon) : Smart Microsystems & Internet of Things, innove depuis 2016 en acquisition et mise en réseaux, traitements embarqués grande autonomie, des masses de données multimodales synchrones et véloces. Les systèmes SMIoT acquis par le CNRS, Parc Nationaux et des ONG concernent le suivi de la faune sous-marine et terrestre joint à la pollution environnementale (son, lumière, chimie), i.e. DAQ avancé 5 voies 2mHz Fe intégrant compas, GPS, transmission radio, GSM, Wifi, accéléromètre, thermo, hygro et luxmètre longue autonomie et haute dynamique pour le suivi nocturne. Plusieurs To sont ainsi traités par semaine.

Proposition 4: *Le CSI préconise de réfléchir au problème de la reproductibilité des expérimentations. Des scénarios et jeux de données pourraient par exemple être définis afin de servir d'étalon aux solutions, et d'identifier plus précisément les cas d'usage et solutions adaptées. Ce problème de reproductibilité se retrouve dans de nombreux domaines de l'IoT, touchant à la science expérimentale.*

6.2.2 GDR

Les GDR suivants abordent de façon partielle la gestion de données dans l'IoT :

GPL : le GDR Génie Logiciel et Programmation étudie les aspects logiciels de l'IoT, notamment au travers des axes *Génie Logiciel pour les systèmes cyber-physiques*.

MADICS s'intéresse aux nouvelles méthodes et outils pour la gestion, l'exploitation et la valorisation des données scientifiques. Les données considérées sont produites en grande quantité (dont la taille est difficilement appréhendable), à un rythme intense, avec des structures variées ou hétérogènes, avec des sémantiques variables et présentant de fait une qualité et précision très peu contrôlables. Les caractéristiques de ces données scientifiques soulèvent de nombreux défis concernant leur traitement qui peut se représenter comme un continuum de tâches non indépendantes, adaptatives et performantes allant de la collecte des données à leur nettoyage, leur intégration, leur analyse et l'interprétation de ces analyses

par des experts. Les Actions mises en place dans MaDICS se focalisent sur une ou plusieurs de ces tâches et adressent un type de données scientifiques. Les reflexions en cours dans ces Actions sur les nouveaux outils algorithmiques et mathématiques (e.g. pour les très grands graphes, l'analyse de flux de données, la reproductibilité via des workflows, etc.) peuvent bénéficier amplement aux données de l'IoT.

6.2.3 Laboratoires & Équipes

Nous avons notamment identifié les laboratoires suivants traitant cet aspect des problèmes :

I3S : l'équipe SPARKS du laboratoire I3S développe des formalismes de composition logicielle pour le déploiement de comportements intelligents dans les réseaux de capteurs hétérogènes et la synthèse des masses de données collectées dans la hiérarchie du réseau. La contractualisation entre objets de l'IoT est également abordée.

INRIA Paris : l'équipe MiMove s'intéresse au middleware pour des systèmes distribués.

LIG : l'équipe SLIDE (ScaLable Information Discovery and Exploitation) développe des algorithmes pour analyser de larges volumes de données.
L'équipe HADAS s'intéresse à la gestion de masses de données (stockage, indexation à la volée, composition de flux) et aux systèmes de requêtage adaptatifs (optimisation par apprentissage, langage de requêtes hybrides, opérateurs sur des données en continu).

LIRIS : l'équipe BD (https://liris.cnrs.fr/bd/) porte des travaux sur l'intégration de données, les données capteurs dans le cadre des bâtiments et de la ville intelligente ainsi que sur le calcul distribué de requêtes continues et la détection de corrélations dans les flux de données.

LIS Toulon : DYNI (Information Dynamics http://sabiod.org/dyni) innove des algorithmes en capture et traitement de masse de données acoustiques (coll. NYU, Cornell Univ...). Elle a co-crée la plateforme SMIoT (suite à JASON, Pôle Information Numérique Prévention UTLN), et centre ses recherches en traitement du signal et apprentissage pour la compression (ondelette, DeepLearning) embarquée et en réseau de masses acoustiques véloces (ultrasoniques). Les applications couvrent le suivi de la faune (terrestre et sous-marine) joint aux pollutions audio-lumineuses, atmosphériques etc.

UVSQ / INRIA : l'équipe SMIS (https://project.inria.fr/smis/) s'intéresse au problème de traitement de gros volumes de données, en utilisant des objets contraints en termes de capacité.

LIPADE : l'équipe DiNo (http://dino.mi.parisdescartes.fr/) s'intéresse aux problèmes d'indexation et d'analyse de séries temporelles à large échelle avec des applications diverses dans la maintenance prédictive et l'astronomie.

LIP6 : l'équipe BD (http://www-bd.lip6.fr/) s'intéresse aux problèmes d'intégration de données pour l'interrogation et l'analyse sémantique de données IoT.

7 Autres Domaines

L'IoT ne se restreint pas aux 4 domaines sur lesquels nous avons souhaité nous focaliser principalement dans ce document. Nous introduisons donc ci-dessous certains domaines à la frontière pouvant intéresser l'INS2I. Nous rappelons que ce document n'a pas pour but d'être exhaustif, un référencement complet du domaine n'étant pas son objectif.

7.1 Ethique

Les discussions lors du séminaire IoT du Conseil Scientifique de l'INS2I ont permis de mettre en exergue des problèmatiques liées à l'éthique des usages. Nous avons notamment identifié :

Vie Privée : les capteurs collectent un volume très important de données, pouvant être exploitées pour en déduire des informations privées (ex : présence de personnes dans un logement pour les compteurs d'eau, identification de comportements habituels). Assurer le service demandé tout en garantissant la vie privée des usagers reste un problème ouvert ;

Santé : de plus en plus de capteurs bio-médicaux voient le jour. S'inscrivent donc les problématiques liées à la bio-éthique.

Automatisation des décisions : l'IoT peut fonctionner de façon autonome, sans humain intervenant dans la boucle de décision. Il est donc nécessaire d'introduire une formalisation des règles, incluant les critères éthiques si besoin est.

Proposition 5: *Le CSI pense qu'une action avec l'INSHS serait pertinente pour développer les aspects inter-disciplinaires liés à l'éthique dans l'IoT et ses usages. Elle pourrait prendre la forme d'une année thématique, ou d'un appel PEPS.*

Les GDR suivants peuvent s'intéresser à cette thématique :

NoST s'intéresse aux normes sciences & Techniques. Ce GDR pourrait aider à encadrer juridiquement l'usage de l'IoT.

Economie & Sociologie pourrait regarder les problématiques liées à la monétisation des données.

7.2 Capteurs matériels

Naturellement, l'Internet des Objets comprend également les dispositifs matériels le composant, chargés de générer ou traiter les données. Les premières applications intégraient des capteurs *basiques* de température, humidité, luminosité haute dynamique, acoustique infra, ultra ou sonique...

Ce thème est à l'interface au CNRS entre INS2I et INSIS. Les GDR suivant sont concernés :

Soc-SIP : les trois axes du GDR trouvent des applications dans le domaine de l'IoT (SoC, CPHS, systèmes fiables, etc.).

MADICS : l'axe EADM du GDR couvre la capture par capteurs haute dynamique et basse consommation dans le domaine de l'IoT.

Nous avons identifié les défis suivants :

7.3 Interagir avec l'IoT

L'IoT requiert de reconcevoir l'interface avec ces machines : aucun écran n'est disponible, les périphériques tels que clavier ou la souris n'existent plus. Ainsi, la sécurité ou la mise en réseau reposent sur une auto-configuration. Cependant, il est également nécessaire de reconcevoir la façon d'interagir avec cette masse d'équipements connectés.

Nous pouvons notamment identifier les défis suivants :

Reconnaissance vocale : le domaine a fait des progrès spectaculaires, et devrait continuer à se développer (ex : bâtiments intelligents) ;

Human Interface Device : des équipements spécifiques commencent à être étudiés, telles que les lentilles connectées, permettant la réalité augmentée.

Réalité augmentée : il s'agit d'ajouter l'information numérique aux images captées.

Eco-monitoring : il s'agit de proposer au citizen scientist de placer des objets IoT dans son environnement pour observer la biodiversité qui l'environne.

Les GDR suivants ont été identifiés à l'interface avec l'IoT :

I3 : le thème 5 "Interaction et coopération" pourrait s'appliquer à l'IoT.

IG-VR s'intéresse aux problématiques de la réalité augmentée ;

8 Propositions du CSI

Le caractère transverse de l'IoT rend difficile l'identification de problématiques de recherche propres à l'IoT en particulier dans les domaines de la transmission, de la sécurité ou des données. Faudrait-il de grands projets pluridisciplinaires allant des capteurs aux données ?

Proposition 6: *Le CSI a identifié un besoin d'augmenter le dialogue et les échanges entre académiques & industriels. Il préconise des réunions mixtes organisées par les GDR considérés.*

Quel que soit le domaine, l'Europe, dans le cadre des ESFRI, demande à ce que les données collectées soient librement diffusées, à des fins de vérification, d'étude, d'exploitation. Dans le cadre de l'IoT, il s'agit donc tant des données collectées sur les plateformes que sur des déploiements réels. L'opendata va plus loin que la reproductibilité scientifique.

Proposition 7: *Le CSI demande d'insister sur le problème de l'opendata dans le cadre d'expérimentations pour l'IoT. A chaque publication / projet devrait être adossé a minima un lieu public ou récupérer les données, préalablement anonymisées si la mise à disposition des données pose des questions de vie privée.*

L'Internet des Objets représente un domaine pluridisciplinaire, dont les défis scientifiques sont tirés des contraintes propres au domaine d'application. Le thème demande donc des compétences pratiques afin de pouvoir développer une recherche expérimentale.

9 Récapitulatif des préconisations

Proposition 1: *Le CNRS devrait étudier l'opportunité de participer activement dans l'IoT Valley. Il existe une opportunité intéressante pour favoriser à travers cette structure les collaborations entre académiques et industriels.*

Proposition 2: *Le CSI propose aux GDR d'organiser régulièrement des journées thématiques inter-GDR sur le problème de l'IoT, comme l'ont fait RSD & SoC-SiP en 2016.*

Proposition 3: *Le CSI encourage l'émergence de collaborations focalisées sur l'Internet des Objets, au sein du pre-GDR sécurité. Ce domaine recouvre plusieurs des axes du pré-GDR. Ces interactions pourraient être favorisées par exemple par des journées thématiques, la création d'un axe, etc.*

Proposition 4: *Le CSI préconise de réfléchir au problème de la reproductibilité des expérimentations. Des scénarios et jeux de données pourraient par exemple être définis afin de servir d'étalon aux solutions, et d'identifier plus précisément les cas d'usage et solutions adaptées. Ce problème de reproductibilité se retrouve dans de nombreux domaines de l'IoT, touchant à la science expérimentale.*

Proposition 5: *Le CSI pense qu'une action avec l'INSHS serait pertinente pour développer les aspects inter-disciplinaires liés à l'éthique dans l'IoT et ses usages. Elle pourrait prendre la forme d'une année thématique, ou d'un appel PEPS.*

Proposition 6: *Le CSI a identifié un besoin d'augmenter le dialogue et les échanges entre académiques & industriels. Il préconise des réunions mixtes organisées par les GDR considérés.*

Proposition 7: *Le CSI demande d'insister sur le problème de l'opendata dans le cadre d'expérimentations pour l'IoT. A chaque publication / projet devrait être adossé a minima un lieu public ou récupérer les données, préalablement anonymisées si la mise à disposition des données pose des questions de vie privée.*